What is Area?

Despite being such a fundamental notion of geometry, the concept of area is very difficult to define. As human beings, we have an intuitive grasp of the idea which suffices for our everyday lives. However, as mathematicians, we've only relatively recently been able to identify what we actually mean by the word. Efforts to define and generalise the notions of length, area and volume have led to the development of the branch of mathematics known today as *measure theory*.

Part of the difficulty in defining area lies in the fact that subsets of the plane can be quite wild in comparison to the objects we encounter in the physical world. In particular, there exist certain sets for which we cannot ascribe an area without disturbing one or more of the fundamental premises that govern measure theory. These so-called non-measurable sets make any efforts to characterise area rather complicated. To avoid such intricacies, let's propose the more modest task of defining area for polygons in the plane.

A seemingly elementary approach to the problem is to use the method taught to children in school. This involves drawing the shape on graph paper and then counting the number of squares which lie within the figure. This, of course, gives a lower bound for the area, while counting the squares which contain some part of the shape in question will yield an upper bound. The idea now is to consider graph paper with smaller and smaller squares, thus giving more and more accurate lower and upper bounds on the area in the hope that they will converge to the same number. It's this number which is defined to be the area. It turns out that this definition of area, although unsatisfactory for more complicated subsets of the plane, is well-defined for every polygon in the plane. In this way, we have constructed the area function which, given a polygon P, returns a real number $a(P)$.

It's quite unfortunate that this definition of area, even for polygons, should involve infinite processes and continuity arguments. Could we perhaps find an alternative definition, one which requires purely elementary techniques? A common trend in contemporary mathematics is to define an object by the properties which it satisfies. For example, we may like to extract enough properties of the area function defined above that there can be only one such function which satisfies them all. In fact, the four properties listed below perform such a task and we present them as the *area axioms*.

- *Non-negative.* If P is a polygon, then $a(P) \geq 0$.
- *Additive.* If P_1 and P_2 are polygons with no interior points in common, then $a(P_1 \cup P_2) = a(P_1) + a(P_2)$.
- *Invariant.* If P is a polygon and f is an isometry, then $a(P) = a(f(P))$.
- *Normalised.* If S is the square with sides of length one, then $a(S) = 1$.

I think you'd have to agree that any notion of area must satisfy these four axioms. A careful look at them leads us naturally to the following definition. We say that two polygons are *scissors congruent* if one can be cut into finitely many polygons which can be rearranged to give the other. In other words, scissors congruent polygons are common solutions to a jigsaw puzzle with the same set of pieces. Let's denote the fact that polygons P and Q are scissors congruent by $P \sim Q$. Also, note that we can extend the definition to unions of polygons with no interior points in common. In this case, we denote the union by $P_1 + P_2 + \cdots + P_n$, where $P_1, P_2, \ldots, P_n$ denote the individual polygons. The notion of scissors congruence, first for polygons, and then for polyhedra, will be our main topic of investigation.

Scissors Congruence in the Plane

Our exploration into the areas of polygons has naturally led us to the notion of scissors congruence. It's a simple consequence of the area axioms that if two polygons are scissors congruent, then they have the same area. Now it's only natural to ask the converse — if two polygons have the same area, then are they necessarily scissors congruent?

This question was answered in the affirmative by Bolyai in 1832 and independently by Gerwien one year later. Before embarking on the proof, let's begin by presenting a few important lemmas on scissors congruence.

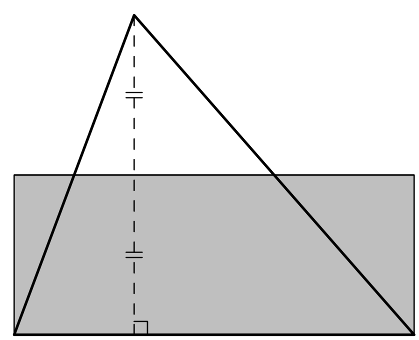
Lemma. *Scissors congruence is an* equivalence relation, *which means that*

- *for all polygons P, $P \sim P$;*
- *if $P \sim Q$, then $Q \sim P$; and*
- *if $P \sim Q$ and $Q \sim R$, then $P \sim R$.*

Proof. The first two statements are immediately evident from the definition of scissors congruence. For the third statement, suppose that we trace out the cuts required to decompose Q into pieces which rearrange to give P as well as the cuts required to decompose Q into pieces which rearrange to give R. Then cutting along all of these lines will yield a finite set of pieces which can be rearranged to produce either P or R. It follows that P and R are scissors congruent. □

Lemma. *Every triangle is scissors congruent with some rectangle.*

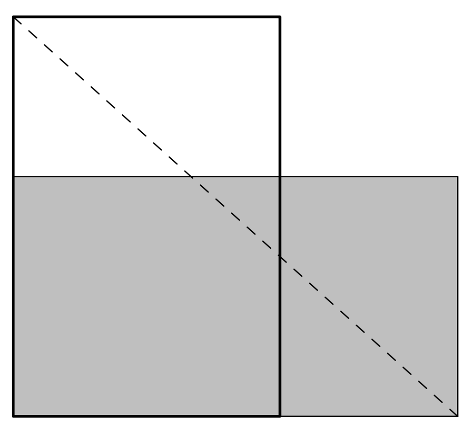
Proof. Cut along the altitude to the longest side of the triangle as well as along the perpendicular bisector of this altitude. This will decompose the triangle into two triangles and two quadrilaterals which can be rearranged to give a rectangle with the same base and half the height of the triangle. □



Lemma. *Any two rectangles with the same area are scissors congruent.*

Proof. Place the two rectangles in the plane so that they have one vertex in common, with two adjacent sides aligned as shown in the diagram. Cutting one of the rectangles along the lines shown produces a pentagon, a quadrilateral, a large triangle and a small triangle which can be rearranged to give the other rectangle.

