Equivalence

Contents:

1. "History" of sets

R

Everybody knows that points are the basic concept of geometry, which was the main concern of the greek or euclidian mathematics. Less known is that the concept of sets were only introduced at the end of the 19-th century by Cantor and this revolutionized the whole mathematics. This didn't go smoothly! Some quite influential mathematicians attacked Cantor and the recognition came very late and he died in a mental hospital in 1918. At this time it was already unquestionable that he made the most original steps in the history of mathematics. By today the word "set" became almost household expression and the idea that we can talk about sets in general without specifying what they are is accepted. In one sense the stupid Venn diagrams that appear in education books are the most annoying misrepresentation of Cantor's ideas. Indeed the whole breakthrough was the special relations of infinites and these diagrams only show the trivial relations of containing. Still, the wider context, namely the fact that all sets obey the same rules induces a subconscious awakening of our a-priori intuitions. This might sound freaky but I already expressed these views and want to repeat again, I believe: We possess all future intuitions and there are no paradoxes only yet unawakened intuitions that resolve them. Everything that is true will become natural for us too. From this, it is only one step to say that God created us to understand the world he created, but I didn't say that. The real problem is to see in retrospect why we had the delays and sticking to our subject then the question is obvious: Why couldn't Euclid discover the concept of sets? After all, regarding the geometrical objects as sets of points is so natural. When lines or circles cross they share common points. The answer in full detail can be researched by looking at Hilbert's geometry. Hilbert, the biggest supporter of Cantor's discoveries put his mind to where his mouth was and actually re-axiomatized Euclid's geometry based on sets. In fact he made the first exact axiomatic geometry because it turned out that Euclid's system was not logically perfect. And yet Hilbert's geometry is a monstrosity! It couldn't be used in schools! It hides more basic principles then it reveals. This contradiction is at the heart of any attempt for total axiomatization, or in a wider sense is a contradiction of formalism. Returning to the question why Euclid didn't discover sets, there are two reasons. One was a true limitation of his level of abstractions, while the other exactly oppositely his unlimitation to jump through the sets to a higher abstraction. To tell how space is a collection of points you would have to specify how the points are located on a line or plane and so on. The obvious facts like that between two points there are always others don't tell explicitly how they are "following each other". Clearly there can't be a next point to any, because as we just said between any two there are new ones. But this feature in itself is not enough. Indeed, if we look at the dividers of a distance that is its half, third, two third, **. . .** or in short the rational points, then among these we also have the feature of "density". For example the halving point of the interval is obviously approached by an infinity of other fractions from both sides:

$\frac{1}{3}, \frac{2}{5}, \frac{3}{7}, \frac{4}{9}, \ldots \to \frac{1}{2}$ similarly from above. But these rational dividers are not the totality of points and the greeks knew this very well. Indeed, they proved that $\sqrt{2}$ is not a fraction, so by the Pythagoras theorem the hypotenuse of a symmetrical right angle unit triangle can't be a fraction either. They were puzzled by this fact and probably they also felt that these non rational or irrational distances are not merely holes on the line. Indeed, these must be at least as many as the rationals, because any fractions or multiples of such irrationals are again irrationals.

For example if $\sqrt{2}\,\frac{n}{m}$ were a fraction $\frac{N}{M}$ then $\sqrt{2}$ were too, namely $\frac{N}{M}\,\frac{m}{n}$.

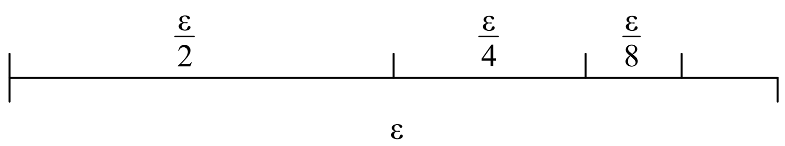
This indirect argument is mainly appealing to our logic but not to our vision and gives the impression that rationals and irrationals are two similar "set of points".

I will show that this impression is false but first I want to show in physics how puzzles don't go away just because we are in possession of our abstractions today. In fact the abstractions make it easier to get into paradoxes. We are now 500 years in the history of Newton's mechanics but still everybody must start from the level before Newton. The education system has no intention helping in this process because its goal is reproduction, spreading formalism. My favorite pastime was to ask freshly graduated physics teachers: "Why the moon doesn't fall to the earth?" For a pre-newtonian this is not a provocation of paradox because he can ignore gravitation. But if we learned that the earth attracts the moon just as it attracts an apple, then our logic says that it should also fall similarly. The physics teacher has the disadvantage that his mind was filled in with formal junk as well, so he will search his memory for resolving the problem and come up with answers like: "the centrifugal force compensates the gravitational pull". This even "feels right" because the moon is circling the earth while the apple is not. Then of course I usually said: "so the total force is zero on the moon, then why doesn't it fly away in straight line as Newton's first law requires? Then usually aggression awakens and hostility to hide stupidity. But there are some graduates who tell the correct answer, namely that: "I'm sorry but your question was wrong because the moon does fall to the earth. It has been falling for a very long time and will keep on." Not all who tell this really understand this but most do. Newton tried his best when put the drawing in his book with a cannon on mount Everest, shooting further and further away and gradually changing "falling to the earth" into "circling the earth".

Returning to math, if I ask what is the difference between the ordering of rationals and irrationals, then by common sense, the student will try to find some local difference in how the points are separated. But there aren't such local differences, so instead of the admission of failure and awakening of true questions, some learnt junk jumps in about how "the irrationals are more than the rationals", which are just a "dense sequence" and so on. The total set of rationals and irrationals that is the continuous full distance is on the other hand definitely different <u>locally</u> from both the rationals and irrationals because no matter where we cut it into two segments, either the left or the right must have an edge "last" point. Or to put it in an other way, there is no hole on the distance. In spite of this being obvious, we can't even regard it as an axiom without sets. Indeed, only Dedekind, the first supporter of Cantor stated it as a rule that can define the points of the line. The greeks didn't clear this vital difference between the full line and their subsets that have holes. Not only because they were more interested in the relation of rationals and irrationals but because they were even more interested in how a further special length, the constructible distances relate to these. The above mentioned hypotenuse of the triangle is obviously constructible with a ruler and compass, so clearly there are irrationals that are constructible, but are they all? They didn't raise this question either and instead only looked at special requirements like to find a side that gives a square with equal circumference or area, with a given circle's. So they asked: Can we construct this side from the radius of the circle? Yet even here, the assumption that there is such distance at all, was not conscious. That's how deep the concept of continuity was buried yet. Today we would say that by changing continually the side of a square, its circumference or area changes continually too, so it must take up the in between values. In spite of these "excuses" it is still a mystery to me how, not only the greeks but all the great mathematicians before set theory couldn't find the following very simple yet amazing argument that shows the drastic difference between rational and irrationals: The rationals or the dividers of any d distance can be covered with an arbitrary small ε distance! For simplicity regarding $d = 1$, we have to cover the simple fractions.

These can be sequenced by increasing denominators: $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6} \ldots$

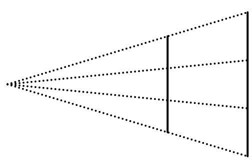
Lets cover them successively with the: $\frac{\varepsilon}{2}, \frac{\varepsilon}{4}, \frac{\varepsilon}{8}, \frac{\varepsilon}{16}, \frac{\varepsilon}{32}, \frac{\varepsilon}{64} \ldots$ which add up to ε:



Even if these covering intervals wouldn't overlap we would have $1-\varepsilon$ length uncovered which contain only irrationals. In short: the rationals in total are less than any ε while the irrationals are more than $1-\varepsilon$. This argument of difference would be of course much less impressive if the remaining irrationals could also be covered by an ε interval. But this would mean that the full 1 length could be covered by a smaller 2ε length. Though this seems impossible, soon we'll see that it is not that far fetched.
The other side I mention, that is the abstraction beyond sets was the "straightness" of lines. Since the line was the basic concept of geometry, they had to struggle how to define it and they realized it is impossible from within. We just have to accept it as basic, undefined and try to list the rules that apply to it. We can't even smart our way out of this by regarding the line as a set of points. Indeed, a curved line is also a set of points! What makes the line straight has nothing to do with it being a set. So Hilbert had to inject his set theoretical geometry with the old "life juice" taken from Euclid. Okay, so geometry is not the perfect candidate for applying set theory but still there should have been some people hitting upon the right ideas even if not fully. Well there were such people and one famous such naïve wonderer was Galileo. Instead of going from the right approach to the law of falling and measuring the fallen distance from the beginning, he pursued his idea of fallings in consecutive equal time intervals. As we know today, the fallen distance is g = 9.81 times the square of the falling time that is $d = g\,t^2$. So if in a first t time the fall is $d_1 = g\,t^2$ then in the second t time it will be $d_2 = g\,(2\,t)^2 - g\,t^2 = g\,4\,t^2 - g\,t^2 = 3d_1$ then under the third t time $d_3 = g\,(3\,t)^2 - g\,(2\,t)^2 = 5\,d_1$ and so on we get the odd number multiples of d_1. This is obvious because $(n+1)^2 - n^2 = n^2 + 2\,n + 1 - n^2 = 2\,n + 1$. But Galileo instead of going back from the odds to the squares, tried to "figure out" why the odd numbers appear and somewhere along these wonderings he "realized" that the odd numbers are just as many as the naturals:

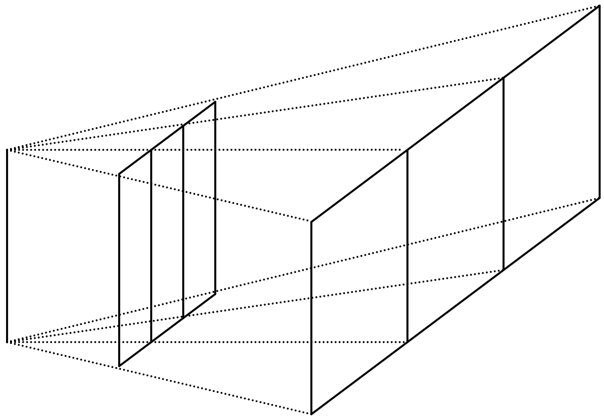
1	2	3	4	5	. . .
↕	↕	↕	↕	↕	
1	3	5	7	9	. . .

Galileo discovered neither the law of gravitation nor the equivalence of infinite sets but he sure had a spin with both. In the end this is all that counts!
The fact that all mathematics can be based on sets but not all fields are becoming natural candidates might indicate that something is missing. When mathematical logic stepped in marriage with set theory then this duality became exact too. Even point sets became abstract in a decade and topology was born. This is the field where formalism did the most harm. Everybody wants to be "general" so the pictures disappear and words have second meanings. Actually a split happened. In the so called set theoretical topology spaces are defined and conditions are narrowed gradually to imitate the real space. The other direction is called geometrical topology but here the word geometry means "shapes". The field that we present in the followings is not only the common base of the two topologies but even of general Set Theory because point sets represent the basic examples of sets. The common feature of all these four subjects is an oppositeness to geometry in the sense that while geometry is purposely dealing with distances, in these fields we try to go beyond the exact measure of sets.
One obvious way to avoid distances is to just order the sets to each other point by point. This one to one correspondence is called equivalence of the two sets, like Galileo's odds with the full set of naturals. Such point by point ordering can clearly change the sizes, as the following projection of a distance into a bigger one shows it:



A natural step from this is the following approach: We can keep some rigid subsets of a set with their distances but try to move these different subsets and reassemble them.

We expect that this “jigsaw” principle would clearly keep the full size of the sets but we are wrong. Indeed, without telling what kind of subsets should remain fixed, this method can lead to the same change of size as the above projection. For example, a rectangle in space can be projected to one with the same height but different width:



Here the vertical heights are the rigid subsets that are reassembled into the larger rectangle.
Amazingly, this failure of the jigsaw or reassemblence principle to keep the size can happen even when we use only finite many pieces. The Sierpinski-Mazurkiewicz paradox will show this for a strange plane set, while the even more surprising Banach-Tarski paradox, for a simple sphere or ball. This might suggest that in one dimension, that is on the line everything is alright but actually there are strange sets here too. In fact one of these was historically the first, discovered by Vitali.
The size paradoxes have a connection with probabilities that makes them even more striking. The fundamental idea is this: If we shoot points to a target point set and it has two identical sub sets then these must have the same probabilities of hitting into them. Hitting into the combined two sets must have twice of this probability. But if this combined set can be moved into just one half then the half must be the same as the whole. This still seems just as a contradiction in measures but if we regard it as the frequency of hits in a shooting sequence then it seems physically impossible.

R

Equivalence as physically practical use for finite sets

Imagine that in a ballroom we want to know whether there are more boys or girls. Instead of the tedious way of counting them both, we can just ask them to form pairs and see at once which sex remains without pair. How this argument never initiated the more general concept of equivalence, is again a mystery, just as the earlier mentioned late discovery of the ε cover for rationals.

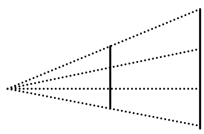
2. “Details” of sets

R

The basic principle of abstract sets is a strange duality.
On one side we disregard all the structure, in other words the relationships that exist among the elements and merely look at them as a collection. Even mathematical language can’t cope with this level of abstraction because when we give a set element by element like $S = \{s_1, s_2, s_3\}$ we list it in a given order. The unimportance of order of course can be expressed by saying that $\{s_1, s_2, s_3\} = \{s_2, s_1, s_3\} = \ . \ . \ .$ With an analogy, this side means that a motorcycle Set Theoretically is the same as all of its parts thrown in a basket. So it’s the total oppositeness of what we always hear about the organic nature of all things, that the whole is more than just the mere collection of its parts. In Set Theory we intentionally want to be inorganic!
On the other hand we allow sub collections to be elements. So we can put the two tires of the motorcycle in a basket and then this can be put with other parts in a larger basket. This will be regarded as different collection. So $\{s_1, s_2, s_3\} \neq \{\{s_1, s_2\}, s_3\}$. This second side is the correction of the stupidity of the first, a minimal flexibility to allow some primitive structuring.
The big result of Set Theory is then that: All complicated situations that arise in mathematics can be reduced to this primitive structuring. For example, the build up of a motorcycle should be regarded as a multi leveled collection of parts in larger and larger baskets. Then we see that this is not so stupid! Though we can’t talk about interrelations between parts, the mere collections can show the interrelating parts, even their relation to other sub collections. Medical students know that anatomy is the foundation of all organic functions. So the mentioned non organicness of Set Theory can be rephrased as: Set Theory is the anatomy of mathematics.
The described basic duality of sets in practice also means a technical challenge when changing real problems into the language of sets. Yet the biggest new concept, equivalence can be seen without the technical details. That’s exactly how we introduced it in the previous section.
As we mentioned, even Galileo stumbled upon the fact that the natural numbers are equivalent to their “half” subsets:

$$\begin{array}{ccccc} 1 & 2 & 3 & 4 & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 1 & 3 & 5 & 7 & \dots \end{array}$$

We also saw that distances are equivalent to each other regardless of their lengths:



These strange stretchings of infinites hide a much more amazing fact. Namely, we can prove that some infinite sets can not be ordered to each other element by element. In other words there are nonequivalent sets that is different infinities! This is the heart of Set Theory. Equivalence is a one to one ordering between the elements of two sets. But in proper terminology the word “ordering” here should be replaced by relation because the word order is used for the internal ordering of one single set. Thus, relation is the most general term for “ordering” elements to each other without any restriction. The exact definition of relation thus should be any set with elements that are pairs of two elements. This has two problems! First of all we should have an order of the two elements, to know which one is the first and which is the second. Secondly we should be able to allow relating an element to itself. For example in the above equivalence of naturals we had the $1 \leftrightarrow 1$ pairing. But how could we form a {s,s} pair when it is the same as {s}. The solution is simple! We should differentiate between the two elements of a pair by adding one of them, say the first as an element. So instead of {s,t} we regard {s,{s,t}} which is called the ordered pair of s and t and is abbreviated as (s,t). This solves both of our problems:
We’ll know that the set standing alone is the first. Also, (s,s) = {s,{s,s}} = {s,{s}} $\neq$ {s}, so we can form pairs from a same element. This {s,{s,t}} definition of ordering suggests an instant generalization to triplets or any n-tuples as the increasing or widening collection of subsets.

But there is an other simpler way too, namely to use the pairs: $(s_1, s_2, s_3) = (s_1, (s_2, s_3))$.
Strangely, this definition, which is the accepted custom, still will support the previous more logical widening idea. This can happen because we'll use this to define the most basic order, the natural numbers: 1 , 2 , 3 , . . . A nice definition would be if the n natural number were a concrete set having n many elements and the smaller numbers were used as elements. But then 1 should have a single element as what? Lets introduce the 0 empty set that has no element at all, then: $1 := \{0\}$, $2 := \{0,1\} = \{0,\{0\}\}$, $3 := \{0,1,2\} = \{0,\{0\},\{0,\{0\}\}\}$, . . .
The really nice thing were if this representation of naturals would also give the operations, + , • , power. This is impossible to achieve. Set Theory is only the anatomy of mathematics! The physiology and pathology still remains outside. But there are good analogies for the operations. The two basic set operations are the combining of all elements of two sets, that is the $S \cup T$ union and collecting of the common elements, that is the $S \cap T$ common part or intersection. Here again we see use of the empty set because, while the $S \cup T$ always has elements if S,T had, the $S \cap T$ can be empty if S,T have no common ones. So, using 0 means that $S \cap T$ is always defined. $S \cup T$ is simpler in an other sense too, namely it really corresponds to the addition. Indeed, 3 + 5 = 8 means that a 3 element set combined with a 5 element one gives an 8 element one. But there is some difference too because 5 + 5 = 10 while $S \cup S = S$. More important is that $\cap$ can't correspond to the multiplication at all because $S \cap T$ is smaller than $S \cup T$ while the product should be bigger than the sum. The clue to what set operation the product should really be, can come from our memories when we multiplied bracketed sums. As we remember every thing had to be multiplied with everything, so:
$3 \cdot 5 = \{0,1,2\}\ \{0,1,2,3,4\} = \{(0,0), (0,1), (0,2), (0,3), (0,4), (1,1), \ldots, (2,4)\}$
Indeed, this has 15 elements. So this, forming of all possible pairs, should be the product of sets, $S \cdot T$. To multiply more sets we have to define the n-tuples and here shows the advantage of reducing these to pairs, because then products can be simply built up:
$2 \cdot 3 \cdot 5 := 2 \cdot (3 \cdot 5)$. Of course, this way the bracketing can only be omitted as abbreviation.
Lets remember that the pairs were introduced to define the fundamental concept of equivalence. Now it turned out that the set of all pairs from two sets give this beautiful algebraic meaning. There is a three step narrowing from this algebraic meaning right into to the crucial equivalence. If we just regard any R subset of ST, then it can be called a relation between the S and T sets. Indeed, R tells those pairs that are related. If R is such, that for any $(s,t) \in R$ there is no other u that $(s,u) \in R$, that is the second member is unique, then we call the relation a function and usually denote it with f. The unique second member can be written as $t = f(u)$. Finally, if the first members are unique for the second too, then f is a one to one function or equivalence. Though the relation, function and equivalence were obtained as subsets of an ST product, it doesn't mean that all S or T elements appear. The actually used S and T elements are called the domain and the range of the relation, function or equivalence. After these remarks, the exact definitions speak for themselves.

D

1.) Subset

$S \subseteq T$ S is subset of T if $s \in S \rightarrow s \in T$
$S = T$ S is equal to T if $s \in S \leftrightarrow s \in T$
$S \subset T$ S is proper subset of T if $S \subseteq T$ and $S \neq T$
$0 := \{s ; s \neq s\}$ $0 \subseteq T$ for every T.

2.) Operations

$S \cup T := \{s ; s \in S \vee s \in T\}$
$S \cap T := \{s ; s \in S \wedge s \in T\}$
S and T are disjoint if $S \cap T = 0$.
$(s,t) := \{s,\{s,t\}\}$
$S \cdot T = ST := \{ (s,t) ; s \in S \wedge t \in T \}$
This can be visualized as an S set in which every s element is replaced with a T set, but to distinguish these, all t elements are preceded with the s, it was replaced in.
$S_1 S_2 \ldots S_n := S_1 (S_2 \ldots S_n)$

3.) Functions

R is relation if $R \subseteq ST$

f is function if it is relation and $(s,t) \in f \wedge (s,u) \in f \rightarrow t = u$.

f is equivalence if it is function and $(s,t) \in f \wedge (u,t) \in f \rightarrow s = u$.

Do R = domain of R : = $\{s \,;\, (s,t) \in R\}$

Ra R = range of R : = $\{t \,;\, (s,t) \in R\}$

$S \sim T$ if there is an f equivalence with Do f = S and Ra f = T.

S^T : = $\{f \,;\, \text{Do } f = T \wedge \text{Ra } f \subseteq S\}$

This can be visualized as S sets multiplied repeatedly T many times, that is s elements are picked with t "time" specifiers preceding.

4.) Naturals

$0 := \{s \,;\, s \neq s\}$

$1 := \{0\}$, $2 := \{0,1\}$, $3 := \{0,1,2\} = \{0,\{0\},\{0,\{0\}\}\}$, . . .

$\omega := \{0, 1, 2, 3, \ldots\}$

S^n can also be imagined as the set of all n-tuples picked from S.

$S^{<\omega} := S \cup S^2 \cup S^3 \cup \ldots =$ The set of finite tuples picked from S.

S^{ω} can also be imagined as the set of all $\{s_1, s_2, \ldots\}$ sequences picked from S.

$2^{\omega} = \{0,1\}^{\omega}$ has a special meaning because the 0 , 1 sequences are the infinite binary numbers that can define the points of the [0,1] interval. This is the Continuum.

5.) Subsets

P (S) = power set of S : = the set of all subsets of S including S itself.

$P_n(S)$: = set of all n element subsets of S.

For example $P_1(S) = \{\{s_1\}, \{s_2\}, \ldots\}$ where $s_1, s_2, \ldots$ are the elements of S.

