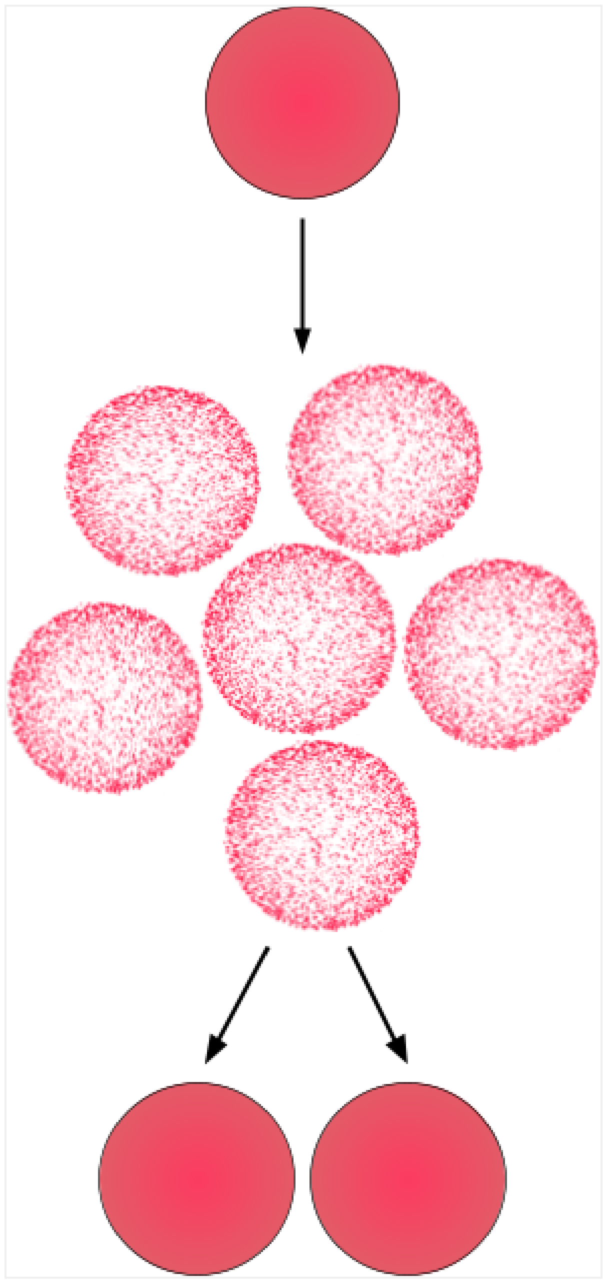
The Banach-Tarski Paradox

How 1 = 1 + 1



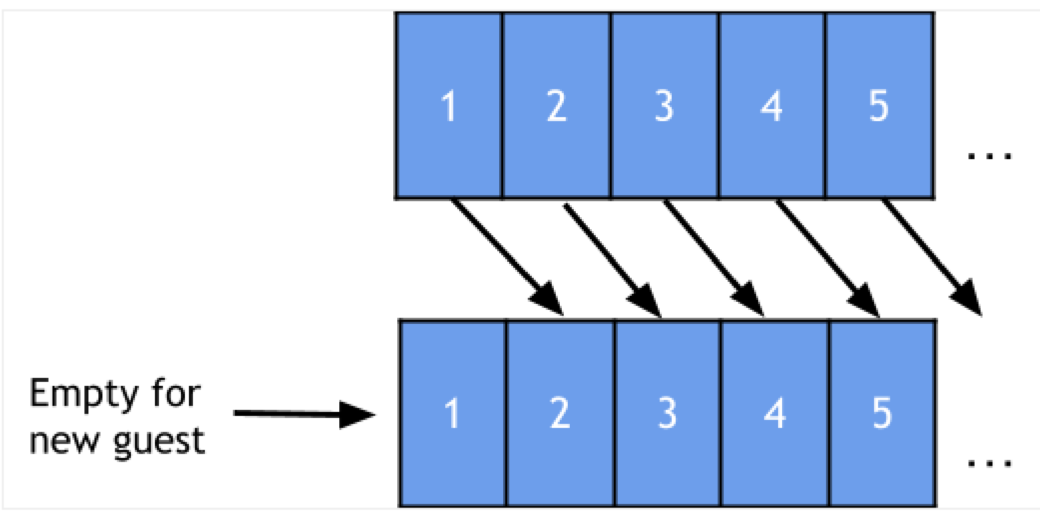
By Hassan Aly

1. Introduction

The Banach-Tarski paradox proposed by Stefan Banach and Alfred Tarski in 1924, proved that a sphere can be split into a finite number of pieces, which, after being only rotated and translated, can be reassembled into two identical copies of the original sphere. This mind-bending theorem counters our intuition through the manipulation of infinity and by seemingly violating the conservation of volume.
To better understand how it works, we will first discuss simpler concepts: Hilbert's Hotel, points on circles, sizes of infinity, and the Hyperwebster.

