Cambridge University Press
978-1-316-62024-3 — Measures, Integrals and Martingales
René L. Schilling
Excerpt
More Information

1

Prologue

The theme of this book is the problem of how to assign a size, a content, a probability, etc. to certain sets. In everyday life this is usually pretty straightforward; we

- count: $\{a, b, c, \ldots, x, y, z\}$ has 26 letters;
- take measurements: length (in one dimension), area (in two dimensions), volume (in three dimensions) or time;
- calculate: rates of radioactive decay or the odds of winning the lottery.

In each case we compare (and express the outcome) with respect to some base unit; most of the measurements just mentioned are intuitively clear. Nevertheless, let's have a closer look at areas in Fig. 1.1.

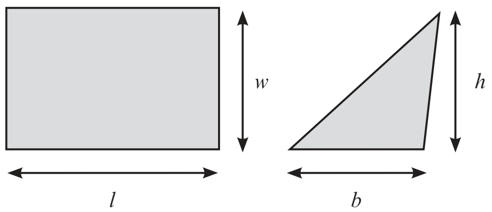


Fig. 1.1. Here area(▭) = length $(\ell) \times$ width (w) and area(△) = $\frac{1}{2} \times$ base $(b) \times$ height (h).

Triangles are more flexible and basic than rectangles since we can represent every rectangle, and actually any odd-shaped quadrangle, as the 'sum' of two non-overlapping triangles. In doing so we have *tacitly* assumed a few things. In Fig. 1.2 we have chosen *a particular* base line and the corresponding height arbitrarily. But the concept of *area* should not depend on such a choice and the

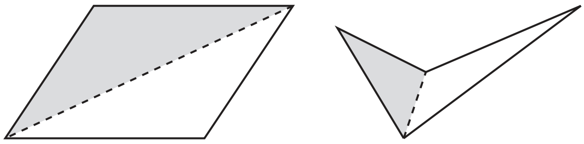


Fig. 1.2. Here area = area(shaded triangle) + area(white triangle).

calculation this choice entails. Independence of the area from the way we calculate it is called *well-definedness*. Plainly, we have the choices shown in Fig. 1.3. Notice that Fig. 1.3 allows us to pick the most convenient method to work out the area. In Fig. 1.2 we actually use two facts:

- the area of non-overlapping (disjoint) sets can be added, i.e.

$$\text{area}(A) = \alpha, \text{area}(B) = \beta, A \cap B = \emptyset \implies \text{area}(A \cup B) = \alpha + \beta;$$

- congruent triangles have the same area, i.e. area(◸) = area(◿).

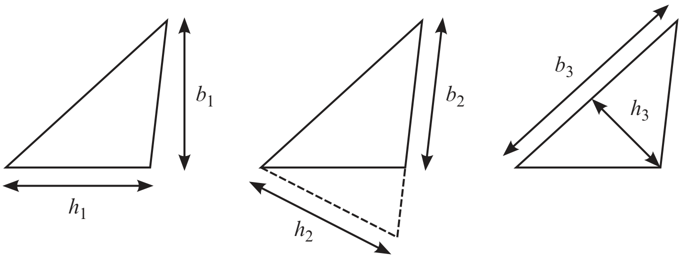


Fig. 1.3. Here area(△) $= \frac{1}{2} \times h_1 \times b_1 = \frac{1}{2} \times h_2 \times b_2 = \frac{1}{2} \times h_3 \times b_3$.

This shows that the least we should expect from a reasonable measure μ is that it is

$$\text{well-defined, takes values in } [0, \infty], \text{ and } \mu(\emptyset) = 0; \tag{1.1}$$

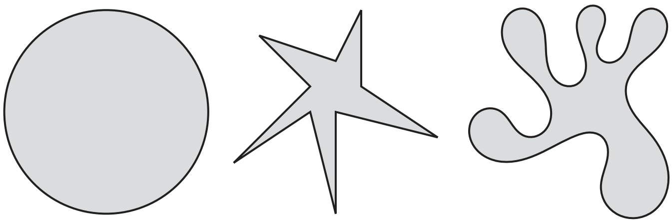
$$\text{additive, i.e. } \mu(A \cup B) = \mu(A) + \mu(B) \text{ whenever } A \cap B = \emptyset. \tag{1.2}$$

The additional property that the measure μ

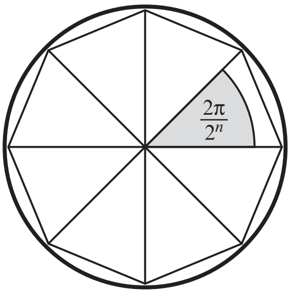
$$\text{is invariant under congruences and translations} \tag{1.3}$$

turns out to be a characteristic property of length, area and volume, i.e. of *Lebesgue measure* on $\mathbb{R}^n$.

