Pigeonhole and Double Counting

Anant Asthana, Tristan Liu, Mark Takken

Spring 2022

1 Introduction

The two principles of this paper, the pigeonhole principle and double counting, are extremely simple and obvious, yet they are crucial in a variety of non-obvious and beautiful results. In this paper, we discuss miscellaneous uses of the principles, followed by applications to theoretical computer science, card tricks and games, number theory, graph theory and a remarkably useful result called Sperner's Lemma. Each section, describing one genre of applications, is approximately independent from the other sections, so the reader could read the sections he or she finds more interesting and skim or skip the other sections. We will assume readers are familiar with the definition of a limit (for the Bolzano-Weierstrass Theorem), the definition of a group (for showing elements of finite groups have finite order), the definition of modular arithmetic and Euclids Lemma (for several number theory results), and the definition of a graph (for several graph theory results). With that said, let us state the principles discussed by this paper.

Theorem 1.1 (Pigeonhole principle)**.** *If n items are put in m containers, with $n > m$, there must be a container with more than one item.*

More generally, we have the following.

Theorem 1.2 (Generalized pigeonhole principle)**.** *If $km + 1$ items are put in m containers, there must be a container with at least $k + 1$ objects.*

The infinite analogue of this principle is

Theorem 1.3 (Infinite pigeonhole principle)**.** *If infinitely many items are put in finitely many containers, there must be a container with infinitely many objects.*

This principle was first introduced by Dirichlet, and is also known as *Dirichlet's Box Principle*. It was used in the proof of Dirichlet's Approximation Theorem, which is one of many applications we will cover in this paper. Now for double counting, which is equally trivial:

Theorem 1.4 (Double counting)**.** *Suppose that we are given two finite sets R and C and a subset $S \subseteq R \times C$. Whenever $(p, q) \in S$, then we say that p and q are incident. If r_p denotes the number of elements that are incident to $p \in R$, and c_q denotes the number of elements that are incident to $q \in C$, then*

$$\sum_{p \in R} r_p = |S| = \sum_{q \in C}.$$

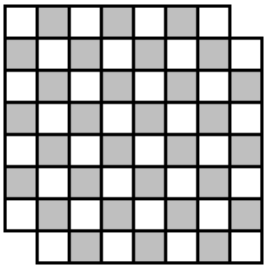
2 Introductory Examples

We show some elementary examples of the pigeonhole principle and double counting to show its versatility.

Example (Putnam Exam 2002 A2). Given any five distinct points on the surface of a sphere, show that some four of them must lie on a closed hemisphere.

Proof. Pick two of the points and consider the great circle they form on the sphere. There are three remaining points, and the circle splits the sphere into two hemispheres, so there must be one hemisphere that contains two points. Thus, four points lie on a closed hemisphere. □

Example (Mutilated Chessboard). A chessboard with two opposite corners removed cannot be tiled with dominoes. Below is an illustration of the situation [17].



Proof. The two removed corners were either both white or both black; without loss of generality, assume they were black. Then there are 32 white squares and 30 black squares remaining. A covering with 31 dominos, by the pigeonhole principle, must then have 2 white squares covered by one domino, which is impossible. □

Example. Pick any $n+1$ numbers from 1 to $2n$. There are two that are relatively prime, and there are two such that one divides the other

Proof. Consider the n pigeonholes $\{1,2\},\{3,4\},\cdots,\{2n-1,2n\}$. Any $n+1$ numbers hence must contain two numbers that are one apart from each other, and thus relatively prime.

Consider the n pigeonholes $\{1,2,4,\cdots\},\{3,6,12,\cdots\},\{5,10,20,\cdots\},\cdots,\{2n-1\}$. Any $n+1$ numbers must then contain two numbers in the same set and consequently two numbers that divide each other. □

Example. A sequence of n integers has a consecutive sum that is a multiple of n

Proof. Consider the sums $a_1, a_1+a_2, a_1+a_2+a_3, \cdots \sum_{i=1}^{n} a_i$ modulo n. If any equal 0, we are done. Otherwise, we have n sums and $n-1$ remainders, so two sums have the same remainder, say $\sum_{i=1}^{l} a_i$ and $\sum_{i=1}^{k} a_i$, where, without loss of generality, we take $l > k$. Then

$$\sum_{k+1}^{l} a_i \equiv \sum_{i=1}^{l} a_i - \sum_{i=1}^{k} a_i \equiv 0 \pmod{n}$$

as desired. □

Example. Elements of a finite group have finite order.

Proof. Let G be a finite group and let $g \in G$. Consider the sequence $e, g, g^2, \cdots$. This is an infinite sequence and there are finite elements in the group, so for some $a \neq b$ we have $g^a = g^b$. Taking inverses, we have $g^{a-b} = e$, and consequently have found a finite order for g. □

Example (Bolzano-Weierstrass Theorem). Any bounded sequence has a convergent subsequence.

Proof. Let a_n be a sequence which is contained in the interval $I_0 = [a, b]$. Pick some a_{i_1} in I_0. We will construct I_n and a_{i_n} iteratively. Let $I_{n-1} = [a_{n-1}, b_{n-1}]$. Then consider the intervals $[a_{n-1}, \frac{a_{n-1}+b_{n-1}}{2}]$ and $[\frac{a_{n-1}+b_{n-1}}{2}, b_{n-1}]$. We have finite intervals and infinite points in I_{n-1}, so one of these must contain infinitely many points. Call this interval $I_n = [a_n, b_n]$, and take some $a_{i_n} \in I_n$ with $i_n > i_{n-1}$.

We claim that the subsequence a_{i_n} converges. For any $\varepsilon > 0$, pick N such that $\frac{b-a}{2^N} < \varepsilon$. For $i, j > N$, we have that $a_i, a_j \in I_n$ and I_n has length $\frac{b-a}{2^N}$, so $|a_i - a_j| < \frac{b-a}{2^N} < \varepsilon$. Thus a_{i_n} is a Cauchy sequence and hence convergent. □

Example (Erdős-Szekeres Theorem). Every sequence $a_1, a_2, \cdots a_{mn+1}$ of distinct real numbers has an increasing subsequence of length $m + 1$ or a decreasing subsequence of length $n + 1$.

Proof. For each i, let t_i be the length of the longest increasing subsequence which starts at a_i. If $t_i \geq m + 1$ for some i, we are already done, so assume that $t_i \leq m$ for all i. We have $mn + 1$ numbers in the sequence t_i, and m possible values (from 1 to m), so by the pigeonhole principle, we must have at least $n + 1$ values of i with the same t_i. let these values of i be $i_1, i_2, \cdots, i_{n+1}$.

We claim $a_{i_1}, a_{i_2}, \cdots, a_{i_{n+1}}$ is a decreasing subsequence. If, for any j, $a_{i_j} < a_{i_{j+1}}$, we could take the longest subsequence starting at $a_{i_{j+1}}$ and append a_{i_j} to the front, which would mean $t_{i_1} \neq t_{i_2}$ which contradicts our assumption.

Thus we have found a decreasing subsequence of length $n + 1$. □

Example.

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

Proof. Consider triples of (x, y, z) where x, y, z are integers between 1 and $n + 1$ and $x, y < z$. We will count the number of such triples.

We first count by value of z. If $z = k$, there are $(k-1)^2$ ways of choosing x and y to be less than it. Summing over all values of k, we get that the number of triplets is $\sum_{k=1}^{n+1}(k-1)^2 = \sum_{k=1}^{n} k^2$. We now count by directly picking values of x, y, z.

- If $x = y$, there are two distinct values for x, y, z and they can be chosen in $\binom{n+1}{2}$ ways, each of which will correspond to a triple.
- If $x \neq y$, there are three distinct values for x, y, z and they can be chosen in $\binom{n+1}{3}$ ways, each of which will correspond to two triples, the one where $x > y$ and where $y > x$.

Thus, $\sum_{k=1}^{n} k^2 = \binom{n+1}{2} + 2\binom{n+1}{3}$, which can be expanded to give our desired equality. □

Similar, though slightly more involved arguments can give formulas for sums of other powers.

Example. If a rectangle can be tiled by rectangles all of which have at least one side of integer length, then the rectangle has one side of integer length.

Proof. Without loss of generality, the rectangle has vertices $(0,0),(a,0),(0,b),(a,b)$. Consider the set $A = \{(p,R)|\ p \text{ is an vertex of rectangle } R \text{ and } p \text{ has integer coordinates}\}$. We will count $|A|$ in two different ways.

We count by rectangles. Suppose a rectangle has an integer vertex. Then, depending on whether one or both sides of the rectangle have integer length, it will have either 2 or 4 integer vertices. Thus, counting by rectangle, we have an even number of vertex-rectangle pairs.

We now count by vertices. Every vertex that is not one of the four corner vertices is part of 2 or 4 rectangles. Thus, there must be an even number of integer vertices out of $(0,0),(a,0),(0,b)$, and (a,b). Because $(0,0)$ is an integer vertex, we must have at least one other integer vertex, so a or b must be an integer, as desired. □

3 Computer Science

3.1 Sorting

Theorem 3.1 (Non-Messing-Up Theorem)**.** *Let m,n and r be nonnegative integers, and let* $\mathbf{x} = (x_1, \ldots, x_{m+r})$ *and* $\mathbf{y} = (y_1, \ldots, y_{n+r})$ *be any sequence of numbers such that* $y_i \leq x_{m+i}$ *for each i from* 1 *to r. Then this condition continues to hold if* $\mathbf{x}$ *and* $\mathbf{y}$ *are sorted independently.*

Proof. For purposes of visualization, let us create a $2 \times (m+n+r)$ rectangle, place $\mathbf{x}$ in the first $m+r$ entries of the top row and place $\mathbf{y}$ in the last $n+r$ entries of the bottom row, like so:

x_1	$\ldots$	x_m	x_{m+1}	$\ldots$	x_{m+r}			
			y_1	$\ldots$	y_r	y_{1+r}	$\ldots$	y_{n+r}

Now, assume for the sake of contradiction that, after sorting $\mathbf{x}$ and $\mathbf{y}$, there exists a k such that $y_k > x_{m+k}$. Define $X_< := \{x_1, x_2, \ldots, x_{m+k}\}$ and $Y_> := \{y_k, y_{k+1}, \ldots, y_{n+r}\}$. If $x \in X_<$ and $y \in Y_>$, then $x \leq x_{m+k} < y_k \leq y$; that is, any element of $X_<$ is strictly less than any element of $Y_>$. Note that the union of $X_<$ and $Y_>$ has $(m+k)+(r-k+1+n) = m+n+r+1$ elements, but there are only $m+n+r$ columns in our table. Thus, by the pigeonhole principle, for any permutation of $\mathbf{x}$ and $\mathbf{y}$, there must exist a column that contains an element from $X_<$ and an element from $Y_>$, that is, there must exist a k' such that $x_{m+k'} < y_{k'}$. But this contradicts the given initial permutations of $\mathbf{x}$ and $\mathbf{y}$, which satisfied $y_i \leq x_{m+i}$ for all i from 1 to r. □

Corollary 3.2. *If, after sorting each column of a two-dimensional array, the rows are sorted, then the columns remain sorted.*

Proof. This follows immediately by applying the previous theorem to each consecutive pair of rows. □

Definition 3.3. An array is said to be *k-ordered* or *k-sorted* if each subarray with gaps of k between adjacent indices is sorted.

