The *Fog* Covering Cantor's *Paradise*[1]: Some Paradoxes on Infinity and Continuum

Zhang Ke

October 3, *2021* [2]

Abstract. We challenge Georg Cantor's theory about infinity. By attacking the concept of "countable/uncountable" and diagonal argument, we reveal the uncertainty, which is obscured by the lack of clarity. The problem arises from the basic understandings of infinity and continuum. We perform many thought experiments to refute current standard views. The results support the opinion that no potential infinity leads to an actual infinity, nor is there any continuum composed of indivisibles statically, nor is Cantor's theory consistent in itself.

MSC 2020: 03A05, 03B30, 03E10, 03E17, 00A30

Keywords: Cantor, paradox, infinite set, infinity, diagonal argument, countable, uncountable, continuum, dense, completeness

Opening Words: Some Flowers in the *Fog*

Is this paper accessible and interesting? Let's observe the set of natural numbers and a sequence of variations on it: $\{1, 2, 3, \dots\}$, $\{2, 3, 4, \dots\}$, $\{3, 4, 5, \dots\}$, ... Each variation set is a little bit different from its predecessor, and each is thought to be predictable whilst the general trend obvious. Well, what set is the target aimed at by this evolution? Simple as the question may be, the answer deserves careful thought. Our illustration is in **Thought Experiment 5.3**. Next, consider another perhaps simple question. If infinitely many glasses of salt water with the same concentration add together, they form an infinitely large salt lake. What is the concentration of the lake water? Why is that a question? A shock is waiting for you in **Thought Experiment 5.2**. Now, turn to one more such amazing question. Every point in $[-1, 1]$ has a position. Which one takes the mean position (or, is of the arithmetic mean value)? Certainly point 0 does. Well, think about $(-1, 1]$. Which point this time? We face such an issue in **Thought Experiment 4.2**. Hopefully, by now, we have answered the question at the beginning of this section — there is a simple, familiar and yet strange world ahead for all to rediscover and enjoy.

1. Introduction

Following Georg Cantor's key ideas in set theory[1, Chap. 41, sec. 7−8], we meet the puzzle of logic uncertainty while approaching actual infinity. Surprisingly though, the problem has long been being avoided, or at least insufficiently noticed; and nowadays it's just too easy (but no use) to blame human intuition for all this mess. We design a series of concise thought experiments first to highlight the logic disaster, then to expose the cause — a

---

[1] We borrow the words "*fog*" and "*paradise*" from Hermann Weyl and David Hilbert respectively. Weyl regards Cantor's hierarchy of transfinite cardinals as "*a fog on a fog*"[1, p. 1003]. Hilbert comments[1, p. 1003], "*No one shall expel us from the paradise which Cantor created for us.*"

[2] This is an updated version of the original article. We made some effort to improve the readability, especially of Section 2.

misconception of potential infinity, a half-baked myth of actual infinity and, quite related to the former two, a misperception of continuum.

Outline of the paper. In **Section 2**, we give a counterexample to Cantor's diagonal argument, provided all rational numbers in (0, 1) are countable as in Cantor's theory. Next, in **Section 3**, to push the chaos to a new high, we present a plausible method for putting all real numbers to a list. Then, to explore the cause of the paradoxes we turn to some basic and primitive issues. In **Section 4** and **5** we reexamine the prevailing opinion on continuum and infinity. After that, in **Section 6** we discuss at the basic logic level the origin of all the confusion.

2. Facing the *Fog*: Countability and Diagonal Argument

Cantor's idea of countability/uncountability is the starting point of his theory of transfinite numbers. Relatedly, his diagonal argument is the most important method in set theory and mathematical logic. But we do not think his reasoning is clear.

We run a thought experiment with the rational numbers in (0, 1), all of which, in Cantor's view, can be listed as an infinite sequence (with no duplicate elements; the same hereinafter). Given such a sequence $\langle q_1, q_2, q_3, \ldots\rangle$, we rewrite each element of it as a repeating decimal and denote the result by $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$. For the rationals that are thought to have two expansions, we choose the ones ending with 0s.

Thought Experiment 2.1 Out of the Picture

Let q be an arbitrarily given element of the sequence. We notice that there are special relatives of q, each of which differs from q at only one decimal digit (after the decimal point; the same hereinafter). Go through all decimal digits of q in natural order. For the nth decimal digit of q, swap positions of the nth element of current sequence and the nearest following element that differs from q at only the nth decimal digit. For each n, the swap is executed exactly one time. Before the nth swap, considering that 0.000... and the current nth element may happen to be such special relatives of q, there are at least 7 possible candidates in other positions and all of them are in the after-n part of the current sequence, so we can always find one for the two-position swap. We describe the operations in step-by-step form:

Step 1: In $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$, search from q_{02} successively until find an element differing from q at only the 1st decimal digit. Swap positions of the search result and q_{01}; and denote the rearranged sequence by $\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$.

Step 2: In $\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$, search from q_{13} until find an element differing from q at only the 2nd decimal digit. Swap positions of the search result and q_{12}; and denote the rearranged sequence by $\langle q_{21}, q_{22}, q_{23}, \ldots\rangle$, which is the same as $\langle \boldsymbol{q_{11}}, q_{22}, q_{23}, \ldots\rangle$.

Step 3: ..., search from q_{24} ... only the 3rd .... Swap ... and q_{23}; and ... by $\langle q_{31}, q_{32}, q_{33}, \ldots\rangle$, which is the same as $\langle \boldsymbol{q_{11}}, \boldsymbol{q_{22}}, q_{33}, \ldots\rangle$.

Go on throughout the rest of the fractional part of q.

Aside: *We adopt essentially the same trick as that exploited by Cantor — taking full advantage of infinity, just present the robust body of an argument and leave the untouchable tail to fog. Our goal is to beat Cantor at his own game to clarify that his logically unclear game is only a paradox maker.*

The stage result sequences are distinct from one another, for the results of nth and $(n+k)$th steps (n, $k = 1, 2, 3, \ldots$) differ at least at the $(n+1)$st position.

Write all the stage result sequences as *list A*:

$$\begin{array}{l} \langle\ \boldsymbol{q_{11}},\ q_{12},\ q_{13},\ q_{14},\ \cdots\ \rangle\ , \\ \langle\ q_{21},\ \boldsymbol{q_{22}},\ q_{23},\ q_{24},\ \cdots\ \rangle\ , \\ \langle\ q_{31},\ q_{32},\ \boldsymbol{q_{33}},\ q_{34},\ \cdots\ \rangle\ , \\ \langle\ q_{41},\ q_{42},\ q_{43},\ \boldsymbol{q_{44}},\ \cdots\ \rangle\ , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{array}$$

Aside: *The sequences $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$, $\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$, $\langle q_{21}, q_{22}, q_{23}, \ldots\rangle, \ldots$ are based on the same set and each two of them differ at finitely many positions. Although they are mutually rearranged sequences, each of them, if exists, could stand on its own (that is, as for existence, none of them is prior to another).*

The above mentioned step-by-step operations tell us *list A* is the same as

$$\begin{array}{l} \langle\ \boldsymbol{q_{11}},\ q_{12},\ q_{13},\ q_{14},\ \cdots\ \rangle\ , \\ \langle\ \boldsymbol{q_{11}},\ \boldsymbol{q_{22}},\ q_{23},\ q_{24},\ \cdots\ \rangle\ , \\ \langle\ \boldsymbol{q_{11}},\ \boldsymbol{q_{22}},\ \boldsymbol{q_{33}},\ q_{34},\ \cdots\ \rangle\ , \\ \langle\ \boldsymbol{q_{11}},\ \boldsymbol{q_{22}},\ \boldsymbol{q_{33}},\ \boldsymbol{q_{44}},\ \cdots\ \rangle\ , \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots . \end{array}$$

Each item of *list A* is a rearranged sequence of $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$. The orderly aligned elements of all sequences form a matrix, which is divided in half by diagonal $(q_{11}, q_{22}, q_{33}, \ldots)$. Does the diagonal cover all the columns? In Cantor's opinion, it does, for all the elements on the diagonal can be indexed by natural numbers and so can all the columns. On the other hand, another piece of equally (un)clear reasoning shows his is a one-sided view. We denote the part below (and to the left of) the diagonal, including the diagonal, by L. It is plain that L only contains a minority of q's special relatives that each differs from q at only one decimal digit; and infinitely many other elements of $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$ stay outside. Each row contains all these "other elements", thus the part L, which contains the whole of the so-called diagonal, misses many columns.

Note: *For those who notice the faint shadow cast by the concept of order type, we add a **discussion** near to the end of this paper.*

(We also include an optional variation of the above procedure in **Appendix A**.)

Diagonal argument works on the premise that the diagonal covers both horizontally and vertically. Now, make a comparison between our case and that of Cantor. The number of the rows of *list A* is exactly that of all the decimal digits of q, and the number of the columns is exactly that of all the rationals in $(0, 1)$. Obviously, the columns of the proposed list in Cantor's diagonal argument[1, p. 997, Cantor's second proof] match the rows of our *list A*, whilst the rows of the supposed list there contain but not limited to all the contents of our columns. How can we believe that his diagonal covers well?

While dealing with finitely many items, Cantor's diagonal argument is clear. As for infinitely many items, his argument is also hoped to be rigorous, but that is not the case before us. And once logical uncertainty has crept in, everyone may "prove" whatever at will.

3. Under Cover of the *Fog*: Listing all Real Numbers

According to Cantor's theory, all real numbers or even the reals in a bounded interval cannot be listed as a sequence. We are going to list them to reveal more confusion.

We just create a sequence for all the reals in $[0, C)$, $C \in \mathbb{R}^+$, and leave the rest job to some well-known solutions. As there is a mapping between real numbers and points on straight line, numbers and points are interchangeable. And it is also the same for intervals and line segments (including open or half-open segments; the same hereinafter).

Preparation: Bend $[0, C)$ into a circle (henceforth, we straighten or bend it on demand without declaring; and to ensure clarity, we do not use any chord). Denote point 0 by P_0. Suppose that there is a light beam emitted from P_0, aimed at and reflected by another point, which is denoted by P_1, on the circumference. If the line segments $[0, P_1)$ and $[0, C)$ are incommensurable, and P_0 is the only point that has no reflective feature, then the light beam produces a series of reflection points, $P_1, P_2, P_3, \ldots$, and does not make a repeat or return to P_0 within finitely many reflection steps. We refer to P_0 and all the reflection points as the *bright points*.

Lemma 3.1 *Bright points* are dense on the circumference.

(That is to say, for any two points on the circumference, however close they may be to each other, there is a *bright point* in between. It is essentially the simplest case of Kronecker's Theorem. For the sake of self-containedness of this paper, we include a proof in **Appendix B**.)

Note: *We use the character of ideal light beam for offering a logic clue to thread all the relevant points. However, from another view angle, whether or not a point is a* bright point *is determined by itself (its specific position) once C, P_0 and P_1 given. A point $x \in [0, C)$ is a* bright point *if it makes an indefinite equation ($x + mC = nP_1$, where m and n are unknown non-negative integers) solvable. (And for a positive n, point x turns out to be the nth reflection point.)*

A natural question is: Are all the points on the circumference *bright*?

> **Thought Experiment 3.1** A Sequence of All the Real Numbers in $[0, C)$
>
> If the answer to the above question is yes, the sequence $P_0, P_1, P_2, P_3, \ldots$ contains all the real numbers in $[0, C)$. If the answer is no, we may classify all the points into two sets — set B for all the *bright points* and set D for the rest, which we call the *dark points*. Naturally, we have a sequence of all *bright points*; and we aim to add all *dark points* to it.
>
> For visualization, attach $[C, 2C]$ to $[0, C)$ as a "handle". Hold the "handle", cut $[0, 2C]$ at all *dark points* simultaneously while keeping each of the *dark points* as the right-hand endpoint of its own fragment. Then throw away the "handle" together with the perhaps existing remainder of $[0, C)$.
>
> ***Note:*** *The scheme of disintegration can be easily described in formal logic. As mentioned above, set B is $\{P_0, P_1, P_2, P_3, \ldots\}$, and set D is $\{x : 0 \leq x < C\} \setminus B$. We denote $\{x : 0 \leq x \leq 2C\}$ by A, and $D \bigcup \{2C\}$ by D'. The elements of D' determine a partition of A. For $d_j \in D'$, the equivalence class $[d_j]$ is $\{x \in [0, d_j] : \forall d_i \in D(d_i < d_j \Rightarrow d_i < x)\}$. Then ignore the equivalence class [2C], just consider the others ....*
>
> Now, we focus on the fragments, each of which contains exact one *dark point*. According to **Lemma 3.1** , any two *dark points* are separated from each other by some *bright points*. Considering also that the leftmost point of $[0, C)$ is a

bright one, (thus, intuitively, within the whole segment each *dark point* may look to the right or left like a human being, and the scene coming into its sight is the same as if it were located in the middle of an otherwise wholly *bright* segment,) none of the fragments can be free of *bright point* (since, otherwise, if the unique point of an exceptional fragment looks to the left while located in the whole segment, the scene coming into sight would be the same as if it looked to the left while located in the middle of a wholly *dark* segment. Why such an awkward interpretation? We come back to this issue later in **Thought Experiment 4.4**). In this sense, the number of the fragments, which equals the number of *dark points*, is not greater than that of *bright points*. Now that all the *bright points* are in $\{P_0, P_1, P_2, P_3, \ldots\}$, according to the *well-ordering principle* (which states that *every nonempty subset of the natural numbers has a least member*), in each fragment there is a *bright point* with the least index number. Pair up each *dark point* with such a *bright* one of the same fragment. Then execute a series of insertions based on $\langle P_0, P_1, P_2, P_3, \ldots\rangle$ — for each *bright point*, if it is an unpaired one just skip it, else insert right after it its *dark* partner. The target sequence appears.

Up to this point everything is seemingly normal. But the assumption of the existence of the *dark point* still leads to confusion. If there is a *dark point*, then we can find an endless chain of *dark points* by reasoning backwards repeatedly — tracing the *dark point* back iteratively along the circumference with the same step-length as the reflection of the light beam (but in the reverse direction) to other *dark points*. Consequently the number of the *dark points* could not be smaller than that of the *bright* ones, and in this case the structure of each fragment is unimaginable. (Moreover, an even annoying question is: How many isolated *dark* chains are there? We choose to walk around this muddy place.) Needless to say, the *dark points*, if exist, are also dense on the circumference. Therefore, in each fragment the *dark point* would be the only point, or another *dark point* would thereupon exist in the same fragment. This means no *bright point* could exist, but it is absurd. Therefore, the assumption of the existence of *dark point* is false.

A much more direct way to show the confliction led to by *dark points* is adopting another cutting rule — cutting $[0, C)$ at each *bright* point while keeping the *bright point* as the left-hand endpoint of its fragment. Then similar reasoning as above indicates that there is no room for any *dark point*.

We are familiar with the saying that a number system, for example the rational numbers, can be dense on real line without completely filling the line. That means "dense" does not need to be "without any gap". But is that clear in logic?

Thought Experiment 3.2 The Puzzling "Dense"

The property of "dense" implies that between any two members, however close they may be, there is always a third one. The rational number system has this property. Let's observe an infinite sequence of nested rational pairs: $\langle -1, 1\rangle$, $\langle -1/2, 1/2\rangle$, $\langle -1/4, 1/4\rangle$, $\langle -1/8, 1/8\rangle, \ldots$. If a third number is embraced by an inner pair, it is also embraced by every outer one. Considering the property of "dense", we naturally expect a common rational number that is between the two partners of each pair; and sure enough, 0 is such a number. However, a sticky fact is that 0 is the only one. Does that mean if it is taken away, "dense" would no longer hold? If the answer is yes, then the existence of 0 as a rational is a must for the denseness property. But what if the thoroughly wrapped number is not a rational (actually, there are such examples for some other sequences of rational pairs)? And if the answer is no, then to ensure that “dense” still holds, should there be another rational that is wrapped as deeply as 0 is? Neither of the answers is doubtless.

