SPERNER'S LEMMA

MOOR XU

ABSTRACT. Is it possible to dissect a square into an odd number of triangles of equal area? This question was first answered by Paul Monsky in 1970, and the solution requires elements from two seemingly disjoint areas of mathematics: number theory and combinatorial topology. Along the way, we will see p-adic numbers and a combinatorial proof of Brouwer's fixed point theorem.

1. INTRODUCTION

We will discuss a simple theorem that is connected with some interesting results. The topic is dissection of polygons into triangles. We want triangles of equal areas – equiareal triangulation.

Theorem 1.1. *If a square is cut into triangles of equal areas, then the number of triangles must be even.*

We start with a little history.

Remark. The first person who thought about this was Fred Richman (1965) at New Mexico State. He was preparing a masters exam, and he wanted to include this but couldn't solve it. He asked the problem in the Am. Math. Monthly, and Paul Monsky proved the theorem in 1970.

There are several results of a similar nature, but there are many more open problems than results. The proof of this fact is difficult and makes use of two disjoint areas of mathematics. This is the only known proof so far. There are two ingredients:

(1) combinatorial / topological
(2) number theoretic / algebraic.

2. SPERNER'S LEMMA

We start with the topological part.

Let P be a polygon in the plane, and consider a triangulation of this polygon. Color each vertex of this triangulation by one of three colors 1, 2, or 3. An edge is called a *12-edge* if its endpoints is colored 1 and 2, and a triangle is said to be *complete* if each of its vertices is colored using a different color.

Lemma 2.1 (Sperner, 1928)**.** *If we have a polygon whose vertices are colored by three colors and a triangulation is given for this polygon, the number of complete triangles (with three different colors) is equal to the number of 12-edges on the boundary of the polygon (mod 2).*

This is a way to predict if we have a complete triangle inside the polygon.

Proof. Put a dot on each side of each 12-segment. We want to count the number of dots in two different ways. Firstly, each interior segment contributes either 0 or 2 dots (depending on whether it is a 12-edge), while each boundary segment contributes 0 or 1 dots. The number of dots is therefore equal to the number of 12-edges on the boundary of the polygon (mod 2).
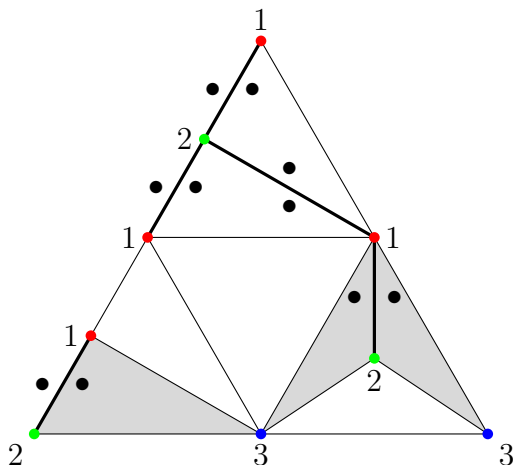
Now count the number of dots in each triangle. Complete triangles contain one dot (check this), while other triangles contain an even number of dots. Therefore, the number of dots is equal to the number of complete triangles (mod 2). This concludes the proof. □

3. Brouwer fixed point theorem

This is an optional section. Later sections do not depend on it.

Theorem 3.1. *Let B^n be the n-dimensional ball. Then any continuous map $f : B^n \to B^n$ has a fixed point.*

Example 3.1. In the one-dimensional case, this is simply a consequence of the intermediate value theorem.

This is much less trivial in higher dimensions. In the two-dimensional case, we can prove this using Sperner's lemma.

We'll need to use a lemma that is a simple corollary of Sperner's Lemma.

Lemma 3.2. *Consider a triangulation T of a triangle ABC. Color the outer three vertices A, B, C as 1, 2, and 3 respectively. Use only colors $\{1, 2\}$ on the edge AB, only colors $\{1, 3\}$ on the edge AC, and only colors $\{2, 3\}$ on the edge BC. This triangulation contains a complete triangle.*

Proof. A 12-boundary edge can only appear on the side with endpoints colored 1 and 2, so there are an odd number of 12-edges on the boundary. □

Remark. This works for all polygons, not just triangles. We will use a related result in our proof of the Monsky theorem.

