Monsky's theorem and p-adic valuations

Mathcamp 2025

Glenn Sun

These notes may contain errors,
and should not be considered an authoritative source.

Contents

1 Introduction **2**

2 Coloring the unit square **3**

2.1 p-adic valuations 3

2.2 Computing the area 7

3 Finding a trichromatic triangle **11**

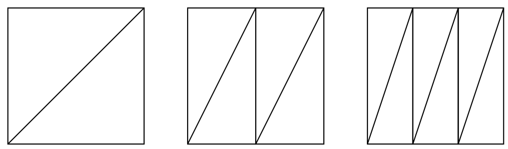
4 Extending to real-valued points **15**

1 Introduction

Every month or so, the MAA (the same organization that writes the AMC, AIME, and Putnam competitions) publishes the *American Mathematical Monthly*, which contains expository and accessible articles, notes, and open questions for its readers. In the March issue of 1967, Fred Richman and John Thomas asked:

> "Let N be an odd integer. Can a rectangle be dissected into N nonoverlapping triangles, all having the same area?"

It's easy to divide a rectangle into N triangles of the same area, when N is even:



But they couldn't find a way to do this with an odd number of triangles.

In the February issue of 1970, Paul Monsky delivered his proof that such a task was impossible, using just a little more than 2 pages. In these notes, we'll describe all of the background you need to understand his proof. The proof has two key components:

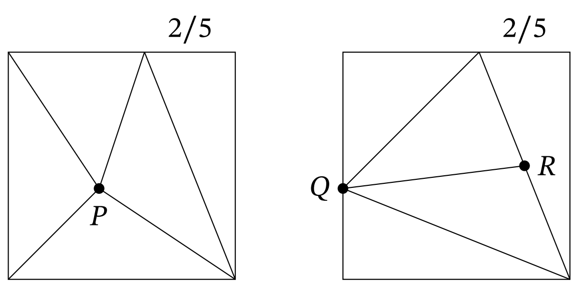
1. Color the points of the unit square with 3 colors, such that any triangle with 3 differently-colored vertices has area with even denominator in simplest terms.
2. Prove that whenever the unit square is divided into an odd number of triangles, at least one of the triangles must have 3 differently-colored vertices.

In what follows, we will prove a simpler version of (1) first, in which we will only color the points of the unit square with rational coordinates (so no coordinates with $1/\sqrt{2}$ or $\ln(2)$, say). Then we'll prove (2), from which we conclude that there is no way to divide a square into an odd number of triangles when the triangles have rational coordinates. Finally, we will sketch how to extend (1) to all real numbers.

Practice

Do these problems if you want to reinforce ideas and motivations from the main lesson.

1. You attempt to dissect the unit square into 5 pieces of equal area in two ways:



   In each diagram, the triangle on the right has area $1/5$, and the remaining 4 triangles have areas that depend on the labeled points. For each diagram, prove that no choices for the labeled points, with edges as depicted in the picture, results in every triangle having area $1/5$.

2 Coloring the unit square

2.1 p-adic valuations

Our goal is to color the unit square so that something has area with even denominator. The following concept from algebraic number theory will help us talk about this much more easily:

Definition 2.1 (p-adic valuation)**.**

- Let $n \in \mathbb{N}$ be a positive integer and let p be a prime number. Let the prime factorization of n look like $n = p^k q$ where q is not divisible by p. Then the p-adic valuation of n is
$$v_p(n) = k.$$
- Let $n \in \mathbb{Z}$ be an integer. Then for $n \neq 0$, we define
$$v_p(n) = v_p(|n|),$$
and $v_p(0) = \infty$.
- Let $a/b \in \mathbb{Q}$ be a rational number with $a, b \in \mathbb{Z}$. Then we define
$$v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b).$$ ⌟

In other words, a p-adic valuation of a number says how many times p is a factor, which may be negative if it has p in the denominator. So $v_2(100) = 2$ and $v_2(3/8) = -3$. A rational number $x \in \mathbb{Q}$ has even denominator when written in lowest terms if and only if $v_2(x) \leq -1$.

Introducing this notation allows us to prove properties that help us calculate whether a number has odd or even denominator. In particular, we have the following:

Proposition 2.2 (basic properties of valuations)**.** *Let p be a prime and $x, y \in \mathbb{Q}$.*

1. $v_p(xy) = v_p(x) + v_p(y)$.
2. $v_p(x + y) \geq \min(v_p(x), v_p(y))$, *with equality guaranteed when* $v_p(x) \neq v_p(y)$. ⌟

Proof. Exercise. □

This result should be fairly intuitive: it should follow from your long-standing intuition about multiplying and adding fractions. For example, we have the following calculations:

$$-1 = v_2\left(\frac{75}{2}\right) = v_2\left(100 \times \frac{3}{8}\right) = v_2(100) + v_2\left(\frac{3}{8}\right) = 2 - 3 = -1$$

$$-3 = v_2\left(\frac{35}{8}\right) = v_2\left(4 + \frac{3}{8}\right) = \min\left(v_2(4), v_2\left(\frac{3}{8}\right)\right) = \min(2, -3) = -3$$

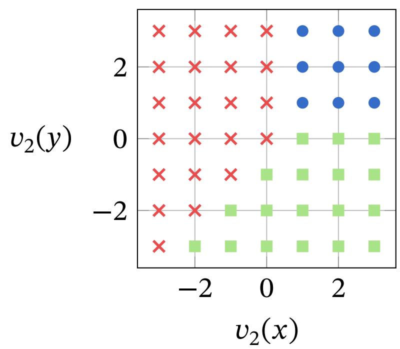
The inequality can be either equality or strict inequality when $v_p(x) = v_p(y)$. For example, $v_2(1/4) = -2$ and $v_2(3/4) = -2$, but $v_2(1/4 + 3/4) = v_2(1) = 0$. For $p = 2$, it turns out that the inequality is always strict when $v_2(x) = v_2(y)$, essentially because odd plus odd is even. However, when $p \neq 2$, we can have equality for addition, for instance $v_3(1/3) = -1$ and $v_3(1/3 + 1/3) = v_3(2/3) = -1$ still.

With that said, we now have all the tools we need to define Monsky's tricoloring of the unit square.