We are bothered by the question: Do the rational numbers (or, every “dense” collection of numbers with no bound) fill up the whole real line? For example, we may design a strictly increasing rational sequence $p_1, p_2, p_3, \ldots$ and a strictly decreasing rational sequence $q_1, q_2, q_3, \ldots$, and they have the same limit, say $\sqrt{2}$. (See a numerical example in the **Appendix C**). Now we have an infinite sequence of nested pairs: $\langle p_1, q_1\rangle, \langle p_2, q_2\rangle, \langle p_3, q_3\rangle, \ldots$; apparently $\sqrt{2}$ is the only number to penetrate into all the pairs. If we take away $\sqrt{2}$ (for it is an irrational number, which is unrelated to the denseness property of the rational system), then, does the property of “dense” still hold for the rational system? If the answer is yes, should we take on the hopeless task of finding a rational that is located in the same depth as $\sqrt{2}$ is? If the answer is no, should we accept $\sqrt{2}$ as a rational number?

4. Is a Continuum Composed of Indivisibles Statically?

In comparing “dense” with “without any gap”, we get curious about the relationship between line segment (as continuum) and point (as indivisible). Now we turn to thinking about the prevailing opinion — a continuum is a collection of (stationary) indivisibles. Is that clear in logic? Why have there always been some people standing on the opposite side since ancient times?

In mathematics a point is thought to have no extension in any direction. As questioned by many, how can zero-magnitude points accumulate to a positive-magnitude segment? An intuitive example adding to the doubt is: The segment [0, 1] can cover [0, 1), but the latter cannot cover up the former. On the ground of this difference, the size of one point does matter and cannot be zero. On the other hand, if someone tries to assign a nonzero value to the size, which value could work? Why not, say, half the value? Well, to suspend the debate over the detail value, let us just say one point is a point-measured entity.

We devise a group of thought experiments to illustrate that a fixed aggregation of points can never reach the status of a continuum, in other words, the static-indivisible-composed model of a continuum, which we simply refer to as the *indivisible-model* (of a continuum) hereafter, is untenable. What we aim to argue against is not the existence of continuum or indivisible, but the reasonability of the *indivisible-model*.

First of all, what is a continuum? It is said to be a continuous entity; but if only “continuous” were not equally in need of clarifying. Here we just mention some understanding about continuum: A straight line is a simple example (with real line being its *indivisible-model*). Informally, it is so perfect with regard to uniformity as to have no structure or detail (so to speak) — while looking at it, no matter which part or what a scale to focus on, one would never find more than the first glance tells. The perfect uniformity requires the *indivisible-model* of a straight line to be either free of gap or full of gap — strictly unified throughout. If there is no gap in the *indivisible-model*, as expected, each possible magnitude value would have its point representation on a real line, and each reasonably deduced point would be found in the *model*. It follows that the *model* has reflectional and translational symmetries.

The notation [0, 1) refers to the set of all real numbers between 0 and 1, including 0 but excluding 1. However, what we want to illustrate below is it is impossible for such a stationary aggregation to be "without any missing number". So, in our discussion, "without gap" is not a presupposed or known property. From now on, we use a modified notation, $\underset{\cdots}{[0,\ 1)}$ — the dotted underline stands for the uncertainty about its completeness property. And, for convenience, we usually omit the underline if without causing any confusion.

Thought Experiment 4.1 The *Enclosed Point*

Between any two distinct real points there is always another one, which we call an *enclosed point* between the two given points. If $0 < X_1 < X_2 < \cdots < X_n$, and Y is an *enclosed point* between 0 and X_1, then Y is also an *enclosed point* between 0 and X_i for $i = 2, 3, \ldots, n$. Accordingly, we call Y an *enclosed point* between 0 and $\{X_1, X_2, \ldots, X_n\}$. Further onwards, 1 is an *enclosed point* between 0 and (2, 3].

Is there an *enclosed point* between 0 and $[x, 1]$ for each $x\ (0 < x < 1)$? Everyone would answer yes without hesitation, since $(0 + x)/2$ is an example. Then, is there an *enclosed point* between 0 and $(x, 1]$ for each $x\ (0 \leq x < 1)$? Is the answer just the same as the former? Or, can "$(x, 1]$ *with* $0 \leq x < 1$" (as a whole of various possibilities) stretch even closer to 0? As far as can be told from the concept of "open/closed intervals", it has advantage in approaching to 0 — "$[x, 1]$ *with* $0 < x < 1$" and "$(x, 1]$ *with* $0 \leq x < 1$" may be regarded as two sets of *indivisible-model* of segments that are each right-ended by point 1, and (0, 1] is the only element that can cover any other element of the two sets. (But is that clear? We continue the discussion in **Thought Experiment 4.4**.) However, considering that $\bigcup\{[x, 1] : 0 < x < 1\}$ and $\bigcup\{(x, 1] : 0 \leq x < 1\}$ are the same set of real numbers (points) — both of them are $\{x : 0 < x \leq 1\}$, the aforementioned view needs careful thinking. The identity of the two point-level sets suggests that the two original sets cover exactly the same range, and the two answers for the relevant questions, on this account, should be the same.

Actually, what we want to ask in the second question is, whether there is an *enclosed point* Y between 0 and $\underset{\cdots}{(0,\ 1]}$? Now, take $\underset{\cdots}{(0,\ 1]}$ as an immutable aggregation of points and face the question directly. If the answer is yes, $\underset{\cdots}{(0,\ 1]}$ is not as complete as expected, for missing at least Y. If the answer is no, there is no gap between the point 0 and the segment of (0, 1] (in other words, they touch each other). In this case, if a segment is composed exclusively of the static points, the consequent question is: Which point of (0, 1] are so close to the point 0 that they bear no gap in between (or, which point of the half-open segment has direct contact with the point 0)?

If a segment could be explained as a set of points with each point having a position value, then there is an arithmetic mean for all the values. And the position of the centroid is of this very value.

Thought Experiment 4.2 The Absent Centroid

Observe the segment (–1, 1), obviously 0 is the arithmetic mean. But what if the object is $(–1, 1)\bigcup\{1\}$? Does the segment (–1, 1] have a centroid? Of course, just like every object has its center of mass (and the remaining question is whether the center is within the body. For some objects, say, a doughnut, it is not). And certainly, the "center of mass" for each bounded connected part of a straight line should be within the "body" and identical with the centroid. Now, where is the centroid for (–1, 1]? (For those who think that one more point is not enough to make any difference, we have some questions: Whether a one-point-figure has a centroid? What is the centroid of $\{0\}$? And then what is the centroid of $\{0\}\bigcup\{1\}$?) Intuitively the newly joined point 1 would "drag" it to the right a tiny little bit. Another simple fact is that, for every bounded geometric figure, any part (but not whole) that is centrally symmetric about the centroid may be omitted without affecting the position of the centroid. So we can easily examine any supposed centroid of (–1, 1]. Unfortunately, no known point can stand the test. Is it an *enclosed point* between 0 and (0, 1] discussed in **Thought Experiment 4.1**?

How about switching back to the concept of the arithmetic mean? Well, because the segments (–1, 1) is just a left part of (–1, 1], the two mean positions cannot coincide. For those interested in calculating the precise mean position for all the points of the latter, the results tend to favor $0+$ (*half a point-measured length*). Sure enough, no known point takes this position.

If, as we are taught, there are no adjacent points on a real line, it is impossible for any bounded entity to make an exact point-measured displacement along it. But on second thought, if a point is an individual entity, and taking away one point from a real line would result in a point-measured gap, we can certainly expect the existence of a point-measured distance and a motion of such a distance.

Thought Experiment 4.3 Moving a Point-measured Distance

Observe [0, 2) and (0, 2]. From the view angle of linear motion, they may be taken as each other's displacement result — a motion of a point-measured distance comes to light. By the way, the two mean positions, referring to **Thought Experiment 4.2**, would be 1– (*half a point-measured length*) and 1+ (*half a point-measured length*), which differ by one point in position.

The points of a real line somewhat resemble the sheets of paper of a book, which can be divided into two groups by a bookmark without disrupting the order. (Next, if something recalls a classical thought to mind, just forget the famous idea for a moment.) Suppose that a virtual bookmark partitions a real line into $(-\infty, b)$ and $[b, +\infty)$, and we notate the "bookmark" together with its position by ϒ**b**. And a fellow notation, **b**ϒ, is for the "bookmark" that is "one sheet of paper" after ϒ**b**. If there is no gap between, say, $(-\infty, b)$ and $[b, +\infty)$, the "thickness" of a "bookmark" is 0 and it is possible to make a virtual insertion immediately before or after each real point. That being the case, the virtual bookmark can move a point-measured distance — the distance between ϒ**b** and **b**ϒ is exactly a point-measured long. Moreover, if there is a point immediately before the "bookmark", according to the property of symmetry, there should be another point located after the "bookmark" symmetrically — a pair of adjacent points comes to light. In fact, the trouble might spring out even earlier — before the above discussion we should ask in advance whether an extra point can be put at the very position of the "bookmark". This reminds us of **Thought Experiment 4.1**.

> ***Aside****: All the above examples seem to describe an abrupt "movement". Indeed, our goal is to refute the* indivisible-model *of continuum by showing the existence of some embarrassing tiny distance. For those who care much more about the continuity of a movement, we provide an example: When point x moves continuously from position* 2 *to* 4*, the arithmetic mean position for all the points of* $(0, 2)\bigcup\{x\}$ *moves continuously from 1+ (half a point-measured length) to 1+ (half a point-measured length)+ (one point-measured length).*

Remember "why such an awkward interpretation" in **Thought Experiment 3.1**? And the question arises in **Thought Experiment 4.1**: How can "$[x, 1]$ *with* $0 < x < 1$" and "$(x, 1]$ *with* $0 \leq x < 1$" (as two sets) cover exactly the same range, while the super element $(0, 1]$ existing in the latter but not in the former? For a definitive answer, we come to another question: Is an open interval really open?

In standard opinion, [0, 3] could be "cut off" into [0, 2) and [2, 3]. But why [0, 2) has no right endpoint (whereas [2, 3] has its left endpoint)? If it has one, that one would be the left adjacent of point 2. We are taught that no adjacent points exist on a real line. So, there is nothing strange. However, such an explanation can never eliminate our doubt. From the viewpoint of uniformity of straight line, each point of real line has nothing special except for its unique position. That means if any of them can act as an endpoint, so can others. As concerns the edge of a bounded connected part, adding or removing a point would cause nothing more than a shift of position to the related boundary — there would be no change in its form or style.

Thought Experiment 4.4 The Suspicious Open Interval

Suppose that, along a real line orbit, an antimatter object $[-1+t, 1+t]$ is sliding towards a normal object [2, 4], and its position varies directly with time, t. A special collision would be inevitable and the two would cancel each other out during the process, for at any moment they match each other symmetrically.

But what would happen if there are two antimatter points $-1+t$ and $1+t$, instead of the object $[-1+t, 1+t]$, moving towards [2, 4]? A stage result would be that one of the moving points has disappeared together with the left endpoint of [2, 4], and the other point is sliding towards (2, 4] from left side. What would happen next? Which point of (2, 4] would vanish?

Now turn to another case. Suppose that a point x moves from 1 to 3 continuously along a real line. If [1, 3] has no gap, all points of the interval would be exactly all the possible positions of x. Therefore, "*taking the position of each of the points in orderly accordance with time*" is "*moving continuously through the interval*". Then observe a growing segment $[0, x]$ with x running through all the points of [1, 3]. The growing segment would apparently run through all the connected parts (of [0, 3]) that contains [0, 1], for the growth is a continuous one. Thus there exists a point X_i ($1 \leq X_i \leq 3$) such that $[0, X_i]$ is exactly [0, 2), which is surely a connected part of [0, 3] and contains [0, 1]. That means [0, 2) is right-ended by X_i. Denying the existence of X_i means a gap in the *indivisible-model* of [1, 3]. But the existence of X_i violates the standard opinion about adjacent points.

There is a motion-independent variant of the experiment: Now that real line is assumed to be without gap, it is naturally the most accurate ruler with each of its point working as a scale mark. The "most accurate" means, while making a measurement, there are always plenty of marks such that the concepts of tolerance and estimation can be dismissed — every finite length value can be read off directly from some point on the "ruler" and even a difference of one point would be reflected in the scale readings without additional description. If not, the accuracy can be improved simply by adding (to the original model as new point, new scale mark) every known exceptional length value (that cannot be directly read off yet) and hence the original model cannot be "without gap". From another aspect, every bounded connected part of a straight line is a definite length value in itself, no matter there is a ruler (that can directly measure it) or not.

5. The "Final Result" of a Potentially Infinite Process

The *indivisible-model* of continuum cannot stand further thought. Now we are faced with the sequential question (a classical one): What is the final result of iterative bisection of a line segment?

To discuss this question, we begin with clarifying what does "final result" mean. While describing the consequence of a potentially infinite process we use "after all steps", "final result", and the like. What is their origin in intuition? Here are two examples:

Example 1: For a potentially infinite collecting task, the initial status is an empty set; and the stage results are $\{1\}$, $\{1, 2\}$, $\{1, 2, 3\}$, and so on; and the "final result" is thought to be $\{1, 2, 3, \ldots\}$.

Example 2: For the geometric series related to the first one of Zeno's paradoxes[2, p. 349], the initial status is an empty record; and the stage results are $\frac{1}{2}$, $\frac{1}{2}+\frac{1}{4}$, $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}$, and so on; and the "final result" is thought to be $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots$.

For a potentially infinite process, we denote the first n steps by $S_1, S_2, \ldots, S_n$, the initial status by C_0, the nth stage result by C_n and the "final status" C. Now we execute a collecting task — collect the notations of executed steps successively. If the main process can reach completion, our collecting task can thereupon reach completion and we can get an actual infinite set $\{S_1, S_2, S_3, \ldots\}$, that is to say C_0 can become C after experiencing $\langle S_1, S_2, S_3, \ldots\rangle$. From the viewpoint of transformation function, the first n step(s) as a group determines a function T_n such that $C_n = T_n(C_0)$, whilst all the steps as a group determine a function T such that $C = T(C_0)$.

Nevertheless, we believe that all the so-called final results of the potentially infinite processes are just imaginations of human beings, or more precisely, some preexisting or supposed existing objects are assigned to act as the nonexistent "final results". Hence, no wonder, sometimes people are puzzled by different versions of "final result" of one potentially infinite process. Usually it is difficult to take sides between (or among) the different versions, as shown in the case of the Ross–Littlewood paradox[3](which has offered a good chance to rethink infinity).

Thought Experiment 5.1 The "Final Results"?

There is an endless row of road lamps indexed by natural numbers. We begin with all the lamps lit, and denote the status by $\langle \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots\rangle$, a sequence of period 1. Here, "of period m" means for all n the status of the $(n+m)$th lamp is the same as that of the nth lamp (m, $n = 1, 2, 3, \ldots$).

Step 1: For all the lit ones, from left to right, switch off the 1st, skip the 2nd, switch off the 3rd, skip the 4th, and so on. (Those have the form $2k+1$ for non-negative integer k are switched off, and all the multiples of 2 remain.) Then denote the result by ⟨~~1~~, **2**, ~~3~~, **4**, ~~5~~, **6**, . . .⟩, a sequence of period 2.