The astute reader may have noticed that this particular decomposition may not always work if one of the rectangles is "too long". More explicitly, suppose that the rectangles have dimensions $\ell_1 \times h_1$ and $\ell_2 \times h_2$ where we may assume without loss of generality that $h_1 < h_2 \leq \ell_2 < \ell_1$. Then the construction works only when $\ell_1 \leq 2\ell_2$. However, this can be arranged by repeatedly cutting the longer rectangle in half and stacking the two pieces on top of each other until the condition $\ell_1 \leq 2\ell_2$ is satisfied. □



With these three lemmas under our metaphorical belts, we're now ready to prove the Bolyai–Gerwien Theorem.

Theorem. *Two polygons are scissors congruent if and only if they have the same area.*

Proof. First, note that any polygon P can be cut into finitely many triangles. Furthermore, each of these triangles is scissors congruent to some rectangle and each of these rectangles is scissors congruent to a rectangle, one of whose sides has length one. Therefore, we can write

$$P \sim R_1 + R_2 + \cdots + R_n,$$

where each R_i is a rectangle with one side of length one.

Concatenating these rectangles together produces a single rectangle R which has one side of length one. Of necessity, the other side of R must have length equal to the area of P. Therefore, if P and Q have the same area, they are both scissors congruent to the rectangle R, so they are scissors congruent to each other. □

Philosophically, the Bolyai–Gerwien Theorem tells us that it's possible to see whether two polygons have the same area or not using a finite process of cutting and pasting. This is mathematically more elementary — and hence, more satisfactory — than the usual definition of area which requires an infinite process involving graph paper with smaller and smaller squares.

What is Volume?

It's simple enough to develop a theory of volume for polyhedra in an analogous manner to the theory of area for polygons. The area axioms transform naturally into volume axioms to give a rigorous definition of the volume of a polyhedron. However, it had been noted by Gauss that proofs for the volume of a tetrahedron had all used in some way or another infinite processes and continuity arguments, rather than entirely elementary methods. Such an elementary proof would require that polyhedra with the same volume be scissors congruent — if two polyhedra have the same volume, then are they necessarily scissors congruent?

Hilbert considered this question of such importance that he included it in his famous address to the International Congress of Mathematicians at Paris in 1900. As is well known, his address included a list of twenty-three unsolved problems in mathematics which he considered to be of great significance. This question of scissors congruence in three dimensions was the third on his list.

It's clear from Hilbert's exposition on the matter that he didn't expect the Bolyai–Gerwien theorem to carry over from polygons in the plane to polyhedra in space — and he was exactly right. Hilbert's third problem was answered by his own student Max Dehn in 1900, the very year in which Hilbert had announced his list of problems. The crux of the proof lies in constructing ingenious invariants which have since been named after Dehn. Using these, Dehn managed to prove that the cube and the regular tetrahedron of unit volume are not scissors congruent. Philosophically, Dehn's result tells us that it's impossible to see whether two polyhedra have the same volume or not using a finite process of cutting and pasting.

Cantor and Cardinality

Naively, it might seem that counting and measuring are two of the simplest and most fundamental concepts in mathematics. On the other hand, if you examine them closely, you'll find that exploring these two areas can throw up a lot of seemingly nonsensical but mathematically true statements.

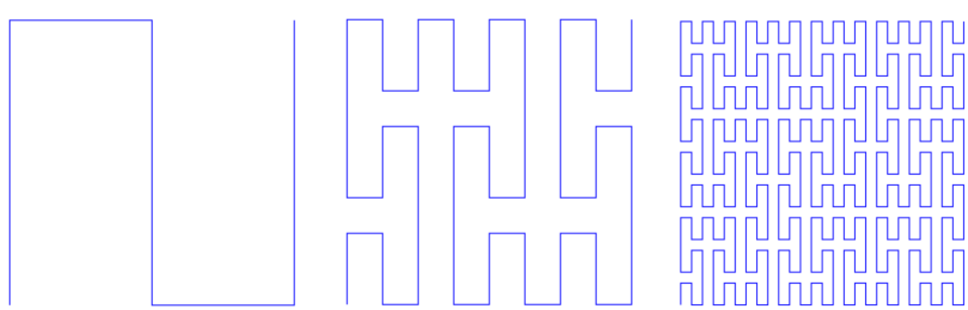
When you count the number of fingers of your left hand, you're just finding a bijection — that is, a one-to-one correspondence — from the fingers to the set $\{1, 2, 3, 4, 5\}$. If you wanted to convince someone that you had the same number of fingers on each hand, then you could simply count the number of fingers on your left hand and your right hand and find that the two answers are the same. Of course, this means that you would essentially be finding a bijection from the fingers of your left hand to the set $\{1, 2, 3, 4, 5\}$ and a bijection from the fingers of your right hand to the same set. Obviously, a far more economical way to perform the same task is to simply line the fingers from your hands together, so that every finger from one hand matches up with a unique finger from the other. In other words, you can convince someone that you have the same number of fingers on each hand by finding a bijection between the two sets. Therefore, we say that two sets have the same *cardinality* — in other words, the same size — if there exists a bijection from one to the other. It turns out that not only is this a sensible definition, but it's essentially the only sensible one possible.