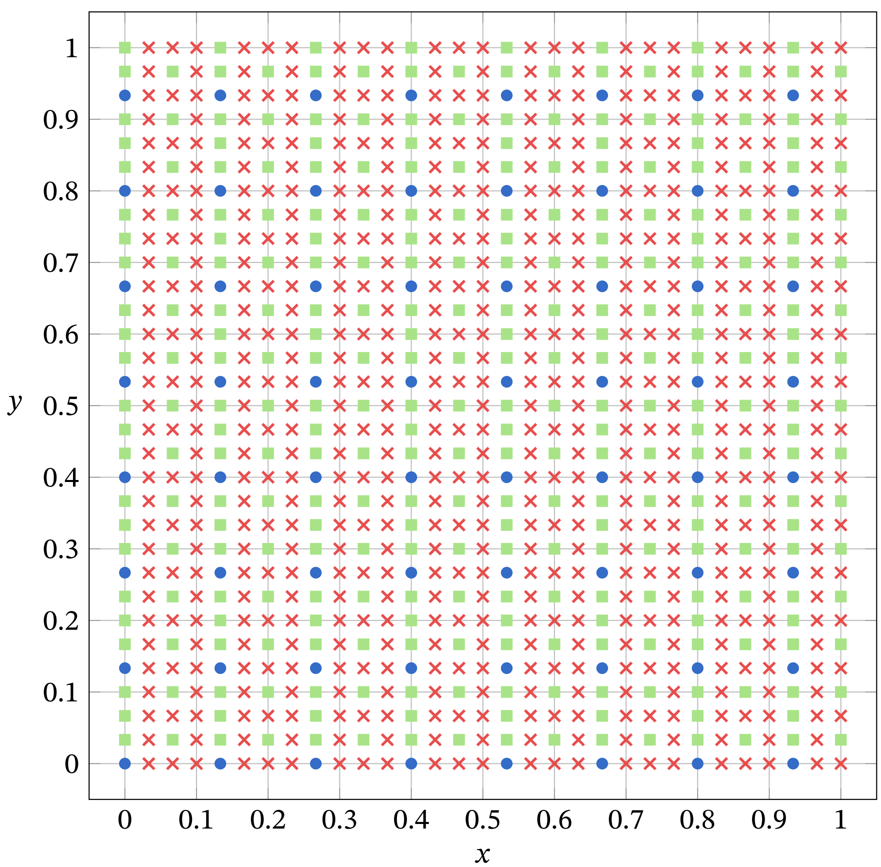
Definition 2.3 (Monsky's tricoloring)**.** We assign every point $(x, y) \in \mathbb{R}^2$ with rational coordinates a color:

- red/$\times$, if $v_2(x) \leq v_2(y)$ and $v_2(x) \leq 0$,
- green/■, if $v_2(x) > v_2(y)$ and $v_2(y) \leq 0$,
- blue/●, if $v_2(x) > 0$ and $v_2(y) > 0$. ⌟

The following picture clarifies how this coloring gives every rational point a unique color, as the conditions span all possible values of $v_2(x)$ and $v_2(y)$:



Because the points are colored based on their valuations, i.e. the number of 2s in the denominator, it does not look nearly as neat when visualized in the standard x-y plane. In the picture below, we visualize all points of the form $(m/30, n/30)$ in the unit square.



Before we end, we make a few useful observations about this coloring that will be useful later on, explaining the symmetries seen above.

Lemma 2.4 (symmetries in Monsky's tricoloring)**.**

1. *Every* $(x, y) \in \mathbb{R}^2$ *has the same color as* $(-x, y)$, $(x, -y)$, *and* $(-x, -y)$.
2. *Suppose* $(a, b) \in \mathbb{R}^2$ *is blue/●. Then for all points* $(x, y) \in \mathbb{R}^2$, *the point* (x, y) *has the same color as the point* $(x + a, y + b)$. ⌟

Proof.

1. This is obvious, since valuations were defined such that $v_p(x) = v_p(-x)$.
2. This is a direct application of Proposition 2.2. We show the case of (x, y) being red/× and leave the remaining two colors as exercises.

   We are to show that $v_2(x + a) \leq v_2(y + b)$ and $v_2(x + a) \leq 0$. First, we claim that $v_2(x + a) = \min(v_2(x), v_2(a))$ (the equality case of Proposition 2.2.2). This is just because $v_2(x) \leq 0$ and $v_2(a) > 0$, from the tricoloring definition. This also gives $v_2(x) < v_2(a)$, so actually $v_2(x + a) = v_2(x) \leq 0$.

   Meanwhile, $v_2(y + b) \geq \min(v_2(y), v_2(b))$. By the tricoloring definition again, $\min(v_2(y), v_2(b)) \geq \min(v_2(x), 0)$, and this is equal to $v_2(x) = v_2(x + a)$ by above, as was to be shown. □

Note that neither of these properties depended on fact that Monsky's tricoloring used $p = 2$. These properties would have been true for all prime p, but we want $p = 2$ to eventually argue that a certain area has even denominator.

Practice

Do these problems if you want to reinforce ideas and motivations from the main lesson.

1. Suppose $v_p(x) < v_p(y)$. Simplify $v_p(x^2 - y^2)$.
2. Prove the proofs that we omitted.
   (a) Proposition 2.2 (basic properties of valuations)
   (b) The remaining two cases of Lemma 2.4.2 (symmetries in Monsky's tricoloring)
3. Prove that the equality condition in Proposition 2.2.2 is an if and only if, assuming $p = 2$. That is, if $v_2(x) = v_2(y)$, then $v_2(x + y) > \min(v_2(x), v_2(y))$.

Extensions

Do these problems if you want to explore new ideas related to the main lesson.

4. Some resources phrase Monsky's theorem in terms of *p-adic absolute value*, rather than p-adic valuations. We are all familiar with the standard absolute value on $\mathbb{R}$, but a *generalized absolute value* is any function $|\cdot| : \mathbb{R}^2 \to [0, \infty)$ satisfying for all $x, y \in \mathbb{R}$,
   - $|x| = 0$ if and only if $x = 0$
   - $|xy| = |x||y|$
   - $|x + y| \leq |x| + |y|$ (triangle inequality).