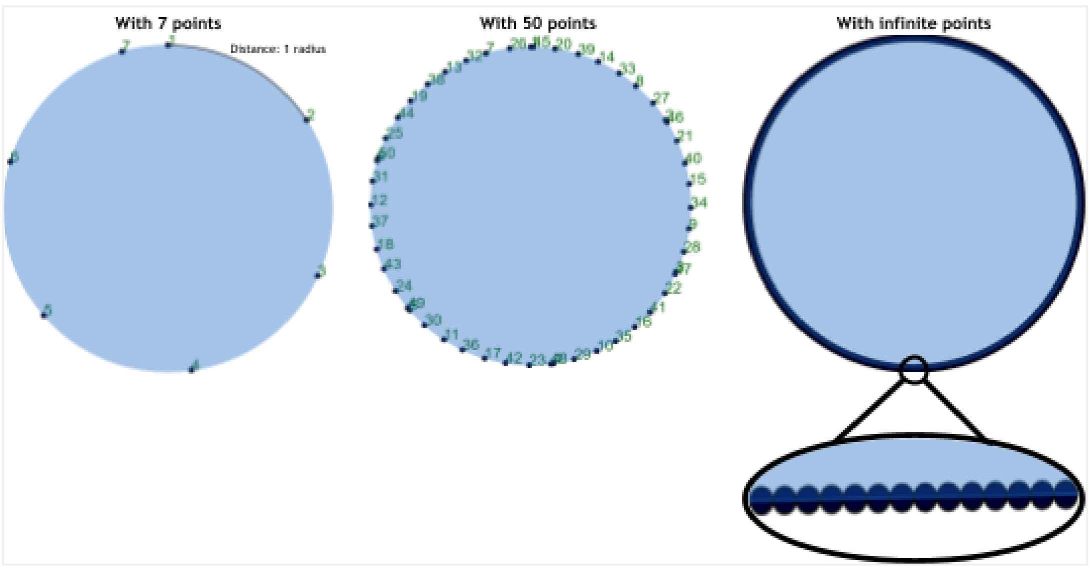
2. Hilbert's Hotel

This thought experiment proposed by David Hilbert presents the counterintuitive behavior of infinity through the concept of a fully occupied hotel with an infinite number of rooms. What's surprising is that any new guests that check in can still be accommodated. To do this, each guest simply moves to the next room, so the guest in Room 1 moves to Room 2, the guest in Room 2 moves to Room 3 and so on. If each guest moves from Room n to Room n+1, all current guests will have a room, and Room 1 will become free for the new guest. If a guest leaves, each guest can move down to the previous room, after which every room will still be occupied.