$P_{<\omega}(S)$: = set of all finite subsets of S, that is: $P_1(S) + P_2(S) + \ldots$

$P_{\omega}(S)$: = set of all $\{s_1, s_2, \ldots\}$ sequencable subsets of S.

6.) Content and sample

$\bigcup S$ = content of S : = the combining of all elements of S, that is:

$t \in \bigcup S$ if there is $s \in S$, so that $t \in s$.

f is choice function if it is a function and $f(s) \in s$.

f is choice function on S, or sample from S if it is choice function and Do f = S.

$\prod S$ = sample space of S : = set of all samples from S.

The elements of $\prod(T^{\omega}) = \prod\{S_1, S_2, \ldots\} = S_1 S_2 \ldots$

can be imagined as the $\{s_1, s_2, \ldots\}$ sequences picked from $S_1, S_2, \ldots$

7.) Size

$S \leq T$ if S is equivalent to a subset of T.

$S < T$ if S is equivalent to some (proper) subset of T but can not be equivalent to T.

$S \lhd T$ if $S < T$ and there is no R set so that: $S < R < T$.

T is jump set if there is S so that $S \lhd T$.

T is limit set if there is no such S.

$S \ll T$ if $S < T$ and for all $s \in S$, $s < T$ too.

S is exceeding if $S \ll \bigcup S$.

R is accessible from S if $S \ll \bigcup S \sim R$.

R is accessible if it is accessible from some S, otherwise inaccessible.

3. Basic theorems

T

Equivalences

1.) $\left(R^{S}\right)^{T} \sim R^{ST}$.

2.) $2^{S} \sim P(S)$, $[0,1] \sim P(\omega)$.

3.) Bernstein: $S \le T \wedge T \le S \rightarrow S \sim T$.

4.) Cantor finite diagonal: $\omega^2 \sim \omega$, $P_2(\omega) \sim \omega$, $[0,1]^{\omega} \sim [0,1]$

5.) $\omega^{n} \sim \omega$, $P_n(\omega) \sim \omega$.

6.) $\omega^{<\omega} \sim \omega$, $P_{<\omega}(\omega) \sim \omega$.

P

1.)

Every element of $\left(R^{S}\right)^{T}$ is a function that orders to every $t \in T$ a function from S to R.
Thus, it is also a function that orders to every $t \in T$ and $s \in S$ an $r \in R$.

2.)

$\{0,1\}^{S}$ is the set of functions that have the 0 or 1 values on S. If 0 is regarded as not taking, while 1 as taking the element into a subset, then these functions give exactly the subsets.

The $\{0,1\}^{\omega} = [0,1]$ Continuum is thus equivalent to the subsets of naturals.

3.)

$S \le T$ means $S \underset{f}{\sim} T_1 \subseteq T$ and $T \le S$ means $T \underset{g}{\sim} S_1 \subseteq S$. Then:

$S \underset{f}{\sim} T_1 \underset{g}{\sim} S_2 \underset{f}{\sim} T_3 \underset{g}{\sim} S_4 \underset{f}{\sim} T_5 \ . \ . \ .$ and $T \underset{g}{\sim} S_1 \underset{f}{\sim} T_2 \underset{g}{\sim} S_3 \underset{f}{\sim} T_4 \underset{g}{\sim} S_5 \ . \ . \ .$

These give a single chain of equivalences for the differences:

"difference chain" : $(S - S_1) \underset{f}{\sim} (T_1 - T_2) \underset{g}{\sim} (S_2 - S_3) \underset{f}{\sim} (T_3 - T_4) \underset{g}{\sim} \ . \ . \ .$

Expressing S and S_1 with differences:

$$S = [(S - S_1) \cup (S_2 - S_3) \cup (S_4 - S_5) \cup . . .] \cup [(S_1 - S_2) \cup (S_3 - S_4) \cup . . .] \cup S_1 \cap S_2 \cap . .$$

$$S_1 = [(S_2 - S_3) \cup (S_4 - S_5) \cup . . . .] \cup \quad \text{same} \quad \cup \quad \text{same}$$

Then these two are equivalent by ordering the same parts to themselves and for the first part using the S-members of the difference chain: $(S - S_1) \sim (S_2 - S_3) \sim (S_4 - S_5) \ . \ . \ .$

Then of course: $S \sim S_1 \sim T$.

4.)

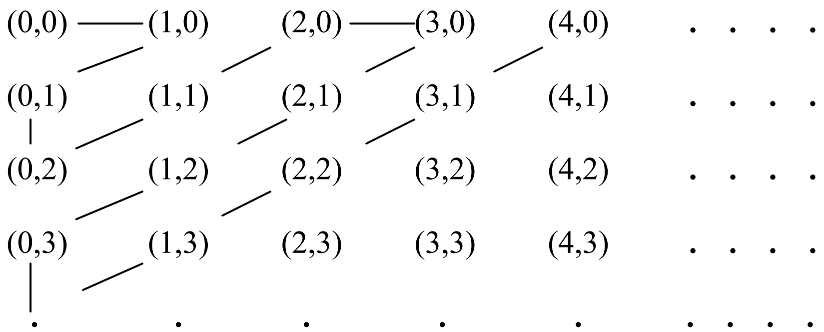
It's easy to list $P_2(\omega)$ by the increasing larger members and sublisting all the smaller members:

{0 , 1} {0 , 2} {1 , 2} {0 , 3} {1 , 3} {2 , 3} {0 , 4} {1 , 4} {2 , 4} {3 , 4} {0 , 5} . . .

We can list the ordered pairs similarly, by repeating the pairs in opposite order plus include the self pairings: (0 , 0) (0 , 1) (1 , 0) (1 , 1) (0 , 2) (2 , 0) (1 , 2) (2 , 1) (2 , 2) (0 , 3) . . . .

We can go by increasing sums too: (0 , 0) (0 , 1) (1 , 0) (0 , 2) (2 , 0) (1 , 1) (0 , 3) (3 , 0) . . .

These sequencings become much more surprising with the $\omega^2 = \omega \cdot \omega$ meaning as the single sequencing of a sequence of sequences:



The lines of "finite diagonals" represent the previously mentioned listing by increasing sums.
Placing a minus sign in front of the second non 0 members, we get the grid points of the right lower, quadrant, in a Descartes coordinate system.
All four quadrants can be listed at once in a spiral walk, starting from the origin.

$[0,1]^\omega \sim \left(2^\omega\right)^\omega \sim 2^{\omega^2} \sim 2^\omega \sim [0,1]$. This is amazing, because:

The elements of $[0,1]^\omega$ can be also regarded as the points of an infinite dimensional space.

5.) Follows from 4.) by repeated application.

6.) Follows from 5.) because it is a sequence of sequences.
Directly can be proved too by ordering to an $n_1, n_2, \ldots, n_k$ sequence of naturals, the $2^{n_1} 3^{n_2} \ldots p_k{}^{n_k}$ natural number, where p_k is the k-th prime.

T Inequivalences

1.) König

a.) Increasing sum versus product: If $S_1 < S_2 < \ldots$ then: $S_1 \cup S_2 \cup \ldots < S_2 S_3 \ldots$

b.) Element by element reduced content versus the sample space:
If f is a function that has smaller values, that is for every $S \in$ Do f, f (S) < S, then: the content of its range is also smaller than the sample space of its domain: $\bigcup$ Ra f < $\prod$ Do f

c.) Set without single element versus its sample space:
If every element of the S set has at least two elements, then $S < \prod S$.

2.) Generalized Cantor anti diagonal: $S < 2^S \sim P(S)$.

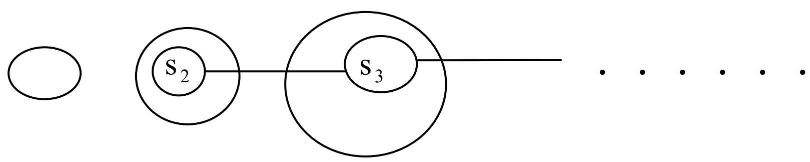
3.) Cantor anti diagonal: $\omega < 2^\omega \sim P(\omega)$.

4.) Generalized Continuum Hypothesis: If S is infinite, then $S \lhd 2^S$.

P 1.)

a.)

Enough to show, that if g is a function that orders elements of $S_2 S_3 \ldots$ to elements of $S_1 \cup S_2 \cup \ldots$, then there must be at least one element of $S_2 S_3 \ldots$ that is not a value of g. The g values of the S_1 elements relate a unique element of S_2, that appear in the value. But since $S_1 < S_2$, there must be at least one $s_2 \in S_2$, that is not in any g value. Similarly, there is $s_3 \in S_3$ that is not in any g value for S_2 elements. And so on, we get a sequence of elements that together is an element of $S_2 S_3 \ldots$ but can't be any g value.



b.) We'll use D and R for Do f and Ra f.
Enough to prove it for f, where the elements of D are disjoint. Indeed, we can replace every $s \in S$ with (S , s). Then R can remain the same and $\prod D$ equivalent.
First we show $\bigcup R \leq \prod D$. Since $f(S) < S$, none of the S is 0, so we can pick an element of each as a B basic sample. Leaving out this chosen element from an S, we still get more or equivalent with f (S). So, $\bigcup R \leq \bigcup D - B$. But $\bigcup D - B \leq \prod D$, because to every s element of $\bigcup D$ outside B, we can order an altered B by replacing s for the B element in the same $S \in D$. Thus, we only have to show that $\bigcup R \sim \prod D$ is impossible.
For this, it is enough to show that if g is a function that orders elements of $\prod D$ to some t elements of $\bigcup R$, then there must be at least one element of $\prod D$ that is not a value of g.
Let g [S] denote the set of all S elements that the g (t) samples pick from S if we go through all t elements of f(S). Since $g[S] \leq f(S) < S$, there must be S elements that are not in g[S]. Let's pick one of these from each S! This is a sample that is definitely none of the g (t)-s!
Finally we mention that the element by element reduction of D is necessary to guarantee an inequality. Indeed, if D contains finite many ω sets then $\prod D \sim \omega^n \sim \omega \sim n\omega \sim \bigcup D$. For infinite D, a counter example can be obtained by the D = {X , Y , Z , . . .} sequence of coordinate lines. The sample-space is then the set of points in the infinite dimensional space and as we saw by 4.) of previous theorem, this is equivalent to a single line and thus also to the sum of the coordinates.

c.)
Let f be the function that: Do f = S and f (s) = {s}. Then, f (s) < s, so a.) can be applied and $S = \bigcup \text{Ra } f < \prod \text{Do } f = \prod S$.

2.)
Let S_0 be the set that contains the {s,0} pairs as elements for every s element of S!
Then we can use 1.) b.) and every sample will define a unique subset of S if we only regard the non 0 choices. We give a direct proof without using 1.):
Enough to show that for any f function that orders subsets of S to some elements of S, there is a subset of S, that is not value of f.
If $s \in f(s)$ for all f (s) values then the {s} single subsets can only be values as f (s).
So either f (s) = {s} for all values and then the non single subsets are not taken as values, or $f(s) \neq \{s\}$ for at least one s and then {s} is not taken as value.
If $s \in f(s)$ is not true for all f (s) values then we can collect all the $s \notin f(s)$ elements as an S_0 subset. We claim that $f(s) = S_0$ is impossible. Indeed, for any $s \in S_0$, $s \notin f(s)$ so, $f(s) = S_0$ is impossible, while for any $s \notin S_0$, $s \in f(s)$ so, $f(s) = S_0$ is impossible again.

3.) Trivial by 2.)
We give a direct proof too, because this was Cantor's original result that initiated Set Theory:
Enough to show that $\omega \sim 2^{\omega}$ is impossible. Suppose we could list all infinite binary numbers as:

. 0 0 1 1 0 0 1 1 1 0 1 0 1 1 0 1 0
. 1 0 0 1 1 1 1 0 1 0 1 1 0 0 0 1 1
. 0 1 1 0 1 0 1 0 0 0 1 1 1 0 1 0 0
.

Lets make a number from the diagonal digits, that is take the first digit of the first, second of the second, third of the third, and so on: . 0 0 1 . . . This number could very well be in our list. But lets change every digit to its opposite! This "anti diagonal": . 1 1 0 . . . number can definitely not be in our list because it differs from the first in the first digit, from the second in the second, and so on. By the way, we could have easily used infinite decimals and then the anti diagonal digits only have to be any digit different from the diagonal.

4. Listing of sets

R The original Continuum Hypothesis of Cantor was merely $\omega \triangleleft 2^{\omega}$, so it meant that the continuum is the next infinite after a sequence. Or with $\omega \triangleleft P(\omega)$ it meant that the subsets of the naturals is the next infinite after the naturals. The puzzle of this might shadow our mind about a more fundamental question, namely whether the size of sets is really meaningful at all for all sets. Indeed, how can we even be sure if for two sets the $S < T$ or $T < S$ or $S \sim T$ choices are necessary. Obviously these three exclude each other by their definitions. If we could prove that: $S \leq T$ or $T \leq S$ is always true, that is there is at least one equivalence that compares any set to any other, then we could be sure that the above three possibilities can be decided. As it will turn out, we are mistaken! The word "decided" will only mean a theoretical necessity without being able to actually decide $\omega \triangleleft P(\omega)$. So, we can know that all sets of subsets from the naturals is either merely sequencable or is equivalent to the set of all possible subset, or is in between. Yet, we still may not be able to decide if there are such last third possibilities. This means that the sets of natural numbers are beyond our grasp. We can pick some of them and we may feel that all pickings are objectively determined sets, but if they are undescribable forever, then sooner or later we have to accept that they are undetermined. Usually, when seemingly objective phantoms are unmasked, we replace them with something new, tangible reality. But as the number of phantoms increased, their replacements became less and less tangible too:

The feeling of vacuum that sucks in everything turned out to be the outside pressure. This is completely proved when we see that the sucking has its limit. We can't suck up more than 760 millimeters of mercury into a glass tube, no matter how perfect vacuum we create. A one meter tall glass tube filled with mercury and turned over into a tray of mercury, will drop back to 760 millimeter, leaving the rest above it in total vacuum. This vacuum doesn't suck anything! The atmosphere of the earth is pushing the mercury up into our tube through the tray.

The feeling of cold freezing our fingers, the chill coming out of the fridge is also a phantom. In fact, the heat is leaving our fingers and also rushing into the fridge.

Then the Ether filling the universe was much harder to give up, because the new reality of relative references obeying absolute laws was much harder to "feel".

To stay in mathematics, the imperfections in a description of natural numbers can be quite easily accepted: Indeed, for example the operations $+ , \bullet$ can tell a lot about the $1, 2, 3, \ldots$ sequence, but why would they tell everything. This can be best seen if we just look at the ordering sequence. Clearly, $<$ can be defined as $x < y := \exists\, z,\ x + z = y$. Then we can tell that every number has previous one except one and every number has next one. But this only forbids continuation before 1 or restarting after the infinite sequence. However, it still allows: $1, 2, 3, \ldots \quad \ldots, -3, -2, -1, +1, +2, +3, \ldots$ Indeed, here 1 is the only without previous, so we don't contradict our requirements. Of course, we can say that every subset should have a first element. But this, we can only say with using subsets. And if we change our language from numbers to sets of numbers, then a whole new world of even wider possibilities open. We see things about the natural numbers from the outside, that we can not express by simply talking strictly about them. And as we see, it's not merely that $+ , \bullet$ is not enough. More relations still wouldn't describe sets. There is a tricky way to go around this restriction of talking only about numbers which allows to express sets. Namely, we could use $+ , \bullet$ or other relations to make logically built $P(x)$ properties and then say that there is a first among these $P(x)$ numbers. But this is only an approximation to imitate all subsets, so as expectable it will still leave room for weird models.

Set Theory is different in the sense that everything is a set, so we can't go out of it as it happened with the naturals. But this advantage still couldn't lead to an absolute theory, because the same problem of approximation or compromise, appeared here too. Strangely, this compromise was also caused by $P(x)$ property description, but here this was the beginning not the end of the problem. Indeed, for a while it seemed that using $P(x)$ properties for sets is not a compromise, but the perfect way to collect sets. Unfortunately, $\{x \,;\, P(x)\}$ definitions

lead to contradictions, so here the possible P (x) properties had to be restricted themselves. The new axiom for set collection was called the axiom of Replacement and we might think that the name referred to this replacement of the original unrestricted collection, but it's a mere coincidence. So it was obvious that we only have approximative axioms, so our Set Theory can't be perfect. But even more strangely, it turned out that a completely different cause of imperfection is present too. The restrictions to avoid contradictions are concerning the big sets. The undecidability of $\omega \triangleleft P(\omega)$ means that we already don't know exactly what the subsets of naturals are. Of course, the main truth is that we can not separate simple and complicated sets. The fact, that some number theoretical proofs can only be obtained with higher methods is the classical version of this. And indeed, building up sets above the natural numbers lead back to form new subsets of them. Still, the restrictions of build ups can not be the sole cause of our lack of grasping the subsets. So the truth is, we don't know what the problem is. Probably, something about randomness is missing! We should regard randomly picked subsets of naturals, but we don't know exactly what random is.
So, we should get back to our business of showing, at least the theoretical comparability of sets. By the way, even though this won't solve the Continuum Hypothesis, it still solves many details. Most importantly, it will prove a basic missing theorem up until now. Cantor's earliest result was $\omega^2 \sim \omega$ with the finite diagonal sequencing. The generalization, $S^2 \sim S$ is quite expectable, but to prove it, one has to go towards the comparability of sets.
How to compare sets, how to create equivalence between two S , T sets? The natural idea is to pick elements after elements from both until one of them "runs out". The elements picked at the same time will be obviously the ones related to get an equivalence, but at the same time, we are creating a $\prec$ relation within each of the two sets, namely the one that tells the "order of pickings". This is the Set Theoretical representation of "time" as instances. Amazingly, two simple properties will describe it sufficiently. One is that, it is an ordering, so any two elements must be related in one and only one direction. The other is that, there is always a next one. But here, the "always" can not be simply regarded as any instant, because after infinite many picked elements there is no last to follow. We might say then that instead we should require that every set of elements has a next one. But this is not good again, because a set of elements can go on "forever" in the whole listing of the set. Instead of differentiating between such "cofinal" subsets and ones that finish before the whole set, we can simply require that all subsets should have a first element. In fact, an even simpler requirement is enough, namely that there is no backwards going sequence of instances. As is usually the case, such tricky simplifications, only avoid the real messy distinctions temporarily.
Indeed, cofinality will be a real problem of infinites. But now, lets go as simply as possible.

D

1.) An R relation is an <u>ordering</u>, if:
  a.) R doesn't have (s , s) elements.
  b.) For any two different s_1 , s_2 that appear in the domain or range of R, exactly one of (s_1 , s_2) or (s_2 , s_1) is $\in$ R.
  c.) $(s_1 , s_2) \in R$ and $(s_2 , s_3) \in R \rightarrow (s_1 , s_3) \in R$.
  If all elements of S appear in R, then we call it an ordering <u>of</u> S.
  We'll use $s_1 \prec s_2$ instead of $(s_1 , s_2) \in R$, and say that s_1 is before s_2.
  $\ldots \prec s_1 \prec s_2 \prec s_3 \prec \ldots$ means . . . and $s_1 \prec s_2$ and $s_2 \prec s_3$ and . . .

2.) A $\prec$ ordering is <u>well</u>-ordering if there is no $\ldots s_3 \prec s_2 \prec s_1$ backwards sequence.
  A well-ordering <u>of</u> S, is also called an $\widetilde{S}$ <u>list of</u> S.
  Every S_0 subset of S has a first element in a list of S.
  Indeed, if a subset had no first element, then we could pick an s_1 from it, then an earlier s_2, then s_3, and so on, achieving a $\ldots s_3 \prec s_2 \prec s_1$ backwards sequence.
  We'll denote the first element of the whole S as s_0.

3.) $\tilde{S}$ (s) or the <u>beginning</u> of $\tilde{S}$ before s, denotes the relation of $\tilde{S}$, but kept only on the elements of S before s. Since there are no elements before s_0, thus $\tilde{S}\,(s_0) = 0$.
So, 0 is regarded as a beginning in every $\tilde{S}$.
So, 0 is a common beginning of any two $\tilde{S}$, $\tilde{T}$.

4.) Let f be a function! $\tilde{S}$ is an <u>f-list</u> if: $f\,(\tilde{S}(s)) = s$.
If $\tilde{S}$ is non empty, that is there is an s_0, then $s_0 = f(0)$, so f is defined on 0.

T

1.) The union of some beginnings of an $\tilde{S}$ list is a beginning of $\tilde{S}$ too, or the full $\tilde{S}$.

2.) For any two different f-lists, one is beginning of the other.

3.) The union of some f-lists, is also an f-list.

4.) Well-Ordering Theorem: For every S set, there is an $\tilde{S}$ list of S.

P

1.) If all elements of S appear in some of the beginnings, then the union will define $\tilde{S}$.
Indeed, for any $s_1 \prec s_2$ elements, they are in the same relation in any beginning of $\tilde{S}$.
If there are elements of S, that don't appear in the beginnings and the first of these is s^0, then the union is $\tilde{S}\,(s^0)$. Indeed, any two $s_1 \prec s_2$ elements before s^0 must come from B_1 , B_2 beginnings. But then both of them are already in B_2.

2.) Lets regard the common beginnings of the two f-lists!
The U union of these is either a beginning or the full of each by 1.).
It can't be the full of both, because then the two f-lists were the same U.
It can't be the beginning of both, because then U + f (U) were a wider common beginning of the two f-lists, contradicting that U must be the widest.
Thus, U must be a beginning of one of them and the full of the other.

3.) We have to show that the union is an ordering, a <u>well</u>-ordering and an <u>f</u>-well-ordering.
The first follows from 2.), the other two are then trivial.

4.) Let g be a sample from P (S) – 0 and f (T) = g (S – T).
The union of all f-lists is an f-list by 3.). And:
It must be on the whole S, otherwise there were wider f-lists.