Example. Shellsort is an algorithm that sorts an array $\mathbf{a}$ by k-sorting $\mathbf{a}$ by insertion-sort for progressively decreasing k. The following pseudocode [15] details the algorithm, given an array to sort `a` and a decreasing gap sequence `gaps`:

```
foreach (gap in gaps) {
    for (offset = 0; offset < gap; offset++) {
        #Insertion sort of {a[offset],a[offset+gap],a[offset+2*gap],...}
        for (i = offset; i < n; i += gap) {
            temp = a[i]
            j = i
            while (j >= gap and a[j - gap] > temp) {
                a[j] = a[j - gap]
                j -= gap
            }
            a[j] = temp
        }
    }
}
```

This algorithm relies on the following crucial corollary of Theorem 3.1.

Corollary 3.4. *If a k-ordered array* $\mathbf{a}$ *is h-sorted, then it remains k-sorted.*

Proof. For each offset f from 1 to $h-1$, let $\mathbf{y} = (a_f, a_{f+h}, a_{f+2h}, \ldots, a_{f+\lfloor(|\mathbf{a}|-f)/h\rfloor\cdot h})$, $\mathbf{x} = (a_{f+k}, a_{f+h+k}, a_{f+2h+k}, \ldots, a_{f+\lfloor(|\mathbf{a}|-f-k)/h\rfloor\cdot h+k})$, $r = |\mathbf{x}|$, $m = 0$ and $n = |\mathbf{y}| - |\mathbf{x}|$. For each i from 1 to r, we have $y_i \leq x_i = x_{m+i}$ because $\mathbf{a}$ is k-sorted. By Theorem 3.1, when we sort $\mathbf{x}$ and $\mathbf{y}$ as a consequence of h-sorting $\mathbf{a}$, this inequality will still hold. Since this is true for each offset from f from 1 to $h-1$, this means that $\mathbf{a}$ remains k-ordered after being h-sorted. □

3.2 Pumping Lemma

The pigeonhole princple is key in understanding what strings can be mapped to by regular expressions, and what strings cannot. We begin with some definitions.

Definition 3.5 (Formal languages)**.** An *alphabet* is a (typically finite) set of elements called letters that is denoted as Σ. A *word* or *string* is a sequence of letters, and the set of all words over the alphabet Σ is denoted as Σ^*. A (formal) *language* L on Σ is a subset of Σ^*.

Definition 3.6. If L is a language, then L^* denotes the empty string and all finite-length strings that can be generated by concatenating arbitrary elements of L.

Definition 3.7. If w is a word in a language L, then w^n is the concatenation of w with itself n times.

Definition 3.8. The collection of *regular languages* over an alphabet Σ is defined recursively in the following manner:

- The empty language $\emptyset$ is regular.
- If a is a letter in an alphabet Σ, then the singleton language $\{a\}$ is regular.
- If A is a regular language, then A^* is regular.

- If A and B are regular languages, then $A \cup B$ (union) and $A \cdot B$ (concatenation) are regular.
- No other languages over Σ are regular.

That is, a regular language is the set of strings that are mapped to by a regular expression.

Example. The language containing the strings consisting of a sequence of 0s followed by a sequence of 1s and the strings consisting of a sequence of 1s followed by a sequence of 0s is regular. It is mapped to by the regular expression $0^*1^* + 1^*0^*$, where "+" denotes the union operator.

Example. The language with alphabet $\{0, 1\}$ consisting of strings with an even number of zeros is regular. It is mapped to by the regular expression $(1 + 01^*0)^*$ (among others).

Do non-regular languages exist? How could we show that a given language is non-regular? For this, we will need consider regular languages in the context of models of computation called *finite automata*.

Definition 3.9. A *deterministic finite automaton* (DFA) M is a five-tuple $(Q, \Sigma, \delta, q_0, F)$, where

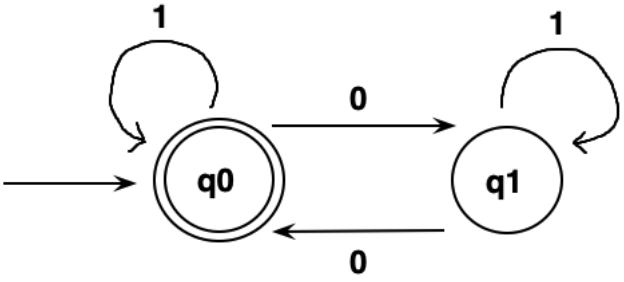
1. Q is a finite set of states,
2. Σ is an alphabet,
3. $\delta : Q \times \Sigma \to Q$ is a transition function,
4. $q_0 \in Q$ is the start state, and
5. $F \subset Q$ is the set of accept states.

We say that M *accepts* a string $w = w_1 w_2 \dots w_n$ if there exists a sequence $r_0, r_1, \dots, r_n$ in Q with

1. $r_0 = q_0$,
2. $\delta(r_i, w_{i+1}) = r_{i+1}$ for $i = 0, 1, \dots, n-1$, and
3. $r_n \in F$.

We say that M *recognizes* a language A if $A = \{w |\ M \text{ accepts } w\}$.

Example. Consider the following deterministic finite automaton.

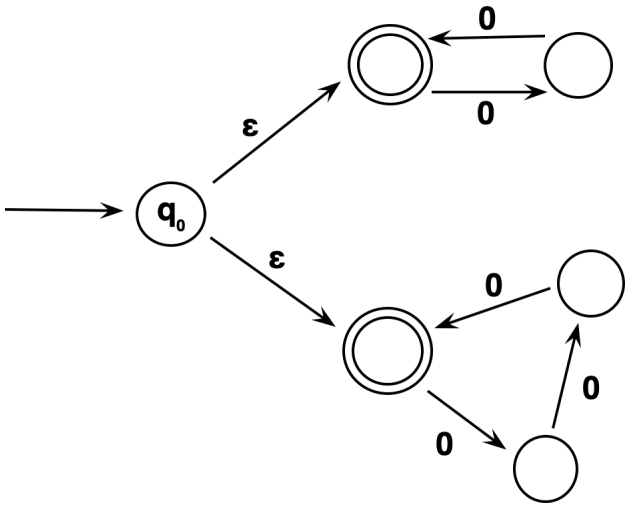


The states $Q = q_0, q_1$ are represented by circles, the alphabet is $\{0, 1\}$, the transition function is represented by the arrows, q_0 is the start state, and the accepts states $F = \{q_0\}$ are represented by double-circles. This automaton recognizes the language of strings with an even number of zeros.

Definition 3.10. A *non-deterministic finite automaton* (NFA) M is a five-tuple $(Q, \Sigma, \delta, q_0, F)$. Q, Σ, q_0 and F are as in the preceding definition, but the transition function δ is different. We now allow simultaneous transitions to, and coexistence in, multiple states. For convenience, we also allow instantaneous transitions, denoted by ε, transitions that occur immediately and independently of the string in question. Thus, δ is a function from $Q \times \Sigma_\varepsilon$, where $\Sigma_\varepsilon = \Sigma \cup \epsilon$, to the power set $\mathcal{P}(Q)$. Note that δ can map to the empty set, in which case the thread in question dies. M will then accept a string w if at least one of the parallel threads from processing w reaches an accepted state. More formally, we say that M accepts a string $w = w_1 w_2 \dots w_n$ if there exists a sequence $r_0, r_1, \dots, r_n$ in Q with

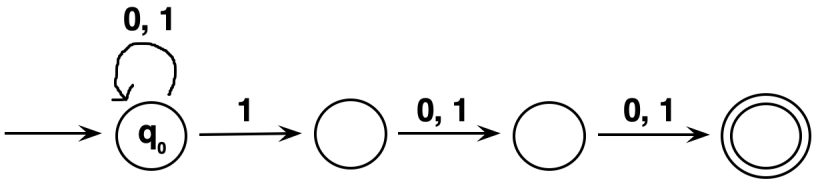
1. $r_0 = q_0$,
2. $r_{i+1} \in \delta(r_i, w_{i+1})$, and
3. $r_n \in F$.

Example. Consider the following NFA:



Once the starting thread enters the initial state q_0, two threads simultaneously and immpediately travel along the two ε edges. Then the input begins to be read. This NFA recognizes the language $L = \{0^k |\ k \text{ is a multiple of 2 or 3}\}$.

Example. Consider the following NFA:



It recognizes the language $L = \{w|$ third entry from right holds a $1\}$.

Clearly, every DFA is an NFA, since NFAs are simply a generalization of DFAs. But is the reverse true—does every NFA have an equivalent DFA representation? The answer is yes! The idea is, given an NFA, we construct a DFA whose states represent subsets of the set of states of the NFA.

Theorem 3.11. *Every NFA can be equivalently expressed as a DFA.*

Proof. Suppose we have an NFA $N = (Q, \Sigma, \delta, q_0, F)$ recognizing a language L. For any subset of states $Q_{\text{sub}} \in Q$, let $E(Q_{\text{sub}})$ denote the set of all states that can be immediately reached via 0 or more ε edges from a state in Q_{sub}. Then we construct a DFA $M = (Q', \Sigma', \delta', q_0', F')$ recognizing L, with

1. $Q' = \mathcal{P}(Q)$,
2. $\Sigma' = \Sigma$,
3. For $R \in Q'$ and $a \in \Sigma'$, define $\delta'(R, a) = \bigcup_{r \in R} \{\, q \mid q \in E(\delta(r, a)) \,\}$,
4. $q_0' = E(\{q_0\})$,
5. $F' = \{q' \in Q' \mid q' \cap F \neq \varnothing\}$.

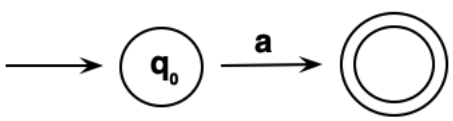
□

Thus, a language is accepted by some DFA if and only if it is accepted by some NFA. We now use what we've learned to prove a very useful property of regular languages.

Theorem 3.12 (Keene's Theorem)**.** *A language L is regular if and only if there exists a finite automaton that recognizes it.*

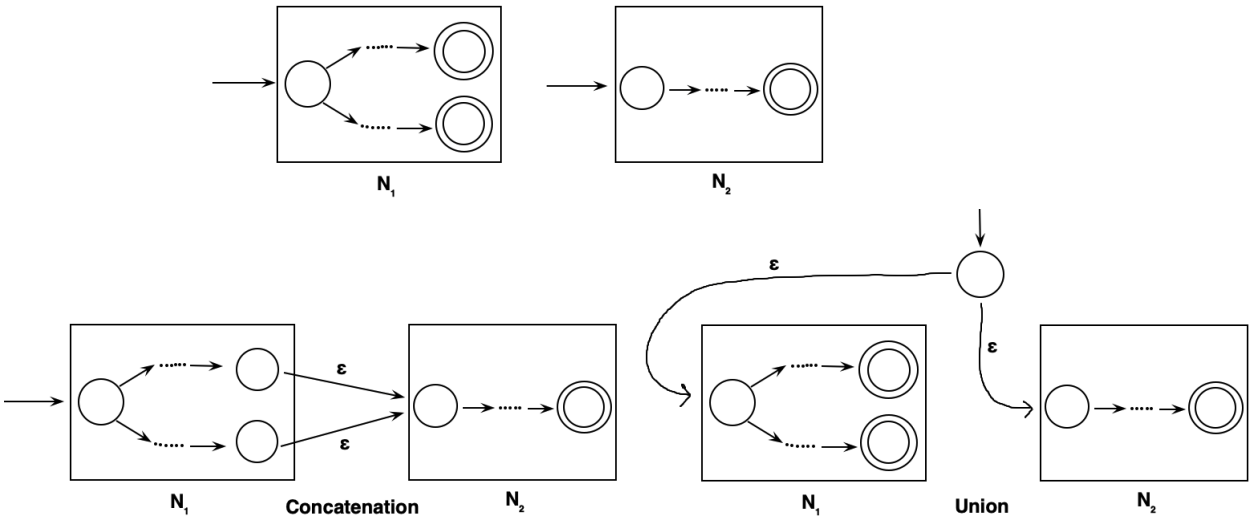
Proof. First, we show that if a language is regular, then there exists a non-deterministic (and hence also a deterministic) finite automaton that recognizes it. It suffices to verify that the set of languages recognized by an NFA satisfies the first four bullet points of Definition 3.8.

