

Equivalence Relations and Functions

October 15, 2013

Week 13-14

1 Equivalence Relation

A **relation** on a set X is a subset of the Cartesian product $X \times X$. Whenever $(x, y) \in R$ we write xRy , and say that x is related to y by R . For $(x, y) \notin R$, we write $x \not R y$.

Definition 1. A relation R on a set X is said to be an **equivalence relation** if

- (a) xRx for all $x \in X$ (**reflexive**).
- (b) If xRy , then yRx (**symmetric**).
- (c) If xRy and yRz , then xRz (**transitive**).

Let X be a set and R an equivalence relation on X . It is quite common to denote the equivalence relation R by \sim if there is only one equivalence relation to be considered. So xRy becomes $x \sim y$. For each $x \in X$ we define

$$[x] = \{y \in X \mid y \sim x\},$$

called an **equivalence class** of \sim ; each element of $[x]$ is called a **representative** of the class $[x]$. The collection all equivalence classes of \sim is called the **quotient set** of X modulo \sim , denoted X/\sim .

Definition 2. A **partition** of set X is a collection $\mathcal{P} = \{A_1, \dots, A_k\}$ of disjoint nonempty subsets of X such that $X = \bigcup_{i=1}^k A_i$.

Proposition 3. Let X be a set. If \sim is an equivalence relation on X , then the collection $\{[x] : x \in X\}$ of equivalence classes of \sim is a partition of X .

Conversely, if $\mathcal{P} = \{A_1, \dots, A_k\}$ be a partition of X , then the relation $R = \bigcup_{i=1}^k A_i^2$ is an equivalence relation on X .

Example 1. For any positive integer m , the congruence relation modulo m is an equivalence relation on the set \mathbb{Z} of integers. The quotient set of \mathbb{Z} modulo m is

$$\mathbb{Z}/\sim = \{[0], [1], \dots, [m-1]\}.$$

Example 2. Let \sim be a relation on \mathbb{R}^2 , defined by $(s, t) \sim (u, v)$ if $v - t = 2(u - s)$.

Example 3. Let $f : X \rightarrow Y$ be a surjective function. Define a relation \sim on X by $x_1 \sim x_2$ if $f(x_1) = f(x_2)$. Then \sim is an equivalence relation on X . Moreover, the function

$$\tilde{f} : X/\sim \longrightarrow Y, \quad [x] \mapsto f(x)$$

is a bijection. For instance, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $f(x, y) = \sqrt{x^2 + y^2}$.

2 Function and Inverse Function

Definition 4. Let X and Y be sets. A **function** from X to Y is a rule f that assigns each element x of X a single element y in Y . We write the rule as $f : X \rightarrow Y$, and say that x is sent to y ; we also write $x \mapsto y$ or $y = f(x)$. Functions are also known as **maps** or **mappings**.

Definition 5. Let X and Y be nonempty sets. A function $f : X \rightarrow Y$ is said to be

- **injective** if distinct elements are sent to distinct elements;
- **surjective** if for every element $y \in Y$, there is an element $x \in X$ such that $f(x) = y$;
- **bijective** if f is both injective and surjective. If so it is called a **one-to-one correspondence**.

Definition 6. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions. The **composition** of f and g is a function $g \circ f : X \rightarrow Z$ defined by

$$(g \circ f)(x) = g(f(x)) \quad \text{for all } x \in X.$$

Proposition 7. Let $f : X \rightarrow Y$, $g : Y \rightarrow Z$, $h : Z \rightarrow W$ be functions. Then the composition functions $h \circ (g \circ f)$ and $(h \circ g) \circ f$ are the same.

Proof. For each $x \in X$,

$$(h \circ (g \circ f))(x) = h((g \circ f)(x)) = h(g(f(x))) = (h \circ g)(f(x)) = ((h \circ g) \circ f)(x).$$

□

Definition 8. A function $f : X \rightarrow Y$ is said to be **invertible** if there is a function $g : Y \rightarrow X$ such that

$$(g \circ f)(x) = x \quad \text{for all } x \in X,$$

$$(f \circ g)(y) = y \quad \text{for all } y \in Y.$$

If so, the function g is called an **inverse** of f , written $g = f^{-1}$.