Note that all these properties are true for the standard absolute value on $\mathbb{R}$.

A stronger version of the triangle inequality is called the *ultrametric inequality*: $|x + y| \le \max(|x|, |y|)$. An absolute value satisfying the ultrametric inequality is called *non-Archimedean*.

(a) Prove that the function $|x| = p^{-v_p(x)}$ (where $0 = p^{-\infty}$) is a non-Archimedean absolute value.

(b) Rephrase Monsky's tricoloring in terms of the 2-adic absolute value.

5. We only considered p-adic valuations for prime p. What would go wrong if we tried to copy the definition for non-prime p?

6. Prove that Monsky's coloring is *dense* in every color: for every color, every point $(x, y) \in \mathbb{R}^2$, and every $\epsilon > 0$, there is a point (x', y') with $|x - x'| \le \epsilon$ and $|y - y'| \le \epsilon$ such that (x', y') is the desired color. In other words, no matter how far you zoom into the picture, you will never find a region that is all colored the same—every tiny patch will always have points of all three colors. This makes the coloring very hard to draw.

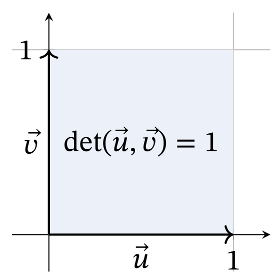
2.2 Computing the area

With Monsky's tricoloring defined, it remains to show that any triangle with three vertices of different colors has area with negative 2-adic valuation (i.e. even denominator). To do this, we need to first to compute the area of a triangle given its three pairs of coordinates. Here are a few that you might think of:

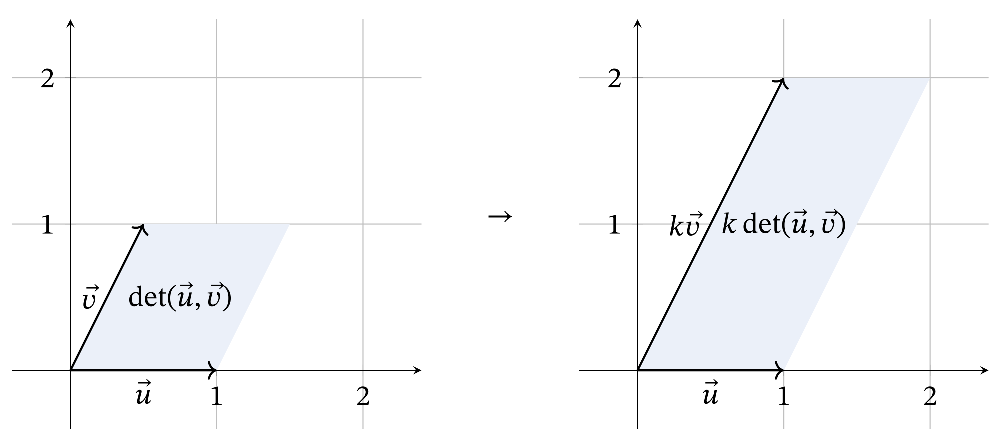
- (very annoying) Compute the side lengths using the distance formula, then apply Heron's formula to get the area.
- (annoying) Compute one side length using the distance formula and regard it as the base of the triangle. Compute the perpendicular line passing through the opposite point, find the intersection with the base, then use the distance formula to compute the height of the triangle. Use $A = bh/2$.
- (advanced knowledge) Directly apply the shoelace formula (a trick that people who do math Olympiads might know, which is a formula to compute the area of any polygon given its coordinates).
- (our choice) Use the *determinant*, a concept from linear algebra.

Definition 2.5 (determinant)**.** Let $\vec{u} = (a, b)$ and $\vec{v} = (c, d)$ be two vectors in the plane. The *determinant* of these vectors, denoted $\det(\vec{u}, \vec{v})$ is the area of the parallelogram that they span, negative if going from $\vec{u}$ to $\vec{v}$ is clockwise. ⌟

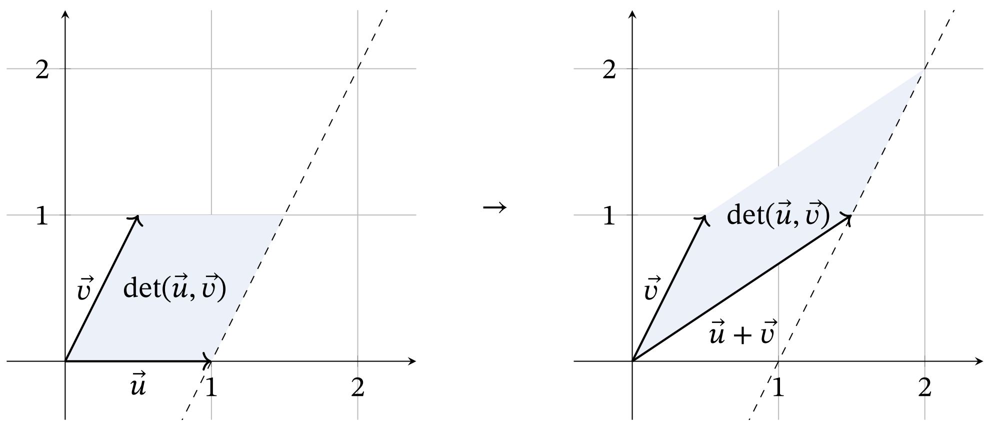
For example, if $\vec{u} = (1, 0)$ and $\vec{v} = (0, 1)$, then $\det(\vec{u}, \vec{v}) = 1$. Swapping $\vec{u}$ and $\vec{v}$ would result in a determinant of -1.



To compute the determinant of (area spanned by) any other pair of vectors, the typical technique is to keep modifying the shape until it becomes the square with area 1 above. For example, one modification that we can make is to multiply one of the side lengths by a factor of k (multiplying coordinatewise, e.g. $2 \cdot (1, 2) = (2, 4)$). This results in the area being changed by a factor of k. In the following example, $k = 2$.



Another modification that we can do is skew the parallelogram, without changing its base or height. We can do this by considering either $\vec{u}$ or $\vec{v}$ as the base. Algebraically, this means adding a multiple of the base to the other vector (adding coordinatewise, e.g. $(1,2)+(3,1)=(4,3)$). This has no effect on the area. In the following example, $\vec{v}$ is considered the base.