R

Now that every set can be listed, we only have to compare lists:

D

$\tilde{S} \approx \tilde{T}$ that is $\tilde{S}$ and $\tilde{T}$ are similar, if there is an f equivalence between them that keeps the ordering too. That is, $s_1 \prec s_2$ in $\tilde{S}$ → $f(s_1) \prec f(s_2)$ in $\tilde{T}$.

T

For any two $\tilde{S}$, $\tilde{T}$ lists: $\tilde{S} \approx \tilde{T}\ \nabla\ \tilde{S} \approx \tilde{T}(t)\ \nabla\ \tilde{T} \approx \tilde{S}(s)$.
For any two S , T sets: $S \sim T\ \nabla\ S < T\ \nabla\ T < S$.

P The first can be shown by four facts:

a.) $\tilde{S} \approx \tilde{S}(s)$ is impossible.

b.) $\tilde{S} \approx \tilde{T}(t)$ and $\tilde{T} \approx \tilde{S}(s)$ is impossible.

c.) $\tilde{S} \approx \tilde{T}(t_1)$ and $\tilde{S} \approx \tilde{T}(t_2)$ is impossible.

d.) $\tilde{S} \approx \tilde{T} \vee \tilde{S} \approx \tilde{T}(t) \vee \tilde{T} \approx \tilde{S}(s)$.

Clearly, d.) with a.) and b.) gives our claim but its easier to prove them as a.) b.) c.) d.).
For a.), first we show that if f is a similarity from $\tilde{S}$ to an $\tilde{S}_0$ sub list of itself, then $f(s) \prec s$ is impossible. Indeed, if there were $f(s) \prec s$ elements and the first of these were s_1, then $f(s_1) \prec s_1 \rightarrow f(f(s_1)) \prec f(s_1) \prec s_1$. So $f(s_1)$ were an earlier one.
Now, if $\tilde{S} \approx \tilde{S}(s)$ were, then to the elements after s we couldn't pair any by the aboves.
For b.), if $\tilde{S} \approx \tilde{T}(t)$ is f and $\tilde{T} \approx \tilde{S}(s)$ is g, then $f(g)$ would give an $\tilde{S} \approx \tilde{S}(s)$.
For c.), if $t_1 \prec t_2$ then $\tilde{T}(t_1)$ is a beginning of $\tilde{T}(t_2)$ contradicting again a.).
For d.), let R (s,t) be the relation that there is similarity between $\tilde{S}(s)$ and $\tilde{T}(t)$. Then:
R is one to one by c.) and keeps the ordering by b.).
Finally, R must be defined or ranging on the whole S and T, otherwise for the first unrelated s_1, t_1 elements, R itself up to s_1, t_1 would give R for s_1, t_1.

The second claim, the comparability of sets, follows from the first by previous T 4.).

R A minor step in our proof was to choose the g choice function or sample from 2^S.
The existence of such can not be proven. This is plausible because it is based on random choices from the subsets. When the axioms for Set Theory were established, this assumption became known as the Axiom of Choice. The proper name should have been the Axiom of Choice<u>s</u>. Indeed, to make one choice is easy, because mathematical logic has a rule for that. The whole point was to go beyond this and make infinite many choices simultaneously. So in a sense, the Axiom of Choice is a bridge from Logic to Set Theory. Before its discovery by Zermelo, it was already used many times by mathematicians in classical mathematics, but without being aware of it. How easy it is not to realize when we use it can be shown by the example of our definition of lists themselves. Indeed, at 1.) in our first definition we "proved" that the requirement of not having backwards infinite sequence implies a first element in every subset. We did this by simply choosing the $s_1, s_2, s_3, \ldots$ earlier and earlier elements. But this was not quite exact. The proper proof should've been to regard an f choice function that picks an element from our S_0 subset and from all beginnings of it. Then s_1 is the element picked from the total subset, s_2 is the one picked from its beginning before s_1, s_3 from the beginning before s_2, and so on. Here again we said "and so on", so this still must be replaced by an explicit set. Indeed, we can call a finite sequence of $(s_1, s_2, \ldots, s_n)$ elements an f-sequence if $s_1 = f(S_0)$ and $s_{k+1} = f(S_0(s_k))$. Then the union of all such f-sequences will be the $S_\omega = \ldots \prec s_2 \prec s_1$. As we see, this line of precise reasoning resembles the whole proof of set listings. Usually it's similar how to turn the imprecise "timely" choices into exact "predetermined" one, but we'll have a more universal tool for this.
The Axiom of Choice involves the random choices. But as I said, randomness is missing from Set Theory. The problem is that the Axiom of Choice doesn't differentiate between the random and non random sets. That's what is missing.
Finally, I will give three classical theorems that are equivalent with the well ordering theorem:

D

1.) An $s \in S$ can be widened in S, if there is a $t \in S$, so that $s \subset t$.
An $S_0 \subset S$ can be widened in S, if there is a $t \in S$, so that $s \subset t$ for all $s \in S_0$.
If s or S_0 can not be widened in S, then we call them maximal.

2.) An S set is widening, if for every $s , t \in S , s \subseteq t$ or $t \subseteq s$.
Every S set has widening subsets, namely the single element subsets are such.

3.) An S set is widening complete if for any S_0 widening subset of S: $\bigcup S_0 \in S$.

4.) An S set is finite complete if: $s \in S \leftrightarrow$ all finite subsets of s are $\in S$.

T

1.) Hausdorff:
Every S set has a maximal widening subset.

2.) Kuratowski-Zorn:
If S is widening complete, then it has maximal element.

3.) Teichmuller-Tukey:
If S is finite complete, then it has maximal element.

P

1.) Lets choose an f continuation function on S, so that if S_0 is widening, and can be widened in S, then $f(S_0) \supset S_0$. This merely specifies the picking of $f(S_0)$ from a smaller subset of $S - S_0$. By the proof of the well ordering theorem, there will be an f well ordering of S. The full S can not be widened obviously, so there will be maximal subsets or S itself. If the first such is S_0, then it must be widening. Indeed, if it weren't, then the union of previous widening and non maximal beginnings would give such.

2.) Let S_0 be a maximal widening set guaranteed by 1.). $\bigcup S_0 \in S$ by our condition, so since S_0 is maximal, $\bigcup S_0$ is maximal too. By the way, $\bigcup S_0$ must be $\in S_0$.

3.) Enough to show that finite completeness implies widening completeness because then 2.) automatically proves it. Suppose S_0 is a widening subset of S.
Any $\{s_1 , . . . , s_m\}$ finite subset of $\bigcup S_0$ has its elements from the $S_1 , . . . , S_m$ elements of S_0. Let the widest of them be S_i. Then this has all the $s_1 , . . . , s_m$ elements, so by one direction of the finite completeness, $\{s_1 , . . . , s_m\} \in S$ too.
Then by the other direction, $\bigcup S_0 \in S$ too.

5. Ordinals

R

Our result that all sets can be listed, raises the next even better idea. When we list a sequence, we use the notation s_1 , s_2 , . . . So the set ω is used as a universal orderer. Why can't we continue this for all listings? Well, we need to continue then the naturals themselves to get a universal list. The method of continuation is quite obvious too. We don't have new elements like in a set when we list it. But indeed, we don't even need this because the natural numbers already were built with the idea of self continuation. Taking as next element the set of all previous ones. So then after 0 , 1 , 2 , 3 , . . . the first infinite number should be
$\{0, 1, 2, \ldots \{0, 1, 2, \ldots\}\}$ which can be abbreviated as ω'. Then, comes:
$\{0, 1, 2, \ldots \{0, 1, 2, \ldots\}, \{0, 1, 2, \ldots \{0, 1, 2, \ldots\}\}\}$ abbreviated as ω''.
Now we might wonder, what's the point of having these universal numbers if they are so complicated that we have to use abbreviations instead of them. The truth is even worse, because we can't keep on adding the apostrophes either, so ω''' must be abbreviated again. Luckily, a much better abbreviation system arises, if we define new operations for these numbers. Both the addition and multiplication is an improvement of the set addition, that is union and the set multiplication, that is set of pairs. But the exponentiation will go into a strange split personality. This is expectable! Even among naturals, the nice and logical + , • can't be repeated completely by exponentiation, in spite of its similar definition of it as multiplication was from addition: $m \cdot n := n + n + \ldots + n$ and $n^m = n \cdot n \cdot \ldots \cdot n$.
Indeed, $m \cdot n = n \cdot m$ but $n^m \neq m^n$.
Here, at infinites already the + and • is not exchangeable in order. Indeed,
$3 + \omega \neq \omega + 3$ and $3\,\omega \neq \omega\, 3$ because $3 + \omega = \omega\, 3 = \omega$.
But the exponentiation is much worse! It actually goes out of our infinite numbers if we try to follow the set exponentiation. A rescue solution can define a new exponentiation, but it will not solve the original one. And all this is related to the Continuum Hypothesis. So we are not just dealing with some fancy new arithmetic.

D

1.) Ordered sets have an $\prec$ relation on them for which: $s = t \ \nabla\ s \prec t \ \nabla\ t \prec s$

2.) Two ordered sets are similar if there is an equivalency between them that keeps the order.

3.) The similarly ordered sets are called the same type and we can use special symbols for their ordering, which is also called their type. We denote these with greek letters.

4.) The listed, that is well ordered types are called ordinals and can be represented as the self continuing extensions of naturals.

5.) $\alpha + \beta$ is the type obtained from α and β, by continuing α with a β.
If α and β are ordinals, then $\alpha + \beta$ is such too.

6.) $\alpha\,\beta$ is the type obtained if we place β type sets into every element of an α type set.
If α and β are ordinals, then $\alpha\,\beta$ is such too.

7.) A sequence product, $\alpha_1\,\alpha_2\,\ldots$ means that we place α_2 types into the elements of an α_1, then α_3 into the elements of these, then α_4, and so on. The final elements after an infinite sequence are undetermined, but our solution is quite simple: We regard all possible e_1 , e_2 , . . . sequences, with e_1 being an element of α_1, e_2 of α_2, and so on. Then any two such element sequence is ordered by their first differing element.

For example, $\omega\,\omega\,\omega\,.\,.\,.$ will have as elements all the possible $n_1, n_2, \ldots$ sequences of natural numbers. Then for example, $7, 3, 5, \ldots \prec 7, 3, 9, \ldots$

$\alpha\,\alpha\,\alpha\,.\,.\,.$ is abbreviated as α^{ω}.

8.) α^{β} can be defined as the β types picked from elements of an α.

If α and β are ordinals, then α^{β} is not such!

For example, 2^{ω} is the type of the Continuum [0,1].

9.) $\alpha^{<\omega}$ denotes the type: $\alpha + \alpha^2 + \alpha^3 + \ldots$
If α is an ordinal, then this limit power is too.

10.) If $\beta + \gamma = \alpha$ then this can also be denoted as:
$\beta \prec \alpha$ and saying that β is a beginning of α or as:
$\alpha - \beta = \gamma$ and saying that γ is the end of α after the β beginning.

11.) β is an omittable beginning of α, if $\alpha - \beta = \alpha$,
that is, cutting off the β beginning doesn't change α.
This of course, also means that $\beta + \alpha = \alpha$,
that is, adding β to the beginning, doesn't change α either.

12.) $\mathrm{beg}\,\alpha$ is the first not omittable β beginning of α.
Since $\mathrm{beg}\,\alpha$ is omittable from $\omega\;\mathrm{beg}\,\alpha$, but not from α, thus, $\alpha \prec \omega\;\mathrm{beg}\,\alpha$ and so: there is an n largest natural number, that $n\;\mathrm{beg}\,\alpha \prec \alpha$ or $n\;\mathrm{beg}\,\alpha = \alpha$.

13.) $\mathrm{end}\,\alpha$ is $\alpha - n\;\mathrm{beg}\,\alpha$ where n is the previous largest natural, if $n\;\mathrm{beg}\,\alpha \prec \alpha$
or $\mathrm{end}\,\alpha$ is 0 if $n\;\mathrm{beg}\,\alpha = \alpha$.

14.) α is a beginning ordinal if $\mathrm{beg}\,\alpha = \alpha$, that is if $\mathrm{end}\,\alpha = 0$.

T

Every α ordinal is a unique $n_1\,\beta_1 + n_2\,\beta_2 + \ldots + n_k\,\beta_k$, with $n_1, n_2, \ldots, n_k$ naturals and $\beta_1 \succ \beta_2 \succ \ldots \succ \beta_k$ decreasing beginning numbers.

P

$\mathrm{beg}\,\alpha = \beta_1$ and $\alpha = n_1\,\beta_1 + \mathrm{end}\,\alpha$. Then:
$\mathrm{beg}\,(\mathrm{end}\,\alpha) = \beta_2$ and $\mathrm{end}\,\alpha = n_2\,\beta_2 + \mathrm{end}\,(\mathrm{end}\,\alpha)$.
And so on, we get beginnings of all ends. But this can only go for up to a natural k, because otherwise we had a backwards infinite subset in α.

R

This seems like a final and perfect characterization of the ordinals, but it's an illusion.
The beginning ordinals are themselves very complicated and different.
Two directions of deeper distinction can be done. One is based on the fundamental relation of equivalence, that is size. The other on the combining of smaller sizes to get bigger ones, which was called "exceeding" at sets. Here the first is called cardinality, the second cofinality.
In spite of this, the unique summing of ordinals from beginning ordinals is enough to prove the mentioned missing claim that $S^2 \sim S$.

T

1.) Generalized Cantor finite diagonal:
If α is an infinite ordinal then $\alpha^2 \sim \alpha$

2.) If S is an infinite set, then

$$S^2 \sim S\ ,\ S^n \sim S\ ,\ P_n(S) \sim S\ ,\ S^{<\omega} \sim S\ ,\ P_{<\omega}(S) \sim S\ ,\ 2^S \sim S^S \sim \left(2^S\right)^S$$

3.) If $\bigcup S$ is infinite jump set, then S can't be exceeding. So:
If R is infinite jump set, then it is inaccessible.

4.) Inaccessible Limit Set Hypothesis (Weakly Inaccessible Set)
There is R limit set that is inaccessible.

P

1.) Enough to show it for beginning ordinals. All the beginnings of these can be represented as unique sums and then for any two:
$m_1\alpha_1 + m_2\alpha_2 + \ldots + m_j\alpha_j$ and $n_1\beta_1 + n_2\beta_2 + \ldots + n_k\beta_k$,
we can order their "natural sum". This is the combining of them, with adding up the common members. Any such sum is still a beginning. Unfortunately, we ordered the same natural sum to different pairs, but only to finite many different ones.

2.) The first five follows from 1.). For the last: $S^S \le \left(2^S\right)^S \sim 2^{S^2} \sim 2^S \le S^S$

3.) If T is an infinite set, so that
$S \le T$ and for every $s \in S, s \le T$ too, then $\bigcup S \le T^2 \sim T$. In negative form:
If $\bigcup S > T$ then $S > T$ or there is some $s \in S$ that $s > T$.
Since $S \le \bigcup S$ and for every $s \in S$, $s \le \bigcup S$ too, thus:
If $\bigcup S \rhd T$ infinite, then $S \sim \bigcup S$ or there is some $s \in S$ that $s \sim \bigcup S$.
So S is not exceeding.

D

1.) car S is the first β ordinal equivalent to S.
This can be used for types or ordinals too in place of S. car α is of course beg α too.

2.) α is a cardinal if car $\alpha = \alpha$. The first infinite cardinal is ω.
The first cardinal after ω is denoted as ω_1, the next ω_2, and so on.
After these, comes $\omega + \omega_1 + \omega_2 + \ldots = \omega_\omega$. Indeed, this must be bigger than all the ω , ω_1 , ω_2 , . . . cardinals, but any beginning of it is equivalent to one of these.

T

König for ω_ω: $\omega_\omega = \omega + \omega_1 + \omega_2 + \ldots < \omega_1\ \omega_2\ \omega_3 \ldots \le (\omega_\omega)^\omega$

Thus, $2^\omega \sim \omega_\omega$ is impossible, because $\left(2^\omega\right)^\omega \sim 2^{\omega^2} \sim 2^\omega$. So the Continuum is not ω_ω.

D

1.) A cofinal of a list, is any subset, not contained completely in any beginning of the list.
If the list has a last element then this element itself is a trivial cofinal.

2.) cof α is the first β ordinal, so that α has β type cofinal.
If α is a non limit ordinal (having last element), then cof $\alpha = 1$.

3.) α is cofinal if, cof $\alpha = \alpha$.

4.) α is a jump cardinal if there is a previous cardinal right before it.
α is limit cardinal if there isn't. For example, ω_ω is such.

T

1.) If $\widetilde{S} \sim \widetilde{T}$, then there is a cofinal of $\widetilde{S}$ on which the $\sim$ equivalence is similarity too.

2.) Every cofinal is cardinal.

3.) An infinite cardinal is not cofinal if and only if it is accessible.

4.) Every infinite jump cardinal is cofinal.

5.) If α is infinite cardinal, then $\alpha < \alpha^{\text{cof}\,\alpha}$.

6.) If S is an infinite set, then $S < \text{cof}\,(\text{car}\ 2^{S})$

P

1.) Let S_0 with those elements of S to which the $\sim$ equivalence relates a t, not preceded by any t' related to an s' after s. Suppose S_0 weren't a cofinal! Then picking an s_1 after S_0, by definition of S_0, there is s_2 after s_1, so that $t_2 \prec t_1$. Also, there is s_3 after s_2, so that $t_3 \prec t_2$. And so on, we had a backwards infinite sequence in $\widetilde{T}$.

2.) If car $\alpha \prec \alpha$ were for an α cofinal, then using 1.) with $\alpha = \widetilde{S}$ and car $\alpha = \widetilde{T}$, we get a similarity between α itself and a subset of car α which is impossible by the unique similarity of ordinals.

3.) If α is cardinal, then cof $\alpha \prec \alpha$ also means cof $\alpha < \alpha$.
If α is infinite cardinal, then every cofinal subset contains only $< \alpha$ elements.
So indeed, α is the union of a smaller size many smaller size sets.
In reverse, an accessible set when listed, gives a list of the union it is equivalent to.
Thus, either there is a member so that's its elements form a cofinal or one element from each will form one. In either case we get a smaller sized cofinal.

4.) We already proved that all infinite jump sets are inaccessible, so by 3.) this means cofinal. With this new cofinal translation, the Inaccessible Limit Set Hypothesis:
There is limit cardinal that is cofinal.

5.) If cof $\alpha = \alpha$, that is α is cofinal, then by Generalized Cantor anti diagonal:
$\alpha < 2^{\alpha} \leq \alpha^{\alpha} = \alpha^{\text{cof}\,\alpha}$. If α is not cofinal, then by 4.)
α is limit cardinal, so $\alpha = \alpha_1 + \alpha_2 + \ldots$ is possible with $\alpha_1 < \alpha_2 < \ldots$ and thus, by König: $\alpha = \alpha_1 + \alpha_2 + \ldots < \alpha_2\, \alpha_3 \ldots \leq \alpha^{\text{cof}\,\alpha}$

6.) $(\text{car}\,2^{S})^{S} \sim \text{car}\ (2^{S})^{S} \sim \text{car}\ 2^{S^2} \sim \text{car}\ 2^{S} < (\text{car}\,2^{S})^{\text{cof}\,(\text{car}\,2^{S})}$

The last $<$ was 5.) used with $\alpha = \text{car}\,2^{S}$.
The beginning and the end implies $<$ for the exponents.

6. Naïve Randomness, An unsuccessful attempt to defy the Continuum Hypothesis

R

The existence of irrational points was reproved by Cantor with his result that the rationals are sequencable, while the points of an interval, the continuum, isn't. But this is just a formalist description because the continuum "being unsequencable" only means that taking out a sequence of points from an interval, some, that is at least one point must always remain. We can of course add as a remark to this formal statement that the remaining points can't be a single one, can't be finite many or can't even be a sequence, but the simple truth is that the full interval is much, much more than any sequence. In short, the continuum is a big pool from which a sequence only takes a minute part, hardly more than a single or finite many points. This is a nice picture but it contradicts an other picture, namely that the dividers or rationals are just a sequence too, even though they are dense that is, are all over the interval. In fact, we don't even have to appeal to a tricky sequencing of the fractions. Indeed, if we just pick a sequence of points from an interval "randomly" then it is quite natural that the picked points will be all over the interval. This means that every point on the interval is approached by some picked point. But of course the unpicked points which are much more, also approach every point. Then the obvious question is if there is some local difference around every point between the picked and the remaining points. Or if we return from the random sequence to the dividers or rationals then the question is: What is the local difference between the rationals and the irrationals? This "local" obsession is even more justified by the simplest continuum paradox, which is the already mentioned projection of intervals into each other. Indeed, in such projection the dividers go into the dividers of the other interval and the non dividers into the non dividers. Then of course every little segment in a rational or irrational ordering is similar to the whole. Giving nice greek letter names to these orderings as η for the rationals and θ for the irrationals, makes us feel that we grasped something when in fact just buried the problems. I want to stress that I don't blame formalism for not solving a problem that is unsolvable, only for not mentioning this problem itself. A formalist would say: "What are you talking about? η is the ordered type of rationals, θ of the irrationals. These are perfectly determined and of course the sequencability or non sequencability is included. But what is your mystical "local" description? You mean, not using subsets and equivalences, only an ordering relation and some restricted statements about this? Well it's not buried at all because it is a basic result in mathematical logic that an axiom system for the dense ordering is complete exactly because all the sequencable models of it are the similar η ordered sets." But then I say to the formalist: I know that the problems are not buried as problems of mathematics! They can be dug out from some journals, but they are not mentioned in their original appearance in textbooks. So the non problematic approach of the basic school system is continuing in the professional too. The real problem is of course in the intentions. A formalist doesn't want to help the reader in his visualization process because visualization itself is not exact. A self-conscious "professional" formalist would even defend his view and say that if something is exactly expressed then the student should sweat out his own visualization. The real reason of course is the fogginess of visualization! If a new didactic logic would become exact, as I believe it will be, then the formalist would follow the new rules. In spite of this moral edge, I admit that formalism is a natural trap, into which we tend to fall back again and again. So the "good" and the "bad" is not simply a difference in morality. To be good we have to first have to admit that we are bad. It sounds like the original sin of Christianity because it is exactly that!

