集合的基数

离散数学

马晓星

南京大学·计算机科学与技术系

提要

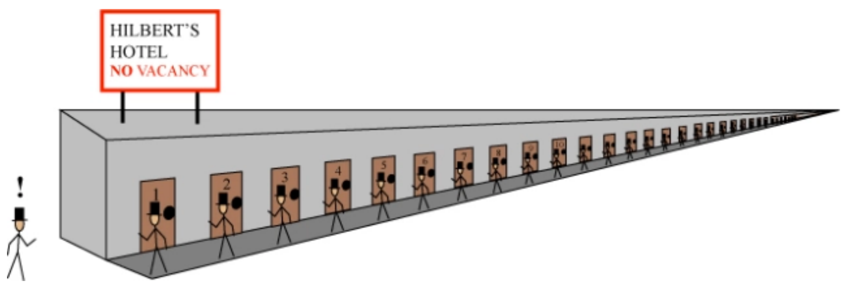
- 集合的大小与等势关系
- 不可数集与Cantor定理
- 优势关系与Bernstein定理
- 基数作为自然数的扩展

集合的“大小”

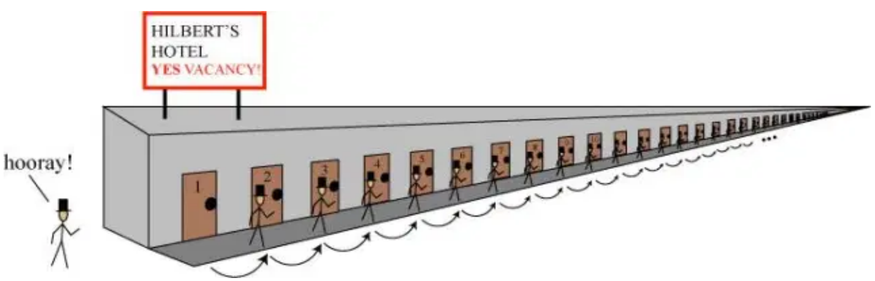
- 集合A的大小, 称为基数(**Cardinality**) ,
  - 记作 |A|, 或者 **card** A.
  - “数得清”的我们就数元素个数
  - “数不清”的咋办?
    - “常识”不一定靠谱

反直觉的无限

Hilbert's Grand Hotel



啊?客满啦?
没关系，我让现在住在 k 号房间 的客人移到 k+1 号. 你就住进第 1号房间吧!



再来一百个、一千个客人?
再来“无限”个?

“整体多于部分”不一定成立.

野人的智慧

大足野人生了一堆儿女，他想知道是女儿多还是儿子多。但他只会数到3...

康托尔老先生的两个集合里都有许多的元素，他想知道哪个大。但他只会数自然数...

等势关系

- **等势(Equipotent)**: 如果存在从集合A到B的双射，则称集合A与B等势.
  - 记为: A≈B. 不等势记 A≉B.
  - 等势的集合被认为“一样大”, 即|A|=|B|, 否则|A|≠|B|.
  - 意味着：A，B中的元素可以“一一对应”.
  - 要证明A≈B，只要找出一个从A到B的双射。

- 等势概念是对“靠数数判断大小一致”的推广
  - 集合A“数得清”意味着存在自然数n, 使得 A≈ n

有限集与无限集

- 称S是**有限集**, 若存在自然数n与之等势. 否则称其为**无限集**或无穷集.

- S是无限集 *iff* 存在S的真子集S'使得S与S’ 等势

⇒ S一定包含一个与自然数集合等势的子集M={$a_0, a_1, a_2, \ldots$}

令S' =S-{a_0}, 可以定义 f:S→S'如下:

对于任意 $a_i \in$M, $f(a_i) = a_i+1$;

对于任意 $x \in$S-M, $f(x) = x$.

显然这是双射，即S与其真子集S'等势。

⇐ 若S是有限集, 令|S|=n, 则对S的任意真子集S', 若|S'|=m,

必有m<n, 因此从S'到S的任一单射不可能是满射。

可数集

- 一个集合称为可数的(**countable**, 也称可列的), 如果它与自然数集**N**的某个子集等势.
  - 换言之, 存在它到**N**的一个单射.
  - 与自然数集**N**等势的集合称为可数无限集.
  - 可数集要么是有限集, 要么是可数无限集.