In other words, we have just shown:

Proposition 2.6 (basic properties of determinant)**.** *Let $\vec{u}$ and $\vec{v}$ be two vectors in the plane, and $k \in \mathbb{R}$ be any constant.*

1. $\det(k\vec{u},\vec{v}) = \det(\vec{u},k\vec{v}) = k\det(\vec{u},\vec{v})$.
2. $\det(\vec{u}+k\vec{v},\vec{v}) = \det(\vec{u},\vec{v}+k\vec{u}) = \det(\vec{u},\vec{v})$. ⌟

Now, we can use these properties to directly calculate the area of a parallelogram spanned by two vectors (a,b) and (c,d). To minimize the amount of parentheses, people typically write $\det((a,b),(c,d))$ as $\det\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. We have

$$\begin{aligned}
\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} &= a\det\begin{bmatrix} 1 & b/a \\ c & d \end{bmatrix} && \text{(by 1)} \\
&= a\det\begin{bmatrix} 1 & b/a \\ 0 & d-bc/a \end{bmatrix} && \text{(by 2)} \\
&= a\det\begin{bmatrix} 1 & b/a \\ 0 & (ad-bc)/a \end{bmatrix} \\
&= (ad-bc)\det\begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix} && \text{(by 1)} \\
&= (ad-bc)\det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} && \text{(by 2)} \\
&= ad-bc.
\end{aligned}$$

The process that we just performed is called *Gaussian elimination* (in which a set of vectors of reduced to the basis vectors $(1,0)$ and $(0,1)$, or their higher dimensional analogues), and this calculation proves the following theorem:

Theorem 2.7 (area of a triangle)**.** *Let $(0,0)$, (a,b), (c,d) be the coordinates of the vertices of a triangle. Then, the area of the triangle is $|ad-bc|/2$.* ⌟

Finally, we can prove the desired final result of this section.

Lemma 2.8 (valuation of a trichromatic triangle)**.** *Let (r_1, r_2), (g_1, g_2), (b_1, b_2) be the coordinates of the vertices of a triangle, and suppose that according to Monsky's tricoloring, (r_1, r_2) is red, (g_1, g_2) is green, and (b_1, b_2) is blue. Let A be the area of this triangle. Then $v_2(A) \leq -1$.* ⌟

Proof. First, by Lemma 2.4, we may translate the triangle by $(-b_1, -b_2)$ without affecting the colors. Thus, assume without loss of generality that $(b_1, b_2) = (0, 0)$. Then,

$$v_2(A) = v_2\left(\frac{|r_1 g_2 - r_2 g_1|}{2}\right) = v_2(r_1 g_2 - r_2 g_1) - 1.$$

By the definition of Monsky's tricoloring, $v_2(r_1) \leq v_2(r_2)$ and $v_2(g_2) < v_2(g_1)$. Therefore, $v_2(r_1 g_2) = v_2(r_1) v_2(g_2) < v_2(r_2) v_2(g_1) = v_2(r_2 g_1)$. Thus, we may continue to simplify

$$\begin{aligned} v_2(A) &= \min(v_2(r_1 g_2), v_2(r_2 g_1)) - 1 \\ &= v_2(r_1) v_2(g_2) - 1 \\ &\leq 0 \cdot 0 - 1 \\ &\leq -1. \end{aligned}$$ □

Corollary 2.9 (lines have two colors)**.** *Every line in Monsky's tricoloring contains points from at most 2 colors.* ⌟

Proof. If a line contained points of 3 colors, one would be able to find a degenerate (area 0) trichromatic triangle along the line, but $v_2(0) = \infty$, which is not ≤ -1, contradiction. □

Practice

Do these problems if you want to reinforce ideas and motivations from the main lesson.

1. Show that the converse of Lemma 2.8 is false, that is, there are triangles with $v_2(A) \leq -1$ but vertices not all different colors.
2. In our computation of $\det \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, we ignored two edge cases to simplify things: each application of Proposition 2.6.1 involved dividing by something, but we avoided checking that it was non-zero. What should we do if they actually are zero? Why doesn't it affect the rest of the proof?
3. Compute the volume of the tetrahedron with vertices $(0, 0, 0)$, $(3, 0, 3)$, $(2, 2, 0)$, and $(0, 2, 4)$. (Hint: First, what is the volume of the tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$?)
4. Let p be any prime and consider Monsky's tricoloring with this new p. Suppose a triangulation of a shape into n triangles of equal area contains at least one trichromatic triangle. Prove that if the area of the whole shape is A, then $v_p(n) \geq v_p(2A)$. (We will use this fact in a future exercise to extend Monsky's theorem to other shapes.)
5. Corollary 2.9 can be proven directly from the definition of Monsky's tricoloring, without talking about area, but it is more complicated. However, it's still worth trying to get some intuition. For these questions, feel free to avoid writing equations. Argue using the $v_2(x)$-$v_2(y)$ chart on page 4.

(a) Prove that if (x, y) and (a, b) are two different colors, then $(x, y) + (a, b)$ is one of the two colors. (In particular, you already know this fact when one of them is blue/●, so let (x, y) be red/× and (a, b) be green/■.)

(b) Prove that for all $k \in \mathbb{R}$, either (x, y) is the same color as $k \cdot (x, y)$, or $k \cdot (x, y)$ is blue/●.

(c) Show that if (x, y) and (a, b) are two different colors, then for all $k \in \mathbb{R}$, the point $(x, y) + k \cdot (a, b)$ is one of those two colors. Conclude that every line has at most 2 colors.

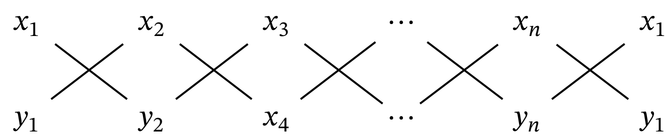
Extensions

Do these problems if you want to explore new ideas related to the main lesson.