The above rules allow us to measure arbitrarily odd-looking *polygons* using the following recipe: dissect the polygon into non-overlapping triangles and add their areas. But what about *curved* or even more complicated shapes, say,



Here is *one* possibility for the circle: inscribe a regular 2^n-gon, $n \in \mathbb{N}$, into the circle, subdivide it into congruent triangles, find the area of each of these slices and then add all 2^n pieces.



In the next step increase $n \rightsquigarrow n+1$ by doubling the number of points on the circumference and repeat the above procedure. Eventually,

$$\text{area}(\bigcirc) = \lim_{n\to\infty} 2^n \times \text{area}(\triangle \text{ at step } n). \tag{1.4}$$

Again, there are a few problems. Does the limit exist? Is it admissible to subdivide a set into arbitrarily many subsets? Is the procedure independent of the particular subdivision? In fact, nothing would have prevented us from paving the circle with ever smaller squares! For a reasonable notion of measure the answer to all of these questions should be *yes* and the way we pave the circle should not lead to different results, as long as our tiles are disjoint. However, finite additivity (1.2) is not enough for this and we have to use instead infinitely many pieces: σ-additivity. Thus

$$\text{area}\Big(\bigcupdot_{n\in\mathbb{N}} A_n\Big) = \sum_{n\in\mathbb{N}} \text{area}(A_n).$$

The notation $\bigcupdot_n A_n$ means the *disjoint union* of the sets A_n, i.e. the union where the sets A_n are pairwise disjoint: $A_n \cap A_m = \emptyset$ if $n \neq m$; a corresponding notation is used for unions of finitely many sets.

Conditions (1.1) and (1.4) lead to a notion of measure which is powerful enough to cater for all our everyday measuring needs and for much more. We will also see that a good notion of measure allows us to introduce integrals, basically starting with the naive idea that the integral of a positive function should stand for the area of the set between the graph of the function and the abscissa.

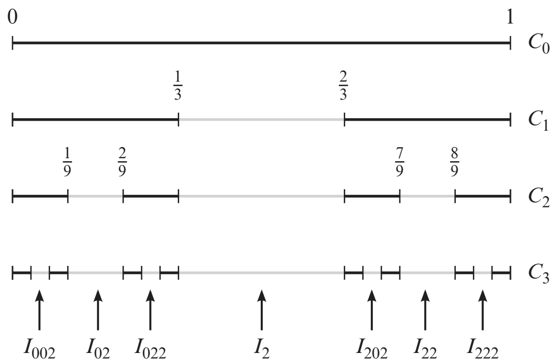


Fig. 1.4. Construction of Cantor's ternary set.

To get an idea of how far we can go with these simple principles, consider *Cantor's ternary set* on the interval $[0, 1]$, see Fig. 1.4. We obtain

1. C_1 by removing from $[0, 1]$ the middle third $I_2 = \left(\frac{1}{3}, \frac{2}{3}\right)$;
2. C_2 by removing from C_1 the middle thirds $I_{02} = \left(\frac{1}{9}, \frac{2}{9}\right)$ and $I_{22} = \left(\frac{7}{9}, \frac{8}{9}\right)$;
3. C_3 by removing from C_2 the middle thirds $I_{002}, I_{022}, I_{202}$ and I_{222};
4. ...

and $C := \bigcap_{n=1}^{\infty} C_n$ is Cantor's ternary set. By construction, C_n consists of 2^n intervals and the endpoints of these intervals will be contained in C. Thus, C is not empty. Let us calculate the 'length' ℓ of the set C. Looking at Fig. 1.4 we see that the length of C_n can be obtained by subtracting from $\ell[0, 1] = 1$ the lengths of the intervals which have been removed in the previous steps:

1. $\ell(C_1) = \ell[0, 1] - \ell(I_2) = 1 - \dfrac{1}{3}$;
2. $\ell(C_2) = \ell[0, 1] - \ell(I_2) - \ell(I_{02}) - \ell(I_{22}) = 1 - \dfrac{1}{3} - 2 \times \dfrac{1}{9}$;