Proposition 9. Let $f : X \rightarrow Y$ be a function.

(a) If f is invertible, then its inverse is unique.

(b) The function f is invertible if and only if f is a bijection.

Proof. (a) Let g_1, g_2 be functions $g : Y \rightarrow X$ such that $(g_i \circ f)(x) = x$ for all $x \in X$ and $(f \circ g_i)(y) = y$ for all $y \in Y$, $i = 1, 2$. Since $y = f(g_2(y))$ and $g_1(f(x)) = x$, then

$$g_1(y) = g_1(f(g_2(y))) = g_2(y) \quad \text{for all } y \in Y.$$

(b) Let g be the inverse of f . For $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$, then $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. So f is injective. For any $y \in Y$, the element $g(y)$ is sent to y , in fact, $f(g(y)) = y$ by definition of inverse. □

Example 4. (a) $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^2$ is neither injective nor surjective.

(b) $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x^3$ is bijective; its inverse $g : \mathbb{R} \rightarrow \mathbb{R}$ is given by $g(x) = \sqrt[3]{x}$.

(c) $f : \mathbb{R} \rightarrow \mathbb{R}_+$ by $f(x) = e^x$ is bijective, where $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$; and its inverse is $f^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}$ is given by $f^{-1}(x) = \ln x$.

(d) $f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ by $f(x) = x^2$ is surjective, where $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} \mid x \geq 0\}$.

(e) $f_1 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ by $f(x) = x^2$ is bijective; its inverse $f^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}$ is given by $f^{-1}(x) = \sqrt{x}$.

(f) $f : \mathbb{R}_{\leq 0} \rightarrow \mathbb{R}_{\geq 0}$ by $f(x) = x^2$ bijective, where $\mathbb{R}_{\leq 0} = \{x \in \mathbb{R} \mid x \leq 0\}$; its inverse $f^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}$ is given by $f^{-1}(x) = -\sqrt{x}$.

- Example 5.** (a) The function $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(\theta) = \sin \theta$ is neither injective nor surjective. However, the function $g : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ is bijective; its inverse $g^{-1} : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ is denoted by $g^{-1}(x) = \arcsin x$.
- (b) The function $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(\theta) = \cos \theta$ is neither injective nor surjective. However, the function $g : [0, \pi] \rightarrow [-1, 1]$ is bijective; its inverse $g^{-1} : [-1, 1] \rightarrow [0, \pi]$ is denoted by $g^{-1}(x) = \arccos x$.
- (c) The function $f : [\frac{\pi}{2}, \frac{3\pi}{2}] \rightarrow [-1, 1]$ by $f(\theta) = \sin \theta$ is bijective. Its inverse $f^{-1} : [-1, 1] \rightarrow [\frac{\pi}{2}, \frac{3\pi}{2}]$ is given by $f^{-1}(x) = \frac{\pi}{2} + \arccos x$.
- (d) The function $f : \bigcup_{n=-\infty}^{\infty} (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ by $f(\theta) = \tan \theta$ is neither injective nor surjective.
- (e) The function $f : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ by $f(\theta) = \tan \theta$ is bijective; its inverse $f^{-1} : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ is denoted by $f^{-1}(x) = \arctan x$. The function $f_1 : (\frac{\pi}{2}, \frac{3\pi}{2}) \rightarrow \mathbb{R}$ by $f(\theta) = \tan \theta$ is also bijective; its inverse $f_1^{-1} : \mathbb{R} \rightarrow (\frac{\pi}{2}, \frac{3\pi}{2})$ is given by $f_1^{-1}(x) = \pi + \arctan x$.

Example 6. The complex exponential function $f : \mathbb{C} \rightarrow \mathbb{C}$ is defined by $f(z) = e^z$, where if z is written as $z = x + iy$ then

$$e^z = e^{x+iy} := e^x e^{iy} = e^x (\cos y + i \sin y).$$

The function f is injective but not surjective.