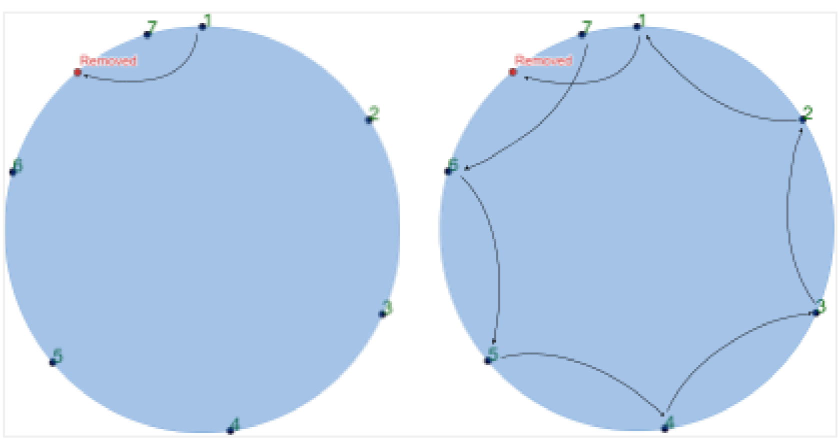


3. Points on a Circle

The circumference of a circle has a length of 2πr, so we can fit 2π radii on the circumference. Placing points on the circumference one radius apart allows us to place six additional points (2π rounded down) in a single revolution. We could repeat this process, adding more and more points. Even if we repeated the process infinitely many times, we would get a new point each time, never landing on a previous point. This is because the only way to land on a previous point would be to have moved by a multiple of 2π, which is impossible as π is irrational.



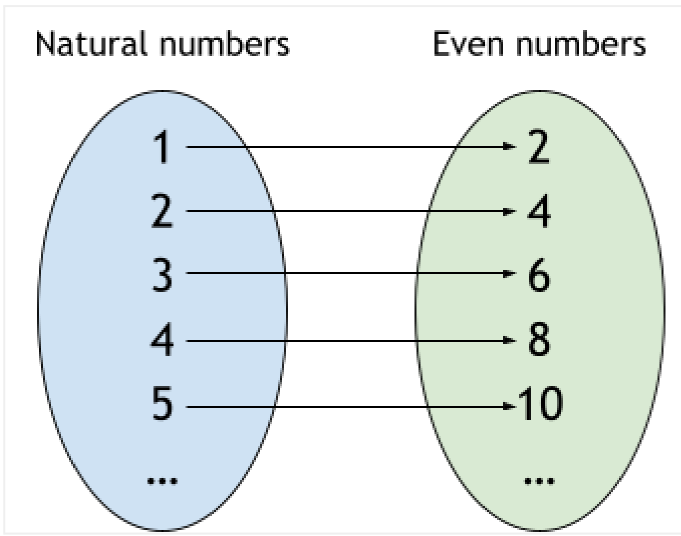
If we removed a point on the circle, for example the point to the left of point 7, we would have a gap in the circumference in the circle. However, this gap can be filled by moving point 1 into it. The gap left by point 1 can be filled by point 2, whose gap can be filled by point 3 and so on.



This scenario is similar to Hilbert's Hotel, as points moving locations are analogous to guests moving rooms. This situation is when a guest checks out, and each guest moves from Room n to Room n-1.

4. To Count or Not to Count - Different Sizes of Infinity

A set is simply a collection of objects, (called elements in set theory) usually using the notation {element 1, element 2, …, final element}. To say sets have the same size means that each element of one set can match 1-1 onto an element of the other set. For example, the sets {1, 15, 92} and {23, 3, 81} are the same size as we can match 1 with 23, 15 with 3, and 92 with 81, and so they must both have the same number of elements. For small, finite sets this seems obvious, however with infinite sets, it can lead to unexpected results.
For example, the elements in the set of all even numbers {2, 4, 6, 8, 10, ...} and the set of all natural numbers (excluding 0) {1, 2, 3, 4, 5, …} can all be matched 1-1.



