The Unit Distance Graph and the Axiom of Choice	Instructor: Padraic Bartlett

Lecture 1: $\chi(\mathbb{R}^2)$ and the Axiom of Choice

Week 5 — *Mathcamp 2014*

(Primary source: "The Mathematical Coloring Book," by Alexander Soifer. It's such a good book!)

1 The Unit Distance Graph Problem

Definition. Consider the following method for turning $\mathbb{R}^2$ into a graph:

- Vertices: all points in $\mathbb{R}^2$.
- Edges: connect any two points (a, b) and (c, d) iff the distance between them is exactly 1.

This graph is called the **unit distance graph.**

Visualizing this is kinda tricky — it's got an absolutely insane number of vertices and edges. However, we can ask a question about it:

Question. How many colors do we need in order to create a proper coloring of the unit distance graph?

So: the answer isn't immediately obvious (right?) Instead, what we're going to try to do is just **bound** the possible answers, to get an idea of what the answers might be.

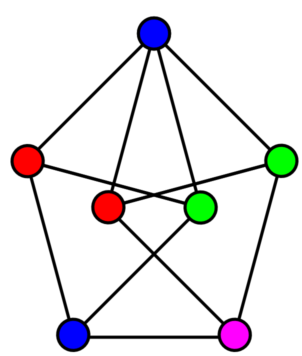
How can we even bound such a thing? Well: to get a lower bound, it suffices to consider finite graphs G that we can draw in the plane using only straight edges of length 1. Because our graph on $\mathbb{R}^2$ must contain any such graph "inside" of itself, examining these graphs will give us some easy lower bounds!

So, by examining a equilateral triangle T, which has $\chi(T) = 3$, we can see that

$$\chi(\mathbb{R}^2) \geq 3.$$

This is because it takes three colors to color an equilateral triangle's vertices in such a way that no edge has two endpoints of the same color.

Similarly, by examining the following pentagonal construction (called a Moser spindle,)



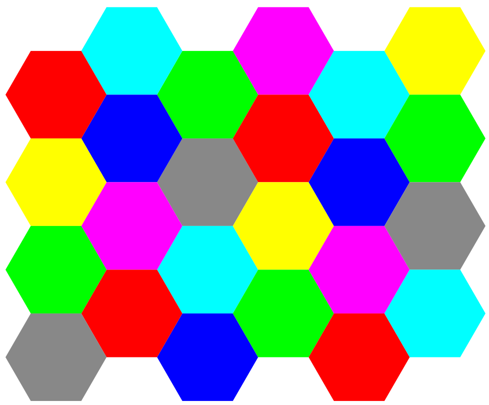
we can actually do one better and say that

$$\chi(\mathbb{R}^2) \geq 4.$$

Verify for yourself that you can't color this graph with three colors!

Conversely: to exhibit an upper bound on $\chi(\mathbb{R}^2)$ of k, it suffices to create a way of "painting" the plane with k-colors in such a way that no two points distance 1 apart get the same color.

So: consider the following way to color the plane!



To be specific: start by tiling the plane with hexagons of diameter slightly less than 1. Then, color the hexagons with seven colors as described above; i.e. repeat the color pattern

gray, red, teal, yellow, blue, green, magenta

on each strip of hexagons, shifted two colors over for each strip. This gives you a mesh of hexagons, so that any two hexagons of the same color are at least more than distance 1 apart. Therefore, any line segment of length 1 cannot bridge two different hexagons of the same color! As well, because the hexagons have diameter slightly less than one, no line segment of length 1 can lie entirely within a hexagon of the same color. Therefore, there are no line segments of length 1 with both endpoints of the same color!

In other words, we have just proven that this is a proper coloring of the plane! So we can color the plane with seven colors: i.e. we just showed that

$$\chi(\mathbb{R}^2) \leq 7.$$

These bounds on $\chi(\mathbb{R}^2)$ were not too crazy to find: it took us no more than two pages to get here, starting from the basic definition of a graph! As a result, we might hope that completely resolving this question is something we could easily finish within a few more pages.