Example 7. Hyperbolic functions

(a) The **hyperbolic sine function** is defined as

$$\sinh : \mathbb{R} \rightarrow \mathbb{R}, \quad \sinh(x) = \frac{e^x - e^{-x}}{2}$$

is bijective. Its inverse is the function

$$\operatorname{arsinh} : \mathbb{R} \rightarrow \mathbb{R}, \quad \operatorname{arsinh}(x) = \ln(x + \sqrt{x^2 + 1}).$$

Let $y = \frac{1}{2}(e^x - e^{-x})$. Then $e^{2x} - 2ye^x - 1 = 0$. Thus $e^x = \frac{1}{2}(2y \pm \sqrt{4y^2 + 4}) = y \pm \sqrt{y^2 + 1}$. Since $\sqrt{y^2 + 1} > y$ and $e^x > 0$, we must have $e^x = y + \sqrt{y^2 + 1}$. Hence

$$x = \ln(y + \sqrt{y^2 + 1}).$$

(b) The **hyperbolic cosine function**

$$\cosh : \mathbb{R} \rightarrow \mathbb{R}, \quad \cosh(x) = \frac{e^x + e^{-x}}{2}$$

is neither injective nor surjective. It has no inverse function! However, the restriction

$$f : \mathbb{R}_{\geq 0} \rightarrow [1, \infty), \quad f(x) = \cosh(x)$$

is bijective and its inverse function is given by

$$f^{-1} : [1, \infty) \rightarrow \mathbb{R}, \quad f^{-1}(x) = \operatorname{arcosh}(x) = \ln(x + \sqrt{x^2 - 1}).$$

Let $y = \frac{1}{2}(e^x + e^{-x})$. Then $e^{2x} - 2ye^x + 1 = 0$. Thus $e^x = \frac{1}{2}(2y \pm \sqrt{4y^2 - 4}) = y \pm \sqrt{y^2 - 1}$. Since e^x goes to ∞ when y goes to ∞ and $\sqrt{y+1} > \sqrt{y-1}$, then $0 < y - \sqrt{y^2 - 1} = y - \sqrt{(y+1)(y-1)} < y - (y-1) = 1$, we must have $e^x = y + \sqrt{y^2 - 1}$. Hence

$$x = \ln(y + \sqrt{y^2 - 1}).$$

The function $g : \mathbb{R}_{\leq 0} \rightarrow [1, \infty)$ by $g(x) = \cosh(x)$ is also bijective. Its inverse function is given by

$$g^{-1} : [1, \infty) \rightarrow \mathbb{R}_{\leq 0}, \quad g^{-1}(x) = -\ln(x + \sqrt{x^2 - 1}).$$

(c) The **hyperbolic tangent function**

$$\tanh : \mathbb{R} \rightarrow (-1, 1), \quad \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

is bijective. Its inverse function is

$$\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}, \quad \operatorname{artanh}(x) = \frac{1}{2} \ln \frac{1+x}{1-x}.$$

Let $y = \frac{e^{2x}-1}{e^{2x}+1}$. Then $(1-y)e^{2x} = 1+y$, i.e., $e^{2x} = \frac{1+y}{1-y}$. Thus $x = \frac{1}{2} \ln \frac{1+y}{1-y}$.

(d) The **hyperbolic cotangent function**

$$\operatorname{coth} : (-\infty, 0) \cup (0, \infty) \rightarrow (-\infty, -1) \cup (1, \infty),$$

$$\operatorname{coth}(x) = \frac{\cosh x}{\sinh x} = \frac{e^x + e^{-x}}{e^x - e^{-x}} = \frac{e^{2x} + 1}{e^{2x} - 1}$$

is bijective. Its inverse function is

$$\operatorname{artanh} : \{x \in \mathbb{R} : |x| > 1\} \rightarrow \{x \in \mathbb{R} : |x| > 0\}, \quad \operatorname{arcoth}(x) = \frac{1}{2} \ln \frac{x+1}{x-1}.$$

Proposition 10. *Let $f : X \rightarrow Y$ be a function from a finite set X to a finite set Y .*

(a) If f is injective, then $|X| \leq |Y|$.