Returning to the sequences in the continuum, if we admit that a direct local difference between the sequencable and non sequencable dense sets is not possible, then at least we should try to give some global ones that go beyond the mere fact that the continuum is not sequencable. The nicest were some space oriented global description of how big the continuum is. We'll do this in the next section and here, only mention the problems with such an approach.

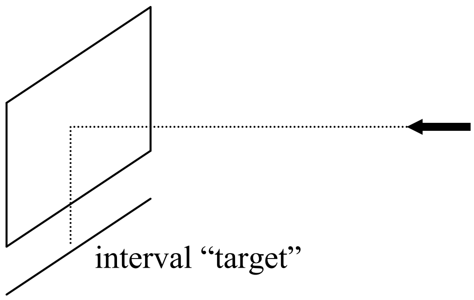
Strangely if we really want to describe what space is like then we have to involve time too. Indeed the distances or more exactly the equality of distances mean subjectively that they can be moved to cover each other. Then this opens up a Pandora's box, namely the question of rigidness comes into much more then the timing of the motion. If we turn to formalism and cut through all these with a euclidian type axiomatization of space then we can exactly talk about the copies of a point set or if we prefer we can even use the earlier terminology that a point set can be moved onto the other. Formalism of course is always lying and the buried problems come out later and bite us in the ass. The problem of time, rigidness of distances came back through Relativity first as in theory of a non euclidian space-time and then in its consequence through the Big Bang as a concrete impossibility that the real physical space could be euclidian. Indeed, no matter how big the Big Bang was, if space were euclidian then looking from far enough it would seem like a tiny flicker in a corner of a big room. Then regardless whether the fate of the universe is a cold or hot end, the infinite space were useless. So the well known argument for extra terrestrial intelligence that other wise “it would be an awful waste of space” here applies in reverse but infinitely stronger. This doesn’t mean that I completely gave up my belief in the a-priori-ness of euclidian space. Indeed, the idea that it is acquired in childhood from perceptual experiences, is so far fetched that any wild assumption sounds more real to me. In fact my whole platonism changed to a more dialectical one that not only assumes that the “world of ideas” involves time but also the roads to intelligent discoveries. Then the world of ideas doesn’t even have to correspond directly to the physical reality. Returning (again) to our subject, we can now almost expect that, moving, that is copying point sets will cause some problems. Just as the extensive infiniteness of a space that could accommodate a sequence of copies like telegraph poles on an infinite road became physically impossible, the intensive infiniteness of point sets determined and copied as we wish, can be controversial. The big difference is that here we’ll only have paradoxes with out intuitions and not with physical reality. I do believe though that sooner or later these paradoxes will have physical meaning too and lets remember that luckily for Relativity even the extensive alternatives to euclidian space came first as possibilities by intuition. Indeed, otherwise Einstein wouldn’t have had the tools for his General Relativity. Dropping the ideal of the infinite sequence of telegraph poles was first just a logical endeavor to test the parallel axiom but then both Lobachevsky and Bolyai fell in love with the new wider possibilities so much that they actually considered a non euclidian physical world. In fact, as it always happens this love made them blind! They failed to see the biggest consequence of their new geometry and only Beltrami realized that regardless what physical reality is, even if it is the old euclidian space, the fact that in this space it is possible to create an imaginary one that obeys the old axioms except the one of parallelity, automatically proves that this axiom is not a logical consequence of the others! Indeed even though the concept of logical consequence was not quite defined at that time it was immaterial just as the physical reality. All we have to accept is that if something is logical consequence, then it must be true in any world where the axioms are true. Now, since in some artificial worlds like on a sphere or other more complicated surfaces with intentionally curved lines, all the axioms are “true” that is valid when applied to these artificial lines but the unique parallel lines of the parallelity axiom fail, therefore this assumption is not a logical consequence of the others.

This idea that mathematics is also a game of consequences and possible realities that is, models was the spark that initiated from one single root the two lines of new mathematics, Logic and Set Theory. The two branches grew and then officially became united again by Gödel’s Completeness Theorem which proved the reverse of the hidden assumption behind Beltrami’s realization, namely that not only if something is logical consequence then it is true in every model but if something is true in every model then it can be derived by logic.

This explains the name “completeness” in the sense that our logic is complete. Interestingly both the original Beltrami principle and the reverse in Gödel’s proof were applied in their negative forms: “There was model with parallel axiom being false so it can’t be derived from the others” and “If a statement is not consequence of some axioms then there is a model where the axioms are true but the statement is false”.

Above I said that a difference between the extensive and intensive infiniteness of space is that the first contradicts physics while the second is paradoxical already within mathematics.

Well, with a little stretch of imagination we can see even more basic paradoxes in the extensive infiniteness than the promised copy paradoxes for intensive infiniteness. These paradoxes don't seem to have anything to do with the physical arguments against extension, that is with Relativity, in fact quite on the contrary, they use probabilistic arguments reminding us of Quantum Mechanics, which was rejected by Einstein exactly for the probabilistic features. A probabilistic argument was already mentioned above, namely that picking a random sequence of points on an interval will give a dense set. The existence of such random point sequence was not even questioned and the only problem we could see is the one dimensionality of our interval. Indeed, a two dimensional area like a disc or rectangle can be imagined as a target for shooting, points onto it. But then, regarding the vertically identical hits as same, turns a rectangle target into an interval one:

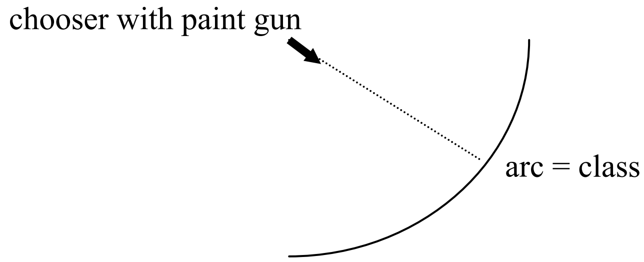


If we want to be really precise then we can't even start with plane targets because these have a slightly higher probability of hits around the directly frontal, that is perpendicular positions than away at the edges. So we should use spherical targets where it's really just the direction of the gun that matters. Then we can still project from these to get an arc as a target, which can be straightened into an interval. This seemingly over detailed idea of spherical targets to achieve perfection becomes a very good way to get a point across, namely that unbounded areas can't be fair targets at all. Indeed, the bigger a plane target becomes so also a bigger sphere is needed to approximate it. But for an unbounded area we would need an infinite sphere. As a consequence, an infinite line or half line can't be a target either. We could also say that, no matter how we choose points from an infinite line, the points toward infinity must have an unfair disadvantage. Wherever we are, the line must have closer points to us. “Subjectivity can not be eliminated from geometry”. A half line or a strip of equal squares towards infinity would be a perfect target representation of the natural numbers:

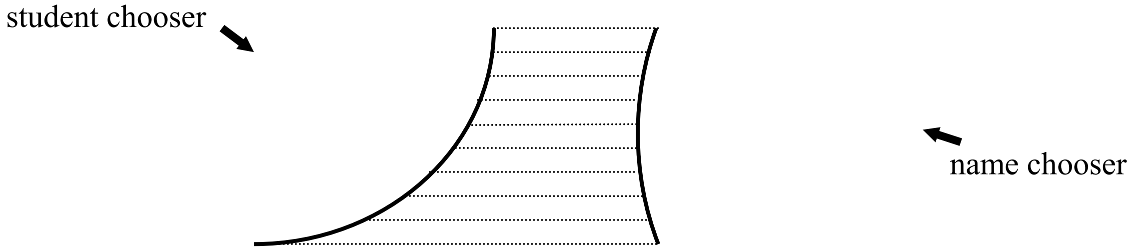
1	2	3	4	5	6	. . . .

but this is unusable as target so we can draw the conclusion:
We can't choose a natural number randomly!
This is quite “natural” even without the whole targeting approach by merely admitting that any numbers taken from the sequence 1, 2, 3, . . . will be only at the very beginning. No matter how big number we pick, it's small compared to the “possibilities”. And yet we somehow still believe that if we could throw all the natural numbers in a big bag, shake it a bit and then put our hand in and take one, then we should get a randomly picked natural number. Are we right or wrong? Before answering this, we show something else that reveals how unreliable our feelings about random choices are.
The picking from the bag reminds us of choosing raffle tickets or name slips. I dare to say that our approval of the process for choosing a member of a class by choosing a name, is a-priori instinctive. Indeed, nobody, regardless of his intellectual level would even question the meaning of the process when we pick a name from the hat. So just as we have a-priori ability to use names, we have some a-priori consequence of name usage. In fact the names themselves are consequence of more basic a-priori abilities, namely associating without learning when recognizing similarities. Indeed, the child recognizes his mother on a picture but knows that it is not the real mother. Later, the learning of names actually hides this instantaneous and sharp difference between human intelligence and animal abilities. The mathematical consequences are a wider and more basic field than all so called “psychologists” realize, but I'm not going to

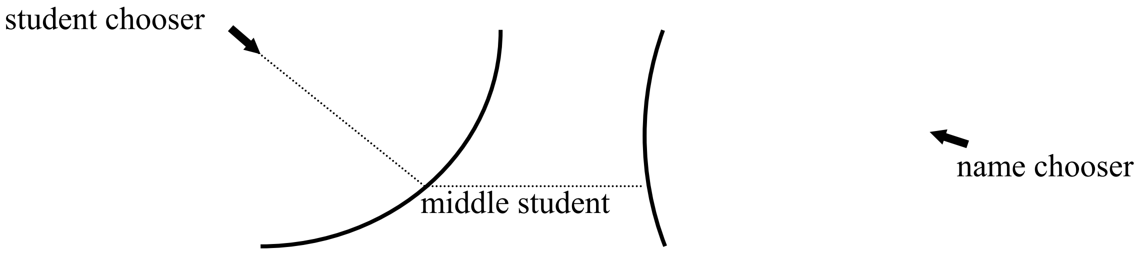
reveal more about this now! Instead, I show that our belief in that a random choice of name is automatically a random choice of the name holder is false when there are infinite many "people". Luckily the transition from finite to infinite is also an a-priori ability so we can all appreciate the surprise. Let the class of students be now an interval of points or for better choosability bend it into arc of circle. Indeed, then a paint gun from the center can choose any "student" with equal chance:



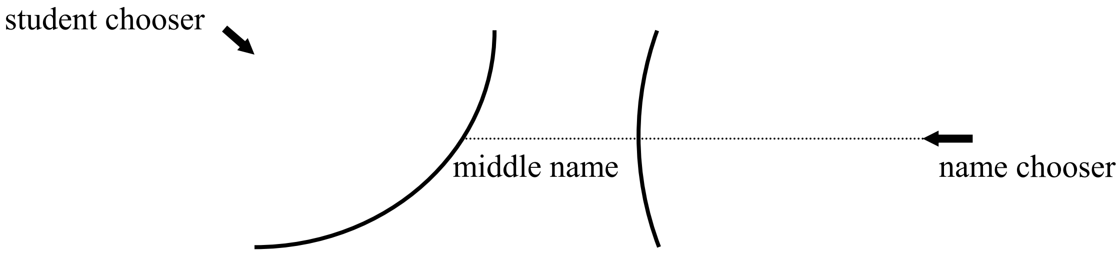
So this is the direct selection of a student which we will replace now with a selection by name. In order to get an easy choice from the names, we use again an arc for the set of names but a much less curved one and simply use horizontal lines to associate the names to each student:



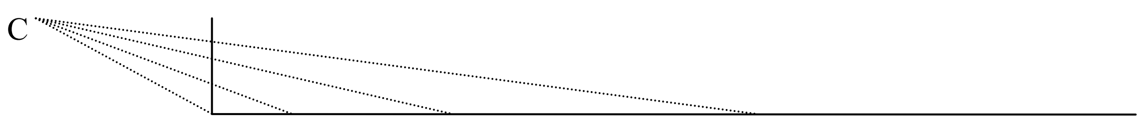
Lets check the middle student in the class:



As we see, the middle student is not middle on the name order but more importantly all the students in the lower half of the class have only a smaller than half chance in the name order. Or looking in reverse a half-half chance on the name order will advantage the upper smaller arc in the student order:



To say mathematically what we realized: <u>The random choice is not equivalence invariant!</u>
This can be shown in an even more drastic way than above by finding equivalence between a randomly choosable and randomly non choosable set. For example, the projection from a C center orders the points of a vertical interval to all points of a horizontal half line:



After one digests these more, he must realize that the basic idea of random choice from a pure set is questionable and we should only talk about random choices from structured sets.
This is of course only for infinite sets because finite sets don't need structure! From a finite set, we can choose a random member and it remains random for all equivalent sets too!
Unlike in infinite sets, where the closeness of the members can "mess up" the equivalence, in finite sets, the "discrete" nature of the elements, guarantee that the choice remains fair by equivalence. But then again, a sequence is discrete too! Of course, it can become dense, like the results of a shooting sequence. Or for that matter, the sequence:

$$\frac{1}{2} , \frac{1}{3} , \frac{2}{3} , \frac{1}{4} , \frac{3}{4} , \frac{1}{5} , \frac{2}{5} , \frac{3}{5} , \frac{4}{5} , \frac{1}{6} , \ldots .$$

is listing all "simple" fractions (less than 1), by increasing denominators, and yet they are obviously dense in [0 , 1]. If sequences keep the fair random choice in general, then our earlier question about choosing a natural from a bag, can be rephrased as:
<u>Is there a sequencable structure from which we can choose a random member?</u>
An amazing, "Yes", answer and the concrete example will be given in the next section.

An other direction for random choices is to generalize the choice itself from members to subsets, like sequences. Then a whole new door opens and formerly randomly unchoosable pools become fair game. For example, from the naturals to pick an infinite sequence, we can simply toss a coin for each number. Since the simple fractions can be sequenced, we can pick a random sequence from these too. Then, an interesting new possibility emerges to pick a random sequence from the half line. First we might think of picking a point from consecutive intervals like all unit intervals between the naturals. But this wouldn't be random because it has the pattern of containing an element in each interval. Instead, we can toss a coin at every interval whether to pick or not from it. But it is still predictable that the sequence would never contain two points closer than the unit. So we have to pick the consecutive intervals randomly too. We can regard a random sequence of rationals for this.
This is still unexplored math, which is the only real math in a sense, but we have to return to the explored sense. As we remember the basic problem was that we can't see the local difference between a sequencable point set and the much more remaining points from the continuum, so we should at least see some global properties that reveal more than the simple fact that the continuum is not sequencable. We aimed for some spatial properties so we needed the concept of distance that involves moving or copying. Before we get to these repeatedly delayed and promised copy paradoxes (in the next section) we have to get into our last and most important detour.
The first attempt for a global though not spatial comparison between the continuum and a sequence was made by Cantor himself. It could be summed up by saying that the continuum is much bigger than a sequence but still not that much bigger if we regard all the possible infinites. The first half was put in an exact form by proving that the possible point sequences taken from an interval is only equivalent to the interval itself. This really means that the possible sequences don't give much more variations than the points themselves so a sequence must be small compared to other subsets because as we proved, the totality of subsets is more than the points of the interval that is the continuum.
The proof of this is amazingly simple by regarding the interval as the [0,1] set of real numbers and then to a sequence of infinite decimals we can order uniquely a single infinite decimal. Indeed, if we list the sequence of decimals under each other, then with the already used diagonal walk we can create a single sequence of digits. This result can also be interpreted in an other way that emphasizes not as much how small a sequence is but rather how big the continuum is. An infinite sequence of points in [0,1] can also be regarded as the infinite many coordinates of the infinite dimensional unit cube, which thus is equivalent to [0,1]. From this it's easy to prove that all the points of the infinite dimensional space is equivalent to any tiny little interval on the line. In short:
<u>The continuum is always continuum regardless of its size and dimension.</u>

The second half was not a proven theorem, only a belief by Cantor namely that: The continuum is the next infinite after a sequence, so if we pick out all the points of an interval then the first subset that is not sequencable in this listing is already equivalent to the whole interval. This has become known as the Continuum Hypothesis. As we saw the continuum is also equivalent to the subsets of the naturals so a Generalized Continuum Hypothesis claims that for every set S between S and P(S) there is no other infinite, that is after S comes P(S). Gödel believed for a while that there is an other infiniteness between the sequencable and the continuum. In 1963 Paul Cohen proved with his new forcing method that the Continuum Hypothesis is undecidable in the present Set Theory and believed that it is not a fault but simply an expression of that the continuum is far beyond the sequencable and not even accessible by any method of listing infinites. The fact that by Zermelo's theorem every set can be listed including an interval could be interpreted as a purely theoretical truth. To accept a theory with such open problems as natural, is pretty hard. Even if it is the case, we would like to get a wider theory to explain why the theory of pure sets must leave exactly this problem open. It's quite clear now that a wider theory must include the concept of random choices and our previous remarks reflected my belief. I am not even reacting to those earlier opinions that questioned the axiom of choice! To pick simultaneously elements from sets is totally acceptable! To regard this as some kind of random selection is false but this doesn't diminish the existence of it. Later we'll see that the copy paradoxes are not weakening the axiom of choice either. A simplified subjective view is that the axiom of choice is true but the listing of the continuum is impossible. This is a bad repetition of Cohen's view. Much better (but still incorrect) to say that the continuum is listable but doesn't have a minimal listing. Indeed, if we keep on picking random points from an interval, this must exhaust it after a while, and there is nothing really problematic in this. It is also acceptable that after a while the picked out set is already equivalent to the full interval. It is quite impossible though to imagine that the full interval is finished exactly at a jump in infiniteness. Instead we must have a real subset from the interval that is already the same size.
The equivalence of the full interval to this subset then of course gives a minimal listing of the interval itself but by our earlier results this doesn't have to be a random picking order. So to be precise:

<u>The continuum is not randomly minimal listable, only artificially.</u>

This distinction saves us from using our plausibilities about random choices for artificial ones. The question of what assumptions we could find about random choices is interesting but would become really amazing if it could link to the non random choices or sets in general and thus decide unknown problems. <u>This hasn't happened yet!!!</u>

The special question of defining random subsets of naturals, that is the possible outcomes of coin tossings was not well publicized in mathematics earlier, in spite of people like Church and Kolmogorov working on it continually. Lately very big connections were made but no new results for basic Set Theory. See Michiel Van Lambalgen's article in Journal of Symbolic Logic 55 "The Aximoatization of Randomness".

So it's quite a claim by Chris Freiling's article also in J.S.L. 51 that he found a simple probabilistic argument that refutes the Continuum Hypothesis. His claim is false and I think he knows it too, but a bit of sensationalizing never hurts a good cause!

Before we turn to Chris Freiling's axiom lets regard an arbitrary relation between people.

It could be the "liking" or "hating" or nomination for an election. So one person can relate to, that is nominate more and it doesn't have to be symmetric, that is reversed. In fact if two people nominate each other it's quite a fishy business. Of course more people can nominate each other in a cycle which is just a trickier way of the same cheating. In a big enough country if we pick some, but not too many people randomly, then it's very unlikely that any of them would nominate, that is relate to any of the others. Of course it depends also on how many persons can be nominated by one. Indeed for example if everybody nominates everybody else then we can't even choose two honest persons at all. If in 50 people everybody nominates 10, then picking 20 people we also should find plenty of nomination pairs. But if among a million people, everybody nominates less than 10 persons then in 5 randomly picked people we would have a very low chance of finding two so that one nominates the other. The smallest possible choice of sample for unrelatedness is of course two and we'll stick to this.

A real advantage comes if we regard an infinite population because then restricting the possible nominations to a non equivalent subset seems to guarantee at once the possibility of easy fair samples of duos. In other words we claim:

T If in an S set every element is related randomly to some elements, but each one only to a subset of elements that is not equivalent to the full S,
then randomly picking two elements from S, none of them will be related to the other.

R Now we might jump from this to the following.

T If in an S set every element is related to some elements, but each one only to a subset of elements that is not equivalent to the full S,
then we can pick two elements from S so that none of them will be related to the other.

R So we omitted the random relations, which is a big widening, to allow artificially constructed relations, but on the other hand to compensate this, we didn't require the random choice of the sample pair either, instead only the existence of at least one such pair. So seemingly we gave up more from the conclusion than we allowed in the conditions. But amazingly our intuition is false, this second theorem is false! The simplest example is a sequence of population, say the natural numbers and every number should nominate the smaller ones! Then indeed every number only nominates finite many that is a much smaller non equivalent subset from the whole population and yet picking any two numbers, the bigger is nominating the smaller. So there are no unrelated pairs! After this we can present Chris Freiling's suggestion:

A Axiom of Symmetry:
If in the continuum every point is related to only a sequence of points then there are two points so that none of them is related to the other.
(The name symmetry refers to that none of the two are related to the other)

T If the Axiom of Symmetry is true then the Continuum Hypothesis is false.

P Enough to show that the Continuum Hypothesis allows a nomination of sequences in an interval so that every two points are related, that is at least one nominates the other.
Indeed, lets regard the minimal listing of the interval! Every point then follows a sequence of points so these can be nominated by it. Then just like at the natural numbers above, for every two points one is the later in the listing and thus nominates the other.

R The above theorem is perfectly correct, the proof is flawless but by the earlier remarks we can see the error in accepting the Axiom of Symmetry! In spite of how minute a sequence is compared to the full continuum, the non random relations could overcome this difference.
We could achieve an artificial ordering without unrelated pairs, even though "normally", that is by random ordering, such would seem impossible to pick. Our earlier remarks also show that the sequencing of an interval already destroys the possibility of randomly choosing the two points for unrelatedness. Thus even a randomized version of the Axiom of Symmetry seems impossible directly.