6. Use what we learned about the determinant to prove the shoelace formula: Suppose the vertices of a polygon are $(x_1, y_1), \ldots, (x_n, y_n)$ listed in counterclockwise order. Then the area of the polygon is

$$\frac{1}{2}|(x_1y_2 + x_2y_3 + \cdots + x_ny_1) - (y_1x_2 + y_2x_3 + \cdots + y_nx_1)|.$$

It is so named because the terms to be multiplied look like tying your shoelaces.



3 Finding a trichromatic triangle

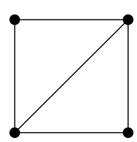
To recap where we are, we have shown that as long as you can find a trichromatic triangle in a triangulation, it is impossible for it to have area $1/N$ for odd N. What remains is to find this trichromatic triangle.

Our tool will be a slight variation of an important result connecting graph theory and topology, called Sperner's lemma. Sperner's lemma is the 2D generalization of the following intuitive fact: if you have a line of points that starts with red/× and ends with blue/●, there must be an odd number of segments with different colors on both ends. In this picture, there are 5 segments with different colors on both ends.

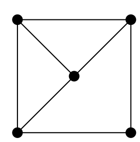
The 2D version will assert the existence of an odd number of trichromatic triangles given appropriate boundary conditions. The main difference between Sperner's original lemma and the version that we will eventually use to prove Monsky's theorem is that Sperner had a slightly different definition of *triangle* than we do.

Definition 3.1 (graph triangulation). A planar graph $G = (V, E)$ has a set of *vertices* $V \subseteq \mathbb{R}^2$ and a set of non-intersecting *edges* E that each connect exactly two vertices, which you can just think of as line segments or curves. The edges cut $\mathbb{R}^2$ into many regions, which we call *faces*, including the unbounded face outside of everything.

A *triangle* in a planar graph is a face formed by exactly three edges. An (internal) *triangulation* of a planar graph adds some vertices and/or edges so that every face is a triangle, except possibly the unbounded face. ⌟



good triangulation



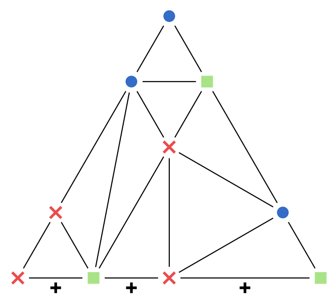
not a triangulation

In the picture on the right, despite every face looking like a triangle as drawn in the plane, it is graph-theoretically not a triangulation, because the larger "triangle" actually consists of 4 edges. (Generally in graph theory, we only care about the structure between vertices and edges, not how they end up laid out in a drawing.)

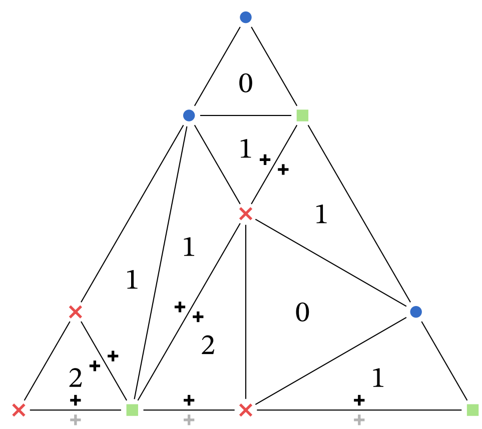
Theorem 3.2 (Sperner's lemma). *Let $G = (V, E)$ be any triangulation of a triangle RGB. Suppose we have assigned colors to the vertices V such that R is red/×, G is green/■, and B is blue/●. Suppose that all vertices along line segment $\overline{RG}$ are either red/× or green/■, and similarly for $\overline{GB}$ and $\overline{RB}$. Then there must be an odd number of trichromatic triangles.* ⌟

Proof. In every face, including the unbounded face, let us count the edges that are red/× on one end and green/■ on the other end. Call these red-green edges. Doing so will actually count every red-green edge twice, once from each side, so we should end up with an even number in total.

First, let us count the red-green edges in the unbounded face. As we noted before, there are odd number of edges along $\overline{RG}$ with different colors on each end. Neither $\overline{GB}$ nor $\overline{RB}$ is allowed to contain red-green edges, therefore the unbounded face contributes an odd number of red-green edges. In the picture below, these counts are marked.



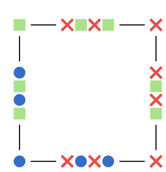
Next, let us count the red-green edges in the bounded faces. Every triangle with at most 2 colors has either 0 or 2 red-green edges, which contributes an even number. Every trichromatic triangle has exactly 1 red-green edge. In the picture below, the number of red-green edges in each triangle is written, and the counts are again marked.



Therefore, in order for the final count to be even, noting that the unbounded face contributes an odd number and each non-trichromatic face contributes an even number, the number of trichromatic triangles must be odd. □

In order to adapt Sperner's lemma to our coloring problem, we have to first make some observations about the boundary of our unit square. These are basic observations that follow directly from the original definition.

Lemma 3.3 (boundary of Monsky's tricoloring)**.** *In Monsky's tricoloring, we have:*



That is,

- $(0,0)$ *is blue/●,* $(0,1)$ *is green/■, and* $(1,1)$ *and* $(1,0)$ *are red/×.*
- *The bottom boundary is entirely red/× and blue/●.*
- *The left boundary is entirely green/■ and blue/●.*
- *The top and right boundaries are entirely red/× and green/■.* ⌟

Proof. Exercise. □

Finally, note that the main difference between Sperner's lemma and our coloring problem is that we allow triangular faces to have more than three edges, as long as enough edges are collinear so that the face still resembles a triangle embedded in the plane. The main idea is to use Corollary 2.9, about how straight lines have only two colors, to recover much of the same argument. The fact that we are now starting with a square, not a triangle, actually does not cause any trouble, because we can treat the top and right boundaries as a single "side" that is red/× and green/■.

Lemma 3.4 (trichromatic triangle in Monsky's tricoloring)**.** *Suppose the unit square is subdivided into triangles (shapes formed by 3 straight lines) and the vertices are all colored according to Monsky's tricoloring. Then there is an odd number of trichromatic triangles (in particular, at least one).* ⌟

Proof. In what follows, to disambiguate, we use the words "vertex" and "side" to refer to the vertices and sides of triangles, and the words "node" and "edge" to refer to the vertices and edges of the underlying graph. In other words, every triangle has exactly 3 vertices but may touch more nodes, and every side may contain several collinear edges.