Proof of Theorem 3.1 in case $n = 2$. Since a disk is homeomorphic to a triangle, it suffices to prove that any continuous map on a triangle has a fixed point.

Let $\triangle$ be the triangle in $\mathbb{R}^3$ with vertices $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$. In other words,
$$\triangle = \{(a_1, a_2, a_3) \in [0, 1]^3 : a_1 + a_2 + a_3 = 1\}.$$
Suppose to the contrary that a continuous map $f : \triangle \to \triangle$ has no fixed point.

For each of these triangulations, color the vertices with the colors $\{1, 2, 3\}$. For each $v \in T$, define the coloring of v to be the minimum color i such that $f(v)_i < v_i$, that is, the minimum index i such that the i-th coordinate of $f(v) - v$ is negative. This is well-defined assuming that f has no fixed point because we know that the sum of the coordinates of v and $f(v)$ are the same.

We claim that e_1, e_2, and e_3 are colored differently. Indeed, they each maximize a different coordinate and hence are first negative at that coordinate. Furthermore, a vertex on the $e_1 e_2$ edge has $a_3 = 0$, so that $f(v) - v$ has nonnegative third coordinate and is hence colored 1 or 2. The same is true for the other coordinates, and we are therefore in the case of Lemma 3.2. Hence, for any triangulation T_k, there is a small complete triangle $\triangle_k$.

Let $\delta(T)$ denote the maximal length of an edge in a triangulation T. It is easy to construct a sequence of triangulations $T_1, T_2, \ldots$ such that $\delta(T_k) \to 0$ as $k \to \infty$.

These triangles $\triangle_k$ don't necessarily have to converge to a point, but every sequence in a compact space has a convergent subsequence. So there is some subsequence of triangles $\triangle_{k_1}, \triangle_{k_2}, \ldots$ that converges to some point x. We know that for each i, there are vertices $v_{i,1}$, $v_{i,2}$, $v_{i,3}$ really close to x such that $f(v_{i,1})$ has its first coordinate less than that of $v_{i,1}$. This means that $f(x)$ has its first coordinate at most that of x. The same is true for the second and third coordinates. This means that therefore $f(x) = x$. $\qquad \square$

4. 2-ADIC VALUATIONS

The second ingredient of the theorem is number theory. We are interested in 2-adic valuations. First, what is a valuation? A valuation measures how big a number is. The absolute value is the most common example of a valuation. A valuation $|\cdot|$ must have the following properties:

(1) $|x| \geq 0$
(2) $|x| = 0$ if and only if $x = 0$
(3) $|x \cdot y| = |x| \cdot |y|$
(4) $|x + y| \leq |x| + |y|$.

Another example of a valuation is the 2-adic valuation.

Definition 4.1. Any rational number $x = p/q$ can be written as $x = 2^n \frac{a}{b}$ where a and b is odd. (Here, n might be negative.) The *2-adic valuation* of x is defined to be $|x|_2 = (1/2)^n$.

Example 4.1. For example, note that $|1|_2 = 1$, $|2|_2 = 1/2$, $|6|_2 = 1/2$, and $|20|_2 = 1/4$.

Remark. Note that an integer has 2-adic valuation less than 1 if and only if it is even. Otherwise, the 2-adic valuation is 1.

Example 4.2. As further examples, $|1/3|_2 = 1$, $|3/20|_2 = 4$, and $|13/16|_2 = 16$.

It is easy to check that this is a valuation. In particular, we will use the following properties:

(1) $|0|_2 = 0$
(2) $|mn|_2 = |m|_2 |n|_2$

(3) $|m + n|_2 \leq |m|_2 + |n|_2$.

(4) if $|m|_2 < |n|_2$ then $|m + n|_2 = |n|_2$.

This defines the 2-adic valuations of rational numbers. What if we wanted to define the 2-adic valuation of the real numbers? We need a magic wand to do this. In some cases, it is clear how to extend from the rationals to the reals. As an example, we have $|\sqrt{2}|_2|\sqrt{2}|_2 = |2|_2 = \frac{1}{2}$, so it makes sense to define $|\sqrt{2}|_2 = \frac{1}{\sqrt{2}}$. What about defining a value for $|\pi|_2$? In fact, this can be done consistently for all real numbers, using the axiom of choice.