Step 2: Execute the same operation as above. (Those have the form $4k+2$ are switched off this time, and all the multiples of 4 remain.) Denote the result by ⟨~~1~~, ~~2~~, ~~3~~, **4**, ~~5~~, ~~6~~, ~~7~~, **8**, ~~9~~, ~~10~~, ~~11~~, **12**, . . .⟩, a sequence of period 4.

Step 3: . . . . (Those have the form $8k+4$ are switched off this time, and all the multiples of 8 remain.) . . . , a sequence of period 8.

Go on iteratively ad infinitum.

We denote the “final result” by ⟨~~1~~, ~~2~~, ~~3~~, ~~4~~, ~~5~~, ~~6~~, ~~7~~, ~~8~~, ~~9~~, ~~10~~, ~~11~~, ~~12~~, . . .⟩, which shows none of the lamps is on. But the strange thing is that the period length turns out to be 1, whereas, to tell from its monotone increasing trend in the process, it cannot be a finite number.

However, if we simulate the above process following another rule — “switch off the 2nd and every other one thereafter”, the stage results would be ⟨**1**, ~~2~~, **3**, ~~4~~, **5**, ~~6~~, . . .⟩, ⟨**1**, ~~2~~, ~~3~~, ~~4~~, **5**, ~~6~~, ~~7~~, ~~8~~, **9**, ~~10~~, ~~11~~, ~~12~~, . . .⟩, . . . with the same period lengths as above, respectively. (Some more detail is: only those in the form $2k+1$ remain to the end of the first step; only those in the form $4k+1$ remain to the end of the second step; only those in the form $8k+1$ remain to the end of the third step; and so on.) We would have ⟨**1**, ~~2~~, ~~3~~, ~~4~~, ~~5~~, ~~6~~, ~~7~~, ~~8~~, ~~9~~, ~~10~~, ~~11~~, ~~12~~, . . .⟩ as the “final result”, which, in accord with the process, has no finite period length.

Now introduce a concept, *relative lighting rate*, which is a ratio like occupancy rate. The *relative lighting rate* of ⟨**1**, **2**, **3**, . . .⟩ is 1, and that of ⟨~~1~~, ~~2~~, ~~3~~, . . .⟩ is 0; and it can be easily figured out for a sequence that follows a simple pattern. So, we may describe the evolution of the *relative lighting rate* in the former process as 1, $1-\frac{1}{2}$, $1-\frac{1}{2}-\frac{1}{4}$, etc., and the “final result” as $(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\cdots)$. For the latter process, we may describe the course as 1, $1-\frac{1}{2}$, $1-\frac{1}{2}-\frac{1}{4}$, etc., and the “final result” as $(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\cdots)$. In this sense, the two processes coincide with each other all the way. But, strangely enough, the two “final states” are different by one lamp. That might be the reason why there are always some people rejecting the opinion that 0.999. . . is exactly 1 — there is a lamp remaining lit somewhere at the back of their minds.

Once again return to the initial status. This time we think of the lamps as an endless row of dots arranged at regular intervals on a rubber band that has been pre-stretched sufficiently. Take the position of the 1st lamp as reference point. Next we are concerned only with the lit lamps.

Step I: For all the lit lamps, switch off the 2nd and every other one thereafter. Then let the rubber band contract around the reference point until the current 2nd lit lamp occupies the initial spatial position of the original 2nd lit one. (As a consequence, the current 3rd lit lamp, in turn, occupies the initial position of the original 3rd lit one. And so forth.)

Repeat ad infinitum.

What is the “final result” this time? To judge from the form of numbers, only 1 would remain, since for a natural number n, to survive the first k steps means

n–1 is divisible by 2^k and vice versa. Nevertheless, it is easy to see, from the duplicate images presented by the steps, that there is no trend towards such a "final result" and there is no hope to escape from the initial status. This gives rise to a growing feeling that all the "final results" do not really exist.

All the potentially infinite processes, when looked at their steps (with no detail in any step being taken into account), share the same endless sequential structure, which we call the *N-structure* because the natural numbers as endless ordinals are abstracted from various examples of it — each example has a unique starting element, which has one and only one successor, which, in the same way, has its own successor, and so forth (there is nothing else — no branch, no loop, and each following element is a finite "distance" away from the starter). A conclusion at this step-sequence level (thus has nothing to do with the details inside any step), no matter which specific process it is drawn from, relates only to the macrostructure and holds equally for all examples of the *N-structure*. Therefore the falsity of any "final result" means none of the examples of the *N-structure* (the processes of potentially infinite steps) can ever reach a completion.

Why there are so many people accept the so-called final results as a given? This is due to that from the beginning their minds are fed with some specially selected cases, each of which suggests an apparent "final result", which is too natural and perfect to be questioned. But, are they really so perfect?

Thought Experiment 5.2 Fresh or Salty?

Take each of the natural numbers as a glass of salt water. We denote the initial status by $\langle \underline{\underline{\mathbf{1}}}, \underline{\underline{\mathbf{2}}}, \underline{\underline{\mathbf{3}}}, \ldots\rangle$. All of them are of the same concentration, and, if added together, can make an infinitely large salt lake of the same salinity.

Step 1: Replace the 1st glass of salt water with a glass of fresh water, which is free of salt. Denote the result by $\langle 1, \underline{\underline{\mathbf{2}}}, \underline{\underline{\mathbf{3}}}, \ldots\rangle$.

Step 2: Replace the next glass of salt water with a glass of fresh water. Denote the result by $\langle 1, 2, \underline{\underline{\mathbf{3}}}, \ldots\rangle$.

Go on ad infinitum.

After all steps, so to speak, we denote the "final result" by $\langle 1, 2, 3, \ldots\rangle$. Obviously all the glasses of water, with none of them containing any salt, can be added together to form an infinitely large lake of fresh water. As to the perfect completion, there seems to be no doubt. Yet, the story does not end here.

On the other hand, we can deduce the concentration of the "final" lake step by step. We denote the original concentration by C_0. Suppose that the result lake of Step 1 has the concentration C_1 (after the uniform state has been reached). Taking into account the effect of the first replacement, we have $C1 > C_0 - C_0/4 = (1-1/4)C_0$. The result lake of Step 2 has the concentration C_2. Taking into account also the second replacement, we have $C_2 > C_1 - C_0/8 > (1-1/4-1/8)C_0$. Accordingly, we have $C_3 > C_2 - C_0/16 > (1-1/4-1/8-1/16)C_0$, and so on. Consequently, the "final" lake has the concentration C, which satisfies $C > [1-(1/4+1/8+1/16+\cdots)]C_0$. Conceivably, by choosing different coefficients on the right side of the inequalities, we can even conclude that C is higher than any given concentration that is below C_0.

We also provide a variant of the experiment from a static viewpoint. There is an infinitely large salt lake that consists of infinitely many disjoint units of salt water indexed by natural numbers. All the concentrations are the same, which we denote by C_0. But, maybe an earnest person would bother to think the relationship between the concentration of the lake and that of each unit of salt water. Evidently, the salinity of the lake is backed up by the salt of all the units; and the contribution of the salt of the nth unit is less than $C_0/2^{n+1}$. As a consequence, the concentration of the lake, which is given as C_0, may be less than $(1/4 + 1/8 + 1/16 + \cdots)C_0$. (And besides, from some angle, one can even conclude that the concentration is higher than C_0.) A contradiction.

A truth may be observed from every angle — there is no contradiction between any two views. An illusion of human beings is another thing. All the confusion we have exposed above reveals that all the "final results" are the work of man.

Thought Experiment 5.3 Poor or Rich?

There are infinitely many piggy banks (indexed by natural numbers) with the 1st one containing 1 coin, the 2nd containing 2 coins, the 3rd containing 3 coins, and so on. We denote the initial status by $\langle 1,\ 2,\ 3, \ldots\rangle$.

Step 1: Take away the piggy bank that contains 1 coin (the original 1st piggy bank). Denote the result by $\langle 2,\ 3,\ 4, \ldots\rangle$.

Step 2: Take away the piggy bank that contains 2 coins (the original 2nd piggy bank). Denote the result by $\langle 3,\ 4,\ 5, \ldots\rangle$.

Go on ad infinitum.

After all steps, so to speak, there would be nothing left over.

Now, return to the initial status and begin a contrastive process.

Step I: Add one coin to each of the piggy banks. Denote the result by $\langle 2,\ 3,\ 4, \ldots\rangle$.

Step II: Add one coin to each of the piggy banks again. Denote the result by $\langle 3,\ 4,\ 5, \ldots\rangle$.

Go on ad infinitum.

After all steps, so to speak, we denote the "final result" by $\langle a_1,\ a_2,\ a_3, \ldots\rangle$. According to Cantor's theory, each of the piggy banks contains $\aleph_0$ coins. However, here we are only interested in whether the result is null or not.

Make a comparison of the stage results between the two processes. In set theory, they are the same respectively. Thus an unavoidable question is: How can the same path (with the same starting point) lead up to two opposite destinations?

No wonder there are always some people accepting only potential infinity and reluctant to go any further — lack of clarity in both logic and intuitive sense proves the critical hurdle in the way to actual infinity.

The "final result" and "total finish" are two sides of the same coin. On this account a contradiction about the "final result" denies the existence of either.

6. A Look Back at the *Fog*

After Cantor's time, in the field of the foundations of mathematics, most mathematicians readily yield to a kind of asymmetrical reasoning — their logic standard differs depending on whether a normally deduced result is for or against the dominating opinion. For example, "*x is different from each element of an infinite set*" is thought to be sufficient to prove "*x is not an element of the set*", whilst "*point 1 is nonzero away from each point of [0, 1)*" is not regarded as adequate to conclude that "*point 1 separates from [0, 1) (there is a gap in between)*", though both assertions are based on the same principle — *a set is determined by its members* (hence a proposition about a set can be verified by testing all its members literally).

However, a mathematical theory can never rest mainly on asymmetrical reasoning for long. Underlying the prosperity and wonderfulness of the theory of transfinite numbers there must be some more or less plausible pillars.

Cantor asserts[4], "*Each potential infinite, if it is rigorously applicable mathematically, presupposes an actual infinite.*" As Morris Kline explains[2, p. 200], "*He argued that the potentially infinite in fact depends upon a logically prior actually infinite.*"

Cantor's words show his acceptance of actual infinity and infinite set. We call this belief the *Belief A*. The actual infinities at this point are still lacking vitality, they cannot promise many exciting stories in such status. Nevertheless, what makes them vivid and substantial, as in many controversial proofs with regard to infinity, is a related but distinct belief: *Each potential infinity, if rigorously applicable, leads up to an actual infinity.* This belief is a bridge across the impossible gulf between finite and infinite. We call it the *Belief B*. With it in mind, everyone may draw a conclusion based on "finishing" a potentially infinite process, and figure out an actual infinity's properties from an associated potential infinity. Like many others, we think of the *Belief B* as a super troublemaker. As Hermann Weyl puts it[1, p. 1200], "*. . . the sequence of numbers which grows beyond any stage already reached . . . is a manifold of possibilities opening to infinity; it remains forever in the status of creation, but is not a closed realm of things existing in themselves. That we blindly converted one into the other is the true source of our difficulties, ...*"

If a potential infinity leads to an actual infinity, then a potentially infinite process could be taken as a function mapping initial status inputs to final status outputs. But many pieces of evidence suggest otherwise. Strictly following the *Belief B* results in the confusion of **Thought Experiment 5.2**, and combining the *Belief B* with the *axiom of extensionality* causes the contradiction in **Thought Experiment 5.3**. (In axiomatic set theory, the *axiom of extensionality* conveys the idea that, as in naïve set theory, *a set is determined by its members rather than by any particular way of describing the set.*) Although the axiom is clear and safe for finite sets, our experiment indicates that it is not reliable while combined with the *Belief B*.

Well, would everything be all right if it were not for the affection of the *Belief B*? No. Current situation is really worrisome. The *Belief A* has been artfully used for presenting a stirring magic show launched by Cantor. And in order to sustain the epic performance, the standard of plausibility and clarity has been distorted; some ingenious arguments only in partial agreement with logic have occupied center stage in the foundations of mathematics ever since. Yet it must be said in fairness to the mathematicians that no one intentionally resorts to trickery or deception.

One-to-one correspondence, the basic principle of set theory, is simple and clear for finite sets. But, when applied to infinite set, it brings about a big shock to humans' intuitive mind. For Cantor "*a set is infinite if it can be put into one-to-one correspondence with part of itself*" [1, p. 995]. However, such a supposed basic property of infinite sets violates *the fifth common notion of Euclid* (which states *the whole is greater than the part*), thereby kicking off the chaos and leading to paradoxes such as Hilbert's Hotel[5].

Let's rethink, from another perspective, the problem mentioned in **Thought Experiment 5.2**. Suppose that there are two infinitely large salt lakes, with one of them having the concentration $C\,(C > 0)$ and the other $C/2$. If the salt and the water are separated, each lake would become an infinite amount of salt and an infinite amount of fresh water, and the two lakes would yield, according to Cantor's theory, just the same amounts of water and the same amounts of salt respectively. That means, reversely, if given an infinite amount of salt and an infinite amount of water, one can obtain a salt lake of concentration C or $C/2$ or some other value. Thus, *the law of contradiction* (a fundamental principle of logic, which states *something cannot be true and not true at the same time*) is violated.

Note: *The unusual concentration is not so abrupt because the paradox of Hilbert's Hotel has already shown a similar ratio, the occupancy rate, which is also without certainty. Given the infinite constant of persons and rooms in Hilbert's Hotel, the occupancy rate can be adjusted as one pleases (with each room containing at most one person).*

How about *the law of excluded middle* (another fundamental principle of logic, which states *any proposition is true or its negation is true*)? We arbitrarily use the law in **Thought Experiment 3.1**, but does it hold for infinite sets? No. To explain our view, we provide a thought experiment adapted from the relevant ideas of **Thought Experiment 3.1**.

Thought Experiment 6.1 Containing or not?

Consider all rational numbers between (and inclusive of) 0 and 1 in the form of reduced fraction (in particular, 0 is taken as 0/1, and 1 as 1/1). We classify these rationals into two sets — set B for those have an even number as its numerator or denominator, and set D for all others. Let set $A = B \bigcup D$. We notice that between any two elements of A there are both an element of B and an element of D.

Note: *The fact may be easily seen. Just observe 2/(2n+1), 4/(2n+1), . . . , 2n/(2n+1), when written in reduced form, each of them is an element of B. They spread on* [0, 1] *with a fixed interval that is smaller than 1/n. Similarly, observe 1/(2n+1), 3/(2n+1), . . . , (2n-1)/(2n+1), . . . of D. They . . . .*

The elements of D determine a partition of A in the way that, for $d_j \in D$, the equivalence class $[d_j]$ is $\{\textit{rational number } x \in [0,\, d_j] : \forall d_i \in D\,(d_i < d_j \Rightarrow d_i < x)\}$. Of course, each equivalence class contains exactly one element of D. Our question is: Does $[d_j]$ contain any element of B? There is a similar question in **Thought Experiment 3.1** (where we include a more intuitive operation — cutting the segment at infinitely many points). The answer cannot be yes, otherwise $[d_j]$ contains another element of D. On the other hand, the answer cannot be no, otherwise, for the same sake, each of the other equivalence classes does not contain any element of B, but that implies $D = A$.

Hilbert has a famous comment on Cantor's work[1, p. 1003]: "*No one shall expel us from the paradise which Cantor created for us.*" However, without basic logic, who can tell a mirage from reality?

7. Some More Words

Once the series of thought experiments reach the public, our mission is fulfilled; we may sit back and see whether the *paradise* will be gone with all the *fog*. Before diverse voices refocus on the "long-settled" controversy respecting Cantor's work, we would like to say some things in addition.