- It is easy to find an NFA that recognizes the empty language: It consists of a single node that is both starting and accepting, with no transitions.
- It is also easy to draw an NFA that recognizes the singleton language $\{a\}$, as shown below:

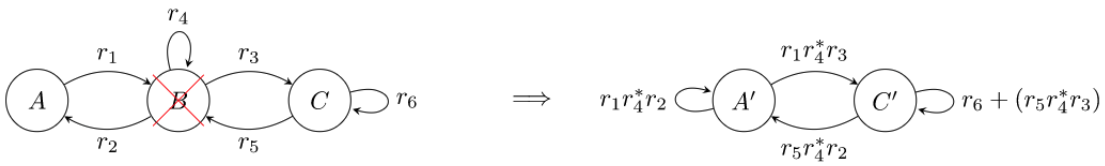


- If L is recognized by an NFA N, then an NFA recognizing L^* is formed by drawing ε-arrows from the accept states of N to the start state of N and by making the start state of N an accept state.

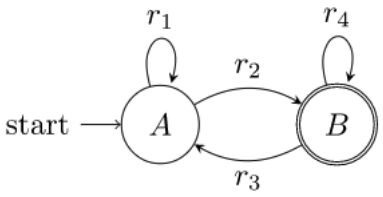
- Suppose L_1 is recognized by an NFA N_1 and L_2 is recognized by N_2. Then an NFA recognizing $L_1 \cdot L_2$ is formed by drawing ε-arrows from the accept states of N_1 to the start state of N_2, and by making the accept states of N_1 non-accept states. An NFA recognizing $L_1 \cup L_2$ is formed by drawing epsilon edges from a new start state q_0 to each of the start states of N_1 and N_2. See below for an illustration.



We now show that if a language is recognized by an NFA N, then it can be mapped to by a regular expression and is thus regular. By creating a new state and mapping all accept states to it via ε-arrows if necessary, we can assume that N has a single accept state. Now, repeatedly remove non-accept and non-start states and replace them with regular expression transitions that capture paths through the removed node. See an example below [8]:

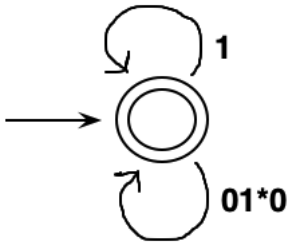


Ultimately, we are left with only the start state and the accept state, as shown below [8]:



If, at this point, the regular expressions associated with each transition are as above, then the regular expression that matches the accepted strings is $(r_1 + r_2 r_4^* r_3)^* r_2 r_4^*$. Thus, we can construct a regular expression that matches the words in any language that is recognized by an NFA, completing the proof. □

Example. Consider the DFA recognizing the language with alphabet $\{0, 1\}$ recognizing all strings with an even number of zeros, given above as the first example of a DFA. We can remove the single non-starting and non-accepting state q_1, obtaining the following diagram:

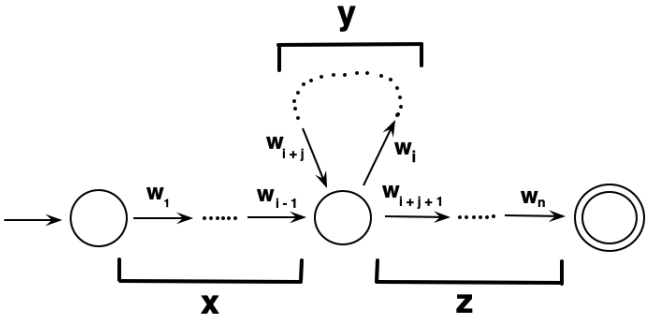


Note that we arrive at only one node because the start state is itself the accept state. This diagram in turn is equivalent to the regular expression $(1 + 01^*0)^*$.

Finally, we come to the Pumping Lemma, the main result of this section. It will allow us to prove that certain languages are not regular, that is, that they cannot be parsed by regular expressions.

Theorem 3.13 (Pumping Lemma for regular languages)**.** *For any regular language L, there exists a constant p such that any string w in L with length at least p can be expressed as a sequence of substrings $w = xyz$, with y nonempty and $|xy| \leq p$, such that xy^nz is in L for all nonnegative integers n.*

Proof. By Keene's Theorem, L is recognized by some DFA D. If L is finite, simply let p be one greater than the maximum word length, and the theorem is vacuously true. Otherwise, let p equal the number of states in D. Now, consider an arbitrary string $w = w_1w_2 \dots w_n$ of length $n \geq p$. Each character w_i is read at some non-start state q_i, and there are only $p - 1$ states other than the start state. Therefore, by the pigeonhole principle, there is a first state q_i that repeats in the processing thread: it reads a character w_i and later some character w_{i+j}. We let $x = w_1w_2 \dots w_{i-1}$, $y = w_iw_{i+1} \dots w_{i+j}$ and $z = w_{i+j+1}w_{i+j+2} \dots w_n$.



Clearly, y is nonempty because w_i and w_{i+j} are distinct. Furthermore, we have $|xy| \leq p$ because every state in the processing thread is distinct until w_{i+j}, the end of y, is read. Lastly, since reading y yields a loop starting and ending at q_i, repeating it any number of times, or omitting it, will not affect the final state reached, so the new string will also be accepted. □

Example. We show that the language $L = \{ 0^k1^k \mid k \in \mathbb{N} \}$ is not regular. Suppose the contrary. Then, by the Pumping Lemma, for each sufficiently long string, we must be able to repeat some middle section y of the string and obtain another member of the language. But clearly this is impossible: If y contains some zeros followed by some ones, then after y is repeated, the ones will alternate with the zeros more than once, producing a string that is not in the

language; and otherwise, if y consists of all zeros or all ones, the number of 0s will no longer match the number of 1s, and thus the resulting string again cannot be in the language. Thus, we have reached a contradiction, implying that the strings in L cannot be parsed by a regular expression.

4 Card Tricks and Games

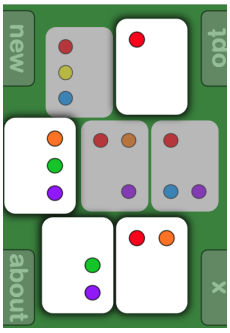
Example. Deal five hands of five cards in the standard way: Deal the first card of each hand, then the second card of each hand, etc. Sort each hand in increasing order from back to front, and assemble the hands in any way you like. Deal five hands a second time, and once again order each hand. Then, assemble the hands by placing the fifth hand on top of the four hand, those two hands on top of the third, etc. Deal five hands a third time. This time, each hand will be ordered!

This card trick works thanks to the Non-Messing-Up Theorem of the previous section. We can consider the 25 cards to be part of a 5-by-5 array. The first time we deal out the cards, each hand is a row in this array. We subsequently order each row. The second time we deal out the cards, each hand is a column in this array. We subsequently order each column. The third time we deal out the cards, each hand is once again a row. By the Non-Messing-Up Theorem, these rows remain ordered, so the each hand is already sorted.

Example. Two magicians named Moe and Maggie perform the following card trick with a standard deck. Moe draws five cards at random. He then chooses four of the five cards and places them in a pile in some order. Maggie, after taking the pile and observing the cards, declares the card that Moe left out. How does this work?

By the Pigeonhole Principle, at least two of the five cards in Moe's hand have to be of the same suit. Thus, by leaving out one of these two cards and placing the other card at the top of the pile given to Maggie, Moe can indicate the suit of the left-out card. Because there are only thirteen ranks, given the rank of one card, it must be possible to obtain the rank of the second card by adding at most six, wrapping around from K to A if necessary. Thus, Moe orders the remaining three cards in one of six ways to indicate the rank of the left-out card. The assignment of a value from one to six to each permutation can be determined via lexicographic ordering.

Example. Projective Set, also abbreviated as Pro Set, is a game implementing a deck of 63 cards, where each card may or may not contain a dot of each of six colors. Each card is distinct, and no card is blank, which is why there are $2^6 - 1 = 63$ cards in the deck. Seven of the cards are dealt onto the table. The objective of the game is to find a subset of cards such that each colored dot appears an even number of times in the subset. The image below shows an example of a deal and a corresponding subset.



But how do we know a winning subset always exists? For this, we use the pigeonhole principle. Define the *sum* of a subset of cards S to be the card c such that if c were appended to S, the result would be a winning subset. Now, there exist 2^7 subsets of the seven cards (including the empty set) but only 2^6 possible sums. Therefore, there must exist two (possibly not disjoint) subsets S_1 and S_2 that have the same sum. But that means that $(S_1/S_2) \cup (S_2/S_1)$ is a nonempty subset that has sum equal to the empty card, that is, it is a winning subset.

5 Number Theory

5.1 Modular Arithmetic

Theorem 5.1 (Chinese Remainder Theorem)**.** *Let m, n be relatively prime and let $0 \leq a < m$ and $0 \leq b < n$. Then there exists an $x < mn$ with $x \equiv a \pmod{m}$ and $x \equiv b \pmod{n}$.*

Proof. Consider the sequence $a, a+m, a+2m, \cdots, a+(n-1)m$. These terms have remainder $a \pmod{m}$ and are less than mn. We will show that one of these terms is $b \pmod{n}$, and will thus have found a value of x satisfying the desired conditions.

Consider their remainders modulo n. Suppose for sake of contradiction, none of the terms have remainder b. There are n terms in the sequence and $n-1$ remainders (excluding b), so by the pigeonhole principle, for some i, j we must have $a + im \equiv a + jm \pmod{n}$. Thus, by the definition of modular equivalence, $(a + im) - (a + jm) = (i - j)m$ is divisible by n. Euclid's Lemma then implies $n|(i - j)$, which is a contradiction, as $|i - j| < n$. Thus there must be some x with remainder $a \pmod{m}$ and $b \pmod{n}$ which is less than mn, as desired. □

Theorem 5.2 (Fermat's Little Theorem)**.** *If p is a prime and a is a positive integer, $a^p \equiv a \pmod{p}$.*

Proof. Consider the strings of length p with a letters. We count all such strings in two ways. First, there are clearly a^p strings. Second, we say that two strings are the same "necklace" if they are cyclic shifts (that is, rotations) of each other. We claim that each necklace corresponds to either 1 or p strings.