  - 直观上说，集合的元素可以按确定的顺序线性排列，所谓“确定的”顺序是指对序列中任一元素，可以说出它“前”、“后”元素是什么。

可数集

- 整数集(包括负数)与自然数集等势:

0, -1, 1, -2, 2, -3, 3, -4, ……

0, 1, 2, 3, 4, 5, 6, 7, ……

$$g(n)=\begin{cases} 2n & n>=0 \\ -2n-1 & n<0 \end{cases}$$

这个技巧可以很容易地推广，用以证明有限个可数集的并集仍然是可数的. 但是无限个可数集的并集呢?

可数个可数集的并集仍然可数

- 所有的自然数对构成的集合可数:

<0,0>	<1,0>	<2,0>	<3,0>	<4,0>	•••
<0,1>	<1,1>	<2,1>	<3,1>	•••	
<0,2>	<1,2>	<2,2>	•••		
<0,3>	<1,3>	•••			
<0,4>					

<0,0>, <0,1>, <1,0>, <0,2>, <1,1>, <2,0>, <0,3>, ...

$$l(m,n)=\frac{1}{2}\sum_{i=1}^{m+n} i+(m+1)=\frac{(m+n)(m+n+1)}{2}+(m+1)$$

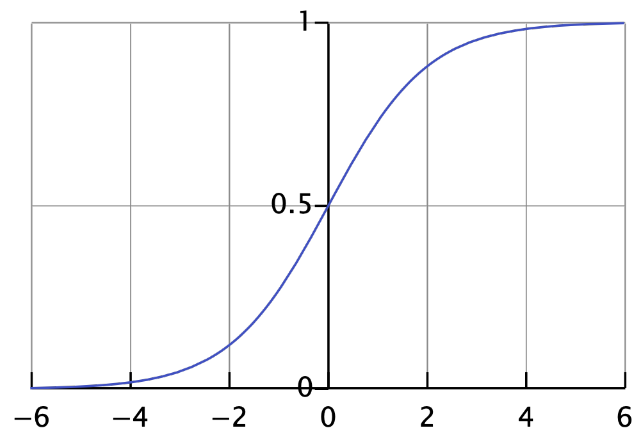
证明无限集等势的例子

- (0,1)与整个实数集等势
  - 双射:

$$f:(0,1)\to\mathbf{R}: f(x)=tan(\pi x-\frac{\pi}{2})$$

还有很多 sigmoid functions, 例如

$$S(x)=\frac{1}{1+e^{-x}}=\frac{e^x}{e^x+1}$$



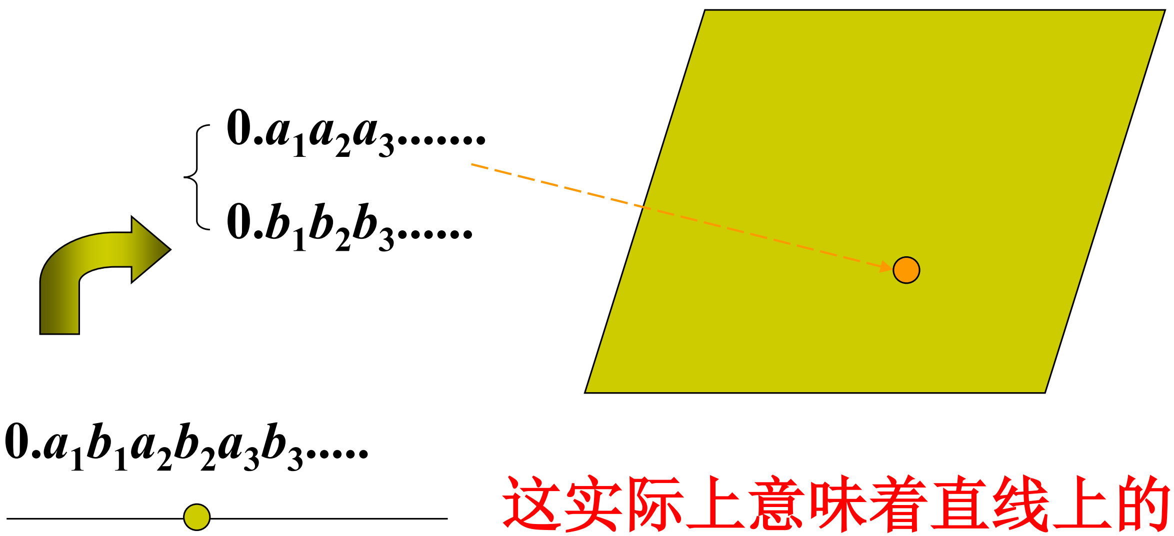
证明无限集等势的例子

- 对任意不相等的实数$a,b(a<b)$, $[0,1]$与$[a,b]$等势
  - 双射:

$$f:[0,1]\to[a,b]:f(x)=(b-a)x+a$$

(这实际上意味着:任意长的线段与任意短的线段等势)

直线上的点与平面上的点一样多



这实际上意味着直线上的点与任意有限维空间的点“一样多”！

那么, 是否所有无限集合都等势?