(b) If f is surjective, then $|X| \geq |Y|$.

(c) If f is bijective, then $|X| = |Y|$.

Proposition 11. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions.

(a) If f and g are both injective, so is $g \circ f$.

(b) If f and g are both surjective, so is $g \circ f$.

(c) If f and g are both bijective, so is $g \circ f$.

Proposition 12. Let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be functions.

(a) If $g \circ f$ is injective, then f must be injective.

(b) If $g \circ f$ is surjective, then g must be surjective.

(c) If $g \circ f$ is bijective, then f is injective and g is surjective.

Proposition 13. Let X be a finite set and $f : X \rightarrow X$ be a function.

(a) If f is injective, then f must be surjective.

(b) If f is surjective, then f must be injective.

Example 8. For each $a \in \mathbb{Z}_n$, define $f_a : \mathbb{Z}_n \rightarrow \mathbb{Z}_n$ by $f_a(x) = ax \pmod n$. Then f_a is invertible if and only if $\gcd(a, n) = 1$.

3 Infinity

Definition 14. Two sets X and Y are said to be **equivalent** if there is a one-to-one correspondence $f : X \rightarrow Y$; written $X \sim Y$. Then \sim is an equivalence relation. When X and Y are finite and equivalent, we say that X and Y have the same **cardinality**.

Definition 15. A set A is said to be **countable** if it is equivalent to the set $\mathbb{Z}_+ = \{1, 2, \dots\}$ of positive integers. In other words, A is countable if it is an infinite set, all whose elements can be listed as $\{a_1, a_2, a_3, \dots\}$. An infinite set that is not countable is said to be **uncountable**.

Proposition 16. Every infinite set contains a countable subset.

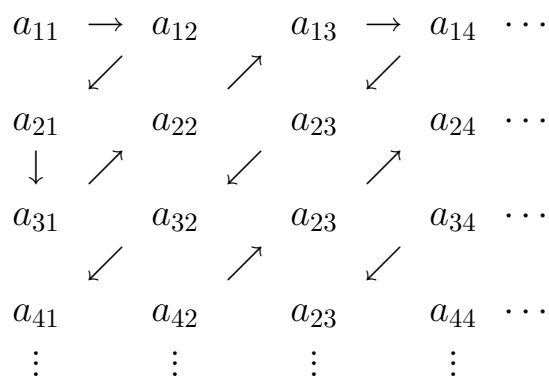
Proof. Let A be an infinite set. Select an element from A , say a_1 . Since A is infinite, one can select an element from A other than a_1 , say a_2 . Similarly, one can select an element a_3 from A other than both a_1 and a_2 . Since the infinity of A , one can continue this procedure by selecting a sequence of elements one after the other to get an infinite countable subset $\{a_1, a_2, a_3, \dots\}$. \square

Theorem 17. *If A and B are countable subsets, then $A \cup B$ is countable.*

Proof. It is obviously true when one of A and B is a finite set. Let $A = \{a_1, a_2, \dots\}$ and $B = \{b_1, b_2, \dots\}$ be countable infinite sets. If $A \cap B = \emptyset$, then $A \cup B = \{a_1, b_1, a_2, b_2, \dots\}$ is countable as demonstrated. If $A \cap B \neq \emptyset$, we just need to delete the elements that appeared more than once in the sequence $a_1, b_1, a_2, b_2, \dots$. Then the leftover is the set $A \cup B$. \square

Theorem 18. *Let A_i ($i = 1, 2, \dots$) be countable sets and $A_i \cap A_j = \emptyset$ ($i \neq j$). Then $\bigcup_{i=1}^{\infty} A_i$ is countable.*

Proof. We assume that $A_i = \{a_{i1}, a_{i2}, a_{i3}, \dots\}$ ($i = 1, 2, \dots$). Then the countability of $\bigcup_{i=1}^{\infty} A_i$ can be demonstrated as



\square

The condition of disjointness in Theorem 18 can be omitted.