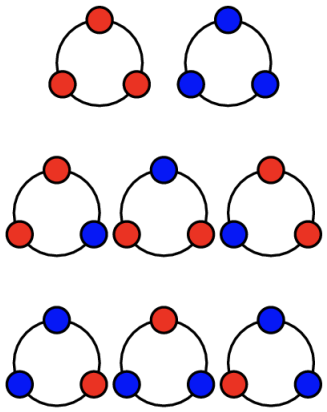


Figure 1: The geometric intuition is that we consider two strings to be the same necklace if, when we connect the ends, one can be rotated to obtain the other. This shows $p = 3$ and $a = 2$. [5]

Consider an arbitrary string and its p cyclic shifts. Suppose two of these strings are the same. Then there is a cyclic shift of m which preserves the string. Because p is prime, there exists some m^{-1} such that $mm^{-1} \equiv 1 \pmod{p}$. If we perform this shift of m a total of m^{-1} times, this preserves the string, so a shift of 1 will preserve the string. Thus, the necklace must have all identical shifts and correspond to 1 string. Otherwise, all the rotations are distinct, so that the necklace corresponds to p strings. This conclusion can be reached more easily if one assumes group theory: The orbit of a string under the group action of cyclic shifts must divide the size of the group, p, and thus must equal either 1 or p.

If a necklace corresponds to 1 string, all its rotations are identical, so it must be monochromatic. There are thus a such necklaces. We then have $a+p$(number of not monochromatic necklaces) total strings. Thus $a^p = a + p$(number of not monochromatic necklaces), so $a^p \equiv a \pmod{p}$ as desired.

□

Theorem 5.3 (Wilson's Theorem)**.** *If p is prime, $(p-1)! \equiv -1 \pmod{p}$.*

Proof. Consider the cyclic permutations of p points, which we will double-count. There are $(p-1)!$ such permutations.

As before, we can consider the cyclic shifts of these permutations, and by a similar argument, get that either all p shifts are distinct, or the permutation is fixed under the shifts. There are $p-1$ cyclic permutations which are fixed (namely the permutations which shift the points by a constant). Thus $(p-1)! = (p-1) + p \cdot$ (number of non fixed permutations), so $(p-1)! \equiv -1 \pmod{p}$ as desired. □

Theorem 5.4 (Sum of Two Squares Theorem)**.** *If a prime p is of the form $4k+1$, then there exist integers x and y such that $p = x^2 + y^2$.*

Lemma 5.5. *If p is of the form $4k+1$, then there exists a such that $a^2 \equiv -1 \pmod{p}$.*

Proof. We claim that $a = (2k)!$ works. We have

$$\begin{aligned} a^2 &\equiv (1 \cdot 2 \cdots 2k)(2k \cdots 2 \cdot 1) \\ &\equiv (1 \cdot 2 \cdots 2k)(-2k \cdots -2 \cdot -1)(-1)^{2k} \\ &\equiv (1 \cdot 2 \cdots 2k)((2k+1) \cdots (4k-1) \cdot 4k) \\ &\equiv (4k)!. \end{aligned}$$

By Wilson's Theorem, this is congruent to -1 as desired. □

Proof. Let a be as defined above, such that $a^2 \equiv -1 \pmod{p}$.

Consider the integers $ax - y$ where $0 \le x, y < \sqrt{p}$. There are $(\lceil \sqrt{p} \rceil)^2 > (\sqrt{p})^2 = p$ pairs of (x, y), and p remainders modulo p, so, by the pigeonhole principle, for some pairs (x_1, y_1) and (x_2, y_2), we have

$$ax_1 - y_1 \equiv ax_2 - y_2 \pmod{p}.$$

Regrouping terms, we have $a(x_1 - x_2) \equiv (y_1 - y_2) \pmod{p}$, which after squaring gives $-(x_1 - x_2)^2 \equiv (y_1 - y_2)^2 \pmod{p}$. We thus have

$$(x_1 - x_2)^2 + (y_1 - y_2)^2 \equiv 0 \pmod{p}.$$

By construction, $(x_1, y_1) \neq (x_2, y_2)$, so $(x_1 - x_2)^2 + (y_1 - y_2)^2 > 0$. Furthermore, $0 < x_1, x_2, y_1, y_2 < \sqrt{p}$, so $(x_1 - x_2)^2 + (y_1 - y_2)^2 < \sqrt{p}^2 + \sqrt{p}^2 = 2p$. Thus, we must have that $(x_1 - x_2)^2 + (y_1 - y_2)^2 = p$, as desired. □

We finish the modular arithmetic section with an elegant theorem: the theorem of Quadratic Reciprocity, a theorem that brought together the minds of Euler, Legendre, and Gauss, three of the greatest mathematicians of all time. To approach this problem, we must first clarify a few definitions.

Definition 5.6. Take any odd prime p. Then, for any $a \not\equiv 0 \pmod{p}$, we say a is a *quadratic residue* modulo p if there exists some $b \not\equiv 0 \pmod{p}$ for which $a \equiv b^2 \pmod{p}$. If a is not a quadratic residue, we call it a *quadratic nonresidue*.

Notice that for any two m, n such that $m^2 \equiv n^2 \pmod{p}$, we have $(m+n)(m-n) \equiv 0 \pmod{p}$. Therefore, we must have either $p|(m+n)$ or $p|(m-n)$, that is, either $m \equiv n$ or $m \equiv -n$. Thus, we have exactly $\frac{p-1}{2}$ quadratic residues, namely $1^2, 2^2, \ldots, \left(\frac{p-1}{2}\right)^2$.

We now introduce some handy notation:

Definition 5.7 (The Legendre Symbol)**.** For any $a \not\equiv 0 \pmod{p}$, the *Legendre symbol* $\left(\frac{a}{p}\right)$ is defined to equal 1 if a is a quadratic residue and -1 if a is a quadratic nonresidue.

We begin with Fermat's Little Theorem, which states that if $a \not\equiv 0 \pmod{p}$, then

$$a^{p-1} \equiv 1 \pmod{p},$$

that is, $a^{p-1} - 1 \equiv 0$. This signifies that all nonzero residues modulo p are roots of the polynomial $x^{p-1} - 1 \in \mathbb{Z}_p[x]$.

Now, notice that we can factor it into $(x^{\frac{p-1}{2}} - 1)(x^{\frac{p-1}{2}} + 1)$. Thus, all nonzero residues are roots of either $x^{\frac{p-1}{2}} - 1$ or $x^{\frac{p-1}{2}} + 1$. For any quadratic residue a, there is some $b \not\equiv 0 \pmod{p}$ such that $a \equiv b^2 \pmod{p}$, so that $a^{\frac{p-1}{2}} \equiv b^{p-1} \equiv 1 \pmod{p}$. Thus, all of the $\frac{p-1}{2}$ quadratic residues must be roots of the first factor, whereas all quadratic nonresidues are roots of the second factor. Notice that this matches precisely the definition of the Legendre symbol. In fact, we have the following, which is also known as *Euler's Criterion*:

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

Euler's Criterion directly leads to the multiplicative property:

$$\left(\frac{i}{p}\right)\left(\frac{j}{p}\right) = \left(\frac{ij}{p}\right).$$

This is very helpful, since now we don't need to calculate the Legendre symbol for every single integer; all we have to do is calculate it for ± 1, 2, and for odd primes $q \neq p$.

Euler and Legendre both were able to prove the theorem of Quadratic Reciprocity for specific cases, but it was Gauss in 1796 who finally provided a general proof. We now reach the one of the pinnacles of modular arithmetic:

Theorem 5.8 (Quadratic Reciprocity)**.** $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$

Proof. The proof we will present here was created by Ferdinand Eisenstein, which makes use of *Gauss's Lemma:*

Lemma 5.9 (Gauss)**.** *Let $a \not\equiv 0 \pmod{p}$. Consider all of the numbers in the form $1a, 2a, \ldots, \frac{p-1}{2}a$, and reduce them to some number between $-\frac{p-1}{2}$ and $\frac{p-1}{2}$, i.e. $ia \equiv r_i \pmod{p}$ where $-\frac{p-1}{2} \leq r_i \leq \frac{p-1}{2}$. Then,*

$$\left(\frac{a}{p}\right) = (-1)^s,$$

where we define $s = |\{i : r_i < 0\}|$.

To prove this, let $u_1, u_2, \ldots, u_s$ be the negative residues, and $v_1, v_2, \ldots, v_{\frac{p-1}{2}-s}$ be the positive residues. Notice that no u_j and v_k can exist such that $-u_j \equiv v_k \pmod{p}$. Consider for the sake of contradiction that such a u_j and v_k existed. Then, there must be some integers m, n such that $u_j \equiv ma \pmod{p}$ and $v_k \equiv na \pmod{p}$. This implies that $u_j + v_k \equiv ma + na \equiv (m+n)a \equiv 0 \pmod{p}$. Therefore, since $p \nmid a$, we must have $p \mid (m+n)$. However, we must have $m+n \leq p-1$ after taking m, n modulo p, leading to a contradiction.

Thus, $\{-u_1, \ldots, -u_s, v_1, \ldots, v_{\frac{p-1}{2}-s}\} = \{1, 2, \ldots, \frac{p-1}{2}\}$, so

$$\prod_j (-u_j) \prod_k v_k = \left(\frac{p-1}{2}\right)!.$$

Now, remembering that all u_j and v_k are residues of multiples of a, we can equate

$$\left(\frac{p-1}{2}\right)! \equiv (-1)^s \prod_j u_j \prod_k v_k \equiv (-1)^s \left(\frac{p-1}{2}\right)! \cdot a^{\frac{p-1}{2}} \pmod{p}.$$

Thus, since $\left(\frac{p-1}{2}\right)!$ is relatively prime to p, we can cancel it out to get

$$(-1)^s a^{\frac{p-1}{2}} \equiv 1 \pmod{p},$$

which rearranging leads to

$$(-1)^s \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

We can replace the equivalence with an equality since both sides of the congruence can only take on values of ± 1.

We now introduce the elegant yet simple logic presented by Eisenstein. Let p, q be odd primes, and consider $\left(\frac{q}{p}\right)$. Then, there exists some integer i such that qi gives a negative residue in Gauss's Lemma. This implies that there is a unique integer j for which $-\frac{p}{2} < iq - jp < 0$. Also note that because $0 < i < \frac{p}{2}$, $0 < j < \frac{q}{2}$. This means that $\left(\frac{q}{p}\right) = (-1)^s$ where s is equal to the number of lattice points (x, y) that satisfy

$$0 < x < \frac{p}{2}, 0 < y < \frac{q}{2}, 0 < py - qx < \frac{p}{2}. \tag{1}$$

In a similar manner, $\left(\frac{p}{q}\right) = (-1)^t$ where t is equal to the number of lattice points satisfying

$$0 < x < \frac{p}{2}, 0 < y < \frac{q}{2}, 0 < qx - py < \frac{q}{2}. \tag{2}$$

Consider the rectangle formed by these lattice points; it has side lengths $\frac{p}{2}$ and $\frac{q}{2}$. Draw the two lines parallel to the diagonal ($py = qx$) described by the equations $py - qx = \frac{p}{2}$ and $qx - py = \frac{q}{2}$.