How to understand infinity? A closely associated question is, How to understand continuum? These questions have been puzzling human beings for thousands of years, and there have been plentiful doctrines formed on unclear arguments since at least Aristotle's time. Our goal is to present some simple and convincing evidence against the current overwhelming opinion, which we have been trying but failing to understand.

Note*: Aristotle writes[6, Book III, Chap. 6], "There will not be an actual infinite." And in his view "... nothing that is continuous can be composed of 'indivisibles': e.g. a line cannot be composed of points, the line being continuous and the point indivisible."[6, Book VI, Chap. 1]*

Our demonstration supports such a view:

(i) No potential infinity leads to an actual infinity. The existence of actual infinities depends upon axiom, and Cantor's theory, though influential, is not reliable. (ii) Connectedly, indivisibles can never statically form a continuum. The existence of continuum depends upon axiom, and its static-indivisible-composed model, though useful, is not a logical one.

Aside*: What does* "exist" *or* "existence" *mean in our discussion? In mathematics, informally, anything that can find its room in a consistent system is said to exist in that system. A thing may exist in one system, while not in another. Although two conflicting things cannot coexist in one system (otherwise the system is inconsistent), they may exist separately in two systems, say, Euclidean geometry and non-Euclidean geometry. And many existing things (e.g., irrational numbers) were once unknown or thought to be impossible. Some others (e.g., infinitesimal) have their ups and downs, though whether something exists or not is objective.*

Gauss states[1, p. 993], "*I protest against the use of an infinite quantity as an actual entity; this is never allowed in mathematics. The infinite is only a manner of speaking, in which one properly speaks of limits ....*" His words indicate he believes that the concept of actual infinity cannot be derived from others and there is no reasonable way of introducing it. However, Cantor does not think so. He points out[4], "*In order for there to be a variable quantity in some mathematical study, the domain of its variability must strictly speaking be known beforehand .... Thus this domain is a definite, actually infinite set of values.*"

In Cantor's judgment, it is a must to accept actual infinity. Normally, next step is to come up with several fresh concepts around this idea and vitalize them by mining some new connections and relationships. The key is to keep a full consistency and clarity, as the lack of which may shake everything. Cantor has done a lot but, disappointingly, left behind a *fog*-wrapped "*paradise*". Worse than that, what threatens his theory is not a leaky roof, but a shaky foundation.

In reality, Cantor's thought has ridiculously won a miraculous victory, but still there are some who hold the opposing perspective. Is there a middle ground between the views of both sides? Perhaps there is. Abraham Robinson's position is based on the two main points or principles[7, p. 507]: "*(i) Infinite totalities do not exist in any sense .... (ii) Nevertheless, we should continue the business of Mathematics 'as usual,' i.e., we should act as if infinite totalities really existed.*" His flexible attitude is helpful for accommodating and accomplishing both of the two sides' goals before actual infinity is recognized just as it is. However, the point is that a tolerance limit should be set for "*as usual*" or "*as if*", otherwise mathematics will always be surrounded by such things as "*a fog on a fog*" ("*Hermann Weyl spoke of Cantor's hierarchy of alephs as a fog on a fog*"[1, p. 1003]), as it is today. (The outline of the limit is not our concern for now.)

At this point, the fate of infinity attracts most attention. Just like Cantor's "*paradise*" is not the final destination, its collapse does not mark the point of terminal decline for infinity. As far as existence is concerned, actual infinity should not be more abrupt than imaginary number (which contributes a lot to the development of mathematics). It is not a matter of proving or disproving; the difficulty for infinity lies in taking on more positive significance than sustaining and entertaining itself. It seems that, in the fading *fog*, actual infinity may walk out of a mythical dream and sip on some fresh air.

Usually, there is more than one belief surrounding the same fact. And those with various opinions are always ready to fight for the truth and continue what is thought to be the right way. Anyway, pushing and hiding the contradictions into the depth of *fog*, instead of eliminating them at the root, is always an option, but never a solution.

***Discussion**: What do we mean by "*. . . the depth of fog*"? We just give a simple example. Let's slip back into Cantor's theory once more. In **Thought Experiment 2.1**, if taking into account the concept of order type, the paradox may be discussed further. Some may argue that infinitely may times of position swaps would change the order type. Others do not agree. They think, in the sequence, the relationship between a position and its element is just like the relation between a fixed seat and the person sitting in it — the seat would remain still whilst the person may not. If there is neither an empty seat nor a standing person, the order type of the persons is just the same as that of the seats, which are stationary all the way. Nevertheless, some may think from another aspect. For the sequence of all natural numbers:* $1, 2, 3, \ldots$, *repeatedly swap positions of the element 1 and its immediate right neighbor. For those who insist on there being a change of the order type, the "final result" would be* $2, 3, 4, \ldots, 1$ *with the order type* $\omega+1$ *in Cantor's theory, hence the paradox of **Thought Experiment 2.1** can be explained. It seems to be reasonable. But what if go a step further along the same lines? Next, repeatedly swap positions of the element 2 and its direct right neighbor. The question is: Whether* $3, 4, 5, \ldots, 2, 1$ *would be a stage result? Confusion is knocking at the door. Some may go on to ask: So what if the element 2 turns out to be in trouble? Does that relate to the realization of the aforementioned* $2, 3, 4, \ldots, 1$*? Well, just compare beginning with* $1, 2, 3, \ldots$ *and with* $1, 2, 3, \ldots, 0$.

Like it or not, the concept of infinity does raise a paradise, which is for imagination. However, even a reasonable imagination, which may be scientifically helpful in virtue of its reasonable aspect(s), still has unreasonable aspect(s), (or it is a fact, a truth). The key is never to indulge in the *fog* while making clever use of imagination. It is not the imagination that hurts, but the mistaking of it for truth.

***Aside**: As the public knows, in physics, a rigid object is commonly treated as if all its mass were concentrated in a single point called the center of mass, though, in fact, the mass is everywhere in the object. This imagination simplifies the solution of problems of force and motion. But it does not always work, for it is not equal to the truth, otherwise there would be no need for the concept of center of gravity. The same is true for the imaginations about infinities and continua in mathematics.*

Acknowledgments

The author is grateful to all those who offered helpful suggestions. Special thanks go to Li Rui, for listening to some rough and raw ideas.

Appendix

A. An Optional Variation of Part of Experiment 2.1

Let q be an arbitrarily given element of $\langle q_{01},\, q_{02},\, q_{03}, \ldots \rangle$. Go through all decimal digits of q in natural order, and execute an operation according to each of the digits:

Step 1: In $\langle q_{01},\, q_{02},\, q_{03}, \ldots \rangle$, search from q_{02} successively until find an element differing from q at at least the 1st decimal digit. Swap positions of the search result and q_{01}; and denote the rearranged sequence by $\langle q_{11},\, q_{12},\, q_{13}, \ldots \rangle$.

Step 2: In $\langle q_{11},\, q_{12},\, q_{13}, \ldots \rangle$, search from q_{13} until find an element differing from q at at least the 2nd decimal digit. Swap positions of the search result and q_{12}; and denote the rearranged sequence by $\langle q_{21},\, q_{22},\, q_{23}, \ldots \rangle$, which is the same as $\langle \boldsymbol{q_{11}},\, q_{22},\, q_{23}, \ldots \rangle$.

Step 3: $\ldots$, search from q_{24} $\ldots$ at at least the 3rd $\ldots$. Swap $\ldots$ and q_{23}; and $\ldots$ $\langle q_{31},\, q_{32},\, q_{33}, \ldots \rangle$, which is the same as $\langle \boldsymbol{q_{11}},\, \boldsymbol{q_{22}},\, q_{33}, \ldots \rangle$.

Carry on throughout the rest of the fractional part of q.

Note: *For each n, the swap is executed exactly one time. We can always find an element for each two-position swap: For base-10 system, every rational number has at least one repeating decimal representation. The same is true for binary system. Moreover, the expression capacity of a base-n numeral system does not depend upon which n distinct symbols to employ, so we can use 0 and 1 to establish a binary system, or 5 and 6 (as two symbols, having nothing to do with their original numerical values) to do the same. In each of the binary systems, every rational number in* $(0,\, 1)$ *has at least one repeating expansion, the fractional part of which shares the same visual appearance with that of some base-10 rational number in* $(0,\, 1)$*. Then in each of the binary systems, there are infinitely many expansions look like base-10 rationals in* $(0,\, 1)$*. Therefore, each search task for the two-position swap cannot fail — at least (after any position of the current sequence) there still are infinitely many choices that look like some type of binary rational numbers and are free of any single digit to be avoided.*

The stage result sequences are different from each other, for the results of nth and $(n{+}k)$th steps ($n,\, k = 1,\, 2,\, 3, \ldots$) differ at least at the $(n{+}1)$st position.

Write all the stage result sequences as *list B*:

$$
\begin{array}{l}
\langle\ \boldsymbol{q_{11}},\ \ q_{12},\ \ q_{13},\ \ q_{14},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ q_{23},\ \ q_{24},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ \boldsymbol{q_{33}},\ \ q_{34},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ \boldsymbol{q_{33}},\ \ \boldsymbol{q_{44}},\ \ \cdots\ \rangle\ ,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots .
\end{array}
$$

Each item of *list B* is a rearranged sequence of $\langle q_{01},\, q_{02},\, q_{03}, \ldots \rangle$. The orderly aligned elements of all sequences form a matrix, which is divided in half by diagonal $(q_{11},\, q_{22},\, q_{33}, \ldots)$. Does the diagonal cover all the columns? We denote the part below (and to the left of) the diagonal, including the diagonal, by L, and notice that $\{$elements that appear in $L\} = \{q_{11},\, q_{22},\, q_{33}, \ldots\}$. Each q is outside L for differing from q_{nn} at the nth decimal digit. Thus q, while appearing in each row, is always outside L. That means part L, which contains the so-called diagonal, does not take over all columns.

B. Proof for Lemma 3.1

Proof. (Here, the term arc refers to inferior arc only.)

1. The light beam does not return to any *bright point* again within finitely many reflections because the line segments $[0, P_1)$ and $[0, C)$ are incommensurable.

2. It suffices to show that for any given arc of the circle, no matter how short it is or where it locates, there is a *bright point* on it.

We denote the length of the given arc by $\varepsilon\,(0 < \varepsilon \leq C/2)$. Let k equal the greatest integer not exceeding C/ε, so $(k + 1) > C/\varepsilon$. The first $k{+}1$ *bright points* divide the circumference into $k{+}1$ disjoint arcs. As $(k+1)\varepsilon > C$, at least one of the $k{+}1$ arc lengths is less than ε; for such a short one, we denote its arc length by $S\,(0 < S < \varepsilon)$; and for the two *bright points* that determine this short arc, we denote the difference between their indexes by $i\,(i \in \mathbb{N},\ 0 < i \leq k)$.

Denote the biggest integer not exceeding C/S by m. As the arc distance between any two *bright points* is determined only by the difference between their indexes, the $m{+}1$ points P_0, P_i, $P_{2i}, \ldots, P_{mi}$ (ignoring all others) divide the circumference into $m{+}1$ disjoint arcs with m of them having the length $S\,(0 < S < \varepsilon)$ and the remaining one being even shorter. That means it is impossible for any ε-lengthed arc on the circumference to avoid all the $m{+}1$ points. ☐

C. A Numerical Example for Thought Experiment 3.2

A strictly increasing rational sequence that has limit $\sqrt{2}$: $7/5, 41/29, 239/169, \ldots$ with general term $\sqrt{2b_n^2 - 1}/b_n$, where $b_1 = 5$, $b_{n+1} = 3b_n + 2\sqrt{2b_n^2 - 1}$ for $n = 1, 2, 3, \ldots$.
And a strictly decreasing rational sequence that has limit $\sqrt{2}$: $3/2, 17/12, 99/70, \ldots$ with general term $\sqrt{2c_n^2 + 1}/c_n$, where $c_1 = 2$, $c_{n+1} = 3c_n + 2\sqrt{2c_n^2 + 1}$ for $n = 1, 2, 3, \ldots$.

References

[1] M. Kline, *Mathematical Thought from Ancient to Modern Times*, Vol. 3, Oxford Univ. Press, 1990. (First published in 1972, in one volume.)

[2] M. Kline, *Mathematics: The Loss of Certainty*, Oxford Univ. Press, New York, 1980.

[3] *Ross–Littlewood paradox.* https://handwiki.org/wiki/Ross–Littlewood_paradox

[4] *Georg Cantor.* https://en.wikiquote.org/wiki/Georg_Cantor

[5] *Hilbert's paradox of the Grand Hotel.* https://wikimili.com/en/Hilbert's_paradox_of_the_Grand_Hotel

[6] Aristotle, *Physics*, Written 350 B.C.E., Translated by R.P. Hardie and R.K. Gaye. http://classics.mit.edu/Aristotle/physics.html

[7] Abraham Robinson, *Selected papers*, Vol. 2, Edited by W.A.J. Luxemburg, S. Körner, North Holland, Amsterdam, 1979.

康托尔乐园的迷雾[1]：关于无穷大和连续体的一些悖论

张科

译自2021年10月3日的英文稿 [2]

摘要. 我们挑战G·康托尔关于无穷大的理论。通过考问“可数/不可数”的概念以及对角线证法，我们揭示了其中隐藏在不清晰性背后的不确定性。问题起源于对无穷大和连续体的根本认识。我们用许多思维实验来反驳当今的正统观念。所有结果支持这样的观点：潜无穷并不导致实无穷，连续体也不能由不可分元素静态地构成，康托尔的理论也并非自身无矛盾的。

MSC 分类号：03A05，03B30，03E10，03E17，00A30

关键词： 康托尔，悖论，无限集，无穷，对角线证法，可数，不可数，连续，稠密，完备

欢迎词：雾里看花

这篇文章有趣易懂吗？先来看一个自然数集合及由此逐步演变生成的一系列集合：$\{1, 2, 3, \ldots\}$、$\{2, 3, 4, \ldots\}$、$\{3, 4, 5, \ldots\}$、$\ldots$。每个衍生集合都较之前有稍许变化，每个又都是可预测的，并且总趋势是明显的。然而，怎样的一个集合是这个演进过程的目标？问题貌似简单，但值得三思。我们的演示在**思维实验5.3**中。接下来，考虑另一个或许简单的问题。如果无穷多杯同样浓度的盐水汇集起来，将共同形成一个无穷大的盐湖。盐湖里的水有怎样的盐度？这还用问？然而**思维实验5.2**会让你吃惊。再转而思考另一个同样具有魔性的问题。$[-1, 1]$ 中的每一个点都有一个位置值。哪个点占据着中位（或者说，其位置值取到算术平均值）？当然是点 0。那么，考虑一下 $(-1, 1]$。这回是哪个点？我们在**思维实验4.2**中面对这样一个问题。至此，相信我们已经回答了开篇提出的疑问 — 在我们面前有一个简单、熟悉而又奇妙的世界，等待着所有人去重新探索和欣赏。

1 引言

顺着康托尔在集合论中的关键思路 [1, 第41章 第7–8节]，当面对实无穷时我们遇到了逻辑不确定性带来的困境。令人惊奇的是，长期以来这个问题被人们回避，或至少没被足够重视。把所有的混乱都说成是人类的直觉出了问题，这在现如今是再容易不过的了（但这并不能解决问题）。我们设计了一系列简明的思维实验，首先把逻辑危机突显出来，然后揭示其成因 — 关于潜无穷的错误观念，围绕着实无穷的不周全的神话，再者是与前两点密切相关的，对连续体的错误认知。