....

n. $\ell(C_{n+1}) = \ell[0, 1] - 2^0 \times \dfrac{1}{3^1} - 2^1 \times \dfrac{1}{3^2} - \cdots - 2^n \times \dfrac{1}{3^{n+1}}$;

and so

$$\ell(C) = 1 - \sum_{n=0}^{\infty} \frac{2^n}{3^{n+1}} = 1 - \frac{1}{3} \sum_{n=0}^{\infty} \frac{2^n}{3^n} = 1 - \frac{1}{3} \frac{1}{1 - \frac{2}{3}} = 0.$$

Thus, the Cantor set has no length in the traditional sense, yet it is not empty.

Problems

1.1. Use (1.4) to find the area of a circle with radius r.

1.2. Where was σ-additivity used when calculating the length of the Cantor set?

1.3. Consider the following variation of Cantor's set: fix $r \in (0,1)$ and delete from $I_0 = [0,1]$ the open interval $\left(\frac{1}{2} - \frac{1}{4}r, \frac{1}{2} + \frac{1}{4}r\right)$. This defines the set I_1 consisting of two intervals, $\left[0, \frac{1}{2} - \frac{1}{4}r\right]$ and $\left[\frac{1}{2} + \frac{1}{4}r, 1\right]$. We get I_2 by removing from each of these intervals the open middles of length $r/8$ and I_3 by removing all open middles of length $r/32$. This defines recursively the sets $I_1, I_2, \dots$. Find the length of the interval I_n and of the generalized Cantor set $I := \bigcap_{n=0}^{\infty} I_n$. Is I empty?

1.4. Let $K_0 \subset \mathbb{R}^2$ be a line of length 1. We get K_1 by replacing the middle third of K_0 by the sides of an equilateral triangle. By iterating this procedure we get the curves $K_0, K_1, K_2, \dots$ (see Fig. 1.5) which defines in the limit Koch's snowflake K_∞. Find the length of K_n and K_∞.

Fig. 1.5. The first four steps in the construction of Koch's snowflake.

1.5. Let $S_0 \subset \mathbb{R}^2$ be a solid equilateral triangle. We get S_1 by removing the middle triangle whose vertices are the midpoints of the sides of S_1. By repeating this procedure with the four triangles which make up S_1 etc. we get $S_0, S_2, S_2, \dots$ (see Fig. 1.6). The Sierpiński triangle is $S_\infty := \bigcap_{n=0}^{\infty} S_n$. Find the area of S_n and S if the side-length of S_0 is $s = 1$.

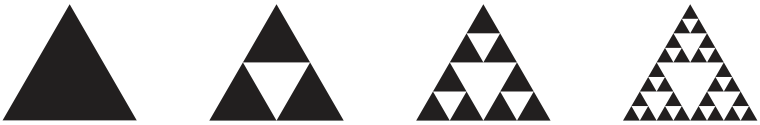


Fig. 1.6. The first four steps in the construction of Sierpiński's triangle.

2

The Pleasures of Counting

Set algebra and countability play a major rôle in measure theory. In this chapter we review briefly notation and manipulations with sets and introduce then the notion of countability. If you are not already acquainted with set algebra, you should verify all statements in this chapter and work through the exercises.

Throughout this chapter X and Y denote two arbitrary sets. For any two sets A, B (which are not necessarily subsets of a common set) we write

$$A \cup B = \{x : x \in A \text{ or } x \in B \text{ or } x \in A \text{ and } B\},$$
$$A \cap B = \{x : x \in A \text{ and } x \in B\},$$
$$A \setminus B = \{x : x \in A \text{ and } x \notin B\};$$

in particular, we write $A \mathbin{\dot\cup} B$ for the *disjoint union*, i.e. for $A \cup B$ if $A \cap B = \emptyset$. $A \subset B$ means that A is contained in B, including the possibility that $A = B$; to exclude the latter we write $A \subsetneq B$. If $A \subset X$, we set $A^c := X \setminus A$ for the *complement* of A (relative to X). Recall also the *distributive laws* for $A, B, C \subset X$