7. Localization, A class of sets where the Continuum Hypothesis is true

D

The title means that we try to find properties of sets that only depend on the features of the sets around single points and thus can be characterized without distances. Of course, every surrounding of a point has a certain size but we can get over this problem, by the heuristic idea of zooming in infinitely. This means that we regard not one small surrounding around a P point, but a whole sequence of smaller and smaller surroundings. Then it turns out that the amount of points contained in the narrowing surroundings is not depending on the chosen sequence. Furthermore, the amount itself "narrows down" to three possibilities: 0 , 1 , ∞ . Then 0 means that no point of the S set is around P, 1 means that only P is from S in a small surrounding and ∞ means that infinite many points of S are around P. Correspondingly to the 0 , 1 , ∞ cases , P will be called: outer, isolated, approached point. While the 0 case automatically means that P is not in S and the 1 case that P is in S, the ∞ case can go either way, so we could talk about an approached element of S or an approached non element. This later is called a disappearing point of S referring to that S is "there, that is nearby but is disappearing in the end".

The points of the space that are not element of an S set, are called the $\bar{S}$ complement of S. The amount of $\bar{S}$ points around a P point can again be split into the 0 , 1 , ∞ possibilities. Here of course, the names should be different. For the 0 case, it should be inner point because it means that in a surrounding there is no $\bar{S}$ element, so P is truly "inside" the S set. The 1 case should be called a hole, because then in a surrounding, P is the only non element. The inner points are obviously approached elements and the holes are disappearing points. The approached elements that are not inner points could be called mixing elements, because these will have both S and $\bar{S}$ points around them. Similarly, the disappearing points that are not holes can be called mixing non elements.

Thus the six types are: outer , isolated , inner , mixing element , hole , mixing non element

(inner , mixing element: approached element; hole , mixing non element: approached non element = disappearing point)

Every point is exactly one of these six possibilities.

Even though our aim is to characterize the point sets locally, using intervals is still important. Being full or empty in an I interval means that all or none of the points of I are elements of S. If the whole S set is contained in an I interval, then we call S bounded.

The "everywhere", "nowhere", "somewhere" expressions will be used to indicate if something is true in every, no or some interval. Thus, for example: Nowhere full means that the set doesn't contain any full intervals. Nowhere empty means that there is no empty interval in the space.

These can again be specified for an I interval, so: Nowhere full in I means not containing a full interval in I. Nowhere empty in I means having point in any interval in I.

The local properties can be automatically used for a set if we agree that they mean the existence of such points. Thus, for example: Outer S means that S has at least one outer point.

Isolated S means that S has at least one isolated point. Disappearing S means that S has at least one disappearing point, that is an approached non element.

The "un" prefix is used to negate such properties. So: Unisolated S means that S has no isolated point. Unouter S means that S has no outer point, which of course simply means that S is nowhere empty. Undisappearing S means that S has no disappearing point, that is all approached points are elements.

Unouter sets are also called <u>dense</u> or everywhere dense. Undisappearing sets are also called <u>closed</u>, while closed and unisolated sets as <u>perfect</u> and closed and bounded sets as <u>compact</u>.

We can use the local derived properties for an I interval too. For example: S unouter in I or S dense in I, means that S has no outer point in I. This of course, can also be said as S being nowhere empty in I. We can also use the interval expressions. For example:

Everywhere outer S means that S has at least one outer point in every interval. Such set is usually called as nowhere dense. But it should be called everywhere somewhere empty, because in every interval there is some empty interval.

R By looking at a point set, we see the isolated points as being apart from the "bulk" of the set. But if we remove the isolated points, then we might end up again with new isolated points that were approached before. So the question is whether there is a meaningful separation of all sets into a bulk = kernel and loose = scattered part. One way is to repeat the removal of the isolated points again and again, until we get a set without isolated points. Luckily, there is a direct way to get this same final unisolated set:

D

1.) S_0 denotes the set of isolated points of S.
2.) S ' = derivative subset of S, denotes $S - S_0$ = approached elements of S.
3.) Totally approached elements of S = kernel of S = S^{Ω} = $S'' \cdots ''\cdots$ =
the first repeated derivative that remains the same, that is $(S^{\Omega})' = S^{\Omega}$
4.) Scattered part of S = $S - S^{\Omega}$. If $S^{\Omega} = 0$, that is, $S = S - S^{\Omega}$ then S is scattered.

T

1.) S^{Ω} is unisolated.

2.) Union of unisolated sets is unisolated.

3.) If T is an unisolated subset of S, then $T \subseteq S'$. So $T \subseteq S''$, . . . , $T \subseteq S^{\Omega}$.

4.) S^{Ω} is the union of all unisolated subsets of S.

P

1.) $(S^{\Omega})' = S^{\Omega} \rightarrow (S^{\Omega})_0 = 0$.

2.) If P is isolated point of $A \cup B$, then P is isolated point of A or B (maybe both).
So if A and B both have no isolated points, then $A \cup B$ can't have either.

3.) If P is isolated in S, that is $\in S_0$, then it would be isolated in T too, so $P \notin T$.
In negative form, $P \in T \rightarrow P \notin S_0$, that is $T \subseteq S - S_0 = S'$.

4.) By 1.) S^{Ω} is subset of the union of all unisolated subsets of S.
By 2.) this union is unisolated and by 3.) is subset of S^{Ω}. So this union = S^{Ω}.

D

1.) A set is sequential if it is empty or finite or sequencable.

2.) P is sequential point of S, if in a surrounding of P, there is only a sequential subset of S.
The non sequential points of S are also called condensation points of S.

3.) $S_{\leq \omega}$ denotes the set of all sequential elements of S.

4.) $S_{> \omega}$ denotes the set of all non sequential (condensation) elements of S.

T

1.) Sequentiality theorem: $S_{\leq\omega}$ is sequential.

2.) Unisolatedness theorem: $S_{>\omega}$ is unisolated.

3.) Scattered part theorems:
$S - S^{\Omega}$: a.) has only sequential elements of S.
b.) has only sequential elements of itself.
c.) is sequential.

P

1.) The intervals of space with rational coordinates can be listed as I_1 , I_2 , I_3 , . . .
Let I_{n_1} be the first, so that $S \cap I_{n_1} = S_1$ is sequential. The next is S_2, and so on.
Then $S_1 \cup S_2 \cup S_3 \cup \ldots$ is sequential too and contains all sequential elements of S.

2.) If P is non sequential element of S, then in all surrounding of it there are unsequencable many S elements. By 1.) only a sequential subset of these can be sequential, so there are non sequentials too. Thus, P is approached by non sequential points.

3.) a.) By 2.) $S_{>\omega} \subseteq S^{\Omega}$ so $S - S^{\Omega} \subseteq S_{\leq\omega}$

b.) A non sequential element of $S - S^{\Omega}$, were such of S too.

c.) By b.) $S - S^{\Omega} = (S - S^{\Omega})_{\leq\omega}$ which is sequential by 1.).

D

P is continuum point of S if in every surrounding of P there are continuum S elements.
$S_{2^{\omega}}$ denotes the set of all continuum elements of S.

R

All non sequential elements of S are in S^{Ω}, but S^{Ω} can have sequential elements too.
So in $S_{2^{\omega}} \subseteq S_{>\omega} \subseteq S^{\Omega}$ we can have proper subset relation in the second.
For example: If S is nowhere outer, that is dense in I, but only contains a sequence there, like the rationals, then $S \cap I$ will be in S^{Ω}. These "dense" are the only type of sequential elements that we can easily visualize to be in S^{Ω}, but they could be more complicated too.
To show some of these, we'll make a remark at the end of this section.
The first $\subseteq$ relation above can only be proper if the Continuum Hypothesis is false. This is undecidable, so we can only look for some special sets, among which all non sequential ones are continuum, that is the Continuum Hypothesis is true. We achieve a bit more because this will be true not only for the infinity of the sets, but of all the points. That is, in the first $\subseteq$, we'll have equality. The special sets are the simplest possible ones, the undisappearing ones.
Strangely, if we restrict S to be undisappearing, that is closed, then not only the non sequential points of S^{Ω} will be all having the biggest non sequencable infinite, that is continuum in their surroundings, but also the sequential points disappear from S^{Ω}.
So we'll have equality in both of the above $\subseteq$ relations.

T

For undisappearing (closed) S sets:

1.) Continuum theorem: If S is unisolated too, that is perfect, then it can only have continuum elements.

2.) Kernel theorems:
   a.) S^{Ω} is undisappearing too.
   b.) S^{Ω} has only continuum elements.
   c.) S^{Ω} has only continuum elements of S.
   d.) S^{Ω} is the set of continuum elements of S.

3.) Splitting theorems:
   a.) S has only continuum and/or sequential elements.
   b.) S is either continuum or sequential. So Continuum Hypothesis is true for S.

P

1.) Let P be an element of the set and I any interval that contains P inside. Since the set is unisolated it must contain infinite many points inside, but we should just pick two of these. Then pick two I_0 and I_1 intervals that are disjoint, are inside I and each contain one of the two points inside of them. Then again, there must be more points of the set in these and we can again pick I_{00} and I_{01} intervals in I_0 and I_{10} and I_{11} in I_1. And so on, we get an infinite sequence or rather tree of nested intervals. Their common points must be in our set because it is disappearing and we have continuum many possible paths.

2.) a.) Suppose S^{Ω} approaches P. Then S does too, so being undisappearing, $P \in S$. S^{Ω} is unisolated subset of S and approaches $P \in S$, so $S^{\Omega} \cup \{P\}$ is also an unisolated subset of S. But S^{Ω} is the widest such, so $P \in S^{\Omega}$ must be.
   b.) Follows from 1.) and 2.) a.).
   c.) Every continuum element of S^{Ω} is such of S too.
   d.) Follows from c.) and previous theorem 3.) a.).

3.) a.) $S = S^{\Omega} \cup (S - S^{\Omega})$ and the first member has only continuum elements of S, while the second, only sequential elements of S.
   b.) If $S^{\Omega} \neq 0$ then there are continuum elements of S, so it is continuum.
   If $S^{\Omega} = 0$ then $S = S - S^{\Omega}$, which is sequential by previous theorem 3.) c.).

T

Cantor Bendixson

If S is undisappearing (closed), then $S^{\Omega} \cup (S - S^{\Omega})$ is the only way that $S = S_1 \cup S_2$ is possible with $S_1 \cap S_2 = 0$ and S_1 = undisappearing, unisolated (perfect) and S_2 = sequential.

P

S_1 must be a subset of S^{Ω}, so $S_1 = S^{\Omega} - \Delta$ and $S_2 = (S - S^{\Omega}) \cup \Delta$.
If Δ would contain a P point, then to keep S_1 undisappearing, Δ would have to contain all S elements in a surrounding of P. But these are continuum many by previous T 3.) and so Δ would make S_2 non sequential.

R Strange sequencable point sets:

Closed sequencable point sets are visualized as isolated points approaching others, that again can approach others, and so on. Indeed, S^{Ω} being 0 for a closed sequencable set makes this reasonable. For a non closed sequencable point set, we don't have a similar uniform visual pattern. Cantor's discovery that the rationals are sequencable made the sequencable dense sets a well accepted image. These are of course, fully in the unisolated kernel S^{Ω} and we might think that that's all what the non closedness can bring about. In other words, dense sequences belong to the kernel and isolated points only lead to points in the scattered part. To be more specific, then we could expect that:

1.) Unisolated sequence must be dense.
2.) Everywhere empty or isolated sequence is scattered.

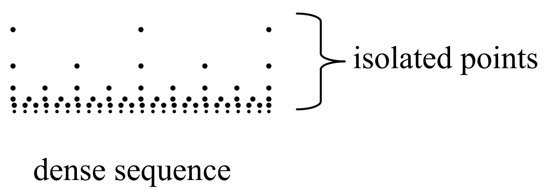
None of these are true, in fact:

T

1.) There is unisolated sequence, that is nowhere dense (everywhere outer).
2.) There is everywhere empty or isolated sequence, that has an unisolated subset, namely 1.).

P

1.) The solution to find such, is amazingly simple.
The rationals or dividers of an interval, are a sequence, because we can spread them successively. For example, just using halvings, P_1 can be the middle point, then P_2 , P_3, the middle points of the two halves, then P_4 , P_5 , P_6 , P_7, again, and so on.
Now all we have to do is to make "double dividers", that is place not one point at each time, but two. So P_1 , P_2 is two points close to each other in the middle of the interval, then P_3 , P_4 , P_5 , P_6 are two-two in the middle of the two halves, and so on. Of course, we have to make the pairs closer and closer, but still, we'll never put points in between these pairs, so these in between intervals remain empty at the end.
2.) Now lets spread "triple" dividers, that is three points! Then the middle one of these three remain isolated at the end, so we get isolated points everywhere where the set is not empty. And yet, the rest of the points, that is the double dividers will be unisolated.
In the plane, we can give an even more visual example of an everywhere isolated sequence with unisolated subset. All we have to do is take a dense sequence for example the rationals, on an interval, and then approach these by isolated points from the plane:



8. Copy paradoxes on the line

D By "copy" we mean identical version of a point set, but this has two possible definitions.

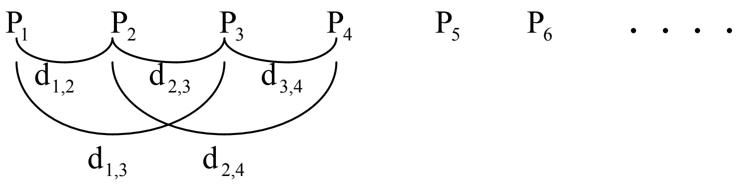
From the "inside", we can define two sets identical if the points can be ordered to each other so that all distances are the same. This definition explains the name "isometric" for copies.
From the "outside" we can move the points of a set together and thus, clearly get a copy of it.
The identity of the isometric copies with the moved copies is a geometrical problem that we ignore here. For example, on a line if we only allow motions inside the line, then we only have shifts. If however, we allow to take a set out of the line and then put it back, then mirrorings would be included too. The isometric definition includes both of these.

R As we saw, a sequence of points chosen from an interval, probabilistically should be dense.

Even though the full line is not a random pool, we can place sequences on it and if we assume that they are dense all over, then we assumed more points than usually the case. If we imagine such dense sequence, then we tend to believe that shifting it will always collide with the original set at infinite many points. Indeed, the dense set has infinitely many inbetween distances all over, so with a lucky moving we can avoid some collisions but so much more remain. To claim that there is a shift that places all the points of the dense sequence into the outer (complement) space of the original set is either a "huge coincidence" for the particular shift or the outer space is huge compared to a dense sequence. But we'll show much more, that indicates that this later is the case. Not only we find such single shift but prove that any fractional, like half, third and so on part of the move, or any multiple of these will put the sequence also into the outer space. And if this is not enough, as a topping on the cake, an obvious consequence of our proof will show that not only all the shifted copies are not colliding with the original ser but they are all disjoint from each other too.

T For any $S=\{P_1, P_2, \ldots\}$ set on the line, there is a d distance so that every $\pm\frac{m}{n}\cdot d$ shifting of S is disjoint from S. The different shifted copies are disjoint from each other too.

P Lets regard all possible distances between two points of S:



This is sequencable simply as:

$$d_{1,2},\ d_{1,3},\ d_{2,3},\ d_{1,4},\ d_{2,4},\ d_{3,4},\ d_{1,5},\ \ldots \quad = \quad d_1,\ d_2,\ d_3,\ d_4,\ \ldots$$

The $\pm\frac{m}{n}$ rationals are also sequencable as we remember, for example by the increasing sum of numerators and denominators with both signs:

$$\tfrac{1}{1},\ -\tfrac{1}{1},\ \tfrac{1}{2},\ \tfrac{2}{1},\ -\tfrac{1}{2},\ -\tfrac{2}{1},\ \tfrac{1}{3},\ \tfrac{2}{2},\ \tfrac{3}{1},\ -\tfrac{1}{3},\ -\tfrac{2}{2},\ -\tfrac{3}{1},\ \tfrac{1}{4},\ \tfrac{2}{4},\ \ldots$$

Leaving out the repetitions like $\frac{2}{2}$, $-\frac{2}{2}$, $\frac{2}{4}$, . . . this can be denoted as $= r_1,\ r_2,\ r_3,\ \ldots$

Then the rational multiples of the inner distances are again sequencable with the usual diagonal method in the infinite square table:

	d_1	d_2	d_3	d_4	. . .
r_1	$r_1 d_1$	$r_1 d_2$	$r_1 d_3$	$r_1 d_4$	. . .
r_2	$r_2 d_1$	$r_2 d_2$	$r_2 d_3$	$r_2 d_4$	. . .
r_3	$r_3 d_1$	$r_3 d_2$	$r_3 d_3$	$r_3 d_4$	. . .
.	.	.	.	.	

By the non sequencability of the line, there must be a distance that doesn't appear in the above squared table, so let one such be d. Since d is not a rational multiple of any distance between two S points, thus neither can any rational multiple of d be a distance between two S points. So any $\pm\frac{m}{n}\cdot d$ move of any $P \in S$ can't fall onto a $Q \in S$. Furthermore the shift between any two S_1, S_2 copies of S is again an $\frac{M}{N}\cdot d - \frac{m}{n}\cdot d = \frac{Mn - mN}{Nn}\cdot d$ that is rational d shift.

R The logical question now is whether in reverse, picking an arbitrary d distance, we can find an S sequence of points so that S can be shifted disjointly by all $\pm\frac{m}{n}\cdot d$ distances.

The construction of such S seems quite simple:

We pick any P_1 point and move it with all $\pm\frac{m}{n}\cdot d$ shifts. Clearly we don't want any of these in our set but these are just a sequencable set so there must be plenty of points outside.

We pick one as P_2. Then we shift P_2 again adding its copies to P_1, P_2 and the P_1 copies.

Then we pick P_3 outside and so on we get the $P_1, P_2, P_3, \ldots = S$ sequence.

No new point can be $\pm\frac{m}{n}\cdot d$ distance away from any already picked one, so our S sequence will have only non rational multiples of d as inner distances. Thus, any rational shift of S will be indeed disjoint from S. Also, any two different rational shifts are disjoint, because otherwise S itself also had to have rational inner distance.

By the way, the possible inner distances in S are only a sequencable set so obviously there are other irrational multiples of d and such shifts of S are still not picked. So we can continue $S = P_1, P_2, P_3, \ldots$ with P_ω , $P_{\omega+1}$, . . . and so on, to get a V set until the whole line would be covered with V or rational shifts of it. We can even restrict V to be in any I interval because if there is a P point outside after an L stage, then there is $\pm\frac{m}{n}\cdot d$ moved copy of P within I too.

The following is a generalization from the rational distances to a more arbitrary set:

D Let D be a set of distances. $\overset{\leftrightarrow}{D}$ denotes the set of all possible finite shift sequences in a line, combined from left or right shifts with any distances from D.

T Vitali: In a line:

For any D set of distances and I interval, that is at least as long as a $d \in D$,
there is an $V \subseteq I$ set, so that:

1.) All $\overset{\leftrightarrow}{D}$ shifts of V are disjoint from V and each other.
2.) V and these shifted copies cover the full line.

P Lets pick a $P_1 \in I$ and shift it with all elements of $\overset{\leftrightarrow}{D}$. Remove these points and pick a next $P_2 \in I$. And so on, P_1 , P_2 , ... , P_ω , $P_{\omega+1}$, . . . will give the V set.

The copies and V will cover I because we kept on picking till all points are obtained.

Thus, the $\pm$ md shifted copies will cover the whole line.

For the disjointness of the copies from V, it's enough to show that for any $P, Q \in V$ they can't be $\overset{\leftrightarrow}{D}$ shifted from each other. Suppose P was picked first, then after this we removed all $\overset{\leftrightarrow}{D}$ shifts of P and so Q can't be such. But also in reverse, $\overset{\leftrightarrow}{D}$ includes both left, right variants of a shift, so if P were a $\overset{\leftrightarrow}{D}$ shifted of Q, then Q were of P too.

For the disjointness of the $\overleftrightarrow{D}$ shifted copies from each other, observe that any R point of such copy is a $\pm d_1 \ldots \pm d_m$ shift of a $P \in V$. Now, if it were common with a $\pm e_1 \ldots \pm e_n$ shift of a $Q \in V$, then the $\pm d_1 \ldots \pm d_m \mp e_1 \ldots \mp e_n$ sequence would move P into Q.

R

Vitali's set was the first "weird" point set, using the new concept of well ordering.

In our definition, it doesn't seem extraordinary at all. In fact, if for example D contains the single d distance, and I is a d long interval, then S will become either [I) or (I] depending on which end of I is picked first and the other end removed. Thus the shifted copies are then these identical d long half open intervals repeating both to the right and to the left. This is the simplest possible covering of a line with infinite many identical sets, so we might wonder why the hocus pocus of picking the points was necessary. Of course, the earlier rational shifts are included in this general theorem, because combined rational shifts are also rational. And above we mentioned how surprising is that S doesn't bump into itself or the other copies. An other, usually more emphasized surprise of a Vitali set becomes apparent from its use in the next proof. In order to appreciate this even more, we won't use our previous general theorem there and rather approach the Vitali set from scratch in an even more modern and visual way. Instead of the step by step picking of the V elements, we'll pick them at once, or rather simultaneously. That's what the Well Ordering theorem does in general, by proving that the successive pickings are avoidable and an actual set exists, where the picking order is merely an ordering of the set. Thus, the concept of time is completely eliminated.

D

A μ function that assigns a 0, positive reals or ∞ to some point sets is called a measure.

The sets for which μ is defined are called measurable.

A μ measure is:

1.) Isometric, if for any measurable S,
and S' that is isometric (congruent, identical or mirror) to S:
S' is measurable too and $\mu(S') = \mu(S)$.

2.) Additive, if for any two disjoint measurable A , B:
$A \cup B$ is measurable too and $\mu(A \cup B) = \mu(A) + \mu(B)$.
Infinite additive, if for any disjoint measurable A, B, C, . . .
$A \cup B \cup C \cup \ldots$ is measurable too and
$\mu(A \cup B \cup C \cup \ldots) = \mu(A) + \mu(B) + \mu(C) + \ldots$

3.) Interval-likely, if intervals have positive real measures, not 0 or ∞.
Interval-faithful, if intervals have their length as their measures.

4.) Differential, if for any two measurable $A \subset B$, $B - A$ is measurable too.
Subtractive if differential and also $\mu(B - A) = \mu(A) - \mu(B)$.
Monotone, if for any two measurable $A \subset B$, $\mu(A) \leq \mu(B)$.
If μ is additive and differential, then of course, it is subtractive too.
If μ is subtractive, then of course, monotone too.