[1] “雾”和“乐园”分别是借用赫曼·魏尔和大卫·希尔伯特的表述。魏尔认为康托尔关于超限数的层级理论是“雾中之雾”[1, 第1003页]。希尔伯特对康托尔的工作赞赏道 [1, 第1003页]，“任何人都不能把我们从康托尔为我们创建的乐园中赶走。”

[2] 这个版本在原始文章的基础上做了整理更新，以期提高（尤其是第 2 节的）可读性。

本文概述。在**第 2 节**，先假定如康托尔的理论所说 (0, 1) 上的全部有理数是可数的，我们对康托尔的对角线证法进行反驳。接下来，在**第3节**，为了将混乱推到一个新的高度，我们给出一个列写全部实数的貌似合理的方法。然后，为了探索这些悖论的成因，我们回归到一些原始基本的主题。在**第4**和**5节**我们重新审视了目前关于连续体和无穷大的主导观点。在此基础上，在**第6节**我们在基本逻辑层面上讨论了所有困惑的根源。

2 初遇迷雾：可数性和对角线证法

康托尔关于可数/不可数的想法是他的超限数理论的起点。相应地，他的对角线证法成为集合论和数理逻辑中最重要的方法。但我们认为他的论证是不清晰的。

我们来做一个关于 (0, 1) 上有理数的思维实验，按照康托尔的观点这部分有理数可以列写成一个无穷数列（不含重复项，以下同）。假设给定了这样一个数列 $\langle q_1, q_2, q_3, \ldots\rangle$，我们把其中各项写成无限循环小数的形式，把结果记为 $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$。对于那些被认为有两种循环小数表示的，选择尾部全是 0 的。

思维实验 2.1 方外之地

令 q 等于数列中任意给定的一项。我们注意到有一类 q 的近亲，它们中的每个与 q 仅在某一个小数位上存在差别。按自然顺序遍历 q 的小数位。对于 q 的第 n 位小数, 找到现时数列中第 n 项之后最靠前的且与 q 仅在第 n 个小数位存在差异的项，交换此项与现时数列中第 n 项的位置。对于每个 n 做且只做一次这样的交换。在第 n 次交换之前，考虑到 0.000... 以及现时的第 n 项可能恰好也是 q 的这种近亲，这样至少还有 7 个可能的候选者存在于其它位置上，而且都在现时数列的第 n 项之后，因此对于每一次的两项交换，我们总能找到一个后项。现将操作按分步形式解说如下：

步骤 1：在 $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$ 中从 q_{02} 开始顺次向后搜索，直到找到某项仅在第 1 个小数位与 q 不同。交换搜索结果与 q_{01} 在数列中的位置；将重排后的数列记为 $\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$。

步骤 2：在 $\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$ 中从 q_{13} 开始向后搜索，直到找到某项其仅在第 2 个小数位与 q 不同。交换搜索结果与 q_{12} 的位置；将重排后的数列记为 $\langle q_{21}, q_{22}, q_{23}, \ldots\rangle$，它等同于 $\langle \boldsymbol{q_{11}}, q_{22}, q_{23}, \ldots\rangle$。

步骤 3： ... 从 q_{24} ... 仅在第 3 个小数位 ... 。交换 ... 与 q_{23} ... ；将 ... 记为 $\langle q_{31}, q_{32}, q_{33}, \ldots\rangle$，它等同于 $\langle \boldsymbol{q_{11}}, \boldsymbol{q_{22}}, q_{33}, \ldots\rangle$。

继续，遍历 q 的小数部分。

旁白：我们的手法在本质上与康托尔的相同 — 充分利用无穷来障眼，仅示人以论证的强健身躯而让那条不可捉摸的尾巴留在云里雾里。我们试图在康托尔自己的游戏中挫败他，从而说明那个逻辑不清的游戏无非是个悖论的温床。

过程中产生的数列彼此不同，因为步骤 n 和步骤 $(n+k)$ $(n,\,k=1,\,2,\,3,\ldots)$ 的结果至少在第 $(n+1)$ 项上是不同的。将所有各步的结果依次列写成表A：

$$\begin{array}{l}\langle\ \boldsymbol{q_{11}},\quad q_{12},\quad q_{13},\quad q_{14},\quad \cdots\ \rangle\ ,\\ \langle\ q_{21},\quad \boldsymbol{q_{22}},\quad q_{23},\quad q_{24},\quad \cdots\ \rangle\ ,\\ \langle\ q_{31},\quad q_{32},\quad \boldsymbol{q_{33}},\quad q_{34},\quad \cdots\ \rangle\ ,\\ \langle\ q_{41},\quad q_{42},\quad q_{43},\quad \boldsymbol{q_{44}},\quad \cdots\ \rangle\ ,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\ 。\end{array}$$

旁白：数列 $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$、$\langle q_{11}, q_{12}, q_{13}, \ldots\rangle$、$\langle q_{21}, q_{22}, q_{23}, \ldots\rangle$、... 都是基于同一个集合的，并且两两之间只有有限个项不同。尽管它们互为重排数列，就存在性而言是彼此独立的（也就是说，如果它们存在的话，并无先后之分）。

从上面的分步骤操作易知表A等同于

$$\begin{array}{l}\langle\ \boldsymbol{q_{11}},\quad q_{12},\quad q_{13},\quad q_{14},\quad \cdots\ \rangle\ ,\\ \langle\ \boldsymbol{q_{11}},\quad \boldsymbol{q_{22}},\quad q_{23},\quad q_{24},\quad \cdots\ \rangle\ ,\\ \langle\ \boldsymbol{q_{11}},\quad \boldsymbol{q_{22}},\quad \boldsymbol{q_{33}},\quad q_{34},\quad \cdots\ \rangle\ ,\\ \langle\ \boldsymbol{q_{11}},\quad \boldsymbol{q_{22}},\quad \boldsymbol{q_{33}},\quad \boldsymbol{q_{44}},\quad \cdots\ \rangle\ ,\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\ 。\end{array}$$

表A的各项都是数列 $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$ 的重新排列。所有数列的项整齐排列成一个矩阵，矩阵被对角线 $(q_{11}, q_{22}, q_{33}, \ldots)$ 一分为二。这个对角线覆盖了所有列吗？在康托尔看来是肯定的，因为这条对角线上的所有项可以用自然数来编号而且所有的列也可以这样编号。另一方面，有一段清晰度与之相当的推理显示他的观点是片面的。我们把对角线及其左下的部分记为 L。显然 L 中只容纳了 q 的近亲中的一小部分，它们中的每个与 q 仅在某一个小数位上存在差别，而 $\langle q_{01}, q_{02}, q_{03}, \ldots\rangle$ 中无穷多的其它项则不在其中。因为每一行都完整地保有全部“其它项”，可见包含了整条所谓对角线的 L 区域无缘许多列。

旁注：有人隐约看到了序型的身影，我们为此在本文临近结尾处加了段讨论。

（我们给出了上述过程的一个可选的翻版，放在**附录A**中。）

应用对角线证法的前提条件是对角线在纵横两个方向上都能覆盖到。现将我们的情况和康托尔的做一个对比。表A的行数正好是 q 的小数位数，而其列数正好是 (0, 1) 上的有理数的总量。显然，康托尔的证明 [1, 第 997 页，康托尔的第二个证明] 里那个假设的实数列表的列对应着我们表A的行，而它的行包含并不限于我们的列的全部内容。这让我们怎么相信他的对角线能完美覆盖？

在应对有限多项时，康托尔的对角线证法是清晰的。对于无限多项，他的证法被寄予了同样的希望，但我们眼前的一幕令人沮丧。而且逻辑不确定性一旦混入，每个人都可以随便“证明”任何想要的结论。

3 雾朦胧：列写全部实数

按照康托尔的理论，所有实数或者哪怕是在有限区间上的实数不可能一个不漏地排成一个数列。为了揭示更多的困惑，我们将给出这样一个数列。

我们将以 $[0, C), C \in \mathbb{R}^+$ 上的全部实数来创建一个数列，余下的工作用众所周知的方法足以应对。由于实数与直线上的点存在着对应关系，我们可以在两者间自由切换。而且对于区间与线段也同样（所称的线段包括除去端点或只含一个端点的，以下同）。

准备工作：将 $[0, C)$ 弯曲成一个圆（以后我们根据需要把它拉直成线段或弯曲成圆，不再专门交待。这里的讨论不涉及弦，所以不至于紊乱）。将点 0 记作 P_0。假设有一束光从 P_0 射出，指向圆周上记作 P_1 的另一点，并经此反射。如果线段 $[0, P_1)$ 与 $[0, C)$ 不可公度，且 P_0 是圆周上唯一不具反射特性的点，那么光束将产生一系列反射点，$P_1, P_2, P_3, \ldots$，并且在有限步反射之内不会重复经过某点或回到 P_0。我们将 P_0 和所有反射点称为亮点。

引理 3.1 亮点在圆周上稠密。

（也就是说，圆周上的任意两点无论多么邻近，总有一个亮点介于它们之间。这实质上是克罗内克定理涵盖的最简单情况。为了本文的自助性，我们在**附录B**给出一个证明。）

旁注：我们借助理想光束的特性将所有相关的点用一个逻辑线索串起来。然而，一旦给定了 C, P_0 和 P_1，某个点是否为亮点取决于其自身（它的特定位置）。点 $x \in [0, C)$ 如果满足一个不定方程（$x + mC = nP_1$，这里 m、n 是未知的非负整数）有解，则它是亮点。（并且对于一个正的 n，点 x 是第 n 个反射点。）

人们自然要问：圆周上的所有点都是亮的吗？

思维实验 3.1 一个由全体 $[0, C)$ 上的实数排成的数列

如果上述问题的答案是肯定的，那么数列 $P_0, P_1, P_2, P_3, \ldots$ 包含了全体 $[0, C)$ 上的实数。如果答案是否定的，我们可将所有点分成两个集合 — 所有亮点归为集合 B，其余的我们称之为暗点，归为集合 D。我们有一个现成的所有亮点排成的数列，接着打算把所有的暗点都插入其中。

为了直观，将 $[C, 2C]$ 作为“手柄”接续到 $[0, C)$ 上。握住“手柄”，将 $[0, 2C]$ 在所有暗点处同时切断并保证每个暗点是其所在碎片的右端点。抛弃“手柄”以及其上可能连带的 $[0, C)$ 的残余部分。

旁注：这个分解方案容易用形式逻辑来描述。如上所述，集合 B 是 $\{P_0, P_1, P_2, P_3, \ldots\}$，而集合 D 是 $\{x : 0 \leq x < C\} \setminus B$。我们将 $\{x : 0 \leq x \leq 2C\}$ 记为 A，将 $D \bigcup \{2C\}$ 记为 D'。D' 的元素决定了 A 的一个划分。对于 $d_j \in D'$，等价类 $[d_j]$ 即 $\{x \in [0, d_j] : \forall d_i \in D(d_i < d_j \Rightarrow d_i < x)\}$。接下来忽略等价类 *[2C]*，仅考虑其它 ...。

现在我们来关注这些碎片，它们每一个包含恰好一个暗点。根据**引理3.1**，任

意两个暗点都被某个亮点隔开。又考虑到 $[0, C)$ 的左端点是亮的，（于是直观上，在整条线段上每个暗点可以像人一样看向左边或右边，映入眼帘的就好像它处在一条原本全亮的线段中间所见到的一样，）没有哪个碎片完全不含亮点（因为, 不然的话，当这样一个特殊碎片中的唯一点位于整条线段中看向左边时，映入眼帘的就好像它处在一条全暗的线段中间看向左边时见到的一样。为什么会有这么费事的表述？我们将在**思维实验4.4**中回到这个话题）。从这个意义上说，碎片总数即暗点的总数不可能超过亮点的总数。既然所有亮点都在 $\{P_0, P_1, P_2, P_3, \ldots\}$ 中，按照良序原理（即每个非空的自然数集的子集都有一个最小元素），在每个碎片中都有一个亮点具有最小的索引号。将每一个暗点与本碎片中的这样一个亮点配对。接下来基于 $\langle P_0, P_1, P_2, P_3, \ldots\rangle$ 做一系列的插入 — 对于每个亮点，如果它是无配对的就跳过，否则将与它配对的暗点插入到其后。目标数列现身了。

至此似乎一切正常。但那个关于存在暗点的假设仍然会导致困惑。一旦存在一个暗点，我们就能通过反复倒推而找到无限多的一串暗点 — 沿着圆周从这个暗点开始迭代地以光束反射的步距（但以相反的方向）回溯就能找到其它的暗点。这样一来暗点在数量上不会少于亮点，而在这种情况下每个碎片的构成是不可思议的。（不仅如此，一个更令人头疼的问题是：有多少条彼此独立的暗点链路？我们选择绕开这滩浑水。）显然, 如果暗点存在，它们在圆周上也同样是稠密的。于是在每个碎片上暗点将是唯一的点，否则另一个暗点将存在于同一个碎片上。这意味着没有亮点存在，但这是荒谬的。因此，关于存在暗点的假设是不成立的。

揭示暗点导致的冲突还有一个更直接的方法，就是采用另一种切断规则 — 将 $[0, C)$ 在每一个亮点处切断，同时保证每个亮点是其所在碎片的左端点。然后用同上的论证说明暗点无处容身。

我们熟知的一种说法是一个数系，例如有理数系，可以是即在实数轴上稠密又不占满整个实数轴的。这意味着“稠密”不必是“无空隙的”。但这在逻辑上是清晰的吗？

思维实验 3.2 令人费解的“稠密”

“稠密”性蕴含着在任意两个成员之间（无论它们多么邻近）总有另一个成员存在。有理数具有这样的性质。我们来观察一个由有理数对构成的无穷序列：$\langle -1, 1\rangle, \langle -1/2, 1/2\rangle, \langle -1/4, 1/4\rangle, \langle -1/8, 1/8\rangle, \ldots$。如果有某个数被其中两个成对的有理数包围，那么它也被更外层的各组成对的有理数分别包围。考虑到有理数的“稠密”性，我们自然会料想有一个数被各层有理数对共同包围。果然，0 就是这样的一个数。然而, 一个无奈的事实是，对于上述序列这样的数只有 0 这么一个。这是否意味着离了这个 0，有理数就失去“稠密”性了？如果回答是肯定的, 那么 0 这样一个有理数的存在对于稠密性来说是必须的。但如果这个被层层包围的数不是有理数（对于其它的有理数对的无穷序列确实存在这样的情况）呢？如果前面的回答是否定的，那么为了“稠密”性仍然能维持，

是否还应该有另一个有理数像 0 一样被彻底包围呢？两种答案都不能令人信服。

困扰我们的问题是：有理数系（或者，每一组"稠密"且无界的数）占满了整个实数轴吗？例如，可以构想出一个严格单调递增的有理数无穷数列 $p_1, p_2, p_3, \ldots$ 和一个严格单调递减的有理数无穷数列 $q_1, q_2, q_3, \ldots$，而且它们收敛于同一个极限值，如 $\sqrt{2}$。（一个数值实例见**附录C**）。现在我们有了一系列无穷多组嵌套的有理数对：$\langle p_1, q_1\rangle$、$\langle p_2, q_2\rangle$、$\langle p_3, q_3\rangle$、$\ldots$。显然 $\sqrt{2}$ 是唯一的能介于各组成对的两数之间的数。如果我们移除 $\sqrt{2}$（因为它是个无理数，与有理数系的稠密性无关），那么有理数系的"稠密"性还在吗？如果回答是肯定的，我们是否要去无望地寻找一个和 $\sqrt{2}$ 一样被被深度包围的有理数呢？如果回答是否定的，我们是否应该把 $\sqrt{2}$ 归为有理数呢？