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C) \end{aligned} \tag{2.1}$$

and *de Morgan's identities*

$$\begin{aligned} (A \cap B)^c &= A^c \cup B^c, \\ (A \cup B)^c &= A^c \cap B^c, \end{aligned} \tag{2.2}$$

which hold also for *arbitrarily* many sets $A_i \subset X$, $i \in I$ (I stands for an arbitrary index set),

$$\begin{aligned}\Big(\bigcap_{i\in I} A_i\Big)^c &= \bigcup_{i\in I} A_i^c,\\ \Big(\bigcup_{i\in I} A_i\Big)^c &= \bigcap_{i\in I} A_i^c.\end{aligned} \tag{2.3}$$

A map $f: X \to Y$ is called

$$\begin{aligned}&\textit{injective} \text{ (or one–one)} &&\iff f(x) = f(x') \text{ implies } x = x',\\ &\textit{surjective} \text{ (or onto)} &&\iff f(X) := \{f(x) \in Y : x \in X\} = Y,\\ &\textit{bijective} &&\iff f \text{ is injective and surjective.}\end{aligned} \tag{2.4}$$

Set operations and *direct images* under a map f are not necessarily compatible: indeed, we have, in general,

$$\begin{aligned}f(A \cup B) &= f(A) \cup f(B),\\ f(A \cap B) &\neq f(A) \cap f(B),\\ f(A \setminus B) &\neq f(A) \setminus f(B).\end{aligned} \tag{2.5}$$

Inverse images and set operations are *always* compatible. The inverse mapping f^{-1} maps subsets $C \subset Y$ into subsets of X and it is defined as

$$f^{-1}(C) := \{x \in X : f(x) \in C\} \subset X \quad \text{for all } \ C \subset Y;$$

it is, in general, multi-valued, the notation $f^{-1}(y)$ is used only if $f^{-1}(\{y\})$ has at most one element, i.e. if f is injective. For $C, C_i, D \subset Y$ one has

$$\begin{aligned}f^{-1}\Big(\bigcup_{i\in I} C_i\Big) &= \bigcup_{i\in I} f^{-1}(C_i),\\ f^{-1}\Big(\bigcap_{i\in I} C_i\Big) &= \bigcap_{i\in I} f^{-1}(C_i),\\ f^{-1}(C \setminus D) &= f^{-1}(C) \setminus f^{-1}(D).\end{aligned} \tag{2.6}$$

If we have more information about f we can, of course, say more.

Lemma 2.1 *$f: X \to Y$ is injective if, and only if, $f(A \cap B) = f(A) \cap f(B)$ for all $A, B \subset X$.*

Proof '⇒': Because of the inclusions $f(A \cap B) \subset f(A)$ and $f(A \cap B) \subset f(B)$, we have always $f(A \cap B) \subset f(A) \cap f(B)$. Let us check the converse inclusion '⊃'. If $y \in f(A) \cap f(B)$, we have $y = f(a)$ and $y = f(b)$ for some $a \in A, b \in B$.

So, $f(a)=y=f(b)$ and, by injectivity, $a=b$. This means that $a=b\in A\cap B$, hence $y\in f(A\cap B)$ and $f(A)\cap f(B)\subset f(A\cap B)$ follows.

'$\Leftarrow$': Take $x,x'\in X$ with $f(x)=f(x')$ and set $A:=\{x\}$, $B:=\{x'\}$. Then we have $\emptyset\neq f(\{x\})\cap f(\{x'\})=f(\{x\}\cap\{x'\})$ which is possible only if $\{x\}\cap\{x'\}\neq\emptyset$, i.e. if $x=x'$. This shows that f is injective. □

Lemma 2.2 *$f\colon X\to Y$ is injective if, and only if, $f(X\setminus A)=f(X)\setminus f(A)$ for all $A\subset X$.*

Proof [1] '$\Rightarrow$': Assume that f is injective. If $f(x)\notin f(A)$, we have $x\notin A$, and so $f(X)\setminus f(A)\subset f(X\setminus A)$ (for this inclusion we do not require injectivity). For the reverse inclusion we pick $x\in X\setminus A$ and observe that by injectivity $f(x)\neq f(x')$ for all $x'\in A$. Thus $f(x)\notin f(A)$, and we have shown that $f(x)\in f(X)\setminus f(A)$. This proves $f(X\setminus A)\subset f(X)\setminus f(A)$ as desired.