We will do the same thing as before, counting the red-green edges in every face, and this process must end with an even number in total. Lemma 3.3 provides us with the boundary conditions needed to count red-green edges in the unbounded face. In particular, there will be a odd number of red-green edges on the top/right boundaries, because the top-left is green/■ and the bottom-right is red/×, and there are no red-green segments on the other boundaries.

The remaining bounded faces are triangles in the sense of being formed by 3 straight lines, each of which contains only vertices of at most two colors by Corollary 2.9. Consider a side of the triangle. If the two vertices are not red/red, green/green, or red/green, then there are no red-green edges on that side. In the case of red/red or green/green, there are an even number of red-green edges. In the case of red/green, there are an odd number of red-green edges.

Thus, the triangle has an odd number of red-green edges if and only if there is exactly one side with red/green vertices, if and only if it is trichromatic. We need an odd number of such triangles, because the total count is even and the count on the unbounded face is odd. Thus, there are an odd number of trichromatic triangles. □

This completes the proof of Monsky's theorem, at least when the triangles all have rational points. Once we extend the concept of valuations to real numbers, the rest of the argument will proceed with no change.

Theorem 3.5 (Monsky's theorem)**.** *There is no way to subdivide a unit square into an odd number of triangles with the same area.* ⌟

Practice

Do these problems if you want to reinforce ideas and motivations from the main lesson.

1. (a) Find a (graph-theoretic) triangulation of a triangle where the three vertices of are red, green, and blue as desired, but there is no trichromatic triangle in the triangulation. (In other words, the boundary conditions on the edges are necessary for Sperner's lemma.)

(b) Now define a triangle as a shape formed by 3 lines, as in Monsky's theorem. Find a triangulation of a triangle that satisfies the boundary conditions of Sperner's lemma, but there is no trichromatic triangle. (In other words, the fact that every line contains at most 2 colors was necessary for our proof.)

2. Prove Lemma 3.3.

3. Write an alternative proof for Sperner's lemma based on the idea that triangles are rooms and red-green edges are doors that you can walk through. (This method is also typically more effective for actually finding a trichromatic triangle, rather than just knowing existence.)

4. Prove the following generalization of Sperner's lemma: Consider a triangulation of any shape. Listing colors counterclockwise, let A be the number of RGB triangles, and let B be the number of RBG triangles. Traversing the boundary counterclockwise, let C be the number of RG edges and D be the number of GR edges. Then $A - B = C - D$.

   (Hint: Mimic our original proof of Sperner's lemma. Name the colors 0, 1, and 2 instead of red, green, and blue. Instead of + signs along red-green edges, place either -1, 0, or 1 along all edges, depending on which one is equivalent to $j - i$ (mod 3), where the edge goes from color i to j counterclockwise.)

5. Use the following steps to prove that a regular hexagon can only be cut into equal-area triangles in multiples of 6.

   (a) Convert any equal-area triangulation of a regular hexagon into an equal-area triangulation of the hexagon with vertices $(0,0)$, $(1,0)$, $(2,1)$, $(2,2)$, $(1,2)$, and $(0,1)$. Thus, we will focus on this hexagon from now on.

   (b) Compute the boundary colors of this hexagon with Monsky's tricoloring with $p = 3$, showing that Sperner's lemma can be applied.

   (c) Apply Problem 2.4 to conclude that the number of triangles n must be divisible by 3.

   (d) Recompute the boundary colors for $p = 2$ and apply Problem 3.4 to conclude that n is divisible by 2.

Extensions

Do these problems if you want to explore new ideas related to the main lesson.

6. Solve the rental harmony problem: Three roommates are moving into a three-bedroom apartment and need to decide how to split the rent. However, the bedrooms are all different and everyone has their own preferences. In other words, each person has a function that takes in the prices for each room and outputs the room they would pick given those prices.

   Suppose that rent is \$3000 and everyone agrees that \$10 is a negligible amount. How can the roommates determine who should take which room, and how much each person should pay? Each person should end up feeling like they got the room they wanted given the final prices for each room, and if there are multiple such configurations, pick any of them. (Hint: Barycentric coordinates!)

4 Extending to real-valued points

The last thing to show about Monsky's theorem is how to rule out triangles with real-valued vertices, not just rational points. This is the most technically challenging part, and most other resources will assume rather advanced abstract algebra knowledge to explain this part, but we will attempt to explain as much as possible concretely.

The goal is to extend v_2 (or v_p more generally) to take real values, while continuing to satisfy Proposition 2.2: the basic properties $v_p(xy) = v_p(x) + v_p(y)$ and $v_p(x+y) \geq \min(v_p(x), v_p(y))$. Here are three illustrative examples of what can happen.

Example 4.1 (v_p for nth roots)**.** For nth roots, there is a clear choice for what we need to pick, in order to satisfy $v_p(xy) = v_p(x) + v_p(y)$:

$$v_p(\sqrt[k]{n}) = \frac{1}{k} v_p(n).$$

This gives $v_2(\sqrt{2}) = 1/2$. ⌟

In particular, we see that valuations can take rational values—that is why we previously insisted on writing "blue/● if $v_2(x) > 0$ and $v_2(y) > 0$" rather than $v_2(x) \geq 1$ and $v_2(y) \geq 1$. Some other irrational numbers can also be computed this way, but not all of them.

Example 4.2 (choices for v_p)**.** Consider $\alpha = 5 + \sqrt{17}$ and $\beta = 5 - \sqrt{17}$. We have

$$\alpha\beta = (5 + \sqrt{17})(5 - \sqrt{17}) = 25 - 17 = 8,$$

hence $v_2(\alpha\beta) = v_2(\alpha) + v_2(\beta) = 3$. Should we just set $v_2(\alpha) = v_2(\beta) = 3/2$ then? No, this does not work. The valuations must satisfy

$$\min(v_2(\alpha), v_2(\beta)) \leq v_2(\alpha + \beta) = v_2(10) = 1,$$

and $1.5 > 1$. In fact, to satisfy both this requirement and the previous requirement, we must have $v_2(\alpha) \neq v_2(\beta)$, thus we are actually in the equality case of the inequality, and one of $v_2(\alpha)$ and $v_2(\beta)$ must be 1, the other equal to 2. Either choice is fine, because it is algebraically impossible to differentiate α and β: any polynomial equation satisfied by α will also be satisfied by β. ⌟

One easier case, however, is extending v_p to *transcendental numbers* like π. Numbers are called *algebraic* if they satisfy polynomials expressions with rational coefficients, and transcendental otherwise. In the previous example, we were restricted in what we could pick for $v_2(\alpha)$ and $v_2(\beta)$ because expressions like $\alpha\beta$ and $\alpha + \beta$ resulted in rational numbers that we already knew the valuations of. However, this is not the case with π, so we are actually free to choose whatever we want for $v_p(\pi)$!