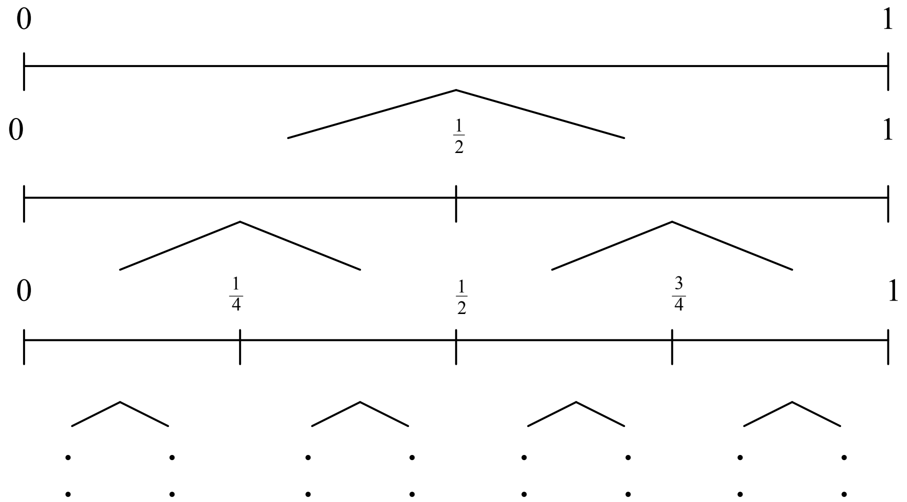
5.) Complete if all sets are measureable.

6.) Continuous if for any measureable widening, $W_1 \subseteq W_2 \subseteq W_3 \subseteq \ldots$ sets,
if all $\mu(W_n) \leq \lambda$ and $\bigcup W_n$ is measureable, then $\mu(\bigcup W_n) \leq \lambda$.
λ-jumps on $W_1 \subseteq W_2 \subseteq W_3 \subseteq \ldots$ if all $\mu(W_n) \leq \lambda$ but, $\mu(\bigcup W_n) > \lambda$.
0-jumps on $W_1 \subseteq W_2 \subseteq W_3 \subseteq \ldots$ if all $\mu(W_n) = 0$ but, $\mu(\bigcup W_n) > 0$.
It's easy to show that infinite additive μ is continuous too.

R All listed properties are expected from a measure. Unfortunately, completeness and continuity exclude each other if 1.) and even the weaker forms of 2.) and 3.) are assumed:

T If μ is isometric, additive and interval-likely, then, μ can't be complete and continuous.. Namely, if it is complete, then it 0-jumps.

P The $W_1 \subseteq W_2 \subseteq W_3 \subseteq \ldots$ sets on which μ 0-jumps, will be created in $[-1, 2]$ as the increasing sums of C_1, C_2, C_3, . . . sets that are all copies of each other.
C_1 will be the crucial original V Vitali set in [0 , 1], while the rest of them will be shifted copies of this.
We first locate the points of [0 , 1] as the limits of repeating halved intervals:



By choosing the left or right halves, we get a narrowing sequence of intervals that determine a point uniquely! The only ambiguity can emerge if we want to locate the end point of the halving intervals themselves. For example, the number $\frac{1}{4}$ could be defined as,
left, right, left, left, . . . or left, left, right, right, . . .
Using 0 for left and 1 for right: $\frac{1}{4}$ = 0 1 0 0 0 . . . or 0 0 1 1 1 . . .
With this same notation, the unique non halving points of [0 , 1] are all infinite binary sequences like 0 1 1 0 0 0 1 1 0 1 1 1 . . . and these are alternating, that is never becoming full 0 or 1 at the end.
This binary representation has a second, much better meaning, than just the mere abbreviation for left and right. Namely, every number in [0 , 1] is a sum of some lengths from the used, $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$ distances and the 1's and 0's can be used to tell which one to use or not.

For example, 0 1 1 0 0 0 1 . . . = $\frac{1}{4} + \frac{1}{8} + \frac{1}{128} + \ldots$

$\frac{1}{2}$ $\frac{1}{4}$ $\frac{1}{8}$ $\frac{1}{16}$ $\frac{1}{32}$ $\frac{1}{64}$ $\frac{1}{128}$ . . .

Even the dual formed halving points are obtained, for example,

$\frac{1}{4}$ = 0 1 0 0 0 . . . = 0 0 1 1 1 . . . = $\frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \ldots = \frac{1}{4}$ indeed.

This is also abbreviated as so called, binary "decimals". In other words, $\frac{1}{4} = .01_2$.

To shift a point with an other point's length from 0, simply means addition of the two numbers. And so it can be performed digit by digit. For us, especially important is the case, where the shift is a halving distance. Then, we merely add a finite long beginning section to an infinite one. For example, 0 1 1 0 0 0 1 1 0 1 1 1 . . . shifted with $\frac{1}{4}$ = 0 1 = 0 1 0 0 0 . . .

$$\begin{array}{l} \underline{0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,\ldots} \\ 1\,0\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,\ldots \end{array}$$

We can easily check that using the alternative form of $\frac{1}{4}$ = 0 0 1 1 1 . . . gives the same.

$$\begin{array}{l} 0\,1\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,\ldots \\ \underline{0\,0\,1\,1\,1\,1\,1\,1\,1\,1\,1\,1\,\ldots} \\ 1\,0\,1\,0\,0\,0\,1\,1\,0\,1\,1\,1\,\ldots \end{array}$$

but we won't use this complicated way.

If we want to allow left shifts, then subtraction must be used instead of addition.
Two points of [0 , 1] are "finitely shifted" if such finite addition or subtraction can move them into each other.
Two infinite binary forms are "commonly ending" if changing some beginnings, they become the same.
Shifting a point is merely adding or subtracting a finite number and every change of beginnings can be achieved by adding or subtracting. Thus, the two concepts, shifting or changing the beginning, are the same.
The commonly ending binary sequences can be imagined to be thrown in a single basket and we'll have such imaginary baskets for all possible endings.
Now, lets pick one representative from every basket!
This chosen set that exemplifies every possible endings, is Vitali's V set.
Now we'll define shifted versions of V, abbreviated as sign subscripted $V_{\pm n}$.
Let < n > denote for an n natural number, the following shift:
1.) We write n in base 2.
2.) We regard this as a halving point. In other words, we put a decimal point in front of it.
3.) We regard the shift of 0, to this number.
For example for, < 9 > :
1.) $9 = 1\,0\,1_2$
2.) $.\,1\,0\,1_2 = \frac{1}{2} + \frac{1}{8} = \frac{5}{8}$
3.) Thus, < 9 > is the shifting of the line that moves 0 to $\frac{5}{8}$.
Using a – in front of < n > means the left directional shift.
Thus, we can define: $V_{\pm n} = V \pm \langle n \rangle$. And the actual sequence of copies is:

$$C_1 = V \;,\; C_2 = V_1 \;,\; C_3 = V_{-1} \;,\; C_4 = V_2 \;,\; C_5 = V_{-2} \;,\; C_6 = V_3 \;,\; C_7 = V_{-3} \;,\; \ldots$$

Finally, the claimed $W_1 \subseteq W_2 \subseteq W_3 \subseteq \ldots$ widening sets are

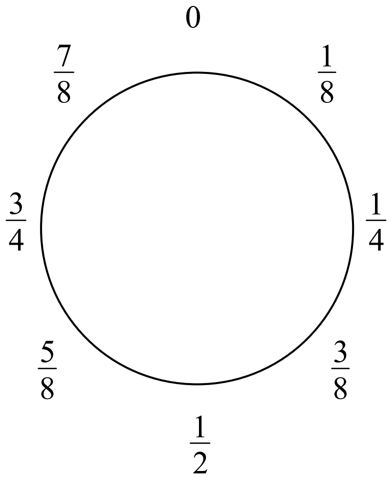
$$W_1 = C_1 = V \;,\; W_2 = C_1 \cup C_2 \;,\; W_3 = C_1 \cup C_2 \cup C_3 \;,\; \ldots$$

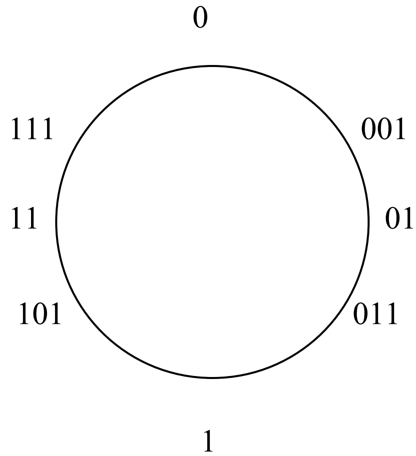
Every point in [0 , 1] can be obtained as a shifted point of V, because every ending appears in V, so we just have to add or subtract a beginning.
In other words, the shifted copies of V, cover [0 , 1] and so $\bigcup W_n \supseteq [0\,,1]$ too.
On the other hand, all shifted copies are within [– 1 , 2] and so $W_n \subseteq [-1\,,2\,]$ too.
μ is isometric, thus $\mu\,(C_1) = \mu\,(C_2) = \mu\,(C_3) \ldots = \mu_0$.
μ is complete, so differential. It is additive, so it is subtractive and monotone too.
Thus, $W_n \subseteq [-1\,,2\,]$ implies $\mu\,(W_n) \le \mu\,[-1\,,2\,]$.
By additivity, $\mu\,(W_n) = n\,\mu_0$ and by interval-likeness $\mu\,[-1\,,2\,] \neq \infty$.
But, $n\,\mu_0 \le \mu\,[-1\,,2\,]$, can only be true for all n if $\mu_0 = 0$ and thus, $\mu\,(W_n) = n\,0 = 0$.

By monotony again, $\bigcup W_n \supseteq [0\,,1]$ implies $\mu\,(\bigcup W_n) \geq \mu\,[0\,,1]$.
And by interval-likeness again $\mu\,[0\,,1] \neq 0$, so $\mu\,(\bigcup W_n) > 0$.

R Lebesgue proved that there is a widest measure that satisfies all our measure properties listed in 1.) to 4.). This so called Lebesgue measure will then be continuous and thus, by our previous theorem, won't be complete. In fact, the used Vitali set will be a non Lebesgue measurable set.
The widest feature might suggest that in some sense, it is still the perfect final measure.
The next example shows that it's not true: Let V be a non Lebesgue measurable set in $[0\,,1]$.
Let $[0\,,1] - V$ shifted with 1, that is into the $[1\,,2]$ interval be V^*.
Then we would expect that $V \cup V^*$ is half of the $[0\,,2]$ interval, so has 1 measure.
Yet, it is also non Lebesgue measurable.
We might even imagine the $[0\,,2]$ interval as a target for "dart throwing".
Then, it's quite convincing that half of the shots would hit and half miss $V \cup V^*$.
Thus, the relative frequencies of hits would give the proper half probability and thus, 1 size.

R By turning the [0, 1] interval into a circle, the halving points or finite binaries are like numbers on a clock.





Points other than these clock numbers can again be located as infinite binaries.
Clockwise turnings with the clock numbers are adding finite beginnings and these are now sufficient. We don't need backwards turnings, because clockwise turnings can reach any point.
In digital form it means that we simply ignore a 1 digit appearing before the decimal point.
A more important result is that here the perimeter is exactly covered by the V Vitali set and its clockwise turnings: $V_1,\ V_2,\ V_3, \ldots$ where $V_n = V + \langle n \rangle$.
Furthermore, we can connect the perimeter with the center and thus, cover a whole disc, except it's center. Then, a dart throwing becomes a real possibility. A hit is an infinite binary sequence. Every hit is landing in a "basket", according to its ending. But it will end in an actual copy of V, according to its beginning.
The members of the $V,\ V_1,\ V_2,\ V_3, \ldots$ sequence, are definitely randomly choosable, thus, give the "Yes" answer and concrete example, we promised in the previous section.
On the other hand: Checking a throw sequence and the $V,\ V_1,\ V_2,\ V_3, \ldots$ copy sequence, the relative frequency of hits landing in $V \cup V_1 \cup V_2 \cup \ldots \cup V_n$ is approaching 0, for arbitrary large n. Yet, landing in $V \cup V_1 \cup V_2 \cup \ldots$, the relative frequency is always 1.
This situation is not that "weird"! An analogy of it could be $V = [0\,,1)$ and $V_n = V + n$.
Then, $V \cup V_1 \cup V_2 \cup \ldots \cup V_n$ is merely a finite interval, while $V \cup V_1 \cup V_2 \cup \ldots$ is the full infinite half line. Of course, the big difference is that we can't throw darts randomly onto the full half line to test the "jump" from finite to infinity.

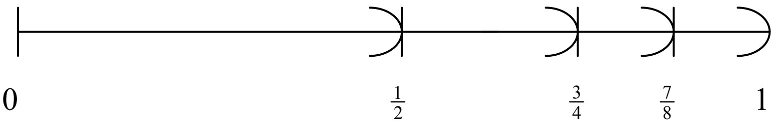
9. Interval paradoxes on the line

D

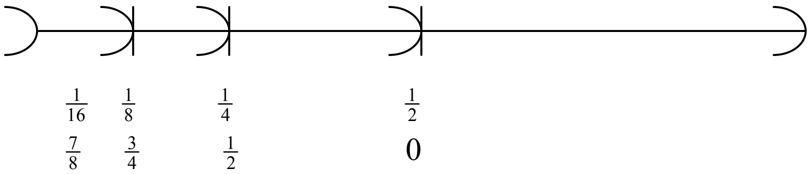
Cantor Set

Cutting an interval into pieces and rearranging them disjointly, is the already mentioned jigsaw puzzle principle. As we promised this principle is paradoxical with even finite many pieces. That of course will happen with arbitrary sets as pieces. Here the pieces are intervals, the simplest possible sets. On the other hand we have infinite many of then and thus the result again is that we can lose points after reassembling the pieces.

The simplest example is if we rearrange the sequence of half decreasing
$[0,\frac{1}{2})$, $[\frac{1}{2},\frac{3}{4})$, $[\frac{3}{4},\frac{7}{8})$, . . . intervals:



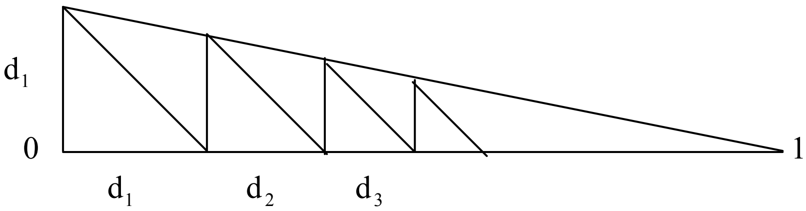
These, in this original order clearly cover the [0,1) interval but all we have to do is put them in opposite order starting from the second half of [0,1) that is as: . . . $[\frac{1}{8},\frac{1}{4})$, $[\frac{1}{4},\frac{1}{2})$, $[\frac{1}{2},1)$.



We put under every fraction "what it was" before the reordering.

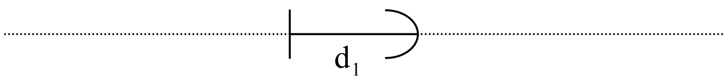
As we see only (0,1) is covered, so we lost the 0 point.

We don't have to decrease the pieces in half of course. In fact a simple proportional decrease can be achieved with any d_1 starting piece:

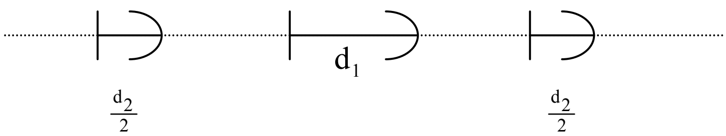


More importantly we can lose infinite many points with the following method:

We place d_1 into the center of [0,1).



Then cut d_2 into two equal pieces and place each into the center of the two windows aside d_1.

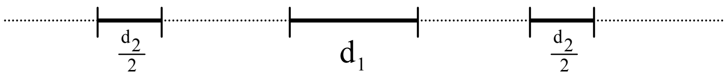


Then cut d_3 into four equal pieces and place each again into the center of the four windows left above. And so on, we feel the $[0,1)$ interval but now not only the 0 point will be missing but also the right end points of all the d_1 , $\frac{d_2}{2}$, $\frac{d_2}{2}$, $\frac{d_3}{4}$, . . . placed intervals.

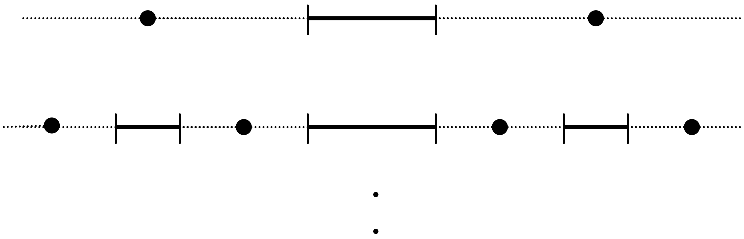
First of all the fill up of $[0,1)$ is itself paradoxical if we start with really small d_1. We tend to believe that if we place small intervals then they can't cover the full interval and thus an effective length of holes will remain. But this is impossible because $d_1 + d_2 + d_3 + . . . = 1$.

What happens is, that even if we start with very small d_1 then the proportional decrease compensates, that is slows down the decrease, so the sum is always 1 independently of d_1.

But the real shocker is this: It's not just the right end points of the placed d_1 , $\frac{d_2}{2}$, $\frac{d_2}{2}$, $\frac{d_3}{4}$, . . . intervals that will be lost! To test this, the simplest way is if we add the right end points, that is we make the d_1 , $\frac{d_2}{2}$, $\frac{d_2}{2}$, $\frac{d_3}{4}$, . . . intervals closed in both ends:



What we claim is that in the end, there will be holes with even this "covering".
To see this, lets regard the center points of the intermediate windows.



As we see, they all become eliminated by the newly placed intervals, but every eliminated one will be "replaced" by a left and right new ones underneath in the two new windows.

Now if we go down in a sequence of such centers always using the left or right "replaced" one as next, then these will approach some points on the line. This is so because they are always contained in the nested smaller and smaller window intervals.

So what are these approached points? Can they be in the placed d_1 , $\frac{d_2}{2}$, $\frac{d_2}{2}$, $\frac{d_3}{4}$, . . . intervals? If so they can only be end points of these. But to get an end point of such, say the right end point of d_1, we always have to choose the left window. Indeed, if our choice sequence of centers from a stage, contains all left or all right choices, then we'll get the end points of the placed d intervals. But of course there are so many other ways to go. For example: left, right, left, right, . . . These will lead to holes not on the d intervals. Using the middle points of the windows, was immaterial at all and only the sequence of left-right decisions determine the approached end points. The shock of our result increases even more if we realize that the left-right choices can be regarded as 0,1 digits of infinite binary fractions:

left, right, left, left, right, . . . = .0 1 0 0 1 . . .

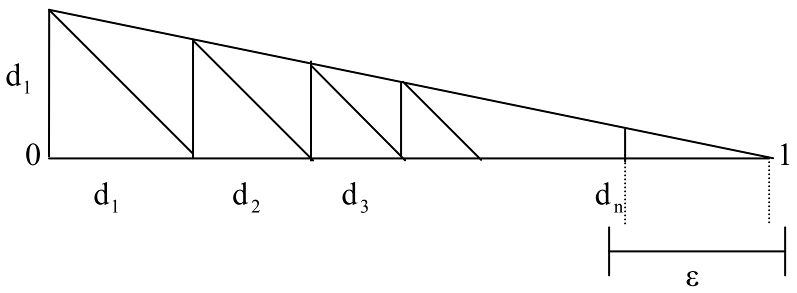
So, only the infinite fractions that end with all 0-s or 1-s lead to end points of d-s. These are merely a sequence of points, while the set of all possible binary fractions is a continuum.

So:

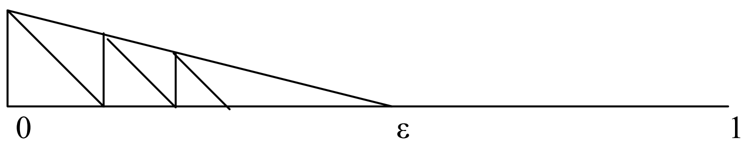
We have continuum many holes missing by our reassemblance of the proportionally cut [0,1)!

We might jump to the conclusion that then the Lebesgue measure is useless because it must regard these holes as 0 measured, being the complements of the d intervals that are clearly 1 together. But this of course must mean by the additivity that the holes are a nil-set.

Seeing this directly, can put our faith back in to the Lebesgue measure and simply make us accept that something that is big as infinite can be small in measure. In fact the holes are not only a nil-set but a zero-set! Indeed, at one stage there are only finite many windows, which cover all the holes and the total length of them becomes arbitrary small. So all we have to do, is go far enough in our proportional d distances, take an ε interval that covers the remaining small interval and then cut this into finite many pieces to cover the windows:



The holes together are called the Cantor Set and since these are much more than the lost end points, we can forget about the strict left closed right open intervals and place any d distances we wish. The total of these d distances doesn't even have to be 1, in fact we can again nicely generate a D set of proportionally decreasing distances with arbitrary ε total as:



D = d_1 d_2 d_3

Then of course the measure of the holes, that is the Cantor Set will be $1-\varepsilon$ positive value, making it a much less shocking set.
Usually the Cantor Set is defined oppositely as I showed it. That is, not as the holes remaining from the wider window holes but instead the d intervals are regarded as cut out windows and then the remaining set will be the Cantor Set. Then regardless of its measure and continuum infinity, the more general striking feature is that it has missing full intervals within any interval where we look at it. This is what we called in section 7 as everywhere outer or nowhere dense set. By the way, in that section we proved the continuum theorem, which at once shows that the Cantor set is continuum, if we use open cut outs.
Indeed, with open intervals removed, we can not lose an approached point, so the final set remains undisappearing. Also, not cutting out two windows right next to each other, that is never creating isolated points in the process, the final set won't have isolated point either, so it remains unisolated. Thus, by the continuum theorem, the final Cantor set remains continuum.