We can conclude that the number of positive even and odd numbers is exactly equal to the number of positive even numbers, despite our intuition leading us to think it would be double. This was proved by a mathematician called Georg Cantor, who described two different sizes of infinity: countable and uncountable infinity. A countably infinite set is one which has a 1-1 correspondence with the natural numbers, such as the set of even numbers. An uncountably infinite set cannot be matched 1-1 with the natural numbers.
An example of an uncountably infinite set is the set of all real numbers. Cantor proved this through his diagonal argument, which demonstrated that just the real numbers between 0 and 1 are uncountably infinite.
Cantor began by assigning reals (written in any order) between 0 and 1 a natural number:
1 ↔ 0.47328914713289473489323242
2 ↔ 0.23940390483409324214393234
3 ↔ 0.89327489468927423987434895
4 ↔ 0.32743289520923874397654294
5 ↔ 0.31947328947912747342384758
…

While it seems we have matched all the real numbers between 0 and 1, Cantor proved that the list is still missing some real numbers by constructing a new number not on the list.
To construct a new number, start by taking the nth digit (after the decimal place) of the nth real:

1 ↔ 0.47328914713289473489323242
2 ↔ 0.23940390483409324214393234
3 ↔ 0.89327489468927423987434895
4 ↔ 0.32743289520923874397654294
5 ↔ 0.31947328947912747342384758
…

This gives us the number 0.43347. We now need to change each digit (e.g. by adding 1), giving us 0.54458. This number cannot be part of the original list, as it is guaranteed to differ from each number listed by at least one decimal place (the nth decimal place for the nth number). This number has not been matched to a natural number, proving that we cannot match all the real numbers to the natural numbers, and hence there are uncountably infinite real numbers, even just between 0 and 1.

5. From Sets to Alphabets - the Hyperwebster

The Hyperwebster is a theoretical dictionary proposed by Ian Stewart in 1996. Stewart described the dictionary as infinitely long, containing every single word that could be constructed from the 26 letters in the alphabet. This includes real words as well as words without any meaning (the vast majority of words in the Hyperwebster). The Hyperwebster is ordered in the following way:

Starting with 1 letter words:
- Start with A
- Cycle through the alphabet until Z is reached
- Increase length of word by 1
- Repeat

The Hyperwebster would look like this:
- A, B, …, Z
- AA, AB, ..., ZZ
- …
- ZZZZZZZZZ… (Zs repeat forever)

The Hyperwebster can be split into 26 volumes, each containing only words that start with a specific letter.