Surprisingly: the answer is no! This problem – often called the Hadwiger-Nelson problem in graph theory literature – has withstood attacks from the best minds in combinatorics since the 1950's, and is still open to this day.

So: it's not too likely that we're going to be able to solve this problem in this class. (Try it, though!) If we were going to try, though, how would we attempt to come up with a solution?

Typically, when presented with an open or difficult problem, mathematicians rarely attempt to **directly** solve the problem; if this was likely to succeed, someone probably would have done it already! Instead, what we do is ~~try to create a related problem to the one we want to study; we either take a special case of the original problem, or remove some conditions from it, or attempt to get a weaker conclusion, or other such things~~ mess[1] with the axiom of choice!

2 The Axiom of Choice

Behold!

The Axiom of Choice : For every family Φ of nonempty sets, there is a choice function

$$f : \Phi \to \bigcup_{S \in \Phi} S,$$

such that $f(S) \in S$ for every $S \in \Phi$.

When this was first proposed as an axiom, mathematicians were opposed to it on several grounds:

- Constructivist and intutionist mathematicians opposed it, on the grounds that it posits the existence of functions without any clue whatsoever as to how to find them!
- Many other working mathematicians just thought it was a true statement; i.e. that AC was a trivial consequence of any logical framework of mathematics.

Surprisingly enough, however, Paul Cohen and Kurt Gödel proved that the axiom of choice is independent of the Zermelo-Fraenkel axioms of set theory, the current framework within which we do mathematics: i.e. that it is its own proper axiom! Pretty much all of modern mathematics accepts the Axiom of Choice; it's a pretty phenomenally useful axiom, and most fields of mathematics like to be able to call on it when pursuing nonconstructive proofs.

There are, however, a number of disconcerting "paradoxes" that arise from working within ZFC, the framework of axioms given by the Zermelo-Fraenkel axioms + the axiom of choice:

- The well-ordering principle: the statement that any set S admits a well-ordering[2] Consequently, there's a way to order the real numbers so that they "locally" look like the natural numbers! Strange.

[1] Week 5 classes are the best classes.

[2] A well-ordering on a set S is a relation $\leq$ such that the following properties hold:

- (antireflexive:) $a \leq b$ and $b \leq a$ implies that $a = b$.
- (total:) $a \leq b$ or $b \leq a$, for any $a, b \in S$.
- (transitive:) $a \leq b, b \leq c$ implies that $a \leq c$.
- (least-element:) Every nonempty subset of S has a least element.

- The Banach-Tarski paradox: there's a way to chop up and rearrange a sphere into two spheres of the same surface area.
- The existence of nonmeasurable sets: There are bounded subsets of the real line to which we cannot assign any notion of "length," given that we want length to be a translation-invariant, nontrivial, and additive function on $\mathbb{R}$.

This third concept — that of measurability and length — is one that I want to pause and go into a little further:

3 A Non-Measurable Set

Loosely speaking, a "measure" on some space X is simply a way to assign a notion of "length" to some of the subsets of that space. For example, you already know how to assign a notion of length to the interval subsets of $\mathbb{R}$: you just set the length of $[a, b]$ to be $b - a$, for any interval $[a, b]$.

Formally, a measure is just the following sort of object:

Definition. A **measure** on a set X is a function μ from some collection of subsets of X (called the "measurable subsets" of X) to $\mathbb{R} \cup \{\infty\}$, such that μ satisfies the following properties:

1. $\mu(\emptyset) = 0$.
2. For any measurable set A, $\mu(A) \geq 0$.
3. For any countable collection of disjoint measurable sets $\{A_i\}_{i=1}^{\infty}$, the set $\bigcup_{i=1}^{\infty} A_i$ is also measurable. Moreover, we have the following equality:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

This definition, however, is missing some of the aspects we really like about measures on $\mathbb{R}$. For example, we could take the "zero" measure, that assigns every set measure 0 — this satisfies our axioms, but isn't really a property that want to think length satisfies! So we have two more properties that we'd like measures on the real line $\mathbb{R}$ to satisfy:

Definition. A measure μ on the real line is called a **Lebesgue measure** if it satisfies the following additional two properties:

4. $\mu([0, 1]) = 1$.
5. μ is **translation-invariant**: in other words, given any measurable set A and real number t, the set $A_t = \{x + t \mid x \in A\}$ is measurable and has the same measure as A.