Example 4.3 (v_p for π)**.** We can freely pick $v_p(\pi) = 0$. However, once we pick this, v_p is also uniquely determined for other numbers, e.g. $v_2(2\pi) = v_2(2) + v_2(\pi) = 1 + 0 = 1$. We can also have expressions like $3 + \pi$, which need to satisfy $v_p(3+\pi) \geq \min(v_p(3), v_p(\pi))$ with equality if $v_p(3) \neq v_p(\pi)$. Although this is an inequality, it turns out that we can just pick $v_p(3+\pi) = \min(v_p(3), v_p(\pi))$ and it works!

It also true that we must have $v_p(\sqrt[k]{\pi}) = 0$ and in general some constraints on roots of polynomials using π as a coefficient, but that is a problem that we'll tackle later, with the rest of the algebraic roots. ⌟

Proposition 4.4 (rational expression). *Given a set of numbers X and a new number t, every number that can be formed using X, t, addition, subtraction, multiplication, and division can be written as*

$$\frac{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n}{b_0 + b_1 t + b_2 t^2 + \cdots + b_m t^m}$$

for some $a_0, \dots, a_n, b_0, \dots, b_m \in X$. (The set of such expressions is often denoted $X(t)$.) ⌟

Proof. Omitted. □

Proposition 4.5 (v_p for transcendental numbers). *Suppose X is a set of numbers satisfying Proposition 2.2, and suppose t is a transcendental number with respect to X (meaning t does not satisfy any polynomial with coefficients in X). Define*

$$v_p\left(\frac{a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n}{b_0 + b_1 t + b_2 t^2 + \cdots + b_m t^m}\right) = \min_i v_p(a_i) - \min_j v_p(b_j).$$

Then Proposition 2.2 continues to be true on $X(t)$. ⌟

Proof sketch. For the sake of illustration, we will just prove Proposition 2.2 for $a + bt$ and $c + dt$, and omit the general case for now.

1. We want to show $v_p((a + bt)(c + dt)) = v_p(a + bt) + v_p(c + dt)$. By the above definition of v_p, this is equivalent to showing

$$\begin{aligned} \min(v_p(ac), v_p(ad + bc), v_p(bd)) &= \min(v_p(a), v_p(b)) + \min(v_p(c), v_p(d)) \\ &= \min(v_p(ac), v_p(ad), v_p(bc), v_p(bd)). \end{aligned}$$

   We know that $v_p(ad+bc) \geq \min(v_p(ad), v_p(bc))$, with equality if $v_p(ad) \neq v_p(bc)$. Thus, the only case that we are concerned about is when $v_p(ad) = v_p(bc)$ and these are actually the minimum of the RHS. That means the following four equations are true:

$$\begin{aligned} v_p(ad) &< v_p(ac) \\ v_p(ad) &< v_p(bd) \\ v_p(bc) &< v_p(ac) \\ v_p(bc) &< v_p(bd). \end{aligned}$$

   Among many possible contradictions, the second equation implies $v_p(a) < v_p(b)$ and the third equation implies $v_p(b) < v_p(a)$, so this case is not possible.

2. We want to show $v_p((a + bt) + (c + dt)) \geq \min(v_p(a + bt), v_p(c + dt))$ with equality if $v_p(a + bt) \neq v_p(c + dt)$. By the above definition of v_p, this is equivalent to showing

$$\begin{aligned} \min(v_p(a + c), v_p(b + d)) &\geq \min(\min(v_p(a), v_p(b)), \min(v_p(c), v_p(d))) \\ &= \min(v_p(a), v_p(b), v_p(c), v_p(d)). \end{aligned}$$

   This is certainly true because we know $v_p(a + c) \geq \min(v_p(a), v_p(c))$ and likewise for $b + d$. For the equality condition, the assumption is equivalent to saying that $\min(v_p(a), v_p(b)) \neq \min(v_p(c), v_p(d))$. Casework on which is actually smaller in these two mins gives exactly the conditions for the equality conditions on $v_p(a + c)$ and $v_p(b + d)$. □

Now, let's turn back to the algebraic case. While it is difficult in general, some clever algebra does let us compute the choices for v_p when given an algebraic number α consisting just of square roots. For example, numbers lie $5+\sqrt{17}$ and $1/3+\sqrt{10/7}$, or even $3-i$ (but we don't particularly care about that). In particular, note that this condition is equivalent to α satisfying a quadratic equation: You all know how to find the solution of a quadratic equation, and it looks $a+\sqrt{b}$ for some $a, b \in \mathbb{Q}$. Conversely, given $a+\sqrt{b}$, we can multiply with $a-\sqrt{b}$ to get $a^2-b \in \mathbb{Q}$. Thus, these two numbers are the solutions to $x^2-2ax+(a^2-b)=0$.

To make notation easier, let α and β be the two roots of $x^2+rx+s=0$. Then we can calculate that:

$$\begin{aligned} v_p(s) &= v_p(\alpha^2+r\alpha) \\ &\geq \min(2v_p(\alpha), v_p(r)+v_p(\alpha)), \end{aligned}$$

with equality if $2v_p(\alpha) \neq v_p(r)+v_p(\alpha)$, in other words $v_p(\alpha) \neq v_p(r)$. That gives us two cases. First, it's possible that $v_p(\alpha)=v_p(r)$. In this case, we are done computing $v_p(\alpha)$.