Exercise 1. If A and B are countable, then $A \times B$ is countable.

Theorem 19. *The interval $[0, 1]$ of real numbers is uncountable.*

Proof. Suppose the set $[0, 1]$ is countable; that is, the numbers in $[0, 1]$ can be listed as an infinite sequence $\{\alpha_i\}_{i=1}^{\infty}$. Write all real numbers α_i in infinite

decimal forms, say in base 10, as follows:

$$\begin{aligned}\alpha_1 &= 0.a_{11}a_{12}a_{13}a_{14}a_{15}\cdots, \\ \alpha_2 &= 0.a_{21}a_{22}a_{23}a_{24}a_{25}\cdots, \\ \alpha_3 &= 0.a_{31}a_{32}a_{33}a_{34}a_{35}\cdots, \\ &\dots\end{aligned}$$

Then we can construct a number $\beta = 0.b_1b_2b_3\cdots$, defined by

$$\begin{aligned}b_1 &= \begin{cases} 1 & \text{if } a_{11} = 0 \\ 0 & \text{if } a_{11} \neq 0, \end{cases} \\ b_2 &= \begin{cases} 1 & \text{if } a_{22} = 0 \\ 0 & \text{if } a_{22} \neq 0, \end{cases} \\ b_3 &= \begin{cases} 1 & \text{if } a_{33} = 0 \\ 0 & \text{if } a_{33} \neq 0, \end{cases} \\ &\dots\end{aligned}$$

The number x is an infinite decimal of 1s and 2s and is a real number between 0 and 1. Since $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, $b_3 \neq a_{33}$, and so on, it follows that $\beta \neq \alpha_1$, $\beta \neq \alpha_2$, $\beta \neq \alpha_3$, etc. Thus β is not in the list $\alpha_1, \alpha_2, \alpha_3, \dots$; that is, β is not a real number between 0 and 1, a contradiction. \square

Note that any finite set cannot be equivalent to a proper subset of itself. However, an infinite set can be equivalent to a proper subset of itself. For instance, $\mathbb{Z}_+ \sim 2\mathbb{Z}_+$, $n \mapsto 2n$ for $n \in \mathbb{Z}$. So $2\mathbb{Z}_+$ is countable.

Theorem 20 (Cantor-Bernstein Theorem). *Given sets X and Y . If there exist subsets $X_0 \subseteq X$, $Y_0 \subseteq Y$ such that $X_0 \sim Y$, $X \sim Y_0$, then $X \sim Y$.*

First Proof. Let $\phi : X \rightarrow Y_0$, $\psi : Y \rightarrow X_0$ be bijections. If there exist subsets $S \subset X$, $T \subset Y$ such that the restrictions $\phi : S \rightarrow T$, $\psi : Y - T \rightarrow X - S$ are bijections, then it is clear that there is a bijection between X and Y .

In fact, for each subset $A \subseteq X$, we have $\phi(A) \subseteq Y_0$. For the subset $Y - \phi(A)$, we have $\psi(Y - \phi(A)) \subseteq X_0$. Now we define

$$\hat{A} = X - \psi(Y - \phi(A)).$$

Note that

$$\text{if } A \subseteq B \text{ in } \mathcal{P}(X), \text{ then } \hat{A} \subseteq \hat{B}.$$

In fact, if $A \subseteq B$, then it is clear that $\phi(A) \subseteq \phi(B)$; subsequently,

$$Y - \phi(A) \supseteq Y - \phi(B) \quad \text{and} \quad \psi(Y - \phi(A)) \supseteq \psi(Y - \phi(B));$$

thus $X - \psi(Y - \phi(A)) \subseteq X - \psi(Y - \phi(B))$, i.e., $\hat{A} \subseteq \hat{B}$.