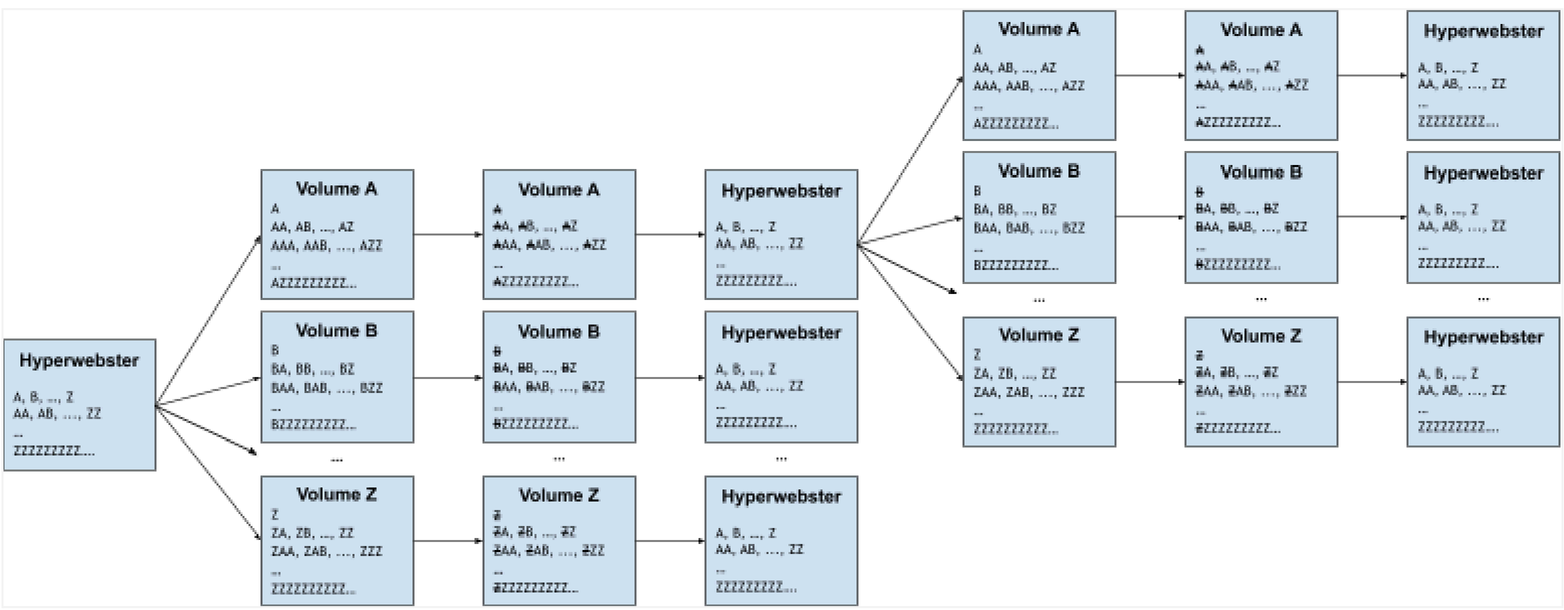
For example, Volume A contains:

- A
- AA, AB, …, AZ
- AAA, AAB, ..., AZZ
- …
- AZZZZZZZZZ… (Zs repeat forever)

The leading A can be removed from all words as the reader knows to add it on.
Volume A now becomes:

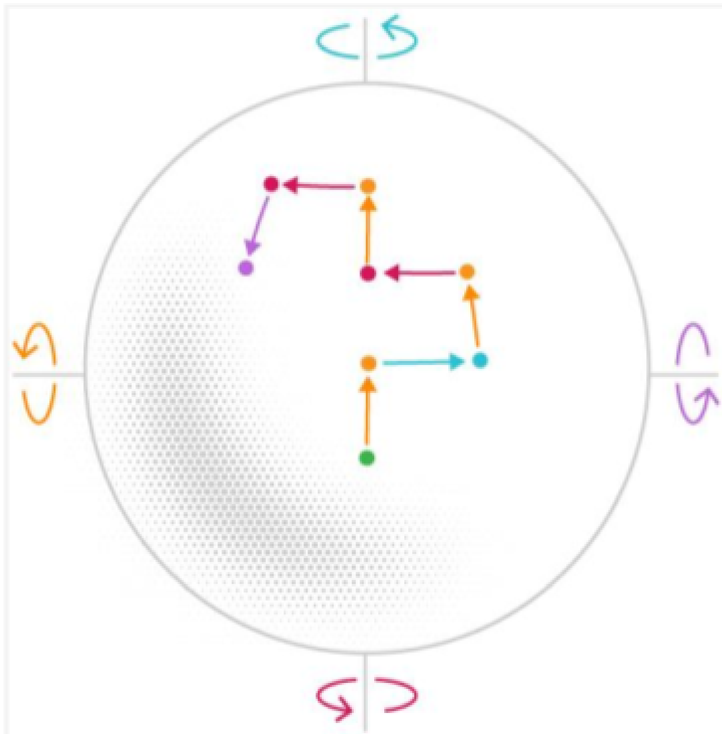
- [blank]
- A, B, …, Z
- AA, AB, ..., ZZ
- …
- ZZZZZZZZZ… (Zs repeat forever)

Volume A is now identical to the complete Hyperwebster. This process can be repeated infinitely, with each volume being split into a further 26 volumes.