实数集合不可数

- (0,1)不是可数集　//注意:(0,1)与实数集合等势
  - “对角线”证明法

    假设(0,1)中的元素可以线性排列:

$$
\begin{array}{l}
\mathbf{0.}\color{red}{\mathbf{b_{11}}}\mathbf{b_{12}b_{13}b_{14}\ldots} \\
\mathbf{0.b_{21}}\color{red}{\mathbf{b_{22}}}\mathbf{b_{23}b_{24}\ldots} \\
\mathbf{0.b_{31}b_{32}}\color{red}{\mathbf{b_{33}}}\mathbf{b_{34}\ldots} \\
\mathbf{0.b_{41}b_{42}b_{43}}\color{red}{\mathbf{b_{44}}}\mathbf{\ldots} \\
\vdots
\end{array}
$$

实数集合的基数记为 $|\mathbf{R}| = \mathfrak{c}$
(mathematical fraktur c)

则可构造 $\mathbf{0.b_1b_2b_3b_4\ldots}$ $(\mathbf{b_i}\neq\color{red}{\mathbf{b_{ii}}})$, 它不在这个序列中.

康托尔定理(Cantor's theorem)

- 任何集合与其幂集不等势, 即: A≠ρ(A)

证明要点:

设 g 是从A到ρ(A)的函数,

构造集合 $B=\{x\in A \mid x\notin g(x)\}$, 则B∈ρ(A),

但不可能存在 $x\in A$，能满足 $g(x)=B$，因为，

如果有这样的 x_0, 则 $x_0\in B \leftrightarrow x_0\notin B$。

因此, g 不可能是满射. ■

右例: f: **N**→ ρ(**N**); T={n∈**N** | n∉f(**N**)}
是一个对角线证明.

$$
\begin{array}{lcllllllllllll}
f(1) & = \{ & 1, & & & & & & & & & & & \} \\
f(2) & = \{ & 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & 11, & \ldots \} \\
f(3) & = \{ & & 2, & 3, & 4, & & 6, & & 8, & & 10, & & \ldots \} \\
f(4) & = \{ & 1, & & 3, & 4, & 5, & & 7, & & 9, & & 11, & \ldots \} \\
f(5) & = \{ & 1, & 2, & & 4, & 5, & 6, & 7, & & 9, & & 11, & \ldots \} \\
f(6) & = \{ & & & 3, & 4, & & 6, & 7, & & 9, & 10, & & \ldots \} \\
f(7) & = \{ & 1, & & & & 5, & & 7, & & 9, & & & \ldots \} \\
f(8) & = \{ & & & 3, & 4, & & & 7, & 8, & & & 11, & \ldots \} \\
f(9) & = \{ & 1, & 2, & & & 5, & 6, & & & 9, & 10, & & \ldots \} \\
f(10) & = \{ & 1, & 2, & & 4, & 5, & 6, & & & 9, & 10, & 11, & \ldots \} \\
f(11) & = \{ & 1, & 2, & & 4, & & 6, & & & 9, & & 11, & \ldots \} \\
\vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}
$$

$$
T = \{ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, \ldots \}
$$

存在不可计算的函数

- 可计算函数

  简单地说就是存在一个计算机程序可以计算这个函数

- 函数集合 $\mathbf{N}^{\mathbf{N}} = \{f | f : \mathbf{N} \to \mathbf{N}\}$ 就不可数.
  - 可用对角线证明法.

- 但是不同的程序的个数是可数的.
  - 长度有限的string的个数是可数的.