So: is there any such measure on the real line that agrees with our notion of length? The answer, as you may have guessed, is yes! It's tricky to check, but you'd proceed by proving the following:

- First, notice that 2 + 3 gives you that for any two measurable sets $A \subset B$, we have $\mu(A) \leq \mu(B)$.
- Now, use the observation above to prove that $\mu(\text{any countable set}) = 0$.
- Now, use 3+4+5+ the above to see that $\mu([a,b]) = b - a$ for any interval $[a,b]$.

From here, you might wonder how much further you can keep going: i.e. how many sets can we define our measure on? Can you make **every** set measurable?

Surprisingly enough, the answer is no! Within ZFC, we can construct sets that cannot have any notion of length in any Lebesgue measure. We sketch a proof of this here:

Theorem. There is a nonmeasurable set.

Proof. Consider the following construction: Take the interval $[0,1]$. Define an equivalence relation on $[0,1]$ as follows: set $x \sim y$ iff $x - y \in \mathbb{Q}$.

Using this equivalence relation, partition $[0,1]$ into equivalence classes $\{E_i\}_{i \in I}$, for some indexing set I. Using the Axiom of Choice, pick one element out of each set, to form some set A.

Assume for contradiction that this set A is measurable. Let A_q denote the translation of A by some number q taken mod 1 (i.e. take A, add q to every element of A, and if any elements are greater than 1 subtract 1 from them.) Then, notice that because the E_i's were equivalence classes, A_q and A_r are disjoint sets for any $q, r \in \mathbb{Q} \cap [0,1]$.

As well, notice that if we take the collection $\{q_i\}_{i=1}^{\infty}$ of all rational numbers in $[0,1]$, we have

$$\bigcup_{i=1}^{\infty} A_{q_i} = [0,1].$$

Consequently, we have

$$1 = \mu([0,1]) = \mu\left(\bigcup_{i=1}^{\infty} A_{q_i} = [0,1]\right) = \sum_{i=1}^{\infty} \mu(A_{q_i}).$$

But all of the A_{q_i}'s are just translations of each other, up to wrapping around $[0,1]$ (which is not an issue, given our translation/union properties.) Therefore they all have the same measure: consequently, our sum on the right must be either infinity or 0, and in particular not 1. This is a contradiction, which completes our proof! $\square$

4 Solovay's System

Motivated by these strange results, Solovay (a set theorist) introduced the following two axioms:

- ($\mathrm{AC}_{\aleph_0}$, the countable axiom of choice): For every **countable** family Φ of nonempty sets, there is a choice function

  $$f : \Phi \to \bigcup_{S \in \Phi} S,$$

  such that $f(S) \in S$ for every $S \in \Phi$.

- (LM, Lebesgue-measurability): Every bounded set in $\mathbb{R}$ is measurable.

Theorem 1 (Solovay's Theorem)**.** *There are models of mathematics in which ZF + LM + $AC_{\aleph_0}$ all hold.*

For brevity's sake, we will denote ZF + the axiom of choice by ZFC, and ZF + LM + $\mathrm{AC}_{\aleph_0}$ by ZFS.

5 $\chi(\mathbb{R}^2)$ in ZFS

This discussion provokes a fairly natural question for this class: does $\chi(\mathbb{R}^2)$ depend on the axiom of choice? In other words, is $\chi^{ZFC}(\mathbb{R}^2)$ different from $\chi^{ZFS}(\mathbb{R}^2)$?

Well: as we currently don't know what $\chi^{ZFC}(\mathbb{R}^2)$ even *is,* answering this question completely seems to be a bit beyond our reach. However, the following example suggests that something weird may indeed happen here:

Theorem 2. *Let G be the graph defined as follows:*

- $V(G) = \mathbb{R}$,
- $E(G) = \{(s,t) : s - t - \sqrt{2} \in \mathbb{Q}\}$.