'$\Leftarrow$': Specialise $f(X\setminus A)=f(X)\setminus f(A)$ to the case $A=\{x\}$ for some $x\in X$. For any $x'\neq x$ we infer that $x'\in X\setminus\{x\}$, and so $f(x')\in f(X)\setminus f(\{x\})$, hence $f(x')\neq f(x)$. □

We can now start with the main topic of this chapter: counting.

Definition 2.3 Two sets X, Y have the same *cardinality* if there exists a bijection $f\colon X\to Y$. In this case we write $\#X=\#Y$.

If there is an injection $g:X\to Y$, we say that the cardinality of X is less than or equal to the cardinality of Y and write $\#X\leqslant\#Y$. If $\#X\leqslant\#Y$ but $\#X\neq\#Y$, we say that X is of strictly smaller cardinality than Y and write $\#X<\#Y$ (in this case, no injection $g:X\to Y$ can be surjective).

That Definition 2.3 is indeed *counting* becomes clear if we choose $Y=\mathbb{N}$ since in this case $\#X=\#\mathbb{N}$ or $\#X\leqslant\#\mathbb{N}$ just means that we can label each $x\in X$ with a unique tag from the set $\{1,2,3,\dots\}$, i.e. we are numbering X. This particular example is, in fact, of central importance.

Definition 2.4 A set X is *countable* if $\#X\leqslant\#\mathbb{N}$. If $\#\mathbb{N}<\#X$, the set X is said to be *uncountable*. The cardinality of $\mathbb{N}$ is called $\aleph_0$, *aleph null*.

Plainly, Definition 2.4 requires that we can find for every countable set some *enumeration* $X=(x_1,x_2,x_3,\dots)$ which may or may not be finite (and which may contain any x_i more than once).

Caution Sometimes *countable* is used only to indicate $\#X=\#\mathbb{N}$, while sets where $\#X\leqslant\#\mathbb{N}$ are called *at most countable* or *finite or countable*. This has the effect that a countable set is always infinite. We do not adopt this convention.

[1] I owe this short argument to Charles Goldie.

The following examples show that (countable) sets with infinitely many elements can behave strangely.

Example 2.5 (i) Finite sets are countable: $\{a,b,\dots,z\}\to\{1,2,\dots,26\}$, where $a\leftrightarrow 1,\dots,z\leftrightarrow 26$, is bijective and $\{1,2,3,\dots,26\}\to\mathbb{N}$ is clearly an injection. Thus

$$\#\{a,b,c,\dots,z\}=\#\{1,2,3,\dots,26\}\leqslant\#\mathbb{N}.$$

(ii) The even numbers are countable. This follows from the fact that the map

$$f\colon\{2,4,6,\dots,2i,\dots\}\to\mathbb{N},\quad i\mapsto\frac{i}{2},$$

is an injection and even a bijection. [✍] This means that there are 'as many' even numbers as there are natural numbers.

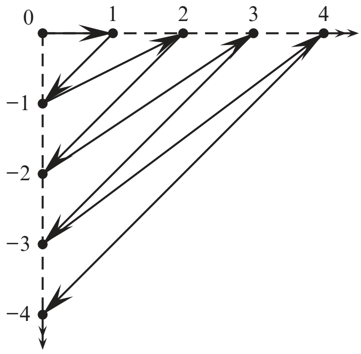
(iii) The set of integers

$$\mathbb{Z}=\{0,\pm1,\pm2,\pm3,\dots\}$$

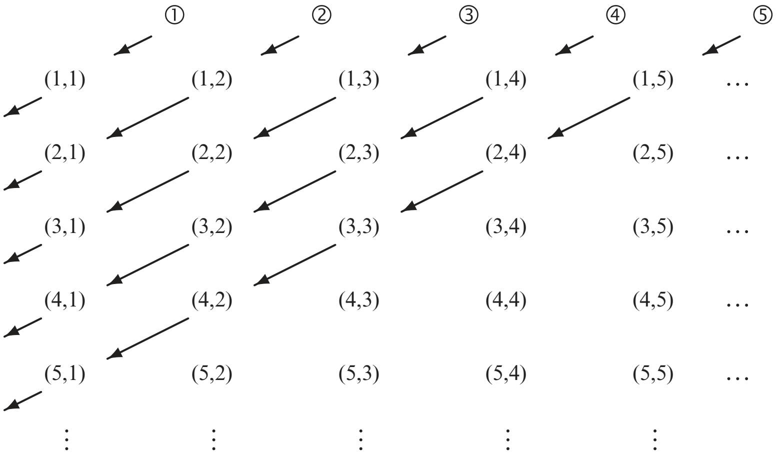
is countable. A possible counting scheme is shown on the right or, more formally,

$$g: i\in\mathbb{Z}\mapsto\begin{cases}2i & \text{if } i>0,\\ 2|i|+1 & \text{if } i\leqslant 0,\end{cases}$$

hence $\#\mathbb{Z}\leqslant\#\mathbb{N}$. [✍]