4 连续体是由不可分元素静态地构成的吗？

在对比"稠密"和"没有空隙"的过程中，线段（作为连续体）和点（作为不可分量）之间的关系引起了我们的好奇。现在我们转而思考这方面的主导观点 — 一个连续体就是一组不可分量的（静态）集合。这在逻辑上是清晰的吗？为什么自古以来总有一些人持与之对立的观点？

在数学上，点被认为在任何方向上都是没有延伸的。我们和许多人一样质疑：大小为 0 的点怎么能堆垒出长度为正的线段？一个直观的例子更增加了这种疑惑：线段 $[0, 1]$ 能够覆盖 $[0, 1)$，但后者并不能完全覆盖前者。基于这种差别，一个点的大小是有影响的，因而不可能是 0。另一方面，如果有人试图给这个大小赋值，什么值合适呢？为什么不是这个值的一半呢？不如先搁置争议，我们姑且说一个点就是具有一个点的大小的实体。

我们设计了一组思维实验来说明一群静态的点的集合达不到能构成连续体的程度，也就是说，连续体的不可分量构成模型，此后简称为（连续体的）*不可分量模型*，是站不住脚的。这里我们要争议的不是连续体或不可分量的存在性，而是*不可分量模型*的合理性。

首先，什么是连续体？据说它就是一个连续的实体（如果"连续的"并非同样需要解释就好了）。在此我们只谈一些对连续体的理解：一条直线就是一个简单的实例（实线是它的*不可分量模型*）。通俗地说，它在一致性方面做到了极致，以至于没有结构和细节（且这么说吧）— 当看到它时，无论着眼于那一部分或多大的范围，瞥一眼就足够了，再怎么看也不可能有更多的发现。这种完美的一致性要求直线的*不可分量模型*或没有空隙或一片空白 — 通体严格地一致。如果*不可分量模型*像希望的那样没有空隙，那么每一个可能的量值在实线上都有一个点与之对应，并且每个合理推断出来的点都能在*模型*上找到。于是*模型*具有镜像对称性和移动对称性。

符号 $[0, 1)$ 表示 0 和 1 之间的全部实数的集合，包括 0 但不包括 1。然而，我们接

下来要阐明的是这样一个静态的集合不可能“没有漏掉的数”。因此，在我们的讨论中“没有空隙”不是一个预设的或已知的性质。从现在开始，我们采用一个经过修改的符号，$\underset{\cdots}{[0, 1)}$ — 点状的下划线代表着其完备性尚在未知。同时为了方便，在不至于引起歧义的情况下我们将省略这个点状的下划线。

思维实验 4.1 内围点

在两个不同的实点之间总存在另一个实点，我们称之为两个给定点之间的一个内围点。如果 $0 < X_1 < X_2 < \cdots < X_n$, 且 Y 是 0 和 X_1 之间的一个内围点，那么 Y 同时也是 0 和 X_i（$i = 2, 3, \ldots, n$）之间的一个内围点。因此，我们说 Y 是位于 0 和 $\{X_1, X_2, \ldots, X_n\}$ 之间的一个内围点。以此类推，1 是位于 0 和 (2, 3] 之间的一个内围点。

对于每个 x $(0 < x < 1)$，在 0 和 $[x, 1]$ 之间是否存在一个内围点？每个人都会毫不犹豫地给出肯定回答，因为 $(0 + x)/2$ 就是一个实例。那么，对于每个 x $(0 \le x < 1)$，是否有一个内围点存在于 0 和 $(x, 1]$ 之间呢？这个问题与前一个相同吗？或者，“$(x, 1]$ 对于 $0 \le x < 1$”（作为各种可能性的总和）可以更逼近 0 吗？单从“开/闭区间”的概念来说，它更有优势去逼近 0 — “$[x, 1]$ 对于 $0 < x < 1$”和“$(x, 1]$ 对于 $0 \le x < 1$”可以看作是两个以线段的不可分量模型为元素的集合，其中每条线段都以 1 为右端点，而 (0, 1] 是两个集合中唯一能覆盖任意其它成员的元素。（但这样就清楚了吗？我们将在**思维实验4.4**中继续讨论。）然而，考虑到 $\bigcup\{[x, 1] : 0 < x < 1\}$ 和 $\bigcup\{(x, 1] : 0 \le x < 1\}$ 是两个由实数（点）构成的集合 — 两者都是 $\{x : 0 < x \le 1\}$，前面的观点需要再推敲。两个归结到点的集合是相同的，这提示我们原本的那两个集合覆盖了完全相同的区域，基于此种考虑，对于先前两个相关问题的回答应该是一样的。

在开头的第二个问题中我们其实想问是否有一个内围点 Y 存在于 0 和 $\underset{\cdots}{(0, 1]}$ 之间？现在我们把 $\underset{\cdots}{(0, 1]}$ 看作是一个以点为元素的确定不变的集合，再来直面问题本身。如果答案是肯定的，那么 $\underset{\cdots}{(0, 1]}$ 并不像希望的那样完备，因为至少缺失了 Y。如果答案是否定的，那么 0 和线段 (0, 1] 之间是没有空隙的（也就是说，它们彼此接触）。这么一来，如果线段是纯粹由点构成的，人们不禁要问：(0, 1] 中的哪个点与点 0 靠的这么近，以致让其它的点无隙可乘（或者，这个单端点线段上的哪个点直接接触点 0）？

如果线段可以解释成一个由点构成的集合，而每个点都有一个位置值，那么这些位置值有一个算术平均值。并且几何形心正是这个平均值的所在。

思维实验 4.2 缺席的形心

观察线段 (–1, 1), 显然 0 是其所有点的算术平均值。但对于 (–1, 1)$\bigcup$\{1\} 来说呢？线段 (–1, 1] 有形心吗？毫无疑问，就像任何物体都有质心一样（问题只是这个质心是否落在了物体的内部。有些物体，例如救生圈，其质心就落到体外

了）。而且理所当然，直线的每个有限的连通部分的“质心”都落在“体”内，并且与形心重合。那么 (–1, 1] 的形心在哪？（对于那些认为多一个点不至于造成影响的人，我们有几个问题：一个单点构成的图形是否有形心？{0} 的形心是哪个点？还有，{0}⋃{1} 的形心呢？）直观上，新加入的点 1 会将它“拖”向右边一点点。另一个简单的事实是，对于一个有界的几何图形，其中任何关于形心对称的部分（但不是全部）都可以被去除而不改变形心的位置。于是我们很容易检验 (–1, 1] 的形心的任何一个待确认点。很遗憾，没有哪个已知点能够通过验证。难道它是一个如我们在**思维实验 4.1** 中所讨论的位于 0 和 (0, 1] 之间的一个内围点吗？

何不回到算术平均值的概念呢？因为线段 (–1, 1) 只是 (–1, 1] 的左边的一部分，所以它们两个的形心不可能重合。对于那些有兴趣计算后者的所有点的精确平均位置值的人来说，结果会聚焦在 0 + (半个点距)。的确没有哪个已知点处在这个位置上。

如果像我们被传授的那样，在实线上不存在相邻点，那么任何一个有界的实体不可能沿着实线做一个距离正好是一个点的大小的位移。但转念一想，如果一个点就是一个实体，并且从实线上移除一个点就会产生一个点大小的空隙，那么我们当然可以料想存在着一个点大小的距离并且存在这样一种距离的位移。

思维 4.3 移动一个点的距离

观察 [0, 2) 和 (0, 2]。从直线运动的角度来看，它们互为位移结果 — 一个点距的移动就此浮现眼前。顺便提一句，参照**思维实验 4.2**，它们两个的位置平均值分别是 1– (半个点距) 和 1+ (半个点距)，在位置上相距一个点。

书中的纸页可以被书签分隔成两组而不打乱原先的顺序，实线上的点类似于此。（接下来，如果某个经典的思想出现在脑海，暂时先忘了它吧。）假设一个虚拟的书签将实线分隔成 $(-\infty, b)$ 和 $[b, +\infty)$，我们将这个“书签”连同它的位置用符号 **ϒb** 表示。另一个同类符号，**bϒ**，表示位于 **ϒb** 后面“一页纸”的“书签”。如果诸如 $(-\infty, b)$ 与 $[b, +\infty)$ 的两部分之间不存在空隙，那么“书签”的“厚度”是 0 并且可以虚拟地插入到每个实点的紧前或紧后。这样一来，这个虚拟书签可以移动一个点距 — **ϒb** 与 **bϒ** 就刚好相距一个点距。而且，如果有一个点位于“书签”的紧前，根据对称性，应该有另一个点对称地位于“书签”的紧后 — 一对相邻点就此现身了。其实，麻烦可以来得更早 — 在这项讨论之前我们应该先问一下是否可以刚好在这个“书签”的位置上放置一个新增点。这让我们想起了**思维实验4.1**。

旁白：上述所有的例子描述的似乎都是一种突变的“运动”。确实，我们的目的在于通过揭示一些小到出人意料的距离的存在来反驳连续体的不可分量模型。对于那些更关心运动的连续性的人群，我们给出一个例子：当点 x 连续地从位

置 2 运动到位置 4 时，$(0, 2)\bigcup\{x\}$ 中的所有点的算术平均值的位置从 *1+(*半个点距*)* 连续地运动到了 *1+(*半个点距*)+(*一个点距*)*。

还记得**思维实验3.1**中的“为什么会有这么费事的表述”吗？还有在**思维实验4.1**中产生的问题：“$[x, 1]$ 对于 $0 < x < 1$”和“$(x, 1]$ 对于 $0 \leq x < 1$”（作为两个集合）两者中的超级元素 $(0, 1]$ 仅存在于后者中，在这种情况下两者怎么可能覆盖了完全一样的区域呢？为了追寻确切的答案，我们需要面对另一个问题：一个开区间真就是“开”的吗？

按标准观点，$[0, 3]$ 可以被“分割”成 $[0, 2)$ 和 $[2, 3]$。但为什么 $[0, 2)$ 没有右端点（而 $[2, 3]$ 有左端点）？如果有，这个点必是点 2 的左邻点。我们接受的教育是，在实线上不存在相邻点。于是没有什么好奇怪的。然而这样的解释并不能消除我们的疑惑。从直线的一致性来看，实线上的每个点除了其自身的独特位置外别无特殊之处。这意味着如果任何一点可以作为端点的话，其它点也可以。对一个有界的连通部分的边界来说，添加或移除一个点除了改变相关边界的位置外别无影响 — 边界的形态或样式不会变。

思维实验 4.4 可疑的开区间

假设有一个反物质实体 $[-1+t, 1+t]$ 沿着实线轨迹滑向普通实体 $[2, 4]$，并且它的位置是随时间 t 均匀变化的。一次特殊的碰撞将在所难免，而且在这个过程中两个实体将同归于尽，因为它们在任意时刻都是镜像对称的。

如果是两个反物质点 $-1+t$ 和 $1+t$ 而不是 $[-1+t, 1+t]$ 移向 $[2, 4]$ 呢？一个中间结果会是其中一个移动点已经和 $[2, 4]$ 的左端点相互抵消了，另一个移动点正从左边滑向 $(2, 4]$。接下来会怎样呢？$(2, 4]$ 中的哪一个点将消失？

转而考虑另一种情况。假设点 x 沿着实线从 1 连续移动到 3。如果 $[1, 3]$ 上没有空隙，那么区间上的所有点恰好就是 x 的所有可能位置。所以“*按照时间顺序相应地处于各点的位置*”即是“*连续运动穿过该区间*”。再来观察一个生长中的线段 $[0, x]$，其中 x 顺次遍历 $[1, 3]$ 中的所有点。因为是连续生长，这个生长的线段显然将遍历（$[0, 3]$ 上的）所有包含了 $[0, 1]$ 的连通部分。因此存在一个点 X_i ($1 \leq X_i \leq 3$) 使得 $[0, X_i]$ 正好是 $[0, 2)$，因为后者显然是 $[0, 3]$ 的一个连通部分且包含了 $[0, 1]$。这就是说 $[0, 2)$ 的右端点是 X_i。否认 X_i 的存在就意味着在*不可分量模型* $[1, 3]$ 上有一个空隙。但 X_i 的存在否定了关于相邻点的标准说法。

再看一个无需借助运动概念的版本：既然实线被假定为无空隙的，用它来当天然的尺子，其上的每个点作为一个刻度，那就是最精确的了。“最精确”意味着总有足够的刻度以应对任何测量，以至于误差和估计这类概念都用不上了 — 每一个有限的长度值都可以从“尺”上的某一点直接读数，哪怕是一个点的长度变化都能从刻度上直接反映出来而无需补充的描述。否则，测量精度还可以提高，只需把每个已知的例外长度值（目前还无法直接读数的那些）直接（在原有的模型上作为新的点、新刻度）添加进来，这也说明原有的模型并非“无

空隙”。另一方面，直线的每一个有界连通部分其本身就是一个确定的长度值，这与是否存在能对它进行直接读数测量的尺子无关。

5 潜无穷过程的“最终结果”

连续体的不可分量模型根本经不起推敲。现在我们要面对一个（经典的）后续问题：对一条线段迭代地进行对半分，最终结果是什么？

要讨论这个问题，我们需要先明确一下“最终结果”的意思。我们在描述一个潜无穷多的操作步序时会用“经过所有步骤后”、“最终结果”或类似的表述。它们在直觉上有着怎样的起源？这里有两个例子：

例 1：对于一个潜无穷次的收集任务，初始状态是一个空集；过程中的中间结果分别是 $\{1\}$、$\{1, 2\}$、$\{1, 2, 3\}$ 等等；而“最终结果”被认为是 $\{1, 2, 3, \ldots\}$。
例 2：对于芝诺悖论之一 [2, 第 349 页] 所涉及的几何级数，其初始状态是一个空记录；过程中的中间结果分别是 $\frac{1}{2}$、$\frac{1}{2}+\frac{1}{4}$、$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}$ 等等；而“最终结果”被认为是 $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots$。

我们将一个潜无穷过程的前 n 步记为 S_1、S_2、$\ldots$、S_n，初始状态记为 C_0，第 n 步产生的中间结果记为 C_n，“最终结果”记为 C。现在来执行一个收集任务 — 连续收集已完成步骤的符号。如果主过程能够结束的话，我们的收集任务也将随之完成，并且能得到一个无穷集 $\{S_1, S_2, S_3, \ldots\}$，也就是说 C_0 在经历了 $\langle S_1, S_2, S_3, \ldots\rangle$ 之后变成了 C。从变换函数的角度来看，前 n 步作为一个整体决定了一个函数 T_n 满足 $C_n = T_n(C_0)$，而所有步骤作为一个整体决定了函数 T 满足 $C = T(C_0)$。

然而，我们相信所有关于潜无穷过程的所谓的最终结果都是人类的想象而已，或更确切地说，一些本来就存在的或是假定存在的东西被拉去充当根本就不存在的“最终结果”。这就难怪人们有时要困惑地面对同一个潜无穷过程的不同版本的“最终结果”。正如在罗斯–利特伍德悖论 [3] 中遇到的情况（这个悖论为人们重新思考无穷大提供了一个良机），想要在不同的版本中选边站往往是困难的。

思维实验 5.1 “最终结果”?