Then $\chi^{ZFC}(G) = 2$.

Proof. Let

$$S = \{q + n\sqrt{2} | q \in \mathbb{Q}, n \in \mathbb{Z}\}.$$

Define an equivalence relation $\sim$ on $\mathbb{R}$ as follows: $x \sim y$ iff $x - y \in S$. Let $\{E_i\}_{i\in I}$ be the collection of all of the equivalence classes of $\mathbb{R}$ under $\sim$. Using the axiom of choice, pick one element y_i from each set E_i, and collect all of these elements in a single set E. Define the function $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) = \text{ the unique element } y_i \text{ in } E \text{ such that } x \sim y_i.$$

Now define a two-coloring of $\mathbb{R}$ as follows: for any $x \in \mathbb{R}$, color x 1 iff there is an odd integer n such that

$$x - f(x) - n\sqrt{2} \in \mathbb{Q};$$

similarly, color x 2 iff there is an even integer n such that

$$x - f(x) - n\sqrt{2} \in \mathbb{Q}.$$

By construction, we know that $x \sim f(x)$; so $x - f(x)$ is always of the form $q + n\sqrt{2}$, and thus we always have exactly one of the two possibilities above holding. As well, if we examine any edge $\{x, y\}$, we have to have $x - y = q + \sqrt{2}$, for some q; i.e. $x \sim y$! So $f(x) = f(y)$, and thus we have that

$$\begin{aligned} & x - y = q + \sqrt{2} \\ \Rightarrow & (x - f(x)) + (y - f(y)) = q + \sqrt{2}; \end{aligned}$$

consequently, if both $x - f(x) - n\sqrt{2}$ and $y - f(y) - m\sqrt{2} \in \mathbb{Q}$, we must have one of n, m be odd and the other be even. $\square$

Theorem 3. *For G as above, $\chi^{ZFS}(G) > \aleph_0$.*

Proof. Consider the following lemma:

Lemma 4. *If $A \subset [0,1]$ and A doesn't contain a pair of adjacent vertices in G, then A has measure*[3] *0.*

Proof. So: consider the following rather large hammer from analysis, which we will use without proof:

Theorem 5. *(Lebesgue Density Theorem) If a set A has nonzero measure, then there is an interval I such that*

$$\frac{\mu(A \cap I)}{\mu(I)} \geq 1 - \epsilon,$$

for any $\epsilon > 0$.

So: choose any set A of measure > 0, and pick I such that

$$\frac{\mu(A \cap I)}{\mu(I)} \geq 99/100,$$

for instance. Then, pick $q \in \mathbb{Q}$ such that $\sqrt{2} < q < \sqrt{2} + \mu(I)/100$, and define $B = \{x - q + \sqrt{2} : x \in A\}$. Then B has been translated by at most 1/100-th of the length of I: so we have that

$$\frac{\mu(B \cap I)}{\mu(I)} \geq 98/100.$$

So, because $(A \cap I) \cup (B \cap I) \subset I$, and both of these sets are almost all of I, we know that they must overlap! In other words, there's an element y in both A and B – but this means that there's an element y in A such that $y = x - q + \sqrt{2}$, with x *also* in A! i.e. there's a pair of elements x, y in A with an edge between them! $\square$

So: with this, our proof is pretty straightforward. Suppose that we could color $\mathbb{R}$ with $\aleph_0$-many colors, and that the collection of colors used is given by the collection $\{A_i\}_{i=1}^{\infty}$. Let $B_i = A_i \cap [0,1]$; then we have that all of the B_i are disjoint and $\bigcup B_i = [0,1]$. Consequently, we have that $\sum \mu(B_i) = \mu([0,1]) = 1$; so at least one of the B_i's have to have nonzero measure! This contradicts our above lemma; consequently, no such $\aleph_0$-coloring can exist. $\square$

[3] The measure of a set S is defined as the infimum of the sum $\sum(b_i, a_i)$, where we range over all collections of intervals $\{(a_i, b_i)\}$ such that $\bigcup(a_i, b_i) \supset S$. We denote this number by writing $\mu(S)$