R Nil-sets

The simplest such set is the set of dividers or rationals. They can be sequenced as $r_1, r_2, r_3 \ldots$ and thus can also be covered by $\frac{\varepsilon}{2}, \frac{\varepsilon}{4}, \frac{\varepsilon}{8}, \ldots$ intervals, which in total is the arbitrary small ε. This was the argument that I already mentioned at the very beginning of this book in the History of sets. I found it a mystery that the greeks didn't realize this whole line of reasoning, even though the question of irrational numbers was their main concern. Probably they already missed to realize, the sequencing of rationals. In fact, this is always taught in math as a discovery of Cantor. If it was really him who realized this first, then that would be an even bigger mystery. In fact, I can't believe that Euler or Gauss wouldn't have been aware of this and I think they simply didn't mention it because they couldn't use it for anything, and as good formalists they didn't speak about "useless" metaphysical truths until they could cash them in as irrefutable proofs. At any rate, if we try to use this nil-set property of a sequence to prove that a whole interval is not sequencable, then we run into some interesting details:

First, the argument seems obvious because we feel that a whole interval can not be a nil set. In fact, we feel that an I interval can not be covered with pieces of a smaller interval. Here "pieces" means the cutting of the smaller interval into arbitrary many even smaller intervals. As we remember with a projection, we can increase an interval into any size, but then we break the interval into its points. Still, this shows that breaking up only into intervals may also lead to a change in size. And indeed, we saw above that we can lose lots of points just by reassembling interval pieces of an interval. In spite of this, one thing is obvious: If we cut an interval only into finite many pieces, then the total sum remains the same. So if we could replace a covering from infinite many pieces with only finite many pieces then that would lead to a correct proof. The easiest "finitization" would be simply to keep finite many pieces from an already existing cover. Amazingly, this can be done if we assume a stronger version of covering. This stronger requirement is that any point that we cover should not only be in the interval pieces, but rather be "inside" of the intervals. By this, we mean that the points shouldn't be covered by the end points of the covering intervals, rather by inner points. Such covering system of intervals could be called an inside cover and then we claim:

T If an I interval is inside coverable by a system of intervals,
then finite many intervals of the system cover I already.

P Suppose I couldn't be covered by a finite subset of the intervals.
Lets cut I into a left and right half and then at least one of them is again not coverable by finite many intervals of the system. Then we cut this half again in two, and so on. We get a narrowing sequence of intervals, each being uncoverable by finite many intervals of the system. But the P point to which the halved intervals narrow down, is covered by our system, so it is inside one of the intervals of the system. Thus the narrowing intervals should be all covered too after a small enough member. This contradicts that they are all uncoverable by finite many of the system, namely a single one covers them already.

R Now we can use this theorem to prove that an I interval is not coverable by pieces of a smaller interval. Indeed, from any covering system, we can make an inside cover just by increasing all the pieces by an arbitrary small proportion both to the left and to the right. Thus, for a cover of I, that is smaller in total than I, we could make an inside cover that is still smaller than I. But then this would have a finite subset already covering I, which is impossible.
In particular, this shows that a full interval can't be not only a zero-set, but a nil-set either.
The difference between the zero sets and the nil sets are related to the everywhere outer sets:
First of all, all zero-sets are everywhere outer or in negative form if a set is nowhere outer, that is dense in an I interval then it can't be zero-set. Indeed, finite many intervals, smaller in total than I, must leave out a sub interval of I and it must contain some point of the set.
The opposite is not true with two easy classes of counter examples:
Firstly, a set combined from finite many intervals can be covered by a single interval that is, it is bounded. So unbounded sets can't be zero sets and thus an unbounded everywhere outer set like the natural numbers are the simplest example of everywhere outer yet non zero-sets.
Secondly, among the bounded sets we have the Cantor sets with arbitrary positive measure so they are non nil and thus non zero yet everywhere outer sets again.
So, the first counter examples are nil but unbounded while the second bounded but non nil.
This clearly provokes the question if bounded <u>and</u> nil everywhere outer set must be zero or not.
The answer is "not" and to find the counter example we have to combine smaller and smaller zero Cantor sets. If we just place one in each of the: $[0\,,\,\frac{1}{2}]\,,\,[\frac{1}{2}\,,\,\frac{1}{4}]\,,\,\ldots$ intervals, then unfortunately the total is still zero-set because the infinite many end intervals themselves can be covered by one single arbitrary small interval. If however we put the smaller and smaller versions the same way as the d distances were at the making of the Cantor Set, that is starting in the middle and then at the sides and so on spreading them all over, then a cover by finite many pieces is impossible.

There is an other reason why this combining of Cantor zero-sets into a non zero only nil-set is useful! This is simply the fact that originally the only nil-sets we could visualize were sequences but this method obviously gives continuum many points because every version in itself has that many.
Actually, this method of combining a sequence of zero-sets seems like a generalization of sequences themselves, regardless of the whole infiniteness result. So we may ask whether it is the final step, that is whether every nil-set can be obtained as a sequence of zero-sets. The answer is no again, but to show a nil-set, that is not such, we first formulate a very general and beautiful theorem. This is a widest generalization of Cantor's most famous theorem, that an interval is not sequencable. To be able to say it in an even more condensed form we can agree to use the term "full" for having a full interval. First of all we repeat that full, outer, dense relate as follows: Being outer and dense are opposites, or rather dense is just a bad name for not being outer. Being full is the most obvious way of not being outer, that is being dense. Being nowhere full means not having any full interval. The non sequencability of an interval can at once be said as: A sequence of points is nowhere full.
Indeed, if a sequence would fill up any interval, that is were full somewhere then the original proofs of an I not being sequencable defy it by simply regarding the subsequence that is in I.
Now, a sequence of everywhere outer sets might be placed always into the previous windows so that we lose the outerness at all. But the good news is that, fullness can't be achieved anywhere. That is:

T A sequence of $S_1, S_2, \ldots$ everywhere outer (nowhere dense) sets together, is still nowhere full.

P The proof also resembles the one for the unsequencability of an I.
Suppose $I \subseteq S_1 \cup S_2 \cup \ldots$
S_1 is everywhere outer, so in I too, so there is an I_1 in I outside S_1.
S_2 is everywhere outer, so in I_1 too, so there is an I_2 in I_1 outside S_2.
And so on we find a sequence: $I \supseteq I_1 \supseteq I_2 \supseteq \ldots$
The common part of these is not empty if we used closed intervals, which we could.
Every point in an I_n was outside S_n, so choosing any P point in the common part will be in all of $I_1, I_2, \ldots$ and thus outside of all $S_1, S_2, \ldots$
But this contradicts the original $I \subseteq S_1 \cup S_2 \cup \ldots$ assumption.

R If we call a set, that is obtainable as a sum of sequence of everywhere outer sets, simply as "obtainable" then this theorem leads to consequences showing that the "obtainable" and "unobtainable" sets are the most important division of point sets. After Baire realized this in 1899 he called the obtainable ones as first category and the unobtainable ones as second.
Of course I won't use these totally formal and unrevealing names.
First of all, the obvious consequence of the above theorem is that no I interval is obtainable, but with a trick the same can be proved for many other sets:

T Baire Unobtainable Complement Theorem
If an $S \subset I$ is obtainable then the $T = I - S$ complement is unobtainable.

P $S = S_1 \cup S_2 \cup \ldots$ with everywhere outer sets.
Now suppose that $T = T_1 \cup T_2 \cup \ldots$ were similarly.
Then $I = S_1 \cup T_1 \cup S_2 \cup T_2 \cup \ldots$ were too, contradicting the previous theorem.

R The reverse is obviously false, that is the complement of an unobtainable set doesn't have to be obtainable. Indeed, we can cut an I into two half intervals.

The easiest example of the theorem is that the irrational points of an I are not obtainable. Indeed, the rationals are obtainable already as a sequence of single points. In this simple situation, the rationals are a nil-set and so the unobtainable irrationals have measure 1. But this is not always necessary! In fact, to get our promised nil-set, that is not the sum of zero-sets, we'll use the Baire complement theorem again but with an opposite measure situation. In other words, we'll make an obtainable set with measure 1, and so we get an unobtainable complement with measure 0. Since a zero-set is everywhere outer too, thus our nil-set complement will be not only unobtainable, that is unobtainable from everywhere outer sets but also unobtainable from zero-sets.

To get the obtainable 1 measured set, lets remember that we made Cantor Sets with D cut out sets having ε total sum. So if in [0,1] interval we use $D_1, D_2, \ldots$ cut out sets with totals of $\frac{1}{2}, \frac{1}{4}, \ldots$ then we get

$S_1 = [0,1] - D_1$, $S_2 = [0,1] - D_2$, $\ldots$ Cantor Sets with measures $\frac{1}{2}, \frac{3}{4}, \ldots$

Thus, $S = S_1 + S_2 + \ldots$ is obtainable and has measure 1.

So $[0,1] - S$ is our nil-set, that is not the sum of any sequence of zero-sets.

This nil-set can be directly defined too, as the common part of the $D_1, D_2, \ldots$ cut out sets.

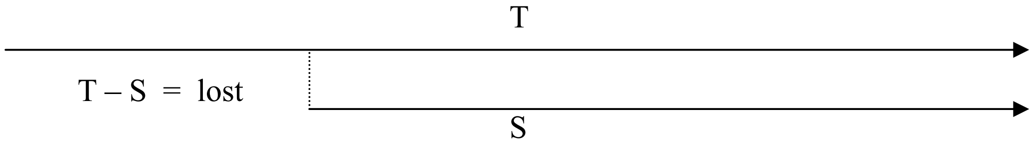
The first $d_1^1, d_1^2, \ldots$ middle distances of the $D_1, D_2, \ldots$ sets are simply centered in each other, but later elements are cutting into each other or can be completely disjoint. So it's very hard to visualize what points will be actually in the common part of $D_1, D_2, \ldots$

Of course, being unobtainable, this common part can't be an everywhere outer set, and neither can be a sequence of points. Being a nil-set it doesn't contain any full interval either. So this set is unsequencable but doesn't "have to have" continuum many points. In fact, this set should be the best prototype to test the Continuum Hypothesis. By Cohen's result nobody can prove it to be less than continuum and as far as I know nobody proved it yet to be continuum. So it might be a concrete set with undecidable infiniteness. Of course, the first non sequencable subset in a listing of an interval is also such undecidable set but this one is obtained without listing, "geometrically" from intervals.

10. Sub-copy paradoxes

1.) S is sub-copy of T, if $S \subset T$ and S is a copy of T.
2.) S is half-copy of T, if both S and $T - S$ are sub-copies of T.

Already a sub-copy sounds paradoxical, but a simple half line can be slided into its subset, leaving out any beginning of it:



In fact, this same works for the individual points of the natural numbers and then we remember Galileo's dilemma about the odds and evens. What puzzled him was, that the odds and evens are not only identical looking to each other, but to the original whole set too. Of course, the "identical" looking is not quite true because the spacing the elements is twice in the half sets. It's an amazing fact that realizing the equivalence to the half is so overwhelming that one misses to even contemplate whether a set could be actually moved into both of its halves. I know, I didn't think about it and none of my friends did either when we learn the basic infinites in the special math high school. This is even more surprising if I tell that the basic ideas of the Theorem of the Domain, Non Standard Analysis, and even forcing, all "came to me" without knowing about them and geometry was my favorite subject. I was more than forty years old when I learnt about the copy paradoxes and never before occurred to me the possibility of these. Pretty embarrassing to me but what can I do? Well, explain them as well as possible.
From the above definition and the examples following it, we might think that sub-copy is the mild paradox because the simple shift to the infinity yields it and half copy is the "real" paradox, which will follow in the next sections. Actually however, the existence of sub-copy can be strengthened in two directions and thus giving four possibilities if we regard these two combined as well. One direction is half-copy but looking at the example of shifting to infinity, we may ask if the extensive infiniteness, in other words whether we could find a <u>bounded</u> set that can be moved into a subset of itself. And then of course, the weirdest set would be a bounded one that can be moved into its two halves. So the four levels of paradoxness should be:

1.) Unbounded set with sub-copy.
Such is the half line or naturals and the sub-copy can be obtained with a simple shift.
2.) Bounded set with sub-copy.
A simple shift clearly can't give a sub-copy so, turning must be involved somehow.
3.) Unbounded set with half-copy.
Such was found in the plane by Sierpinski and Mazurkiewicz.
4.) Bounded set with half-copy.
Such was found by Hausdorff on the sphere and soon led to the famous Banach-Tarski paradox.

To get 2.) the easiest way, we should first still stay with shiftings on the line. Our first paradox was the showing of how vast the line is compared to an S sequence. Namely, S can be shifted by all rational multiples of a d distance disjointly. An obvious special case of this of course is that, S can be shifted with $d, 2d, 3d, \dots$ and all the obtained $S_1, S_2, S_3, \dots$ copies are disjoint. This resembles the shifting to infinity but there we use only one shift, while here a sequence of them and there we move into a subset while here outside the set. So it's totally different! But our hunch was good because with a heuristic trick, the sequence of shifts can be combined into one. Indeed, imagine that the $S_1, S_2, S_3, \dots$ are already on the line and then move the original S and these copies all together with one single d shift.
Clearly, S goes into S_1, S_1 goes into S_2, and so on. Thus, the $S + S_1 + S_2 + \dots$ set moves into its $S_1 + S_2 + \dots$ subset. Big deal you might say. We had that paradox in 1.) already.

But did we? Not really, because there we only lost a beginning of the line or natural numbers. Of course instead of the naturals we can imagine any set in place of them like little ducks standing in a row and then shifting them we can lose some ducks. But still we can only lose a bounded set. Here however S was only assumed to be a sequence, but can be unbounded on the full line. So we obtained a paradox somewhere in between the 1.) and 2.) levels, because the set is not bounded but the lost subset is unbounded too. We might even argue to place this paradox level higher than 1.) because though we were able to lose an unbounded subset, namely S from the $S + S_1 + S_2 + \ . \ . \ .$ set, but S had to be a sequence and so $S + S_1 + S_2 + \ . \ . \ .$ is also a sequence. So the price of the unbounded loss was sequencability and this is at least as big a restriction as boundedness. No matter how we degrade our find, it will be vital not only in creating 2.) but even in 4.) to go to the Banach-Tarski paradox.
The repeating $S \ , \ S_1 , S_2 , S_3 , \ . \ . \ .$ sets reminds us of a cycle and we mentioned how we'll need turnings so the natural idea is to roll up the half line onto a circle. First of all, the S set can be unbounded, so we would only finish with it after rolling up the whole half line. The problem of course is that it can bump into itself. And then we still have to roll up all the infinite many copies of it. We might remember that the inner distances are only a sequence, so some tricky rolling up is possible without collisions, but there is a much better way to go for our goals. We go backwards. That is, imagine a T point set on a circle already and roll this down infinite many times onto the half line to get S_0. So we start with an already periodic $S_0 = T^0 + T^1 + T^2 + \ . \ . \ .$ set, where T^0 is the T set by simply straightening the circle onto the line and then $T^1 , T^2 , \ . \ .$ are the c , 2c , . . shifted copies of T^0, with the c circumference of the circle. Now if T was sequencable then of course S_0 is too and thus we'll have a d distance so that the $S_1 , S_2 , \ . \ . \ .$ copies of S_0 shifted with d , 2d , . . . are all disjoint from S_0 and each other. Now lets roll back to the circle not only S_0 that goes into T, but all the $S_1 , S_2 , \ . \ . \ .$ copies which will go into some $T_1 , T_2 , \ . \ . \ .$
Just as the $T^0 , T^1 , T^2 , \ . \ . \ .$ copies in S_0 roll into each other and form T on the circle, similarly each $S_1 , S_2 , \ . \ . \ .$ will be a compressed set with all the c repetitions becoming one. On the other hand, the $T \ , \ T_1 , T_2 , \ . \ . \ .$ sets will be disjoint! Indeed, two P , Q points from two different S_m, S_n could only become one on the circle if they were at $k \cdot c$ distance away. But then the point being to the right of the other, say Q would have to be in S_m too, because S_m is c-periodical. This contradicts that S_m and S_n are disjoint. Thus, the single d shift of the $S_0 + S_1 + S_2 + \ . \ . \ .$ set that loses S_0 will become a single turn of the $T + T_1 + T_2 + \ . \ . \ .$ total set, losing T. This turning must be done with the same d length, but measured on the circumference of the circle.
The same turn can do this on more than one circles if they are having a common axis:

T Let an A point set be the total points of a set of circles around a common axis. Then:
For any S sequence of points from A there is a $T \supset S$ wider sequence from A and an α angle so that turning T with α around the axis will move T into T – S. That is, we'll lose the S set.

R Of course some points of A will go in place of S so S is not lost from the full A. But if we can find other tricks for A – S with a well chosen S, then getting rid of S first can be useful.

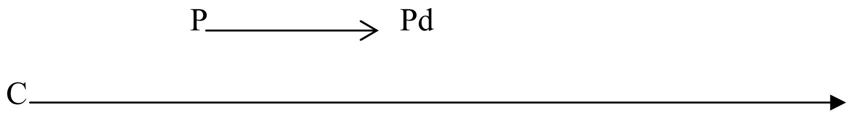
11. Unbounded set with half-copy in the plane, Sierpinski-Mazurkiewicz paradox

D

Lets choose in the plane:

1.) A C center point.
2.) A half line or direction from C.
3.) A d distance.
4.) An α angle.

Then for every P point of the plane:
Pd is a point d distance away from P measured in the same direction as our half line.

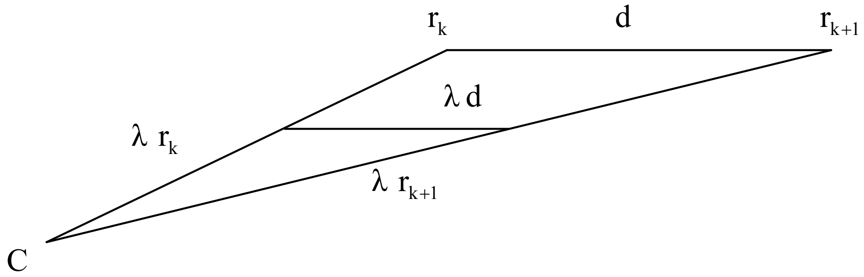


In other words, Pd is the d shifted from P.
$P\alpha$ is a point the same distance from C as P but has α angle between, measured anticlockwise:

In other words, $P\alpha$ is the α turned from P.
These two can be applied repeatedly and mixed. For simplicity we omit the use of brackets, so for example: $Pdd\alpha d\alpha$ means the $((((Pd)\,d)\,\alpha)\,d)\,\alpha$ point. In particular if we start with C then we even omit that from the beginning, so for example $dd\alpha d\alpha$ denotes the point obtained from C after applying these.

R

Every finite sequence of d and α will determine a point, but not each of them, necessarily a different one. In fact, in all sequences that start with one or more α, these can be omitted, because the turning of C is itself. In more general, the question whether there are identical or coinciding points among some sequences can not depend on d.
Indeed, lets look at the r distance of the sequence points from C and what angle these r connectors close with our half line. I claim that changing the used d for λd will change r also to λr, but keep the angle of r. To see this, the best if we regard any particular sequence as one by one increasing and follow how r changes, that is the $r_1, r_2, \ldots r_n$ sequence. For r_1 the claim is obvious. If it's true up to r_k then applying an α as the (k+1)-th step clearly keeps it. Also, if the (k+1)-th step is λd instead of d, then again r_{k+1} becomes λr_{k+1} but with same angle:



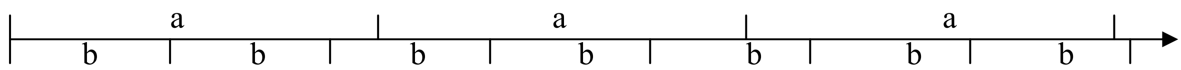
This implies that, if two sequences coincided with using d, then again they will with λd.
So, since d is immaterial and sequences starting with α-s obviously coincide with the ones when omitting these, thus the real question is whether sequences that start with d can be never coinciding:

D 1.) d-sequences are ones that start with d.
2.) α is non coinciding, if all d-sequences determine different points.

T Let α be a non coinciding angle, S be the set of d-sequences and T be the points determined by the S elements plus C. Then T has half-copies! Namely:
If Td denotes the set of points obtained by shifting the elements of T with d and $T\alpha$ denotes the set of points obtained by turning the elements of T with α, then Td and $T\alpha$ are disjoint and together give the full T.

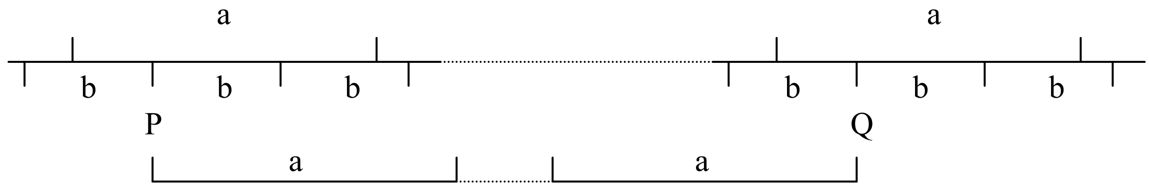
P The proof is at least as paradoxical as the theorem itself because we don't need any geometrical tricks. In fact, we won't even rely on the full meaning of α being non coinciding. All we need is that those sequences that end with d, determine different points from the ones that end with α. Indeed, then Td and $T\alpha$ are disjoint because the first contains all the d-sequences ending with d and the second, all the d-sequences ending with α plus C, because $C\alpha = C$. And of course we get the full T because every sequence in S must end somehow.