设想有一排望不到尽头的路灯按自然数顺次编号。我们从路灯全亮状态开始，并将这个状态记为 $\langle \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots\rangle$，这是一个周期为 1 的序列。这里“周期为 m”指的是对于所有 n 来说第 $n+m$ 盏灯与第 n 盏的状态相同 ($m, n = 1, 2, 3, \ldots$)。

步骤 1：对所有亮着的路灯，从左到右，关掉第 1 盏、跳过第 2 盏、关掉第 3 盏、跳过第 4 盏，以此类推。（那些形如 $2k+1$，k 为非负整数的都被关掉了，而那些 2 的倍数仍保持点亮。）然后将结果记为 $\langle \not{1}, \mathbf{2}, \not{3}, \mathbf{4}, \not{5}, \mathbf{6}, \ldots\rangle$，这是一个周期为 2 的序列。

步骤 2：执行如上一步同样规则的操作。（那些形如 $4k+2$ 的这次都被关掉了，而那些 4 的倍数仍保持点亮。）将结果记为 $\langle \cancel{1}, \cancel{2}, \cancel{3}, \mathbf{4}, \cancel{5}, \cancel{6}, \cancel{7}, \mathbf{8}, \cancel{9}, \cancel{10}, \cancel{11}, \mathbf{12}, \ldots \rangle$，这是一个周期为 4 的序列。

步骤 3：. . . 。（那些形如 $8k+4$ 的这次都被关掉了，而那些 8 的倍数仍保持点亮。）. . . ，这是一个周期为 8 的序列。

如此渐进，以至无穷。

我们将“最终结果”记为 $\langle \cancel{1}, \cancel{2}, \cancel{3}, \cancel{4}, \cancel{5}, \cancel{6}, \cancel{7}, \cancel{8}, \cancel{9}, \cancel{10}, \cancel{11}, \cancel{12}, \ldots \rangle$，看来没有一盏灯是亮的。但奇怪的是序列的周期是 1，而从周期在整个过程中的严格单调递增趋势来看，它不可能最终是一个有限的数。

然而，如果我们按照另一种规则来模仿上面的过程 — “关掉第 2 盏，并且在相邻两盏里面间隔地关掉一盏”，中间结果将分别是 $\langle \mathbf{1}, \cancel{2}, \mathbf{3}, \cancel{4}, \mathbf{5}, \cancel{6}, \ldots \rangle$、$\langle \mathbf{1}, \cancel{2}, \cancel{3}, \cancel{4}, \mathbf{5}, \cancel{6}, \cancel{7}, \cancel{8}, \mathbf{9}, \cancel{10}, \cancel{11}, \cancel{12}, \ldots \rangle$、. . . 与上面的中间结果分别具有相同的周期。（更多的细节是：到第一步结束仅仅那些形如 $2k+1$ 的还亮着；到第二步结束仅仅那些形如 $4k+1$ 的还亮着；到第三步结束仅仅那些形如 $8k+1$ 的还亮着；以此类推。）我们得到的“最终结果”将是 $\langle \mathbf{1}, \cancel{2}, \cancel{3}, \cancel{4}, \cancel{5}, \cancel{6}, \cancel{7}, \cancel{8}, \cancel{9}, \cancel{10}, \cancel{11}, \cancel{12}, \ldots \rangle$，它不具有有限长度的周期，这符合过程中表现出来的趋势。

现在引入一个概念，*相对点亮率*，这是一个类似于入住率的一个比例值。$\langle \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots \rangle$ 的*相对点亮率*是 1，而 $\langle \cancel{1}, \cancel{2}, \cancel{3}, \ldots \rangle$ 的是 0。对于一个呈现出简单模式的序列，这个比值是容易得出的。于是我们可以将*相对点亮率*在前一次实验中的演进过程描述为 1、$1-\frac{1}{2}$、$1-\frac{1}{2}-\frac{1}{4}$ 等等，而“最终结果”为 $(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\cdots)$。对于后一次实验，可以将*相对点亮率*的演进过程描述为 1、$1-\frac{1}{2}$、$1-\frac{1}{2}-\frac{1}{4}$ 等等，而“最终结果”为 $(1-\frac{1}{2}-\frac{1}{4}-\frac{1}{8}-\cdots)$。从这个意义上说，两个实验过程是自始至终保持一致的。奇怪的是，两个“最终状态”相差了一盏灯。这可能就是总有一部分人拒绝认为 0.999. . . 就是 1 的原因 — 在他们的内心深处还有一盏灯亮着。

再次回到初始状态。这次我们将这些灯看作是在一条弹力带上等间距排列的一列无穷尽的点，而弹力带已经做足了预拉伸。将第 1 盏灯的位置作为参照点。接下来我们只考虑亮着的灯。

步骤 I：对所有点亮的灯，关掉第 2 盏，并且在相邻两盏里面间隔地关掉一盏。然后让弹力带相对于参照点做收缩，使得当前第 2 盏亮着的灯在空间上占据初始时第 2 盏亮着的灯的位置。（于是当前第 3 盏亮着的灯占据了初始时第 3 盏亮着的灯的位置。以此类推。）

如此反复，以至无穷。

这次的“最终结果”是什么？单从数的形式上判断，只有 1 还亮着，因为对于自然数 n，想要经历前 k 步而保持点亮状态就意味着 n–1 能被 2^k 整除，反

之亦然。然而，从每个步骤呈现出的重复情形很容易看出，并不存在一种指向“最终结果”的趋势，而且连摆脱初始状态都没指望。这就令人越来越怀疑“最终结果”的真实存在。

所有的潜无穷过程，从它们的步骤上看（不涉及到任何一步的内部细节）都有一个共同的无尽序列结构，我们称之为 N-结构，因为自然数作为无穷尽的序数就是从这种结构的各种实例中抽象出来的 — 每个实例都有一个唯一的起始项，这个起始项有且只有一个后继，这个后继又同样地有且只有一个后继，如此延续（并无其它 — 没有分支，没有循环，而且后面的每项都与起始项相距有限的“距离”）。一个建立在步骤序列层次上的结论（因而无关乎任何一步的内部细节），无论它是从哪个特殊过程得出的，只与宏观结构有关并同样地适用于 N-结构的所有实例。因此，任何一个“最终结果”不成立就意味着 N-结构（潜无穷多步的序列）中没有一个能结束。

为什么总有那么多人认为所谓的最终结果是理所当然存在的？这是因为他们的头脑从一开始就被灌输了一些经过精心挑选的例子，每个例子都指向一个显然的“最终结果”，自然且完美到了毋庸置疑的地步。但它们真的不容置疑吗？

思维实验 5.2 淡水还是盐水？

把每个自然数看作是一杯盐水。我们将初始状态记为 $\langle \underline{\underline{\mathbf{1}}}, \underline{\underline{\mathbf{2}}}, \underline{\underline{\mathbf{3}}}, \ldots \rangle$。它们都有相同的浓度，如果汇集起来能形成一个具有同样盐度的无穷大的盐湖。

步骤 1：将第 1 杯盐水替换成一杯不含盐的淡水。结果记为 $\langle 1, \underline{\underline{\mathbf{2}}}, \underline{\underline{\mathbf{3}}}, \ldots \rangle$。

步骤 2：将下一杯盐水替换成一杯淡水。结果记为 $\langle 1, 2, \underline{\underline{\mathbf{3}}}, \ldots \rangle$。

如此以至无穷。

经过了所有步骤之后（姑且这么说吧），我们将“最终结果”记为 $\langle 1, 2, 3, \ldots \rangle$。显然所有杯子里都是不含盐的淡水，如果它们汇集起来能形成一个无穷大的淡水湖。这样完美的结局似乎是无需质疑的。只是故事并未就此结束。

另一方面，我们可以一步一步地推断“最终”湖水的盐度。我们将初始浓度记为 C_0。假设步骤 1 的结果汇成的盐湖具有浓度 C_1（达到均匀的稳态后）。考虑到第一次替换的影响，我们有 $C1 > C_0 - C_0/4 = (1 - 1/4)C_0$。步骤 2 的结果汇成的盐湖具有浓度 C_2。同时还考虑到第二次替换的影响，我们有 $C_2 > C_1 - C_0/8 > (1 - 1/4 - 1/8)C_0$。同理，我们有 $C_3 > C_2 - C_0/16 > (1 - 1/4 - 1/8 - 1/16)C_0$，以及其它。于是，“最终”的湖水具有浓度 C，满足 $C > [1 - (1/4 + 1/8 + 1/16 + \cdots)]C_0$。可以相信，通过在不等式的右边选用不同的系数，我们甚至能够推断 C 比任何一个低于 C_0 的给定浓度都更高。

我们再给出一个从静态观点出发的实验版本。有一个无穷大的盐湖由无穷多个两两不相交的单元组成，各单元用自然数依次编号。所有的浓度都是相同的，记为 C_0。但是有人可能来了兴致，要思考一下盐湖浓度与每个单元的盐水浓度之间的关系。显然，湖水的盐度是靠所有单元的含盐量来维持的，而其中第

n 个单元中的盐在 C_0 中的贡献值小于 $C_0/2^{n+1}$。结果就成了那个已知为 C_0 的湖水浓度值可以小于 $(1/4+1/8+1/16+\cdots)C_0$。（顺便说一句，从某种角度考虑，有人甚至可以推断湖水的浓度高于 C_0。）这是荒谬的。

一个真相可以接受来自于任何角度的审视 — 不会产生两幅相互矛盾的图景。一个人为的幻象就不是那么回事了。以上我们列举的种种困惑都揭示了所有的“最终结果”尽是人为。

思维实验 5.3 贫穷或富有？

有无穷多只用自然数编号的存钱罐，第 1 只里面有 1 枚硬币，第 2 只里面有 2 枚硬币，第 3 只里面有 3 枚硬币，以此类推。我们将初始状态记为 $\langle 1, 2, 3, \ldots\rangle$。

步骤 1：拿走那只装有 1 枚硬币的存钱罐（原本的第一只存钱罐）。将结果记为 $\langle 2, 3, 4, \ldots\rangle$。

步骤 2：拿走那只装有 2 枚硬币的存钱罐（原本的第二只存钱罐）。将结果记为 $\langle 3, 4, 5, \ldots\rangle$。

如此以至无穷。

当完成了所有步骤（暂且这么说吧），那就什么也不剩了。

现在再回到初始状态，开始一个对比的过程。

步骤 I：往每只存钱罐里添加一枚硬币。将结果记为 $\langle 2, 3, 4, \ldots\rangle$。

步骤 II：再往每只存钱罐里添加一枚硬币。将结果记为 $\langle 3, 4, 5, \ldots\rangle$。

重复以至无穷。

进行完所有步骤（且这么说吧），我们将“最终结果”记为 $\langle a_1, a_2, a_3, \ldots\rangle$。按照康托尔的理论，每只存钱罐里有 $\aleph_0$ 枚硬币。只是，我们在本例仅关注结果是否为空。

对上述两个过程的中间结果做一个对比发现，在集合论中，它们是分别相同的。于是有个无法回避的问题：相同的路径（起点也相同）怎么会通往两个截然相反的目的地？

难怪总有那么多人只接受潜无穷，而就此止步 — 在逻辑上和直觉上都缺乏清晰性成了通往实无穷的道路上的致命障碍。

“最终结果”与“最终完成”是同一枚硬币的两面。基于这一点，一个自相矛盾的“最终结果”表明两者皆是子虚乌有。

6 回首雾区

自康托尔以后，在数学基础领域，大多数数学家轻易屈从于一种非对称的推理 — 他们的逻辑标准是可变的，具体要看一个正常推理的结论是否符合正统观念。例如，“x

*不同于一个无穷集合中的每一个元素”*被认为足以证明“*x 不属于该集合*”，而“*点 1 到 [0, 1) 上的任何一个点都有一个数值为正的距离*”被认为不足以断定“*点 1 与 [0, 1) 是分离的（两者之间有空隙）*”，尽管两个判断都基于同一个原理 — *一个集合是由它的元素唯一决定的（于是，关于一个集合的命题可以通过逐一测试它的所有元素来检验）*。

然而一个数学理论是不可能长期依靠非对称推理来支撑的。在超限数理论繁荣和光鲜的背后，必定有一些或多或少看起来合理的成份作为支柱。

康托尔断言 [4]，“*每个潜无穷，如果在数学上严格可行，应当以一个实无穷作为先决条件。*”莫里斯·克莱因对此解释道 [2, 第 200 页]，“*他认为潜无穷实际上依赖于一个逻辑上在先的实无穷。*”

康托尔的表述说明他接受实无穷和无穷集。我们把这种信念称为信念A。此时实无穷仍然缺乏活力，在这种状态下对它们还不能有太多的热情期待。然而，让它们变得生动鲜活的，就像在许多存有争议的涉及无穷大的证明中那般灵动的，是一个相关的但不同的信念：*每个潜无穷，如果严格可行，将导致一个实无穷*。这个信念在有限与无限之间架起了一座桥梁，跨越无望鸿沟。我们称其为信念B。意识中有了它，每个人都可以通过“完成”一个潜无穷过程来得出一条结论，通过一个相关的潜无穷来推出实无穷的性质。我们和许多其他人一样认为信念B是一个超级麻烦制造者。正如赫曼·魏尔所说 [1, 第 1200 页]，“*... 那些不断发展超越任何阶段已达到的状态的数的序列 ... 面向无穷展现出多种可能性；它永远保持着创建发展的状态，不会作为一个禁锢不变的圈子而独立存在。我们盲目地将一种状态转变成另一种，这种做法成了麻烦之源，...。*”

如果潜无穷通往实无穷，那么一个潜无穷可以被看成是一个函数，它将输入的初始状态映射成输出的最终状态。但许多证据表明实际并非如此。严格按照信念B导致了**思维实验5.2**中的困惑，而将信念B与*外延公理*相结合引发了**思维实验5.3**中的矛盾。（在公理集合论和素朴集合论中，外延公理都同样传递了这样的理念，即*一个集合是由它的元素决定的，而与描述方式无关*。）尽管这个公理对于有限个元素构成的集合是清晰和可靠的，但我们的实验表明当它和信念B结合在一起时是不可信的。

那么是否除去信念B的影响就会一切重归和谐？还没那么乐观。目前的情形着实令人担忧。信念A被巧妙地用来呈现一场由康托尔发起的激动人心的魔术表演。而且为了配合这场史诗级的演出，关于合理性和清晰性的标准已经遭到扭曲；一些仅仅部分符合逻辑的妙证从此占据了数学基础的舞台中央。同时也必须为数学家们说句公道话，没谁诚心想要猾或欺骗。

一一对应，这个集合论的基本原理，对于有限个元素构成的集合是简单清晰的。但将它用于无穷集，结果对人类直觉造成了巨大冲击。对于康托尔来说“*一个集合是无穷集，当且仅当它能与自身的一部分构成一一对应*”[1, 第 995 页]。然而，这样一条想象中的无穷集的基本性质违反了*欧几里德第五公理（即整体大于部分）*，混乱的序幕就此拉开，诸如希尔伯特旅馆 [5] 的各种悖论开始登场。

我们从另一个角度来重新思考**思维实验5.2**中提出的问题。假设有两个无穷大的盐湖，其中一个的浓度是 $C\,(C > 0)$ 而另一个的是 $C/2$。如果将盐和水分离开来，每个湖

都变成了量值为无穷大的盐和淡水，而且根据康托尔的理论两个湖产生了同样多的水和同样多的盐。这就意味着，如果给定了量值分别为无穷大的盐和淡水，人们可以得到一个盐湖，其浓度可以是 C 或 $C/2$ 或某种其它值。这样就违反了*矛盾律*（一个基本逻辑原理，说的是对同一对象不能同时既肯定又否定）。

旁注：*这个异常的浓度来的并不那么突然，因为此前已经有希尔伯特旅馆悖论描述过一个同样不确定的类似比率，入住率。在希尔伯特旅馆问题里先给定了为常数值的无穷大的人数和房间数，这时的入住率可以根据需要来调整（约定每个房间最多住一个人）。*

排中律（另一个基本逻辑原理，说的是*任何一个命题或为真，或其否命题为真*）安然无恙吗？我们在**思维实验3.1**中任意地使用这条原理，可是它对于无穷集还能适用吗？不能。为了解释这个观点，我们给出一个改编自**思维实验3.1**的相关想法的实验。