6. Back to Banach-Tarski

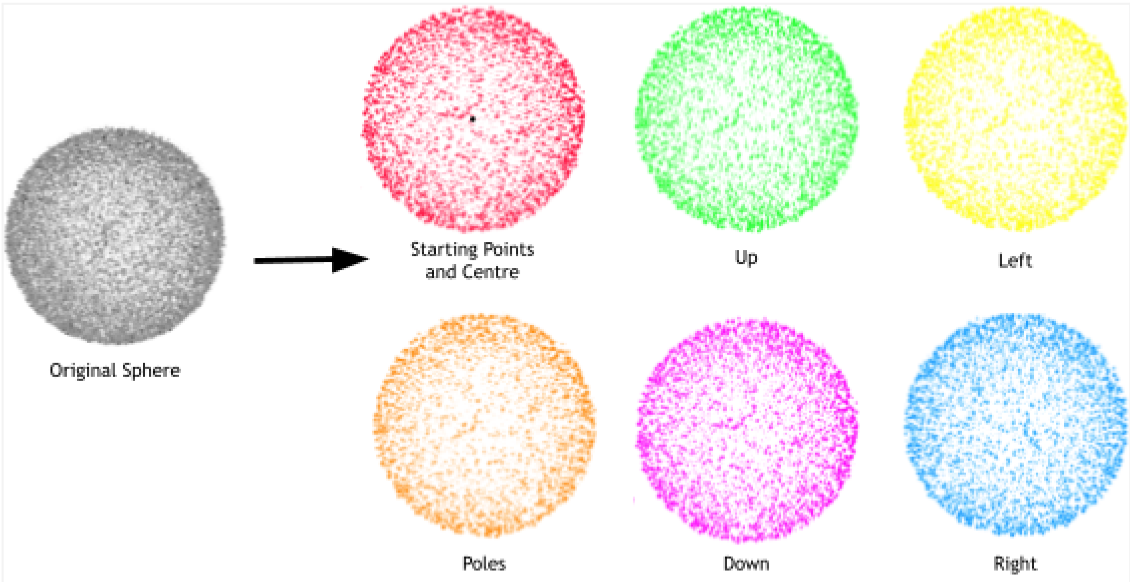
To achieve the Banach-Tarski paradox, each point on a sphere must be named and placed into a group. These groups will then be transformed and reassembled into two copies of the original sphere. To name these points, we choose a starting point on the surface of the sphere, and rotate the sphere either up (U), down (D), left (L) or right (R) by an angle of $\cos^{-1}(\frac{1}{3})$.



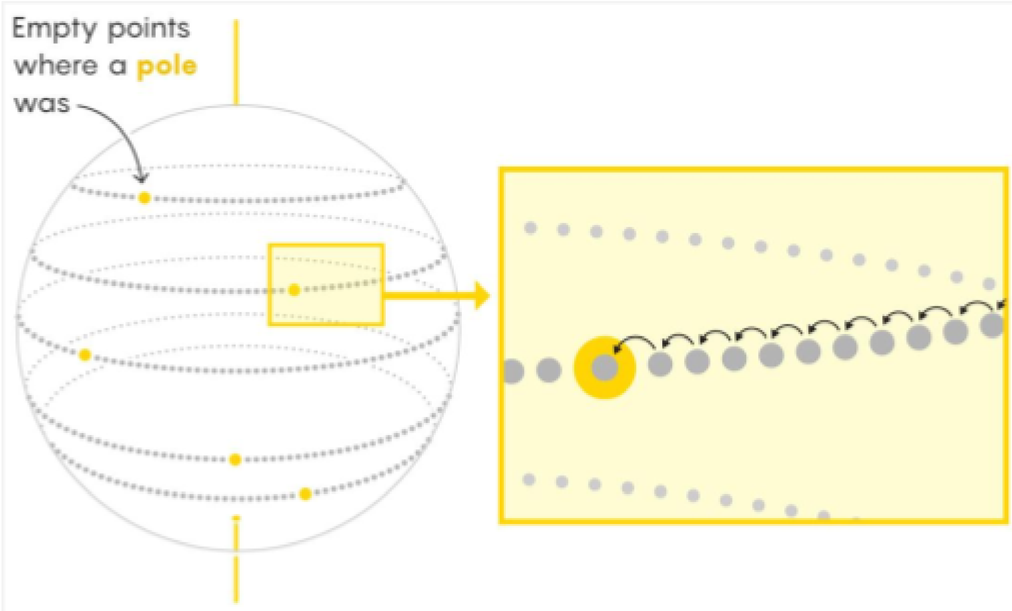
(Source: Quantamagazine.org)