R The T set might seem at first as a twisted grid system but actually the points are dense in the whole plane. This would become obvious if we showed that the d, $d\alpha$, $d\alpha\alpha$, . . . points are dense on the circle with d radius around C. We can better imagine these points by rolling the circle on a half line up to infinity. Indeed if the circle has a perimeter and the α angle determines a b arc on this circle, then we'll get repeated a and b intervals up to infinity:



This way the points appearing after different cycles around the circle will become points in the consecutive a intervals. If b is half, third or any whole m-th division of a, then it returns to the beginning on the circle. Or on the line, the repeated b-s match the end point of the first a.
If b is m times of an n-th division of a, that is $b = m \cdot \frac{a}{n}$ then it is also the n-th of the m a interval, that is $b = \frac{m\,a}{n}$. Then of course m a and n b are equal intervals, so after m repeat of a it catches up with n repeats of b. From then on again everything repeats.
If b is neither a full division of a, nor a multiple of such, then no catch up or coinciding is possible because that would imply backwards that $b = m \cdot \frac{a}{n}$.
This question was the so called commeasurability or rationality problem that bothered the greek geometers the most. The word "commeasurable" refers to that the $\frac{a}{n} = \frac{b}{m} = u$ divisions of a and b give the u common unit interval, in which they both can be measured by simple natural numbers, namely a = n u and b = m u. The word "rational" is less logical because "ratio" is proportion and it can be between any two distances, but here a <u>whole</u> ratio is meant. Usually if one of the a , b distances is taken as a unit then the other commeasurable is called a rational number ratio of the unit.
The greeks realized that there are pairs of distances that are not commeasurable. Or to put it an other way, choosing any unit there are always non rational or irrational numbers. They didn't quite continue the search to see how repeated non commeasurable distances relate to each other. As we said, it's obvious that they never coincide again, but we could still imagine that some patterns appear in how they cut into each other. The best way to visualize this is to take the end points of the repeated b-s in the first a and copy them onto the second, third and so on a, then copy again the end points from the second to the third and so on.

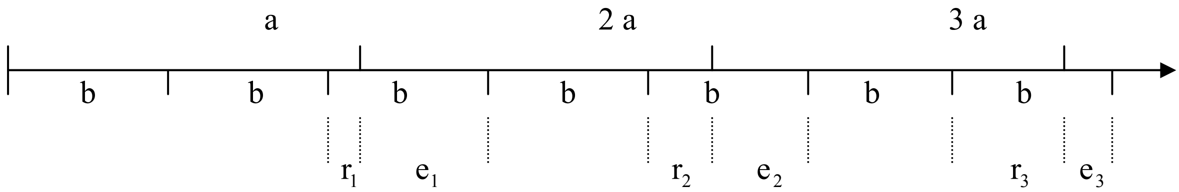
First we might think that a new end point can coincide with a copied old one, but this is impossible because it would imply the coinciding of a and b. Indeed, then the distance between the P point and its Q copy would be a whole multiple of both a and b:



In our original circle picture, this means that if the b arc didn't return to its original beginning after any full cycles, then none of the other end points of the b arcs will be coinciding either. The obvious question is then, how all the appearing distances will behave. Our original claim was that the end points of the repeated b arc will be a dense set on the circle. In our new flattened and repeated cutting point picture this seems quite plausible because the new cuts are indeed always "new", so they must cut in between old ones. Also, they seem to be all over, not leaving any blank full intervals. There is a beautiful way to prove this with a long detour. This will not only show our claim but reveal much more, including the most important theorem of number theory, the Unique Prime Factorization. An other importance of this detour is that the three versions of infinity combine here. Indeed, the original situation of an arc repeated around a circle is the realistic infinity of time in a cycle or periodic motion. The repeated a , b intervals on the half line are examples of the extensive infiniteness of space, while their divisions are the intensive infinity:

D

1.) The extensive infinite investigation of two $a > b$ intervals, is measuring them repeatedly on a half line and looking at how these multiples relate to each other.
This is characterized by the r_m remainders and e_m excesses:
r_m is the distance of the closest b end point to the left or at m a.
e_m is the distance of the closest b end point to the right or at m a.



r_1 is usually just called the remainder of b <u>in</u> a.

2.) "Coinciding" is the special extensive case when: $m\,a = n\,b$ that is $r_m = e_m = 0$

3.) The intensive infinite investigation of a , b is regarding the half, third and so on m-th , n-th part of them and looking at how these dividers relate to each other.

4.) "Commeasurability" is the intensive case when they have common dividers: $\frac{a}{n} = \frac{b}{m}$

T

1.) a , b coincide, that is: $r_m = 0$ or $m\,a = n\,b$. If and only if:
a , b are commeasurable, that is: $\frac{a}{n} = \frac{b}{m}$

2.) If a and b, have common dividers, then these must divide all remainders and excesses.

3.) a.)

$$e_m = \begin{cases} b - r_m & \text{if } r_m \neq 0 \\ 0 & \text{if } r_m = 0 \end{cases}$$

b.) $r_m + r_n < b \rightarrow r_{m+n} = r_m + r_n$, $r_m + r_n = b \rightarrow r_{m+n} = 0$

c.) The difference of two remainders is a remainder or an excess. Namely for $r_m \geq r_n$:

$$r_m - r_n = \begin{cases} r_{m-n} & \text{if } m > n. \\ e_{n-m} & \text{if } n > m. \end{cases}$$

Similarly for excesses.

4.) From coinciding to commeasurability:
a.) The “beginning” section up to a coinciding is symmetrical.
So <u>all</u> remainders are equal to an excess and vice versa.
Thus by 3.) c.) the difference of two remainders is again one.
b.) The remainder of a remainder <u>in</u> an other, is also a remainder.
Thus the smallest u remainder must divide all the others.
c.) u divides $b = r_1 + e_1$ and it also divides $a = n\,b + r_1$.
Every other common divider of a and b divides u by 2.).
So u is the biggest common divider of a and b.

5.) From commeasurability to unique prime factorization:
a.) By 3.) b.), u, 2 u, 3 u, . . . , $b - u = (\frac{b}{u} - 1)\,u$ are all remainders.
$\frac{b}{u}\,a = \frac{a}{u}\,b$ so, $r_{\frac{b}{u}} = 0$ and so $r_1 , r_2 , . . . \; r_{\frac{b}{u}-1}$ contain all remainder values.
Thus, these must be u, 2 u, . . . , $b - u = (\frac{b}{u} - 1)\,u$ in some other order.
So none of $r_1 , r_2 , . . . \; r_{\frac{b}{u}-1}$ is zero and thus the first coinciding is at $r_{\frac{b}{u}} = 0$
b.) Let $a > b$ be natural numbers so that no half, third, or n-th of one is the same as m-th of the other! This means that a , b are distances with the 1 unit as their greatest common divider. Usually these are called relative primes. Then:
a, 2 a, . . is first dividable by b at b a and b, 2 b, . . is first dividable by a at ab.
c.) Let p be a natural number that has no half, third and so on! This is called a prime.
Then: p can only divide a q Q product of naturals if it divides at least one of them.
If a p prime divides a $q_1\,q_2 \, . \, . \, q_k$ product of primes then p is one of the q-s.

6.) From non coinciding to density:
a.) By 3.) c.) all remainders are different: $r_m = r_n$, $m > n \rightarrow r_{m-n} = 0$. Thus:
There are infinite different remainders so there are arbitrary close pairs among them
b.) Then by 3.) c.) again: There are arbitrary small remainders <u>or</u> excesses. Thus:
c.) By 3.) b.) the r , 2 r , . . . , k r, <u>or</u> e , 2 e , . . . , k e approach all values up to b.
So, the copied end points of b intervals will approach any point in the a interval.

R Actually, in the non coinciding case, all remainders <u>and</u> excesses are different too:

By 3.) a.) $r_m = e_n = b - r_n \rightarrow r_m + r_n = b$ so by 3.) b.) $r_{m+n} = 0$
Also by 3.) a.) <u>both</u> the remainders and excesses approach all lengths up to b.
We didn't need these because:
6.) c.) implies that the $d, d\alpha, d\alpha\alpha, \ldots$ points are dense on the d radius circle.
And so:
The T set of points obtained by non coinciding d-sequences will be dense in the whole plane.
But the main part of our original goal is still to be proven:

T There are non coinciding α angles.

P Many times it's easier to show more, or the same but with wider conditions.

Here too, we gradually prove stronger or more general versions:
The first step is similar to the famous Cantor argument for the existence of irrational or transcendental numbers. Indeed, it's enough to prove that the coinciding α values are sequencable because the full range from 0 to 360^{o} is continuum.
The possible different d-sequences, that is finite sequences made from d, α but starting with d, are at once sequencable, if we replace $d = 1$, $\alpha = 0$ and regard every sequence as a binary natural number. Then of course, the pairs of d-sequences are also sequencable and so it would be enough to show that every pair can only coincide at a sequence of angles because then a sequence of sequences is a sequence too. But again, we prove more, namely that there are only finite many coinciding angles for each pair.
We generalize the condition, namely instead of a fix d length, we allow different $d_1, d_2, \ldots$ to be used at each next shift. This seemingly generous change is actually simplifying the situation because by this we can stick to strictly alternating $d_1\alpha d_2\alpha d_3\alpha \ldots \alpha d_n$ sequences. Indeed, repeating $d\, d \ldots d$ can be combined into one d_k and repeating $\alpha\ \alpha \ldots \alpha$ can be written as $\alpha\ 0\ \alpha\ 0 \ldots 0\ \alpha$.
But we go even further and instead of $d_1, d_2, \ldots$ distances along our fixed half line we'll allow any sequence of $D_1, D_2, \ldots$ shifts in the plane. Of course again this leads to an advantage. Indeed, if at an α angle two sequences $D_1\alpha D_2\alpha \ldots \alpha D_n$ and $E_1\alpha E_2\alpha \ldots \alpha E_m$ coincide then it is the same as the difference of these two vectors being a nil vector. A vector is of course a shift of the C center point and the nil vector is the C C "shift".
But then the shifts can be combined, that is:
$D_1\alpha D_2\alpha \ldots \alpha D_n - E_1\alpha E_2\alpha \ldots \alpha E_m = (D_1 - E_1)\alpha(D_2 - E_2)\alpha \ldots \alpha(D_k - E_k)$ where k is the bigger of n or m. With this trick it's enough to prove that one sequence can have only finite many α values with which it becomes a nil vector, that is returns C to C.
Here again we'll show more or we'll be more specific because we prove that the number of such coinciding or returning α values is maximum the number of turns used.
We'll show this by induction and for the first case, that is for one used turning, our claim is trivial because $D_1\alpha D_2 = 0$ is only possible if $\alpha = 0$ and $D_2 = -D_1$.
In order to show the inductive step from $n - 1$ to n we first prove the following identity also by induction:

$$D_1\alpha D_2\alpha \ \ldots\ \alpha D_n\alpha - D_1\beta D_2\beta \ \ldots\ \beta D_n\beta =$$

$$\underbrace{D_1\ \alpha(D_1\ \beta\ D_2)\ \alpha(D_1\ \beta\ D_2\ \beta\ D_3)\ \alpha\ \ldots\ \alpha(D_1\ \beta \ldots \beta\ D_n)\ \alpha}_{} - \underbrace{\ldots\ldots}_{}\ \beta$$

First of all the $n = 1$ case of this identity is trivial and then from the $n - 1$-th case:

$$D_1\alpha \ldots \alpha D_{n-1}\, \alpha - D_1\beta \ldots \beta D_{n-1}\, \alpha = \underbrace{\ldots .}\, \alpha - \underbrace{\ldots .}\, \beta$$

Add D_n to both members of the left side, which doesn't change its value.
Turn both sides with α. This can be done separately for the two vectors, on both sides. So:

$$D_1\alpha \ldots \alpha D_n\, \alpha - D_1\beta \ldots \beta D_n\, \alpha = \underbrace{\ldots .}\, \alpha\alpha - \underbrace{\ldots .}\, \beta\alpha$$

Lets add $D_1\beta \ldots \beta D_n\, \alpha$ to, and subtract $D_1\beta \ldots \beta D_n\, \beta$ from, both sides:

$$D_1\alpha \ldots \alpha D_n\, \alpha - D_1\beta \ldots \beta D_n\, \beta =$$

$$\underbrace{\ldots .}\, \alpha\alpha + D_1\beta \ldots \beta D_n\, \alpha - \underbrace{\ldots .}\, \alpha\beta - D_1\beta \ldots \beta D_n\, \beta =$$

$$\underbrace{D_1\, \alpha (D_1\, \beta D_2)\, \alpha \ldots . \alpha (D_1\, \beta \ldots \beta D_n)}\, \alpha - \underbrace{\ldots \ldots}\, \beta$$

We only used the obvious fact that $\beta\alpha = \alpha\beta$ and summed the commonly turned vectors.
Now returning to our inductive step for the claim, suppose that for $n - 1$ used turn, we can only have maximum $n - 1$ possible α values. We have to prove that using n turns, that is:
$D_1\alpha D_2\alpha \ldots \alpha D_n\alpha D_{n+1} = 0$ can happen for maximum n values of α.
Let one such value be β and then using our identity:

$$D_1\alpha \ldots \alpha D_{n+1} = D_1\alpha \ldots \alpha D_{n+1} - D_1\beta \ldots \beta D_{n+1} =$$

$$\underbrace{D_1\, \alpha (D_1\, \beta D_2)\, \alpha \ldots \alpha (D_1\, \beta \ldots \beta D_n)\, \alpha} - \underbrace{\ldots .}\, \beta = 0$$

implies that $\underbrace{\ldots .} = 0$ which is $n - 1$ used turns, so by our inductive assumption has maximum $n - 1$ values. These with β give maximum n possible values.

12. Halving and doubling of the sphere

R The set in the plane that we produced in the previous section can be moved into its two halves. This set was obtained by first regarding the set of finite shift and turn combinations built from two fixed chosen one and then moving this set into its two halves by applying the shift or the turn. Indeed, if "applying" means to continue all sequences with the shift or the turn then this gives exactly the shift- or turn- ending finite sequences. Then, this halving of the moves gives at once a halvable set if we apply all finite sequences of moves to a starting point. Since the finite sequences of moves is a sequence, so the obtained set is also. If we try to increase our set then the natural idea is to pick a new, second starting point and apply the finite sequences again and thus obtain a new sequence. Then again and again, we can make pickings by the axiom of choice and so it seems we could end up with a set so that its moves by the finite shift, turn sequences fill up the whole plane. And thus of course, the splitting of the finite sequences into the shift- or turn- ending ones, would split the whole plane into two sets. So, the whole plane could be shifted or turned into its two halves. But this is a trivial nonsense because a shift or turn of the whole plane is again the whole plane. The mistake we made is that picking new starting points don't guarantee that them moved by the finite sequences will be new too. In Vitali's construction it was so, because a new point's rational moves had to be new too. Indeed, the old points <u>and</u> all their rational moves were taken together and so if $P + r$ is not new then $P + r - r = P$ isn't either. As we see, the heart of the matter is that $\pm r$ were both in the movings. So the reverse or inverse of every move must be one too. But in our construction of the halvable plane set we aimed almost for the opposite of this, namely that all sequences give different points for any starting point. If the inverses were included then we could always return to a starting point and also make many alternative moves.

So we might sigh and say that Vitali's use of repeated extensions and Sierpinski-Mazurkiewicz's use of the finite move sequences are two opposite directions and can not be combined. Hausdorff showed that this is not so and this also lead to the famous Banarch-Tarski doubling of the ball from finite many pieces. We have to include the inverse moves, there is no compromise about that, but we will still distinguish the finite sequence of moves, so that applying the basic moves to the whole will give important subsets.

If we include the inverse of the two basic moves in the Sierpinski-Mazurkiewicz construction then those sequences that end with one shift or one turn can become turn-ending or shift-ending if we apply the opposite to them. Thus, the splitting by shift- or turn- ending is not working perfectly and so we have to go further in controlling the inverses: We'll not only include them in the basic moves but make them all return to the I identity in a finite repetition. In other words, we have turns that are whole divisions of a full turn. The simplest such is a 180° turn, or mirroring to a diagonal. This returns in one repeat. Shifts can never return, so instead of the plane we must use a sphere with two 180° turns or mirrorings m and M.

Of course, $m\,m$ or $M\,M$ are the I identity, so we never repeat these and thus only regard the alternating m , M , $m\,M$, $M\,m$, $m\,M\,m$, $M\,m\,M$, . . . sequences. These can also be regarded as the possible beginnings of the: $(m)\ M\ m\ M\ m\ M\ \ldots$ infinite sequence, using the first m as optional. Unfortunately, we fail again because applying either m or M to these will <u>both</u> give back the full set except the single m or M. So we need a bit more complicated situation and the obvious step is to use three turns. We might try three mirrorings, but that leads to a bad symmetrical situation, so it's better to regard one m mirroring and a t turn that returns in two steps, that is: $t\,t\,t = I$.

This of course means, that t is a 120° turn. Again we regard only the sequences without containing the obvious $m\,m$ or $t\,t\,t$ identities, that is we regard all possible alternating sequences with m or t and a possible repeat of t. In short, we use the beginnings of:

$(m)\ t\ (t)\ m\ t\ (t)\ m\ t\ (t)\ m\ t\ (t)\ \ldots$ with ( ) as possible variations. The obvious three classes of finite sequences are the m , t or $t\,t$ ending ones and if we abbreviate them by $\{..m\}$, $\{..t\}$, $\{..t\,t\}$ and the application of a move as simply following the set, then:

$\{..m\}\ t = \{..t\} - \{t\}$, $\{..m\}\ t\,t = \{..t\,t\} - \{t\,t\}$

So turning {. . m} will give the turn ending sequences except the two initial ones.
Mirroring {. . m} will get rid of the last m because m m = I. So:
{. . m} m = {. . t} ∪ {. . t t} ∪ {I}.
So mirroring {. . m} gives the union of the two turned copies of {. . m} plus the identity.
Thus, aside from the missing t and t t at the turning and the extra I at the mirroring, we got a very paradoxical {. . m} set of sequences because its mirroring is the same as two of its turnings together. This at once shows that if with repeated extensions we'll obtain the full sphere, then using three moves and thus obtaining three sets gets around the contradiction that the full sphere must remain full when moved. Indeed, instead of moving the full sphere into its halves, which is impossible, we have the next best thing, namely that a "half" is moved into two "halves" of the other half. Then of course if we cut both halves into such "quarters" then these can be moved to be four halves, that is can be reassembled into two full spheres.
Amazingly, to get rid of the minute errors of t , t t , I is quite involved.
A natural idea is to add I to {. . m} because then t and t t are at once obtained by the turnings of {. . m} ∪ {I} = {. . m , I}. But this brings even more trouble because then the mirroring of {. . m , I} must contain m, so then this should be included in {. . t} or {. . t t}. But even if we allow this, still further changes have to be made step by step. In other words, only infinite many changes to the simple m- , t- , tt- ending classification can correct our original t , t t , I glitch. Luckily, there is a quite simple way to make this infinite change:
First of all, we'll use the original ending classification for all beginnings starting with t:

t (t) m t (t) m t (t) m t (t) m . . .

$\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ . . .

The first row above is the infinite sequence of which the beginnings are the possible t starting combinations. The second row shows the three classes according to the endings. So:
$\underline{m}$ = {t (t) m , t (t) m t (t) m , . . .}
$\underline{t}$ = {t , t (t) m t , t (t) m t (t) m t , . . .}
$\underline{tt}$ = {t t , t (t) m t t , t (t) m t (t) m t t , . . .}
Now we extend these classes with the help of an additional second sequence, that contains I and the m starting combinations up to the first single t usage. From this we can continue with the previous infinite sequence to get all possible combinations. These continuations are the ⁄ lines branching up to the previous sequence.

t (t) m t (t) m t (t) m t (t) m . . .

/ / / /

I m t t m t t m t t m t t . . .

$\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ $\underline{t}$ $\underline{tt}$ $\underline{m}$ . . .

Again the last row contains the $\underline{m}$, $\underline{t}$, $\underline{tt}$ classes containing the beginnings up to that point. So:

$\underline{m}$ = {I ; t m , t t m , m t t , m t m ; t m t m , t t m t m , t m t t m , t t m t t m , m t t m t t m , . .}

$\underline{t}$ = {t , m ; t m t , t t m t , m t t m , m t m t ; . . .}

$\underline{tt}$ = {t t , m t ; t m t t , t t m t t , . . .}

Now we truly have: $\underline{m}$ t = $\underline{t}$ $\underline{m}$ t t = $\underline{tt}$ $\underline{m}$ m = $\underline{t}$ ∪ $\underline{tt}$

Applying these three sets of moves to a P starting point we get three points sets, namely: $P\ \underline{m}$, $P\ \underline{t}$, $P\ \underline{t\,t}$. Then $P\ \underline{m}$ if mirrored goes into $P\ \underline{t} \cup P\ \underline{t\,t}$ but if turned once goes into $P\ \underline{t}$ while if turned twice goes into $P\ \underline{t\,t}$. Then of course, $P\ \underline{t} \cup P\ \underline{t\,t}$ is a set with half copy, namely a mirroring and one turn gives the $P\ \underline{t}$ half, while a mirroring and two turns gives the $P\ \underline{t\,t}$ half. Thus, we achieved a bounded half copy as we promised.
But actually we achieved much more: We obtained a Vitali type extension.
So we can choose newer and newer starting points to form an S set, with which the whole sphere becomes the three sets $S\ \underline{m}$, $S\ \underline{t}$, $S\ \underline{t\,t}$. And of course, as we already mentioned above, $S\ \underline{m}$ can be mirrored into the other two together, while turned into the second or double turned into the third.
There were two details we jumped over though:
Firstly, how should the m mirroring and t turn relate to each other? That is, in what angles should be their axis? The only requirement is that the finite sequences never coincide with each other when applied to a point and just as in the previous section this can be achieved because only a sequence of angles give coincidence.
The second detail is that even with such well chosen m , t there are points that obviously coincide, namely the end points of the axises because they don’t even move. So these must be excluded from the sphere and then for all finite sequences this means an infinite sequence of excluded points. So only the sphere remaining from these could be quarter copied or doubled. But we had already a way to get rid of a sequence of points at the end of section 10. So with simply more, but still finite many pieces we can double the full sphere.
Finally, a ball can be done layer by layer as spheres, so we only have to take care of the center, for example by regarding it as a separate piece.