既然有不等势的无限集合，

如何比较其相对“大小”？

优势关系

- 如果存在从集合A到集合B的单射，则称“集合B**优势于**集合A”，记为 $|A| \leq |B|$ 或 $A \preccurlyeq \cdot B$

- 如果集合B优势于集合A，且B与A**不等势**，则称“集合B**真优势于**集合A”，记为$|A| \leq |B|$或$A \prec \cdot B$

- 实数集真优势于自然数集

- 对任意集合A，A的幂集真优势于集合A

集合优势关系的性质

- 根据单射的性质, 可以得到
  - 对于任意集合A, 有$|A| = |A|$ (自反性)
    - 显然.
  - 若$|A| = |B|$ 且 $|B| = |C|$, 则 $|A| = |C|$ (传递性)
    - 单射的复合仍然是单射.
  - 若$|A| \leq |B|$ 且 $|B| \leq |A|$, 则 $|A| = |B|$ (反对称性)
    - 这个需要仔细证明: Bernstein定理

Cantor-Schröder-Bernstein Theorem

- 定理: 若 |A|≤|B| 且 |B|≤|A|, 那么 |A|=|B|.

证明要点: 可设 f:A→B 和 g:B→A 为单射, 我们将构造双射 h:A→B .

考虑反复应用这两个函数, 及其逆(部分函数) 所得的链

$$\cdots \to f^{-1}(g^{-1}(x)) \to g^{-1}(x) \to x \to f(x) \to g(f(x)) \to \cdots$$

该链条往右无限延伸. 但我们考虑往左的各种情况:

(1) 链条进入循环;

(2) 链条无限延伸不循环;

(3) 链条停在A中, 即到达某个x但$g^{-1}(x)$无定义;

(4) 链条停在B中.

同时, 请注意, 每个A或B中的元素都在唯一的链条中.

续上页

构造h: $h(x) = f(x)$ 如果x在(1),(2),(3)种的链中, 否则 $h(x) = g^{-1}(x)$.

注意, 在第(4)种情况下, 是$g^{-1}(x)$有定义的, 否则是第(3)种情况.

现证明h是双射:

$h(x)$是**单射**: 假设$h(x_1) =h(x_2)$. 注意此时x_1和x_2必在同一个链条中.

若在(1),(2),(3)种的链中, 则 $f(x_1) =h(x_1)=h(x_2) =f(x_2)$

又f是单射, $x_1 = x_2$

若在第(4)种链中, 则 $h(x_1) = g^{-1}(x_1)$, $h(x_2) = g^{-1}(x_2)$即

$$x_1 = g(h(x_1))=g(h(x_2))= x_2$$

$h(x)$是**满射**: 对于任意的$y\in$B, 考虑包含y的链条.

若y在(1),(2),(3)种的链中, 则必有某个x使得 $f(x)=y$,

又x和y在同一个链条中, 有$h(x) = f(x) =y$

若在第(4)种链中, 则$h(g(y))=y$.

于是h是双射, |A|=|B|. ■

优势的反对称性用于证明等势

- 证明实数集的两个子集(0,1)和[0,1]等势。直接找双射不太容易

关键是如何安排在[0,1]中但不在(0,1)中的0和1。

想象那个“宇宙旅馆”。我们可以取(0,1)的一个与自然数集合等势的子集(一定有)$\{a_1,a_2,a_3,\ldots\}$,“腾出”前两个位置安排0和1

一种证法：
$$f(x)=\begin{cases}\dfrac{1}{2} & x=0\\[2mm] \dfrac{1}{2^2} & x=1\\[2mm] \dfrac{1}{2^{n+2}} & x=\dfrac{1}{2^n},\ n=1,2,3\ldots\\[2mm] x & x\text{为其它值}\end{cases}$$

麻烦!

优势的反对称性用于证明等势

- 证明实数集的两个子集(0,1)和[0,1]等势。

  分别找两个一对一的映射往往比找一个双射容易

$$f:(0,1)\rightarrow[0,1]: f(x)=x$$

$$g:[0,1]\rightarrow(0,1): g(x)=\frac{1}{2}+\frac{x}{4} \quad \text{注意：} g([0,1])=[\frac{1}{2},\frac{3}{4}]$$

实数集与ρ(N)等势

- $[0,1)\approx\{0,1\}^{\mathbf{N}}$ 从而 $\mathbf{R}\approx\rho(\mathbf{N})$
  - $[0,1)$中的数唯一地表示为$0.b_1b_2b_3b_4\ldots$