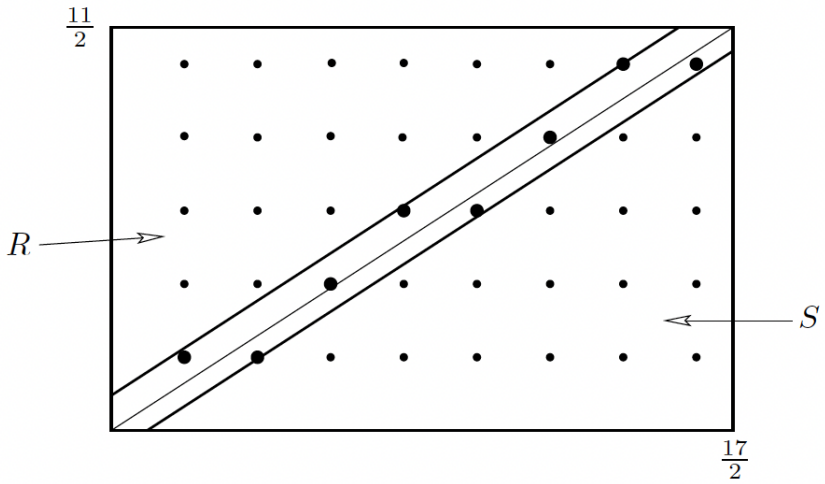


Figure 2: The lattice rectangle for $p = 17$ and $q = 11$.

To complete the proof, we only require a few more observations. The first observation is that there are no lattice points residing on the diagonal and two lines parallel to it. This is because if $py = qx$, since $p \nmid q$, we must have $p \mid x$, which is not possible by the restraints in Gauss's Lemma. For the other two parallel lines, $py - qx$ is an integer, unlike $\frac{p}{2}$ and $\frac{q}{2}$, and we cannot have an integer equal to a noninteger.

The second observation is that all lattice points satisfying the inequalities prescribed in 1 lie in the strip $0 < py - qx < \frac{p}{2}$, whereas the lattice points satisfying the inequalities described in 2 lie in the strip $0 < qx - py < \frac{q}{2}$. Therefore, the total number of lattice points in these two sections, which together make up the middle strip, is $s + t$.

The final observation is that the outer regions R, S of points satisfying $py - qx > \frac{p}{2}$ or $qx - py > \frac{q}{2}$, respectively, contain the same number of points. We can show this by constructing a bijection to take the points (x, y) in R to the point $\left(\frac{p+1}{2} - x, \frac{q+1}{2} - y\right)$ in S. It is not difficult to check that this indeed is a bijection.

Therefore, the parity of the total number of lattice points in the entire rectangle, $\frac{p-1}{2}\frac{q-1}{2}$, is the same as the parity of $s + t$. Thus, we have

$$\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{s+t} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

□

5.2 The Pell Equation

Theorem 5.10 (Dirichlet's Approximation Theorem)**.** *For any real α and positive integer N, there exists integers p, q with $1 \leq q \leq N$ such that $|q\alpha - p| < \frac{1}{N}$*

Proof. Consider the $N+1$ numbers $0, \alpha, 2\alpha, \cdots, n\alpha$ and the N intervals $[0, \frac{1}{N}), [\frac{1}{N}, \frac{2}{N}), \cdots [\frac{N-1}{N}, 1)$. By the pigeonhole principle, the fractional parts of two numbers must be in the same interval. Say those numbers are $i\alpha$ and $j\alpha$ with $i > j$. We claim $p = \lfloor i\alpha \rfloor - \lfloor j\alpha \rfloor$ and $q = i - j$ is a solution.

We can verify

$$|q\alpha - p| = |(i-j)\alpha - (\lfloor i\alpha \rfloor - \lfloor j\alpha \rfloor)| = |(i\alpha - \lfloor i\alpha \rfloor) - (j\alpha - \lfloor j\alpha \rfloor)| = |\{i\alpha\} - \{j\alpha\}| < \frac{1}{N}$$

as $\{i\alpha\}$ and $\{j\alpha\}$ are both in an interval of length $\frac{1}{N}$ (with strict inequality as the interval is half open). □

Corollary 5.11. *For any real α, there exist infinitely many fractions $\frac{p}{q}$ such that $|\alpha - \frac{p}{q}| < \frac{1}{q^2}$.*

Proof. Dirichlet's Approximation Theorem implies that for any positive integer N, there exist $p, q \leq N$ such that $|\alpha - \frac{p}{q}| < \frac{1}{Nq} < \frac{1}{q^2}$. Because for any fixed q, there exists an N such that $|\alpha - \frac{p}{q}| \geq \frac{1}{Nq}$, a finite number of p, q cannot satisfy the equality for infinitely many N. This proves the corollary. □

Theorem 5.12 (Nontrivial solution of the Pell Equation)**.** *For nonsquare integer d, $x^2 - dy^2 = 1$ has an integer solution $(x, y) \neq (\pm 1, 0)$.*

Lemma 5.13. *If $|x - y\sqrt{d}| < \frac{1}{y}$ with $x, y \in \mathbb{Z}^+$, then*

$$|x^2 - dy^2| < 1 + 2\sqrt{d}.$$

Proof. We can first bound x by

$$x \leq |x - y\sqrt{d}| + y\sqrt{d} < 1 + y\sqrt{d}.$$

Thus,

$$|x - dy^2| = (x + y\sqrt{d})|x - y\sqrt{d}| < (1 + 2y\sqrt{d})\frac{1}{y} < 1 + 2\sqrt{d}.$$

□

Proof. By our corollary, there are infinitely many pairs (x, y) satisfying $|x - y\sqrt{d}| < \frac{1}{y}$. Note that y is positive by definition, and $x > y\sqrt{d} - \frac{1}{y} > \sqrt{d} - 1\sqrt{2} - 1 > 0$ is positive. By our lemma, there are finitely many values for $x^2 - dy^2$ for these pairs (x, y) (as it must lie between $-1 - 2\sqrt{d}$ and $1 + 2\sqrt{d}$ which are constants independent of x and y). Thus, the pigeonhole principle states there must be some M which equals $x^2 - dy^2$ for infinitely many pairs (x, y). Note that M is nonzero by the irrationality of $\sqrt{d}$.

Consider these values of (x, y) modulo M. As there are finite pairs (x, y) modulo M and infinitely many (x, y) satisfying $x^2 - dy^2 = M$, by the pigeonhole principle, we must have some pair (x_1, y_1) and (x_2, y_2) that are equivalent modulo M. Say $x_1 - x_2 = Mk$ and $y_1 - y_2 = Ml$.

We then have

$$x_1 + y_1\sqrt{d} = x_2 + y_2\sqrt{d} + M(k + l\sqrt{d}).$$

Substituting $M = (x_2 + y_2\sqrt{d})(x_2 - y_2\sqrt{d})$ and factoring our $(x_2 + y_2\sqrt{d})$ gives

$$x_1 + y_1\sqrt{d} = (x_2 + y_2\sqrt{d})(1 + (x_2 - y_2\sqrt{d})(k + l\sqrt{d})).$$

Note that $(1 + (x_2 - y_2\sqrt{d})(k + l\sqrt{d})$ can be expanded to something of the form $x + y\sqrt{d}$ for some integers x, y, so we can write $x_1 + y_1\sqrt{d} = (x_2 + y_2\sqrt{d})(x + y\sqrt{d})$. Taking conjugates, we also have that $x_1 - y_1\sqrt{d} = (x_2 - y_2\sqrt{d})(x - y\sqrt{d})$. Multiplying these equations, we get $x_1^2 - dy_1^2 = (x_2^2 - dy_2^2)(x^2 - dy^2)$, or $M = M(x^2 - dy^2)$, and hence we have found values for x, y such that $x^2 - dy^2 = 1$.

It can be easily checked that $(x, y) \neq (\pm 1, 0)$, as if $(x, y) = (\pm 1, 0)$, we would have $x_1 + y_1\sqrt{d} = (x_2 + y_2\sqrt{d})(\pm 1 + 0\sqrt{d})$. Equating coefficients would give either give $(x_1, y_1) = (x_2, y_2)$ (contradicting our construction of (x_1, y_1) and (x_2, y_2) as distinct) or $x_1 = -x_2$ (contradicting how x_1 and x_2 as positive). Thus, we have found a nontrivial solution to the Pell equation.

□

6 Graph Theory

The pigeonhole principle and double counting can also be used in numerous graph theory applications. We first prove some relatively simple yet enormously consequential results:

Proposition 6.1. *Any graph G has two vertices with the same degree.*

Proof. Let V_G be the set of all vertices in graph G, and let ΔG denote the maximum degree of any vertex in V_G. Thus, we must have $\Delta G \leq |V_G| - 1$, so for all $v \in V_G$, $\deg(v)$ could take on any value ranging from 0 to $|V_G| - 1$. However, notice that we cannot have two vertices within the same graph such that one has degree $|V_G| - 1$ and the other has degree 0, since the first vertex would be connected to every other vertex. Therefore, $\deg(v)$ can either take on values from $0, \ldots, |V_G| - 2$ or $1, \ldots, |V_G| - 1$. However, in both cases, the total possible number of values of $\deg(v)$ is only $|V_G| - 1$, whereas we have $|V_G|$ vertices. Therefore, by the Pigeonhole Principle, we must have at least two vertices with the same degree. □

Proposition 6.2. $\sum_{v \in V_G} \deg(v) = 2|E_G|$, *where E_G is the set of all edges in G.*

Proof. The proof of this relies on double counting. Notice that the left-hand side counts all edges coming out of all vertices. However, this counts each edge exactly twice. To see why, consider an edge α connected to vertices v and w. On the left-hand side, α is counted twice — once for v and once for w. Thus, each edge is counted twice, leading to the desired equality. □

We now tackle a relatively difficult graph theory problem to prove:

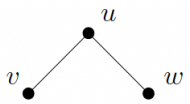
Theorem 6.3. *If G is a graph on n vertices with no 4-cycles,*

$$|E| \leq \left\lfloor \frac{n}{4}(1 + \sqrt{4n - 3}) \right\rfloor$$

Proof. We begin by defining the set S to include all pairs $(u, \{v, w\})$ such that $u, v, w \in V_G$, u is connected to both v and w, and $v \neq w$. Summing over all u, we get

$$|S| = \sum_{u \in V_G} \binom{\deg(u)}{2}.$$

In other words, we are considering all occurrences of the following configuration:



Furthermore, every possible pair of vertices $\{v, w\}$ can share at most one such u, otherwise the 4-cycle condition would be violated. Thus,

$$|S| \leq \binom{n}{2},$$

where $n = |V_G|$. Putting those together, we have

$$\sum_{u \in V_G} \binom{\deg(u)}{2} \leq \binom{n}{2}.$$

For convenience, from now on, we will shorthand $\deg(u)$ to $d(u)$. Expanding and rearranging yields

$$\sum_{u \in V_G} d(u)^2 \leq n(n-1) + \sum_{u \in V_G} d(u).$$

To continue, we apply a common technique used in extremal graph theory problems, namely the Cauchy-Schwarz inequality. Specifically, we apply it to the vectors $(d(u_1), \ldots, d(u_n))$ and $(1, \ldots, 1)$, giving us

$$\left(\sum_{u \in V_G} d(u) \right)^2 \leq n \sum_{u \in V_G} d(u)^2.$$

Combining this with the inequality earlier, we have

$$\left(\sum_{u \in V_G} d(u) \right)^2 \leq n^2(n-1) + n \sum_{u \in V_G} d(u).$$

Now, by Proposition 6.2, we have

$$4|E_G|^2 \leq n^2(n-1) + 2n|E_G|.$$

Rearranging gives

$$4|E_G|^2 - 2n|E_G| - n^2(n-1) \leq 0.$$

Upon solving, the desired inequality is produced.