思维实验 6.1 是否包含？

考虑将 0 和 1 之间（含 0 和 1）的有理数表示成即约分数（特别地，将 0 表作 0/1，而 1 作 1/1）。我们将这些数分成两个集合 — 集合 B 由那些分子或分母是偶数的数构成，其余的都归到集合 D。令 $A = B \bigcup D$。我们注意到介于 A 的任意两个元素之间都存在着一个 B 的元素和一个 D 的元素。

旁注：*这一点容易理解。观察 2/(2n+1)、4/(2n+1)、...、2n/(2n+1)，写成即约分数可以看出它们都属于集合 B。它们在 [0, 1] 上均布，间距小于 1/n。类似地，观察 1/(2n+1)、3/(2n+1)、...、(2n-1)/(2n+1) ... 都属于集合 D。它们 ...。*

集合 D 的元素可以这样来决定集合 A 的一个划分，对于 $d_j \in D$，等价类 $[d_j]$ 定义为 $\{$有理数 $x \in [0, d_j] : \forall d_i \in D\,(d_i < d_j \Rightarrow d_i < x)\}$。显然，每个等价类含且只含集合 D 的一个元素。我们的问题是：等价类 $[d_j]$ 中是否含有集合 B 的元素？在**思维实验3.1**中涉及到一个类似的问题（在那里我们采用了一种更直观的操作 — 同时从无穷多点处分断）。答案不能是肯定的，否则 $[d_j]$ 还会含有集合 D 的其它元素。另一方面，答案也不能是否定的，否则，同样道理其它等价类也不含集合 B 中的任何元素，但这意味着 $D = A$。

希尔伯特对康托尔的工作有一句著名的评价 [1, 第 1003 页]：“*任何人都不能把我们从康托尔为我们创建的乐园中赶走。*”可是，如果抛开了基本逻辑，谁又能区分真实和虚幻呢？

7 一些后话

一旦这一系列思维实验得以呈现给公众，我们的任务就完成了。我们大可静观其变，待迷雾散尽时看乐园是否还在。趁着各种声音尚未重新聚焦到那场“早已尘埃落定的”关于康托尔的论战，我们再多说两句。

怎样理解无穷大？一个密切相关的问题是，怎样理解连续体？几千年来这些问题一直困扰着人类，而且至少从亚里士多德的时代开始就有建立在含糊论证基础上的大量的学说。我们的目的是拿出简单而有力的证据来反驳时下正统的观念，那些我们一直试图去理解但始终无法苟同的说法。

旁注：亚里士多德写道 *[6, 第 III 卷 第 6 章]*，“不会有实无穷。”并且在他看来“... 没有哪个连续体可以由‘不可分元素’构成：例如，直线不能由点构成，直线是连续体而点是不可分元素。”*[6, 第 VI 卷 第 1 章]*

我们的演示支持这样的观点：

(i) 潜无穷并不导致实无穷。实无穷依靠公理而存在，而康托尔的理论尽管影响甚广但不可靠。(ii) 与此相关，不可分元素不可能静态地构成连续体。连续体依靠公理而存在，而它的不可分量静态构成模型尽管实用但不合逻辑。

旁白：在我们的讨论中，“存在”或“存在性”是什么意思？通俗地说，在数学上，任何东西如果在一个相容的系统中能找到立足之地，就说它存在于该系统。一个东西可以在一个系统中存在，而在另一个系统中不存在。虽然两个相冲突的东西不能共同存在于一个相容的系统中（否则该系统就是不相容的），它们可以分别存在于两个系统，如，欧氏几何和非欧几何。并且许多存在的东西（例如，无理数）曾经不为人知或被认为是不可能存在的。尽管存在与否本身是客观的，也有一些东西（例如，无穷小）命运起伏。

高斯说 [1, 第 993 页]，“我强烈反对将无穷大的量当作实体来用，这在数学上从来都是不允许的。无穷大只是一种说话的语气，人们在谈论极限时用到的说话方式 ...。”他的论述表明他相信实无穷的概念不可能从其它概念中导出，也不可能合理引入。但康托尔并不这么认为。他指出 [4]，“为了在一些数学研究中有一个变量，这个变量的定义域必须在严格意义上事先已知 ...。因此这个定义域是一个关于取值的确定的实无穷集合。”

康托尔断定，必须接受实无穷。下一步通常是围绕这个思路找到一些新的概念，并通过发掘一些新的联系和关系来激活这些概念。关键是要完全保证相容性和清晰性，因为它们的缺失足以动摇一切。康托尔做了许多工作，但非常遗憾地留下了一个迷雾笼罩的“乐园”。更糟的是，威胁到他的理论的不是一个漏水的屋顶，而是一个摇晃的地基。

在现实中，康托尔的思想不可思议地取得了非凡的胜利，但仍有一些人持反对观点。在对立的双方之间是否有一个中间地带呢？或许有吧。亚伯拉罕·罗宾逊的立场基于两个要点或原理 [7, 第 507 页]：“*(i)* 无穷大的总体在任何意义上都不存在 ...。*(ii)* 然而，我们应该继续‘照常’做数学，也就是说，我们应该表现的就像它真地存在一样。”他的灵活态度有助于人们在尚未真正认识实无穷之时调和并达成双方的愿望。但关键是对于“照常”或“就像”需要设定一个容忍限度，否则数学将永远如当今这般被“雾中之雾”所笼罩（“赫曼·魏尔把康托尔的超限数的层级理论称为雾中之雾”[1, 第 1003 页]）。（这个限度的轮廓还不是我们此刻所关注的。）

此时，无穷大的命运吸引着更多的关注。就像康托尔的“乐园”并非其最终归宿，“乐园”的崩塌也不意味着无穷大走到了山穷水尽。单就存在性而言，实无穷并不比虚数来

的更突兀（虚数对数学的发展贡献颇多）。这不是什么证明不证明的问题，无穷大的当务之急是发掘更多的正面意义，而不是仅仅自圆其说或自娱自乐。事到如今，迷雾正在逐渐消散，实无穷可以从神话般的梦境中走出来品味一番新鲜空气了。

围绕着一个事实通常会有不止一种信念。而那些持不同观点者时刻准备着为真理而战，并坚持走自认为正确的道路。无论如何，为了隐藏矛盾而将其推向迷雾深处，但并不从根本上消除它们，这样的做法永远是一个选项，但绝非解决之道。

讨论：*上面所说的“... 迷雾深处”指的是什么？在此仅举一个简单的例子。让我们暂时再回到康托尔的理论。在*思维实验*2.1*中，如果考虑序型的概念，其中的悖论还可以进一步讨论。一些人会辩称无穷多次的位置交换将引起序型的改变。另一些人不这么想。他们认为，在序列中位置和成员项之间的关系正如固定的座位与座位上的人的关系 — 座位是静止的而人是活的。如果不存在空着的座位，也不存在未就座的人，那么人的序型即座位的序型，而座位始终是静止的。但有些人从另外的角度看这个问题。对于自然数序列：$1, 2, 3, \ldots$，不断重复调换成员 1 与其紧随右邻的位置。在那些认定了有序型变化的人看来，“最终结果”将是 $2, 3, 4, \ldots, 1$，它的序型在康托尔的理论中是 ω+1，于是*思维实验*2.1*中的悖论能够被解释了。这看起来挺合理。但继续按这个思路再走下去会怎样呢？接下来，再不断重复调换成员 2 与其紧随右邻的位置。问题来了：是否有一步的中间结果是 $3, 4, 5, \ldots, 2, 1$ 呢？困惑又找上门了。有人会接着反问：即使成员 2 陷入了困境，那又怎样？会影响到前面提到的 $2, 3, 4, \ldots, 1$ 的实现吗？好吧，只要对比一下从 $1, 2, 3, \ldots$ 开始和从 $1, 2, 3, \ldots, 0$ 开始就知道了。*

无论人们是否乐见，无穷大的概念确实营造了一个乐园，一个想象力的乐园。然而，即使是一个合理的想象，它会因合理的方面对科学有益，也还有不合理的方面（否则它就是一个事实、一条真理）。关键是要善于利用它，同时又不迷失在雾中。有害的不是想象，而是把想象当作真实。

旁注：*众所周知，在物理学上，通常一个刚体的全部质量被当作集中于一个称为质心的点上，尽管事实上质量分布于物体的各处。这个想象简化了力学和运动问题的求解。但它并非总有效，因为它并不等同于事实，否则也就不需要重心这个概念了。数学上关于无穷大和连续体的想象也同样。*

致谢

作者感谢所有提出宝贵建议的人。特别要感谢李瑞对于一些原始粗陋想法的倾听。

附录

A 思维实验 2.1 其中一部分的并蒂版

令 q 等于 $\langle q_{01}, q_{02}, q_{03}, \ldots \rangle$ 中任意给定的一项。按自然顺序遍历 q 的小数位，根据每一位小数做出一次相应操作：

步骤 1：在 $\langle q_{01}, q_{02}, q_{03}, \ldots \rangle$ 中从 q_{02} 开始顺次向后搜索，直到找到至少在第 1 个小数位与 q 不同的某项。交换搜索结果与 q_{01} 在数列中的位置；将重排后的数列记为 $\langle q_{11}, q_{12}, q_{13}, \ldots \rangle$。

步骤 2：在 $\langle q_{11}, q_{12}, q_{13}, \ldots \rangle$ 中从 q_{13} 开始向后搜索，直到找到至少在第 2 个小数位与 q 不同的某项。交换搜索结果与 q_{12} 的位置；将重排后的数列记为 $\langle q_{21}, q_{22}, q_{23}, \ldots \rangle$，它等同于 $\langle \boldsymbol{q_{11}}, q_{22}, q_{23}, \ldots \rangle$。

步骤 3：... 从 q_{24} ... 至少在第 3 个小数位 ... 。交换 ... 与 q_{23} ... ；将 ... 记为 $\langle q_{31}, q_{32}, q_{33}, \ldots \rangle$，它等同于 $\langle \boldsymbol{q_{11}}, \boldsymbol{q_{22}}, q_{33}, \ldots \rangle$。

继续，遍历 q 的小数部分。

旁注：对于每个 n 做且只做一次这样的交换。我们总能为每次位置交换找到需要的后项：在十进制系统中，每个有理数有至少一个无限循环小数的表示。在二进制系统中也是这样。而且，一个 n 进制系统的计数能力与选取怎样的 n 个符号来表示无关，于是我们能用 *0* 和 *1* 来建立一个二进制系统，也可以用 *5* 和 *6*（作为两个符号，与它们的常规数值无关）来建立。在每个二进制系统中，每个落在 $(0, 1)$ 区间的有理数有至少一个无限循环小数的表示，这种表示的小数部分和某个落在 $(0, 1)$ 区间的十进制有理数的小数部分在外观上是一模一样的。于是在每种二进制系统中都有无穷多个小数，它们看起就像是落在 $(0, 1)$ 区间的十进制有理数。因此，每次为交换位置而做的搜寻都不至落空 —（在现时数列的任何位置之后）至少仍有无限多的选择存在于那些貌似某种二进制数又不含需要回避的数字的数中。

过程中产生的数列彼此不同，因为步骤 n 和步骤 $(n+k)$ $(n, k = 1, 2, 3, \ldots)$ 的结果至少在第 $(n+1)$ 项上是不同的。将所有各步的结果依次列写成表B:

$$
\begin{array}{l}
\langle\ \boldsymbol{q_{11}},\ \ q_{12},\ \ q_{13},\ \ q_{14},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ q_{23},\ \ q_{24},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ \boldsymbol{q_{33}},\ \ q_{34},\ \ \cdots\ \rangle\ ,\\
\langle\ \boldsymbol{q_{11}},\ \ \boldsymbol{q_{22}},\ \ \boldsymbol{q_{33}},\ \ \boldsymbol{q_{44}},\ \ \cdots\ \rangle\ ,\\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\ 。
\end{array}
$$

表B的各项都是数列 $\langle q_{01}, q_{02}, q_{03}, \ldots \rangle$ 的重新排列。所有数列的各项整齐排列成一个矩阵，矩阵被对角线 $(q_{11}, q_{22}, q_{33}, \ldots)$ 一分为二。这个对角线覆盖了所有列吗？我们把对角线及其左下的部分记为 L，注意到 $\{$出现在 L 中的项$\} = \{q_{11}, q_{22}, q_{33}, \ldots\}$。因为 q 的第 n 位小数与 q_{nn} 的不同，所以 q 不在 L 中。于是出现在矩阵的每一行中的 q 总在 L 之外。这说明包含了所谓对角线的 L 区域并未涉及到所有列。

B 引理 3.1 的证明

证明：（名词弧接下来专指劣弧。）

1. 由于线段 $[0, P_1)$ 与 $[0, C)$ 不可公度，光束不可能经过有限步反射重回某亮点。

2. 只需证明对于任何一段给定的弧，无论它多短或位于何处，总有亮点落在其上。

我们将给定的弧的弧长记为 $\varepsilon\,(0 < \varepsilon \leq C/2)$。令 k 为不超过 C/ε 的最大整数，于是有 $(k+1) > C/\varepsilon$。前 $k+1$ 个亮点将圆周分成 $k+1$ 段两两不相交的弧。由于 $(k+1)\varepsilon > C$，这 $k+1$ 段弧中至少有一段的弧长小于 ε。对于这样一个短段，我们将其弧长记为 $S\,(0 < S < \varepsilon)$。对于决定这段弧的两个亮点，我们将其序号的差值记为 $i\,(i \in \mathbb{N}，0 < i \leq k)$。

令 m 为不超过 C/S 的最大整数。由于两个亮点之间的弧距取决于它们两个序号之间的差值，点 P_0、P_i、P_{2i}、...、P_{mi} 这 $m+1$ 个点（忽略其它点）将圆周分成 $m+1$ 段两两不相交的弧，其中 m 段具有弧长 $S\,(0 < S < \varepsilon)$，剩下的一段则更短。这意味着在圆周上任意给定的弧长为 ε 的片段不可能同时避开这 $m+1$ 个亮点。 □

C 思维实验 3.2 的一个数值实例

一个以 $\sqrt{2}$ 为极限的严格递增有理数序列：7/5、41/29、239/169、... 具有通项式 $\sqrt{2b_n^2-1}/b_n$，其中 $b_1 = 5$、$b_{n+1} = 3b_n + 2\sqrt{2b_n^2-1}$，这里 $n = 1，2，3，...$。
还有一个以 $\sqrt{2}$ 为极限的严格递减有理数序列：3/2、17/12、99/70、... 具有通项式 $\sqrt{2c_n^2+1}/c_n$，其中 $c_1 = 2$、$c_{n+1} = 3c_n + 2\sqrt{2c_n^2+1}$，这里 $n = 1，2，3，...$。

参考文献

[1] M. Kline, *Mathematical Thought from Ancient to Modern Times*, Vol. 3, Oxford Univ. Press, 1990. (First published in 1972, in one volume.)

[2] M. Kline, *Mathematics: The Loss of Certainty*, Oxford Univ. Press, New York, 1980.

[3] *Ross–Littlewood paradox.* https://handwiki.org/wiki/Ross–Littlewood_paradox

[4] *Georg Cantor.* https://en.wikiquote.org/wiki/Georg_Cantor

[5] *Hilbert's paradox of the Grand Hotel.* https://wikimili.com/en/Hilbert's_paradox_of_the_Grand_Hotel

[6] Aristotle, *Physics*, Written 350 B.C.E., Translated by R.P. Hardie and R.K. Gaye. http://classics.mit.edu/Aristotle/physics.html

[7] Abraham Robinson, *Selected papers*, Vol. 2, Edited by W.A.J. Luxemburg, S. Körner, North Holland, Amsterdam, 1979.