    不容许连续无数个1，比如1/2=0. 1000... (不用0.0111...)
  - f: $[0,1)\to\{0,1\}^{\mathbf{N}}$ $f(0.b_1b_2b_3b_4\ldots) = b_{1,} b_{2,} b_{3,} b_4\ldots$ 是一个单射.
  - g: $\{0,1\}^{\mathbf{N}}\to[0,1)$ $g(b_{1,} b_{2,} b_{3,} b_4\ldots) = 0.\ b_1b_2b_3b_4\ldots$ 是一个单射 // $0.\ b_1b_2b_3b_4\ldots$ 看作10进制
  - 根据Bernstein定理，得证

回顾：冯·诺伊曼的自然数

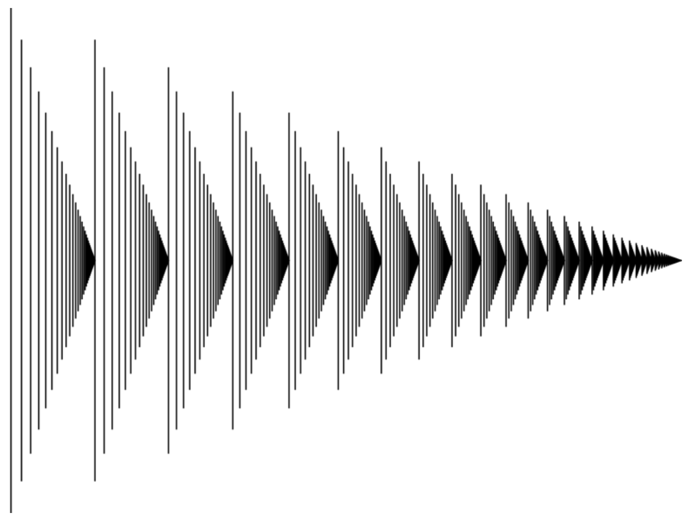
- 自然数集合 **N**：

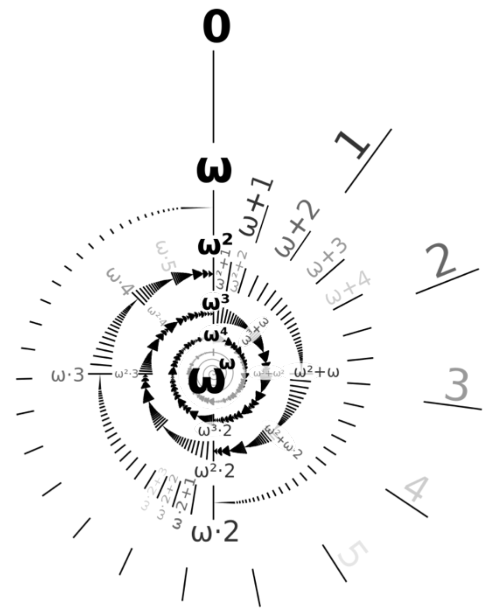
$$0 = \emptyset$$
$$1 = S(0) = S(\emptyset) = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\}$$
$$2 = S(1) = S(\{0\}) = \{0\} \cup \{\{0\}\} = \{0, \{0\}\} = \{0, 1\}$$
$$3 = S(2) = S(\{0, 1\}) = \{0, 1\} \cup \{\{0, 1\}\} = \{0, 1, \{0, 1\}\} = \{0, 1, 2\}$$
$$\cdots$$

0, 1, 2, 3, … ω, ω+1, ω+2, ω+3, … 2 ω, ω+1, ω+2, ω+3, …

自然数的序数扩展

序数





0, 1, 2, 3, … ω, ω+1, ω+2, ω+3, … 2 ω, ω+1, ω+2, ω+3, …

自然数的序数扩展(ordinal numbers)

有兴趣自行了解，
本课程不做要求

基数作为自然数的扩展

- 上面的超穷序数不能用作基数
  - 例如, 0, 1, 2, 3, … ω, ω+1, ω+2, ω+3, … 2ω, 2ω+1, 2ω+2, 2ω+3, …
    显然还是可数的 (换个排队的方法即可)

- Cantor 引入 $\aleph_\alpha$来标记越来越大的超穷基数(cardinal numbers)
  (这里α是序数)
  - 0, 1, 2, 3, ... 是有限集合的基数
  - $\aleph_0$是可数集合的基数
  - $\aleph_1$是下一个更大的基数 (所有可数序数的集合的基数, 因而不可数)
  - $\aleph_2$ ……