□

We would like to now present an elegant double-counting proof to an intriguing theorem by James J. Sylvester, but revised in more graph theory-like terms by Arthur Cayley.

Theorem 6.4 (Cayley's Tree Theorem)**.** *There are n^{n-2} labelled trees on n vertices.*

Proof. The following argument was contrived by Jim Pitman, and can even be used to generalize Cayley's Theorem.

We must introduce the notion of a *rooted forest*:

Definition 6.5. A *rooted forest* on $\{1, 2, \ldots, n\}$ is a group of trees where each tree also has a signified root. We let $\mathcal{F}_{n,k}$ denote the set of all rooted forests with k roots.

For example, $\mathcal{F}_{n,1}$ consists of all rooted trees. Now, we consider every tree $F_{n,k} \in \mathcal{F}_{n,k}$ as a directed graph with edges directed away from the roots.

Definition 6.6. A forest F *contains* another forest G if F contains each rooted tree in G as a directed graph. Thus, if F contains G, then F has fewer or the same number of components as G does.

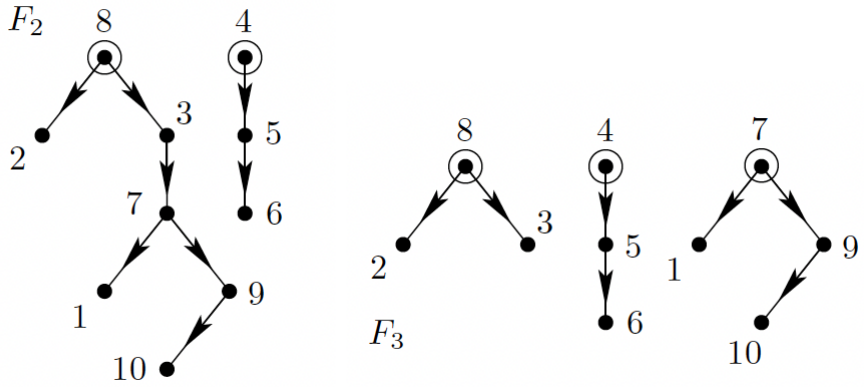


Figure 3: F_2 contains F_3, and as a result, F_2 has fewer components than F_3.

The key idea of Pitman's argument is the idea of a *refining sequence*, which is a sequence of forests $F_1, F_2, \ldots, F_k$ that satisfy $F_i \in \mathcal{F}_{n,i}$ (i.e. the i^{th} forest in the sequence has i components), and for each i, F_{i+1} is obtained from F_i by deleting an edge.

Now, for a given forest $F_k \in \mathcal{F}_{n,k}$, let $N(F_k)$ denote the number of rooted trees with n vertices containing F_k, and let $N^*(F_k)$ denote the number of refining sequences that end in F_k. We will count $N^*(F_k)$ in two different ways. The first way is by constructing the sequence starting with some tree. Let $F_1 \in \mathcal{F}_{n,1}$ such that F_1 contains F_k. There are $N(F_k)$ such starting forests F_1. Then, we obtain F_k from F_1 by deleting the edges of F_1 not found in F_k in any order to create a refining sequence. Since there are $k-1$ edges to remove, we have

$$N^*(F_k) = N(F_k) \cdot (k-1)! \,.$$

Now, we construct our sequence backwards. To create any F_{k-1} out of an F_k, we must draw a directed edge from any vertex to a root of another component tree. (An example of such a connection is shown in Figure 4.)

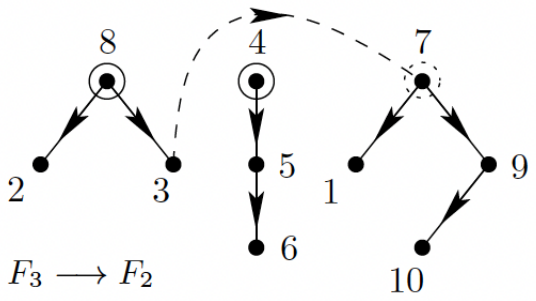


Figure 4: To get from F_3 to F_2, we can pick any vertex and draw a directed edge from that vertex to any root from another component tree. For example, we cannot connect 3 to 8.

Since there are n total vertices, and for each vertex, there are $k-1$ possible other roots to connect to, we could add a new directed edge in $n(k-1)$ ways. Then, to get from F_{k-1} to F_{k-2}, we would have to add another directed connection from any of the n vertices to another root. Thus, we have $k-2$ remaining choices for the next root to connect to, $k-3$ choices for the next, etc. Eventually, we find that for any $F_k \in \mathcal{F}_{n,k}$,

$$N^*(F_k) = n^{k-1}(k-1)!\,.$$

Equating these two expressions for $N^*(F_k)$ and cancelling out $(k-1)!$, we end up with

$$N(F_k) = n^{k-1}.$$

Notice that if we let $k = n$, then our graph consists of n disconnected roots. Then, constructing all possible refining sequences backwards from this ending point, we end up producing *all* rooted trees, so $|\mathcal{F}_{n,1}| = n^{n-1}$. Since we had n original vertices in the graph to choose as our root in the first place, we can divide both sides by n to see that

$$T_n = n^{n-2},$$

where T_n equals the number of labeled trees with n vertices. This completes the proof. □

6.1 Ramsey Theory

Ramsey Theory is generally regarded as the study of order and patterns in substructures within larger structures. One prime example of this is its application to graph theory. Consider the following question:

Question 6.7. *How many people would we need at a birthday party to guarantee that there is a group of 3 people all of whom either know each other or do not know each other?*

It turns out the answer to this question is 6. We can use Ramsey Theory to answer this question, if we rephrase in the terms of the following proposition:

Proposition 6.8. *If we color the edges of the K_6 graph red and blue, there must be a monochromatic triangle.*

Letting the 6 people be the vertices of our graph, we can color the edge between two vertices blue if the people at those vertices if they know each other and red if they do not know each other.

The generalized version of this problem is stated as Ramsey's Theorem, the centerpiece of Ramsey Theory, which is as follows:

Theorem 6.9 (Ramsey's Theorem)**.** *Given two positive integers m and n such that $m, n \geq 2$, there is a minimum positive integer, denoted by $R(m, n)$, such that in any red-blue coloring of $K_{R(m,n)}$, one can find at least one blue clique on m vertices or a red clique on n vertices.*

Thus, for the birthday party question above, we have $R(3, 3) = 6$. It is possible to extend this theorem to an infinitely large graph.

Theorem 6.10 (Infinite Ramsey Theorem)**.** *Consider an infinitely large completely connected graph G that is colored by a finite number of colors. Then there exists an infinitely large clique $\mathcal{C} \subseteq G$ such that all of the edges connecting the vertices of $\mathcal{C}$ have the same color.*

Proof. Let V_0 be the set of vertices in G. Consider one of these vertices $v_0 \in V_0$. By the pigeonhole principle, since we have an infinite number of edges coming from v_0 but only a finite number of colors, at least one of these colors must be coloring an infinite number of the edges coming from v_0. Let one of those such colors be denoted c_0. Now, let V_1 denote the vertices of G such that for all $v' \in V_1$, we have $C(v_0, v') = c_0$. Note that $V_1 \subset V_0$.

We can do the same process again on V_1, since it is also an infinite set of vertices: consider one vertex $v_1 \in V_1$, then by the pigeonhole principle, there must be at least one color, say color c_1, such that there are an infinite number of vertices v'' connected to v_1 such that $C(v_1, v'') = c_1$. Let the set of vertices connected to v_1 by an edge of color c_1 be denoted as V_2. We can repeatedly apply this construction to generate the sets $V_0, V_1, V_2, V_3, \ldots$, where we are considering one vertex $v_i \in V_i$ and then generating the infinite set V_{i+1} from all of the vertices connected to v_i by an edge of color c_i.

Now, we make some observations: for all $i \geq 0$,

1. $v_i \in V_i$,
2. $V_{i+1} \subset V_i$, and
3. $C(v_i, v) = c_i$ for all $v \in V_{i+1}$.

Now, we prove the following lemma:

Lemma 6.11. *For any integers i, j such that $0 \leq i < j$, it is true that*

$$C(v_i, v_j) = c_i.$$

Proof. By property 1, we have $v_j \in V_j$. Then, by property 2, we have $V_j \subset V_{j-1} \subseteq \cdots \subseteq V_{i+1}$, so $v_j \in V_{i+1}$. Finally, by property 3, we have $C(v_i, v_j) = c_i$. □

Revisiting the pigeonhole principle, since we have finitely many colors but infinitely many edges, at least one color, let's say c, occurs infinitely many times in G. Let our clique $\mathcal{C}$ have the set of vertices V where $V = \{v_i : i \geq 0 \text{ and } c_i = c\}$. Then we claim $\mathcal{C}$ is *the* clique we are looking for: $\mathcal{C}$ is infinite, and for any two vertices $v_i, v_j \in V$, by Lemma 4.1 we have $C(v_i, v_j) = c_i = c$. Thus, G has an infinite monochromatic clique $\mathcal{C}$. □

Ramsey Theory can also be applied to more general topics than just graphs. One common area to find patterns in substructures is sequences; thus, the theorem presented below concerns colorings of sequences:

Theorem 6.12 (Van der Waerden's Theorem)**.** *For any two given positive integers s and p, there is a minimum number n such that for any coloring of $\{1, 2, 3, \ldots, n\}$ with p colors, there is an arithmetic sequence of s numbers where all of those numbers have the same color.*

Proof. For our purposes, we will let this smallest possible number n be denoted as $W(s, p)$. Now, we define some key terms:

Definition 6.13. Suppose we have some disjoint arithmetic progressions $A_1, A_2, ..., A_l$. Then, we say the A_i are *color focused* at x if x is the $l + 1^{\text{th}}$ term for all of these sequences.

Definition 6.14. Further, suppose each sequence A_i is monochromatic and has a unique color. Then, we say these sequences are *color-focused* at x.

For this proof, we will be doing double induction, with the outer induction on s and the inner induction on p.

For the base case, it is fairly easy to show that $W(2, p) = p + 1$, since once we have $p + 1$ numbers, by the pigeonhole principle, since there are only p colors, some two of these numbers must have the same color, creating a 2-term arithmetic progression.

Now, for the inductive step, we will assume that $W(s, p)$ is finite for all p, and show that if $W(s - 1, p)$ is finite, then so is $W(s, p)$ for a fixed p. Specifically, we will show that for every $q \leq p$, there exists a number, which we will denote by $V(s, p, q)$, such that among any p-coloring of $[V(s, p, q)]$, there exists one of a

- Monochromatic arithmetic progression of length k, or
- A set of q $(s-1)$-termed color-focused monochromatic arithmetic progressions, which we will denote $A_1, A_2, \ldots, A_q$, along with their common focus.

We begin with the base case, $q = 1$. In this event, we can simply take

$$V(s, p, 1) = 2W(s - 1, p),$$

as we are guaranteed that the sth term of any $(s - 1)$-term arithmetic progression lies within the next $W(s - 1, p)$ integers.