This angle (≈70.53°) is used to ensure that any number of rotations in any direction always reaches a new point on the sphere's surface (assuming we do not backtrack - LR/RL/UD/DU). Each point is named by the rotations needed to reach them from the starting point (the name is written right to left, with the most recent transformation being the first letter).
Listing out all possible finite sequences would give a countable number of points. However the sphere contains an uncountable number of points, so we must choose a missed point to become a new starting point, and get another countable set of points. In order to get every point we would need to choose an uncountable number of starting points.
We have now uniquely named all surface points, except for poles. These are a countable number of points along the axis of rotation, meaning their positions are unaffected by rotations. This means rotations such as R and RR would lead to the same pole, giving them multiple names. To resolve this we move these points into a separate group.
We currently only have points on the sphere's surface. To get all points inside the sphere, each surface point can correspond to a line of points going to the centre of the sphere. We will move the centre to the pole group to prevent it from being repeated.
The rest of the points can be split based on the final rotation into the U, D, L and R groups.

Along with a group of all the starting points, the group of poles and the centre point, we now have 6 groups containing all the points of the original sphere, which we will rebuild into 2 copies of the sphere after a few transformations.



If we chose the U group and rotated all its points down, its points would start with DU…, which can be removed as they cancel out. For example, U, URRDDL, UUUDL and ULDDL would become ~~DU~~, ~~DU~~RRDDL, ~~DU~~UUDL and ~~DU~~LDDL. Doing this gives us a list of points starting with U, R, L or no rotation at all, turning the U group into the U, R, L and starting points group, similar to how we turned Volume A of the Hyperwebster into the full Hyperwebster. Combining the rotated U group with the D group, poles and centre point of the original sphere, we get a complete sphere.
We still have the original group of starting points as well as the L and R groups, from which we can make the second, identical sphere. Applying an R rotation to the L group turns it into the L, D, U and starting points groups combined. However we already have the original group of starting points, so we do not need another copy of them. To prevent copying them, we can move the point L (which would turn into a starting point when rotated) into the R group. However the point LL would turn into L, giving us an L in both the L and R groups. To avoid this, all points which are just a sequence of Ls are moved into the R group. Applying an R rotation now allows us to turn the L group into the L, D and U groups combined. Adding the R group (with the L sequences we added to it), and the group of starting points, we now have a sphere missing its poles and the centre point.
Since there are a countably infinite number of pole points missing, we can choose an axis such that each missing point is on its own circle going across the surface of the sphere. We now have a countable number of circles each with a single point missing. As we saw before, we can simply move other points from the circle to fill in the gaps. The same method can be done with the centre by creating a circle inside the sphere and again shifting points around the circle to fill the gap. And the second sphere is complete.



(Source: Quantamagazine.org)

The Banach–Tarski paradox, which seemingly allows the infinite duplication of objects, is not physically possible. The theorem only holds for ideal spheres as it relies on splitting them into non-measurable sets using infinitely precise divisions. Physical objects - being made of discrete atoms - cannot be split so accurately, and always have a measurable volume. Despite its lack of real-world application and counterintuitive nature, the Banach-Tarski paradox remains a fascinating result of the axioms of mathematics.