We want to find a subset $S \subseteq X$ such that $\hat{S} = S$. If so, we have

$$X - S = X - \hat{S} = X - (X - \psi(Y - \phi(S))) = \psi(Y - \phi(S)).$$

Then $f : X \rightarrow Y$, defined by $f(x) = \phi(x)$ for $x \in S$ and $f(x) = \psi^{-1}(x)$ for $x \in X - S$, is a bijection.

We call a subset A of X *closed* if $A \subseteq \hat{A}$. Clearly, \emptyset is closed. Let H be the union of all closed subsets of X , i.e.,

$$H = \bigcup_{A \subseteq X, A \subseteq \hat{A}} A.$$

We claim that $\hat{H} = H$.

For each element $x \in H$, there is a closed subset D of X such that $x \in D \subseteq \hat{D}$. Then $D \subseteq H$ by definition of H ; subsequently, $\hat{D} \subseteq \hat{H}$. Hence $x \in D \subseteq \hat{D} \subseteq \hat{H}$, i.e., $H \subseteq \hat{H}$. Now the inclusion further implies $\hat{H} \subseteq \hat{H}$. This means that \hat{H} is closed. Therefore $\hat{H} \subseteq H$ by definition of H .

Second Proof. Let $\phi : X \rightarrow Y_0$, $\psi : Y \rightarrow X_0$ be bijections. Since $X \sim Y_0$, we only need to show that $Y_0 \sim Y$. Let

$$X_1 = \psi(Y_0), \quad Y_1 = \phi(X_0).$$

Then $X \supseteq X_0 \supseteq X_1$ and $Y \supseteq Y_0 \supseteq Y_1$. Assume that

$$X \supseteq X_0 \supseteq \cdots \supseteq X_n, \quad Y \supseteq Y_0 \supseteq \cdots \supseteq Y_n,$$

where $X_i = \psi(X_{i-1})$ and $Y_i = \phi(Y_{i-1})$, $i = 1, \dots, n$. Now we define

$$X_{n+1} = \psi(Y_n), \quad Y_{n+1} = \phi(X_n).$$

Since $Y_{n-1} \supseteq Y_n$ and $X_{n-1} \supseteq X_n$, it follows that $\psi(Y_{n-1}) \supseteq \psi(Y_n)$ and $\phi(X_{n-1}) \supseteq \phi(X_n)$, i.e., $X_n \supseteq X_{n+1}$, $Y_n \supseteq Y_{n+1}$. We thus have two sequences of sets

$$X \supseteq X_0 \supseteq X_1 \supseteq \cdots, \quad Y \supseteq Y_0 \supseteq Y_1 \supseteq \cdots,$$

and bijections

$$(Y_{i-1} - Y_i) \stackrel{\psi}{\sim} (X_i - X_{i+1}) \stackrel{\phi}{\sim} (Y_{i+1} - Y_{i+2}), \quad i \geq 0,$$

where $Y_{-1} = Y$. Consider the set

$$Y_\infty = \bigcap_{i=0}^{\infty} Y_i \subseteq Y.$$

Then $(Y_{2i-1} - Y_{2i}) \stackrel{\phi\psi}{\sim} (Y_{2i+1} - Y_{2i+2})$ for $i \geq 0$ and

$$Y = Y_\infty \cup (Y - Y_0) \cup (Y_0 - Y_1) \cup (Y_1 - Y_2) \cup \dots \text{ (disjoint),}$$

$$Y_0 = Y_\infty \cup (Y_0 - Y_1) \cup (Y_1 - Y_2) \cup (Y_2 - Y_3) \cup \dots \text{ (disjoint).}$$

Now the function $F : Y \rightarrow Y_0$, defined by

$$f(y) = \begin{cases} \phi\psi(y) & \text{if } y \in \bigcup_{i=0}^{\infty} (Y_{2i-1} - Y_{2i}) \\ y & \text{otherwise} \end{cases}$$

is a bijection. □

Example 9. Since $[1, 2] \sim [1, 2] \subset (0, 3)$ and $(0, 3) \sim (1, 2) \subset [1, 2]$ as intervals of real numbers, then $[1, 2] \sim (0, 3)$.