For the inductive step, we assume that $V(s, p, q - 1)$ is finite. Then, we claim that

$$V(s, p, q) \leq 2V(s, p, q - 1)W(s - 1, p^{V(s,p,q-1)}).$$

To prove this, say we have a positive integer $n = 2V(s, p, q - 1)W(s - 1, p^{V(s,p,q-1)})$, and we are provided a p-coloring of $[n]$. Now, we divide this coloring up into $2\mathcal{W}$ blocks of size $\mathcal{V}$, where $\mathcal{W} = W(s - 1, p^{V(s,p,q-1)})$ and $\mathcal{V} = V(s, p, q - 1)$. For the outer induction on s, notice that since there are $p^{\mathcal{V}}$ ways to color each block, by the construction, there exists a sequence of $s - 1$ identically colored blocks $B_l, B_{l+m}, \ldots, B_{l+(s-2)m}$ among the first $\mathcal{W}$ blocks, whose sth term lies among one of the blocks.

Now, consider any specific block B_{l+jm} for some $0 \leq j \leq s - 2$. For the inner induction on q (and hence on p), within this block, by the inductive hypothesis, we know that there exist $q - 1$

color-focused monochromatic arithmetic progressions of length $s-1$, along with their focus. For a given color c, where $1 \leq c \leq q-1$, the arithmetic progression is

$$A_c = \{a_c + jm\mathcal{V}, a_c + d_c + jm\mathcal{V}, \ldots, a_c + (s-2)d_c + jm\mathcal{V}\},$$

and it is focused at $f + jm\mathcal{V}$. Now, we have two possibilities: if $f + jm\mathcal{V}$ is also colored with color c, then we have a monochromatic arithmetic progression of length s, and we are done. Otherwise, say the focus $(f + jm\mathcal{V})$ is colored with color c' (which is also the qth color). Then, if we let the monochromatic arithmetic progression be

$$A_i = \{a_i, a_i + (d_i + jm\mathcal{V}), a_i + w(d_i + jm\mathcal{V}), \ldots, a_i + (s-2)(d_i + jm\mathcal{V})\}$$

for all $1 \leq i \leq q-1$, and

$$A_q = \{f, f + m\mathcal{V}, f + 2m\mathcal{V}, \ldots, f + (s-2)\mathcal{V}\}$$

then $A_1, A_2, \ldots, A_q$ form a set of q $(s-1)$-term color-focused monochromatic arithmetic progressions all focused at $f + (s-1)m\mathcal{V}$ (which is indeed less than or equal to n). Thus, our inner induction on q is complete, and thus the induction on p.

To complete the proof and finish the outer induction on s, we can use a similar argument presented at the start of the proof, it follows that $W(s,p) \leq V(s,p,p)$, and thus $W(s,p)$ is finite.

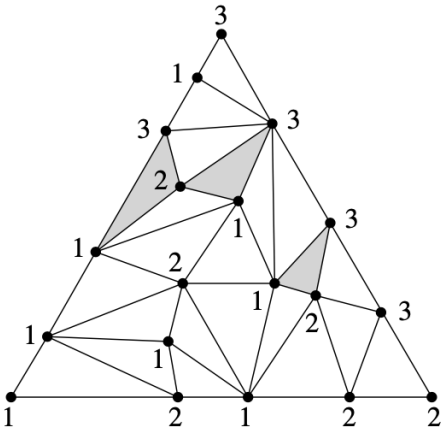
□

Notice that the color-focusing arguments presented here rely on the Pigeonhole Principle, so that we have "enough" integers to get the desired number of monochromatic arithmetic progressions.

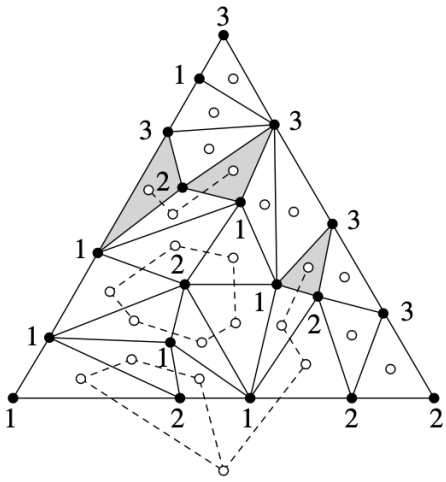
7 Sperner's Lemma

Theorem 7.1 (Sperner's Lemma)**.** *Suppose that a triangle with vertices at V_1, V_2 and V_3 is triangulated. Assume that the each vertex in the triangulation is assigned a color from the set $\{1,2,3\}$, such that each vertex V_i is colored i, and each vertex along the side connecting V_i and V_j is colored only either i or j, while each interior vertex is assigned any color arbitrarily. Then in the triangulation, there must exist a small "tricolored" triangle, which contains all three different vertex colors.*

See below an illustration of the setup with the tricolored triangles shaded in grey [11].



Proof. We will prove a stronger statement: There exists an odd number of tricolored triangles. Consider the dual graph with a vertex in each face (including the outside face) and with an edge between two dual vertices if the edge in the original graph that it crosses connects two vertices colored 1 and 2, as illustrated below [11].



Every interior dual vertex has degree at most 2, and an interior dual vertex has degree 1 if and only if it corresponds to a tricolored triangle. Therefore, equivalently, we have that an interior dual vertex has odd degree if and only if it corresponds to a tricolored triangle. Because the sum of the degrees of each vertex is twice the number of edges, as shown with a double-counting argument in the previous section, the dual graph must have an even number of vertices with odd degree. Now, the outside vertex must have odd degree, since the colors along the bottom side of the big triangle must alternate between 1 and 2 an odd number of times. Thus, there is an odd number of interior dual vertices with odd degree, which implies the desired result. □

Remark 7.2. One can easily generalize Sperner's Lemma to an arbitrary dimension via induction. In n dimensions, we have a triangulated n-dimensional simplex with vertices $V_1, V_2, \ldots, V_{n+1}$, and we color each vertex with one of the colors from $\{1, 2, \ldots, n+1\}$, with the restriction that V_i is colored i and each $(n-1)$-dimensional side of the simplex that excludes V_i cannot contain the color i. Then we wish to prove that there exists an odd number of fully colored small simplices whose vertices are colored with every color from $\{1, 2, \ldots, n+1\}$.

The argument is similar to the case of $n = 2$. First, we assume that Sperner's Lemma is true in $n-1$ dimensions. Then, given a triangulated simplex in n dimensions, we create a new dual graph with a vertex in every small simplex as well as in the space outside the large simplex, with an edge between two dual vertices if the small $(n-1)$-simplex that it crosses connects vertices colored $1, 2, \ldots, n$. Consider a small simplex with n of its vertices colored $1, 2, \ldots, n$. If the last vertex is colored $n+1$, then the small simplex is fully colored, and the degree of the corresponding dual vertex is 1. Otherwise, if the last vertex is (without loss of generality) colored 1, then there are two $(n-1)$-simplices with vertices colored $1, 2, \ldots, n$—one excluding the last vertex and one including the last vertex;—thus, the corresponding dual vertex has degree 2. All of this is to show that a dual vertex has odd degree if and only if it is located in a fully colored small simplex.

Now, by the inductive hypothesis, the dual vertex outside the large simplex has odd degree. But since every graph must have an even number of vertices with odd degree, this means that

there must be an odd number of interior dual vertices with odd degree. By the above reasoning, this implies that there exists an odd number of fully colored small simplices.

This will give us a new proof of the rectangle tiling problem.

Theorem 7.3. *If a rectangle can be tiled by rectangles all of which have at least one side of integer length, then the rectangle has one side of integer length.*

Proof. We will assume, as before, that the rectangle has vertices $(0,0),(a,0),(0,b),(a,b)$. Suppose for sake of contradiction that a and b are not integers.

We will think of our tiling as a graph, with the vertices and edges of the rectangles being the vertices and edges of the graph, respectively. Add in edges to connect a diagonal of each rectangle to make it a complete triangulation. We will color the vertices as follows. If the x-coordinate of the vertex is an integer, color it red. Otherwise, if the y-coordinate is an integer, color it blue. Otherwise, color it green. We will use Sperner's lemma by thinking of the rectangle as a triangle with vertices $(a,0),(b,0),(a,b)$, with the edge from $(a,0)$ to $(b,0)$ being made up of the edge from $(a,0)$ to $(0,0)$ with the edge from $(0,0)$ to $(0,b)$.

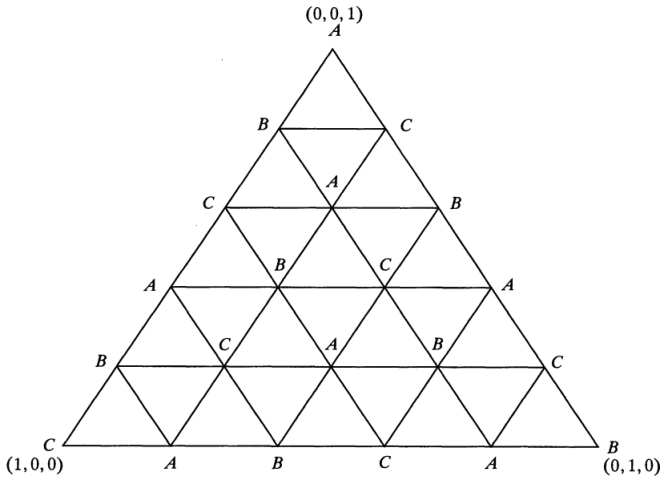
Note that, by definition, $(0,b)$ is red, $(a,0)$ is blue, and (a,b) is green. Furthermore, note that any vertex on the edges $(a,0),(0,0)$ or $(0,0)$ $(0,b)$ cannot be green as they have a coordinate 0. Also, any vertex on the edge $(a,0),(a,b)$ cannot be red as they have non-integer x-coordinate a, and similarly any vertex on $(0,b),(a,b)$ cannot be blue.

Thus, by Sperner's Lemma, we have that there exists a triangle with all three colors. Such a triangle must fall on a rectangle, which by hypothesis must have at least one integer side, so we have either a red point or a blue point that differs from a green point by an integer amount in the x or y coordinate. This is a contradiction, as any point which differs from a green point by an integer amount in the x or y coordinate must be green. Therefore, a or b must be an integer. □

Theorem 7.4 (Envy-free cake division)**.** *Given a cake and n people, there exists a division of the cake into n pieces and a bijection assigning each piece to a person such that no person prefers a piece other than his own, assuming that*

1. *Each person prefers a piece of nonzero size to an empty piece, and*
2. *Preference sets are closed. That is, any piece that is preferred in a converging sequence of cake divisions is also preferred in the limiting cake division.*

Proof. For simplicity of visualization, let us for the moment consider the case of $n = 3$. Then we can imagine cutting the cake with two parallel vertical slices, resulting in three pieces of length (along some arbitrary axis) x, y and z, with $x + y + z = 1$, where 1 is the length of the cake, without loss of generality. The graph of $x + y + z = 1$ is a triangle, which we triangulate in a regular grid pattern and assign "ownership" to each vertex such that every triangle has vertices owned by the three distinct people, as illustrated below [6].