有兴趣自行了解，
本课程不做要求

连续统假设

- 不存在一个集合, 其基数处于$\aleph_0$和$\mathfrak{c}$之间

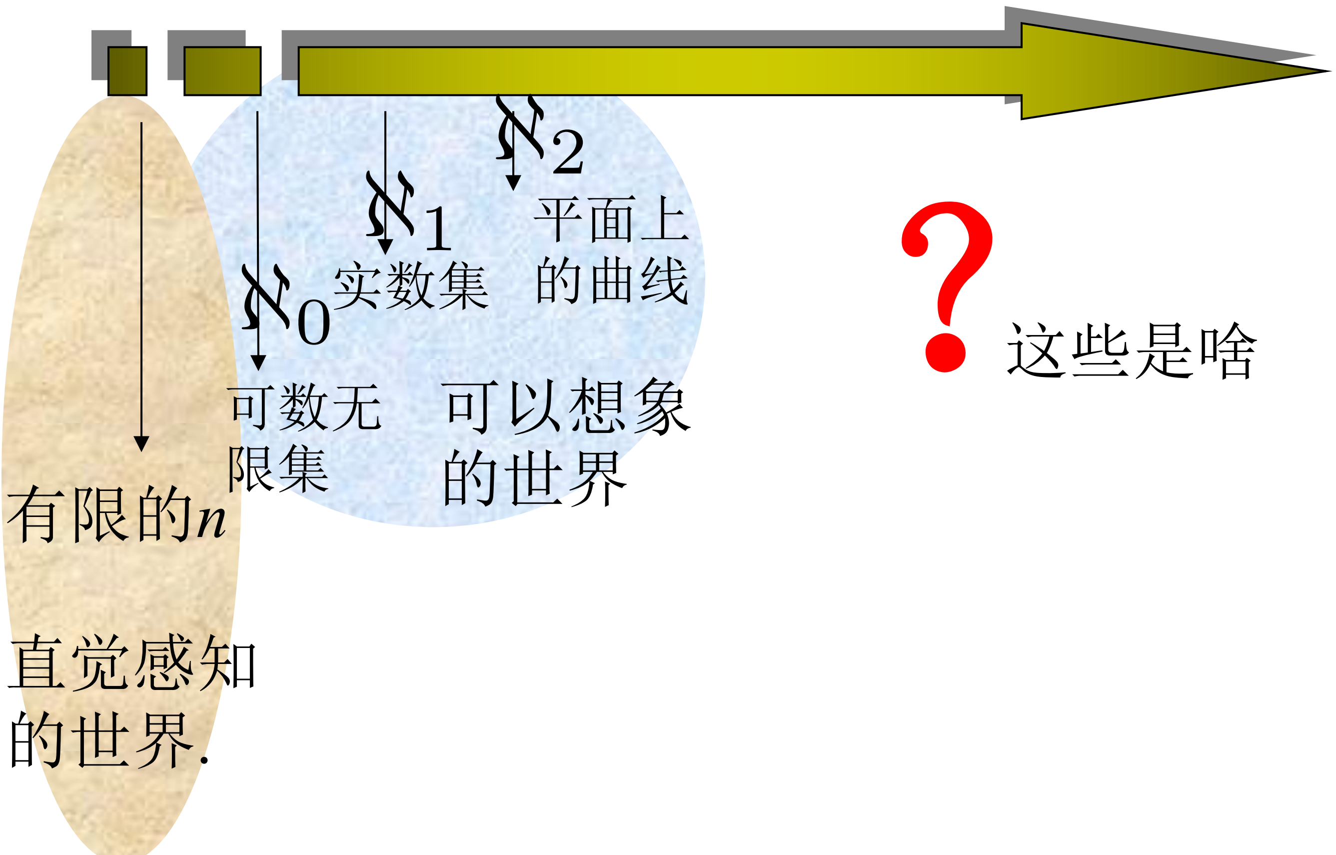
$$\aleph_0 < |A| < \mathfrak{c} \quad ???$$

其中 $\mathfrak{c}$ 是实数集的基数.

有趣的事实: 在ZFC下这既不能证实也不能被证伪.

如果我们承认它, 则$\aleph_1 = \mathfrak{c} = 2^{\aleph_0}$.

基数作为自然数的扩展



$\aleph_2$
平面上的曲线
$\aleph_1$
实数集
$\aleph_0$
可数无限集
可以想象的世界
有限的n
直觉感知的世界.
?
这些是啥

小结

- 无限集比大小的工具: 等势与优势
- 有限集、无限集、可数集、不可数集
- Cantor定理, Bernstein定理
- 基数与序数在超穷情况下不一致