(iv) The Cartesian product $\mathbb{N}\times\mathbb{N}:=\{(i,k):i,k\in\mathbb{N}\}$ is countable. To see this, arrange the pairs (i,k) in an array and count along the diagonals:



Notice that each line contains only finitely many elements, so that each diagonal can be dealt with in finitely many steps. The map for the above counting scheme is given by

$$h\colon (i,k) \mapsto \frac{(i+k)(i+k-1)}{2} - k + 1 \in \mathbb{N}, \qquad (i,k) \in \mathbb{N}\times\mathbb{N}. \tag{2.7}$$

(v) The rational numbers $\mathbb{Q}$ are countable: set $Q_\pm := \{q \in \mathbb{Q} : \pm q > 0\}$. Every element $\frac{m}{n} \in Q_+$ can be identified with at least one pair $(m,n) \in \mathbb{N}\times\mathbb{N}$, so that

$$Q_+ \subset \Big\{ \underbrace{\tfrac{1}{1},}_{①} \underbrace{\tfrac{1}{2}, \tfrac{2}{1},}_{②} \underbrace{\tfrac{1}{3}, \tfrac{2}{2}, \tfrac{3}{1},}_{③} \underbrace{\tfrac{1}{4}, \tfrac{2}{3}, \tfrac{3}{2}, \tfrac{4}{1},}_{④} \dots \Big\};$$

in the set $\{\dots\}$ on the right we distinguish between cancelled and uncancelled forms of a rational, i.e. $\frac{6}{18}, \frac{1}{3}, \frac{2}{6}, \frac{3}{9}, \dots$, etc. are counted whenever they appear. The numbers ⓘ refer to the corresponding diagonals in the counting scheme in part (iv). This shows that we can find injections $Q_+ \xrightarrow{\imath} \{\dots\} \xrightarrow{\imath} \mathbb{N}\times\mathbb{N}$; the set $\mathbb{N}\times\mathbb{N}$ is countable, thus Q_+ is countable [✍] and so is Q_-. Finally,

$$\mathbb{Q} = Q_- \uplus \{0\} \uplus Q_+ = \{r_1, r_2, r_3, \dots\} \uplus \{0\} \uplus \{q_1, q_2, q_3, \dots\}$$

and $p_1 := 0,\ p_{2i} := q_i,\ p_{2i+1} := r_i$ gives an enumeration $(p_1, p_2, p_3, \dots)$ of $\mathbb{Q}$.

Theorem 2.6 *Let $A_1, A_2, A_3, \dots$ be countably many countable sets. Then their union $A = \bigcup_{i\in\mathbb{N}} A_i$ is countable, i.e. countable unions of countable sets are countable.*

Proof Since each A_i is countable we can find an enumeration

$$A_i = (a_{i,1}, a_{i,2}, \dots, a_{i,k}, \dots)$$

(if A_i is a finite set, we repeat the last element of the list infinitely often), so that

$$A = \bigcup_{i\in\mathbb{N}} A_i = \big(a_{i,k} : (i,k) \in \mathbb{N}\times\mathbb{N}\big).$$

Using Example 2.5(iv) we can relabel $\mathbb{N}\times\mathbb{N}$ by $\mathbb{N}$ and (after deleting all duplicates) we have found an enumeration. □

It is not hard to see that for cardinalities '$\leqslant$' is *reflexive* ($\#A \leqslant \#A$) and *transitive* ($\#A \leqslant \#B,\ \#B \leqslant \#C \implies \#A \leqslant \#C$). *Antisymmetry*, which makes '$\leqslant$' into a partial order relation, is less obvious. The proof of the following important result is somewhat technical and can be left out at first reading.

✻Theorem 2.7 (Cantor–Bernstein) *Let X, Y be two sets. If both $\#X \leqslant \#Y$ and $\#Y \leqslant \#X$, then $\#X = \#Y$.*