We can analogously define p-adic valuations, but we don't need that now.

5. Proof of Monsky's theorem

We can color points in $\mathbb{R}^2$ according to the following sets:

$$S_1 = \{(x, y) : |x|_2 < 1, |y|_2 < 1\}$$
$$S_2 = \{(x, y) : |x|_2 \geq 1, |x|_2 \geq |y|_2\}$$
$$S_3 = \{(x, y) : |y|_2 \geq 1, |y|_2 > |x|_2\}.$$

Notice that classes 2 and 3 are translation invariant under points in class 1.

Lemma 5.1. *Let a triangle with vertices in $\mathbb{R}^2$ be complete. Then its area satisfies*

$$|area|_2 > 1.$$

This means that it has two in the denominator.

Proof. By translation invariance, we can assume that the triangle is at the origin, in which case we can write its area as a determinant. Let (x_2, y_2) and (x_3, y_3) be colored 2 and 3. Then,

$$\frac{1}{2} \begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix} = \frac{x_2 y_3 - x_3 y_2}{2}.$$

Furthermore, we have $|x_2|_2 \geq |y_2|_2$ and $|y_3|_2 > |x_3|_2$, where $|x_2|_2, |y_3|_2 > 1$. Therefore, $x_2 y_3$ dominates, having $|x_2 y_3|_2 > |x_3 y_2|_2$. This means that

$$|\text{area}|_2 = \left| \frac{1}{2} \right|_2 \cdot |x_2 y_3 - x_3 y_2|_2 = 2 \cdot |x_2 y_3|_2 = 2 |x_2|_2 |y_3|_2 \geq 2.$$

This proves the lemma. $\qquad\square$

We can now finish the proof of the Monsky theorem.

Theorem 5.2. *Given a dissection of a square S into m triangles of equal area, m is even.*

Proof. Without loss of generality, let the vertices of the square S be $(0, 0)$, $(0, 1)$, $(1, 0)$, $(1, 1)$. Consider a triangulation T of S, and just as in the above way, we can color the vertices of T. Observe that on the boundary ∂S, 12-edges can only occur on the edge connecting $(0, 0)$ and $(1, 0)$; furthermore, vertices colored 3 cannot occur on this edge, so there must be an odd number of 12-edges.

Therefore, by Sperner's Lemma, there is a complete triangle in T. So its area A has a 2-adic valuation $|A|_2 > 1$. But the area of S is $mA = 1$, so $|m|_2 < 1$, which means that m is even. $\qquad\square$

6. Generalizations and Closing Thoughts

Why do we care about equiareal triangulation? Yes, it is a cool problem with a cool solution. However, this problem has no clear applications to anything in the real world, or even in other parts of mathematics. Instead, we are interested in this problem because it connects seemingly disjoint areas of mathematics and inspires us to study things that we might otherwise ignore.

We want to finish with some results of the same spirit.

(1) We partition the n-dimensional cube and partition into simplices. The number of simplices must be a multiple of $n!$.
(2) We partition regular n-gons for $n > 4$. The number of triangles is divisible by n.
(3) For a centrally symmetric polygon, the answer is the same as for the square. This was conjectured by Stein and proven by Monsky in 1990.
(4) This seems somewhat paradoxical. Take a polygon and try to dissect it into triangles of equal areas. There are some polygons for which this can't be done. An example is the trapezoid with vertices $(0,0)$, $(0,1)$, $(1,0)$, and $(a,1)$, where a is not algebraic. This is similar to what we were doing before.

Now we mention some open problems.

(1) Suppose we have a centrally symmetric polyhedron in any dimension. Is it true that when we decompose it into simplices, there must be an even number of them?
(2) A polygon is called balanced if for every direction, the vector sum of edges of that direction is zero. An example is polygons with sides of only two directions. Triangles are not balanced. A balanced polygon is dissected into triangles of equal areas, there are an even number of them. This has been proven only in special cases: rational edges or edges with only two directions.

All of the known results have been done in essentially the same method. It would be interesting to find a new method to approach this.

E-mail address: moorxu@stanford.edu