Next, we ask the owner of each vertex which piece he would prefer if the cake were split according to the coordinates of that vertex, and we color it with the response (1, 2 or 3). Because every person prefers a piece of nonzero size to an empty piece, no vertex on side AB will be colored 1, no vertex on CA will be colored 2, and no vertex on BC will be colored 3. Thus, the coloring of the triangulated triangle satisfies the conditions of Sperner's Lemma, so that there must exist a tricolored small triangle. Such a tricolored triangle will always exist for an arbitrarily dense grid pattern in a sequence $G = G_1, G_2, G_3, \ldots$. Because the triangle is a compact set and the size of each small triangle approaches zero, there must exist a subsequence of G such that the vertices of the tricolored small triangle converge to a point. Since preference sets are closed, if we cut the cake according to the coordinates of this limit point, each person will prefer a different piece.

With the n-dimensional case of Sperner's Lemma, this argument immediately generalizes to any number of people. □

We can use Sperner's Lemma to prove the Brouwer fixed point theorem.

7.1 Brouwer fixed point theorem

Theorem 7.5 (Brouwer fixed point theorem)**.** *Let $S \subset \mathbb{R}^n$ be convex and compact, and let $T : S \to S$ be a continuous transformation. Then T has a fixed point.*

Proof. Let Δ be the simplex in $\mathbb{R}^{n+1}$ with vertices $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_{n+1} = (0, 0, \ldots, 1)$. Because Δ is homeomorphic to S, it suffices to show that any continuous transformation $T : \Delta \to \Delta$ has a fixed point.

Triangulate Δ with an arbitrarily fine mesh as in the cake-splitting scenario of the previous section. Assign to each vertex v the color $\min\{i : T(v)_i < v_i\}$, which, in other words, is the smallest index i such that the i^{th} coordinate of $T(v) - v$ is negative. If such an i does not exist, then we are done: Because the coordinates of v sum to 1, $T(v) \neq v$ necessarily implies that at least one coordinate of $T(v) - v$ is negative, and at least one is positive. We check that this coloring satisfies the conditions of Sperner's Lemma. First, the vertex e_i must receive color i, since the only possible negative component of $T(e_i) - e_i$ is the i^{th} component. Furthermore, if v lies on the $(n-1)$-simplex opposite to e_i, then $v_i = 0$, so that $T(v)_i - v_i \geq 0$, and thus v cannot receive color i.

By Sperner's Lemma, for any arbitrarily fine mesh in a sequence $G = G_1, G_2, G_3, \ldots$, there exists a small simplex whose vertices contain all $n+1$ colors. Because the large simplex is a compact set and the size of each small simplex approaches zero, there exists a subsequence of G such that the vertices of this small simplex converge to a single point. At that limit point v, because T is continuous, the i^{th} component of $T(v) - v$ is negative for every i. This is a contradiction because the components of any point in the simplex must sum to 1. Therefore, there exists a fixed point of T. □

7.1.1 Nash Equilibrium

Definition 7.6. A finite n-player strategic game has a set of players, $\{1, 2, 3, \cdots, n\}$, with each player i having a finite set S_i of pure strategies and a utility function u_i which gives the payoff for an n-tuple of pure strategies.

Definition 7.7. A mixed strategy for player i is some probability distribution s_i over S_i. The probability of picking strategy $j \in S_i$ is $s_{i,j}$. The set of all mixed strategies for player i is Δ_i. A strategy profile is an n-tuple of (mixed) strategies, one for each player. The set of all profiles is denoted Δ. We use the notation (q, s_{-i}) to represent the strategy profile we get by taking the profile s and changing player i's strategy from s_i to q, and we'll write $u(q, s_{-i})$ as shorthand for $u((q, s_{-i}))$.

Definition 7.8. We can extend the definition of a utility function to mixed-strategy profiles to be the expected utility under the probability distribution for the strategy profile, or

$$u_i(s) = \mathbb{E}_{a \sim s}(u_i(a)).$$

Definition 7.9. A Nash Equilibrium is a strategy profile $s^* = (s_1^*, s_2^*, \cdots, s_i^*, \cdots, s_n^*)$ such that for every player i,

$$u_i(s^*) \geq u_i(s_i', s_{-i}^*)$$

for any $s_i' \in \Delta_i$. In other words, the Nash equilibrium is a strategy profile where no player can alter their strategy to improve their utility.

Theorem 7.10 (Nash Equilibrium)**.** *Every n-player strategic game has a Nash equilibrium.*

Proof. For a mixed strategy profile $\sigma \in \Delta$, and a pure strategy j for player i, we define $\text{Gain}_i(s, j) = \max\{0, u_i(j, s_{-i}) - u_i(s)\}$, or the increase in payoff for player i obtained by switching to a pure strategy (or zero if there is no increase in payoff), given a profile.

We define $f : \Delta \to \Delta$ to send the mixed strategy profile $s = (s_1, s_2, \cdots, s_n)$ to the mixed strategy profile with

$$f_{i,j}(s) = \frac{s_{i,j} + \text{Gain}_i(s, j)}{1 + \sum_{k=1}^{|S_i|} \text{Gain}_i(s, k)},$$

where the denominator is a normalizing factor so that each $f_i(s)$ is a mixed strategy, and hence $f(s)$ is a mixed strategy profile. Each f_i is continuous, so f is continuous. Each Δ_i can be thought of as the $(|S_i| - 1)$-dimensional simplex $\{(x_1, x_2, \cdots, x_{|S_i|}) | \sum x_i = 1, \forall i : x_i \geq 0\}$ by taking the vector with components corresponding to the probabilities of each pure strategy, and hence is a compact and convex subset of $\mathbb{R}^{|S_i|}$. As the Cartesian product of finitely many Δ_i,

Δ is also a compact convex subset of some $\mathbb{R}^n$. Thus, by Brouwer's fixed point theorem, f has a fixed point s^*, which we claim is a Nash Equilibrium.

Suppose for sake of contradiction that for some i, $\sum_{k=1}^{|S_i|} \text{Gain}_i(s^*, k) > 0$. By definition of a fixed point, we have that for all j, $s^*_{i,j} = \frac{s^*_{i,j} + \text{Gain}_i(s^*, j)}{1 + \sum_{k=1}^{|S_i|} \text{Gain}_i(s^*, k)}$. Solving for $s^*_{i,j}$ yields

$$s^*_{i,j} = \frac{\text{Gain}_i(s^*, j)}{\sum_{k=1}^{|S_i|} \text{Gain}_i(s^*, k)}. \tag{3}$$

As a second intermediate result, we claim that for all $j \in S_i$, we have

$$s^*_{i,j}\big(u_i(j, s^*_{-i}) - u_i(s^*)\big) = s^*_{i,j}\text{Gain}_i(s^*, j). \tag{4}$$

If $\text{Gain}_i(s^*, j) > 0$, this is true by definition. If $\text{Gain}_i(s^*, j) = 0$, then $s^*_{i,j} = 0$ by (1), making the equality trivially true as both sides equal 0.

Thus, we have

$$\begin{aligned}
0 &= u_i(s^*_i, s^*_{-i}) - u_i(s) \\
&= \left(\sum_{j \in S_i} s^*_{i,j} u_i(j, s^*_{-i})\right) - u_i(s^*) && \text{by definition of } u_i \text{ for mixed strategies} \\
&= \sum_{j \in S_i} s^*_{i,j}(u_i(j, s^*_{-i}) - u_i(s^*)) && \text{because } \sum_{j \in S_i} s^*_{i,j} = 1 \\
&= \sum_{j \in S_i} s^*_{i,j}\text{Gain}_i(s^*, j) && \text{by equation (2)} \\
&= \left(\sum_{k=1}^{|S_i|} \text{Gain}_i(s^*, k)\right) \sum_{j \in S_i} (s^*_{i,j})^2 && \text{by equation (1)} \\
&> 0 && \text{because } s^*_i \text{ is nonzero,}
\end{aligned}$$

which is a contradiction. Therefore, $\sum_{k=1}^{|S_i|} \text{Gain}_i(s^*, k) = 0$ for all i. In other words, $\text{Gain}_i(s, k) = 0$ for all players i and $k \in S_i$, or, by definition of utility, $u_i(s^*) \geq u_i(k, s^*_{-i})$ for any $k \in S_i$. By the linearity of u_i, we then have $u_i(s^*) \geq u_i(s'_i, s^*_{-i})$ for any $s'_i \in \Delta_i$, which is the very definition of a Nash Equilibrium.

□

References

[1] Amites Sankar. *Van der Waerden's Theorem*. Western Washington University.

[2] Carina Letong Hong. *Proof-writing Workshop: Discrete Math, Day 2*. MIT Undergraduate Mathematics Association, 2021. http://uma.mit.edu/static/media/Slides5.e701e4e1.pdf.

[3] David Guichard. *The Pigeonhole Principle*. Mathematics LibreTexts, 2021. https://math.libretexts.org/Bookshelves/Combinatorics_and_Discrete_Mathematics/Combinatorics_and_Graph_Theory_(Guichard)/01%3A_Fundamentals/1.07%3A_The_Pigeonhole_Principle

[4] Donald Knuth. *Art of Computer Programming*, Volume 3: Sorting and Searching (Second Edition). Addison-Wesley, 1998.

[5] *Fermat's Little Theorem*. Art of Problem Solving. https://artofproblemsolving.com/wiki/index.php?title=Fermat%27s_Little_Theorem.

[6] Francis Edward Su. *Rental Harmony: Sperner's Lemma in Fair Division*. Mathematical Association of America. https://www.maa.org/sites/default/files/pdf/upload_library/22/Hasse/00029890.di011943.01p0581t.pdf.

[7] Jayadev Misra. *A Proof of Infinite Ramsey Theorem*. UT Austin Computer Science, 2012.

[8] *Lecture 27: Kleene's Theorem*. Cornell CS 2800, Spring 2017. https://www.cs.cornell.edu/courses/cs2800/2017sp/lectures/lec27-kleene.html.

[9] Ira M. Gessel. *Combinatorial Proofs of Congruences*. 2010. https://people.brandeis.edu/~gessel/homepage/slides/comb-cong.pdf.

[10] Keith Conrad. *Pell's Equation, II*. https://kconrad.math.uconn.edu/blurbs/ugradnumthy/pelleqn2.pdf.

[11] Martin Aigner, Günter M. Ziegler. *Proofs from THE BOOK* (Fifth Edition). Springer, Berlin, 2014. DOI: 10.1007/978-3-662-44205-0.

[12] Martin Gardner. *The Last Recreations: Hydras, Eggs, and Other Mathematical Mystifications*. Springer-Verlag, New York, 1997.

[13] Mia Smith. *Modelling Computation.* Course at Canada/USA Mathcamp 2020.

[14] *Nash Equilibrium*. Wikipedia. https://en.wikipedia.org/wiki/Nash_equilibrium.

[15] *Shellsort*. Wikipedia. https://en.wikipedia.org/wiki/Shellsort.

[16] Shailesh A. Shirali. *Counting Your Way to the Sum of Squares Formula*. Indian Academy of Sciences, 2015. https://www.ias.ac.in/article/fulltext/reso/020/10/0880-0892.

[17] *Tiling and Dissection*. Lecture 41, Math 348, Monash University. https://users.monash.edu/~normd/documents/MATH-348-lecture-41.pdf.

[18] *Two Proofs of Fermat's Theorem on Sums of Two Squares*. Chaitanya's Random Pages, 2011.

[19] Yishay Mansour, Nataly Sharkov, Itamar Nabriski. *Lecture 5: March 30*. Computational Game Theory, Tel Aviv University, 2003/04. http://www.math.tau.ac.il/~mansour/course_games/scribe/lecture5.pdf