This definition for cardinality was first proposed by Georg Cantor in the nineteenth century and used to demonstrate a number of pretty counter-intuitive facts.

- The set of positive even integers has the same size as the set of positive integers.
- The set of integers has the same size as the set of positive integers.
- The set of rational numbers has the same size as the set of positive integers.
- The set of numbers in the interval $[0, 1]$ has size greater than the set of positive integers. There is a famous and beautiful proof of this fact which uses a trick known as *Cantor's diagonal argument*.
- The set of numbers in the interval $[0, 1]$ has the same size as the set of points in a square with sides of length one.

More Nonsense

- *Peano curves.*
  Cantor's work shows that the set of numbers in the interval $[0, 1]$ has the same size as the set of points in a square with sides of length one. However, the usual bijection used to prove this fact is a rather crazy function with no nice properties. In the late nineteenth century, Giuseppe Peano managed to find a function from the interval $[0, 1]$ to the set of points in a square with sides of length one which is surjective — in other words, contained every point of the square in its range — and also continuous. This is an amazing fact, because it means that it's possible to draw a curve which passes through every point in the square at least once. One of the easiest ways to describe such a curve is as the limit of the following sequence of diagrams.



- *Continuum hypothesis.*

  Cantor managed to find many infinite sets which have the same size as the set of positive integers. We say that these sets have cardinality $\aleph_0$, which is pronounced "aleph-nought". He also managed to find many infinite sets which have the same size as the set of numbers in the interval $[0,1]$. We say that these sets have cardinality c, which stands for "continuum". In fact, Cantor found an infinity of infinities by showing that for any set with cardinality x, there is a set with cardinality greater than x. On the other hand, his search didn't turn up any sets whose cardinality was between $\aleph_0$ and c, which led Cantor to conjecture that none existed. This conjecture is now known as the *continuum hypothesis*.

  Rather amazingly, due to work by Kurt Gödel in 1940 and Paul Cohen in 1963, we now know that the continuum hypothesis can neither be proved nor disproved using the usual axioms — known as *Zermelo–Fraenkel set theory* — which form the foundation of modern mathematics. In other words, we may be able to add the continuum hypothesis or its opposite to our axioms and obtain a consistent system of mathematics. For this piece of work, Cohen was awarded the Fields Medal, often viewed as the top honour a mathematician can receive.

- *Banach–Tarski paradox.*

  In 1924, Stefan Banach and Alfred Tarski showed that it's possible to split a solid sphere into a finite number of non-overlapping pieces which can then be put back together in a different way to yield two identical copies of the original sphere. The reassembly process involves only moving the pieces around using isometries. However, the pieces themselves are complicated — they are not usual solid figures but look more like infinite scatterings of points.

  The Banach–Tarski theorem is often referred to as a paradox because it contradicts basic geometric intuition. Doubling the sphere by dividing it into parts and moving them around by isometries — without any stretching, bending, or adding new points — seems to be impossible, since all these operations preserve volume. This paradoxical decomposition shows that the notion of volume certainly can't apply to arbitrary sets of points in space.

Hilbert

David Hilbert was a German mathematician who lived from 1862 to 1943 and is recognised as one of the most influential and universal mathematicians of his day. He discovered and developed a broad range of fundamental ideas in mathematics and established rigour in much of the mathematics used in modern physics. Hilbert famuosly presented a list of twenty-three unsolved problems at the 1900 International Congress of Mathematicians in Paris. This is generally considered to be the most successful and deeply considered compilation of open problems ever to be produced by an individual mathematician.

Hilbert managed to put geometry on far more rigorous foundations than Euclid did. He proposed a much larger set of axioms which avoided the weaknesses which had been identified in the work of Euclid. His approach signalled the shift to the modern axiomatic method which is now considered fundamental to mathematics. The main idea is that although geometry may treat things which we have powerful intuitions about, it's not necessary to assign any explicit meaning to the undefined concepts. For example, points, lines, planes and other concepts could be substituted, as Hilbert says, by tables, chairs, glasses of beer and other such objects.

In 1920, Hilbert proposed a grand research project to reformulate all of mathematics on a complete logical foundation. Hilbert dreamed that all of mathematics could follow mechanically from a finite set of axioms and that one could prove that this system contained no contradictions. Unfortunately, this program was doomed to fail because Kurt Gödel remarkably managed to prove that Hilbert's dream was impossible.