The other case is that $v_p(s)=\min(2v_p(\alpha), v_p(r)+v_p(\alpha)) = v_p(\alpha)+\min(v_p(\alpha), v_p(r))$. A bit of algebra will show that:

$$v_p(\alpha) = \begin{cases} v_p(s)/2 & \text{if } v_p(s)/2 \leq v_p(r) \\ v_p(s)-v_p(r) & \text{if } v_p(s)/2 > v_p(r) \end{cases}$$

Lastly, the other solution β must satisfy the same equations. Furthermore, $v_p(\alpha)+v_p(\beta)=v_p(\alpha\beta)=v_p(s)$. That means we are restricted to two possibilities:

- If $v_p(s)/2 \leq v_p(r)$, then there is only one option and $v_p(\alpha)=v_p(\beta)=v_p(s)/2$. This was the case of $v_2(\sqrt{2})$, since $x^2-2=0$ has $v_2(s)=1$ and $v_2(r)=\infty$.
- If $v_p(s)/2 > v_p(r)$, then there are two options: $v_p(\alpha)=v_p(r)$ and $v_p(\beta)=v_p(s)-v_p(r)$, or vice versa. This was the case of $v_2(5\pm\sqrt{17})$, since it satisfies $x^2-10x+8$ where $v_2(s)=3$ and $v_2(r)=1$.

In other words, we have proven the following, which is the converse of what we really wanted, but illustrates the main difficulty of this task:

Proposition 4.6 (v_p for quadratic extensions). *Let X be a set of numbers satisfying Proposition 2.2 and let α and β be the two solutions of $x^2+rx+s=0$, where $r, s \in X$. If Proposition 2.2 were also true for $v_p(\alpha)$ and $v_p(\beta)$, then*

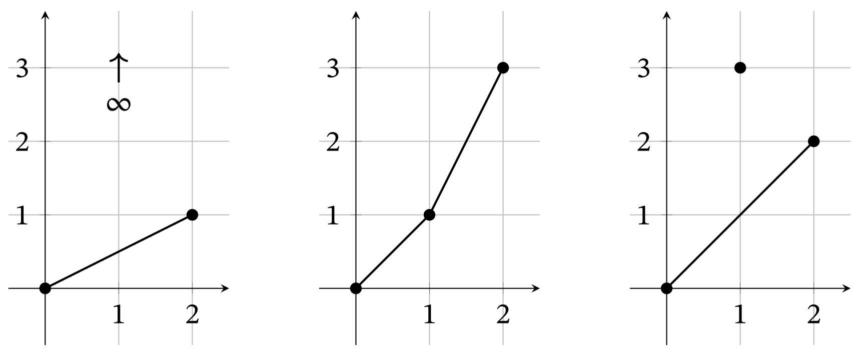
- *If $v_p(s)/2 \leq v_p(r)$, we have $v_p(\alpha)=v_p(\beta)=v_p(s)/2$.*
- *If $v_p(s)/2 > v_p(r)$, we have $v_p(\alpha)=v_p(r)$ and $v_p(\beta)=v_p(s)-v_p(r)$, or vice versa.* ⌟

There are two difficulties: first, we need to actually pick one option for all quadratics with $v_p(s)/2 > v_p(r)$. Then, we need to also make sure that all of our choices are consistent with each other. For example, once we've picked an option for $v_2(5\pm\sqrt{17})$, we certainly no longer have freedom to choose for $v_2(10\pm\sqrt{68})$, and there are more complicated restrictions with addition again. It's a little beyond the scope of this course to entirely understand the method to pick these consistently, but it can be done and relies on Axiom of Choice, or more specifically Zorn's Lemma.

The last thing to show you is a method for computing the options for larger degree polynomials. This is a method known as *Newton polygons*.

Definition 4.7 (Newton polygon). Let p be a prime. The Newton polygon of a polynomial $a_0 + a_1 x + \cdots + a_n x^n$ with respect to p is the lower convex boundary of the points $(0, v_p(a_n)), (1, v_p(a_{n-1})), \ldots, (n, v_p(a_0))$. ⌟

The lower convex boundary means that you should imagine a rubber band stretching around the points, and we look only at the lower half. In particular, the Newton polygon is not a polygon. It is a connected sequence of line segments. In the below examples, we use $p = 2$.



$x^2 - 2$ $\qquad$ $x^2 - 10x + 8$ $\qquad$ $x^2 - 8x + 12$

Since we're already familiar with the first two of these examples, notice the slopes of these segments correspond exactly to the valuations of their roots! Moreover, the segment of horizontal length 2 in the Newton polygon for $x^2 - 2$ corresponds with 2 roots of valuation $1/2$. In general, we have the following theorem.

Theorem 4.8 (fundamental theorem of Newton polygons). *Let $a(x)$ be a polynomial and fix a prime p. For every segment in the Newton polygon of $a(x)$ with respect to p, if the segment has horizontal length ℓ and slope m, then ℓ of the roots of $a(x)$ must have valuation v_p equal to m.* ⌟

Proof sketch. Again for the sake of brevity, we will only show the proof for quadratics. Let $x^2 - ax + b = 0$ and let α and β be the two solutions. Recall that $a = \alpha + \beta$ and $b = \alpha\beta$.

- Suppose that $v_p(\alpha) \neq v_p(\beta)$. Then we have plotted $(0, 0)$, $(1, \min(v_p(\alpha), v_p(\beta)))$, and $(2, v_p(\alpha) + v_p(\beta))$. Because $\min(x, y) \leq (x + y)/2$, the Newton polygon will look like a V and have two segments. The first segment has slope $\min(v_p(\alpha), v_p(\beta))$ (we don't know which one), and the second segment has slope $(v_p(\alpha) + v_p(\beta)) - \min(v_p(\alpha), v_p(\beta))$, which is the other one.
- Suppose that $v_p(\alpha) = v_p(\beta) = k$. Then we have plotted $(0, 0)$, $(1, \ell)$, $(2, 2k)$, where $\ell = v_p(\alpha + \beta) \geq k$. Thus, ℓ lies above the lower convex boundary and the slope of the line is k, as we wanted. □

To reiterate, this is just a more enlightening way, to conclude what $v_p(\alpha)$ must be for algebraic α. We have still not discussed exactly how to ensure that all the choices we make are consistent, but that proof is beyond the scope of this lesson.