1 Basic ideas

1a From Cantor’s uncountability theorem to Baire category theorem

By the famous Cantor’s uncountability theorem, $\mathbb{R}$ is not countable. Here is one of the proofs. Let $a_1, a_2, \cdots \in \mathbb{R}$; we need $x \in \mathbb{R}$ such that $\forall n \ x \neq a_n$. To this end we first take $b_1 < c_1$ such that $a_1 \notin [b_1, c_1]$. Then we take $b_2 < c_2$ such that $[b_2, c_2] \subset [b_1, c_1]$ and $a_2 \notin [b_2, c_2]$. And so on; $[b_1, c_1] \supset [b_2, c_2] \supset [b_3, c_3] \supset \cdots$ Their intersection is not empty, and contains no $a_n$.

Can we generalize it to some sets $A_1, A_2, \cdots \subset \mathbb{R}$ proving that $\bigcup n A_n \neq \mathbb{R}$? Yes, provided that these sets satisfy the following.

1a1 Definition. A set $A \subset \mathbb{R}$ is nowhere dense if every nonempty open interval contains some nonempty open subinterval that does not intersect $A$.

1a2 Exercise. A set $A \subset \mathbb{R}$ is nowhere dense if and only if $\text{Int}(\text{Cl}(A)) = \emptyset$. Prove it. (Here “Int” stands for interior, and “Cl” for closure.)

1a3 Theorem (Baire). If $A_1, A_2, \cdots \subset \mathbb{R}$ are nowhere dense then $\text{Int}(\bigcup n A_n) = \emptyset$.

1a4 Exercise. Prove the theorem.

Equivalently: $\mathbb{R} \setminus \bigcup n A_n$ is dense; that is, $\text{Cl}(\mathbb{R} \setminus \bigcup n A_n) = \mathbb{R}$.

In particular, $\bigcup n A_n \neq \mathbb{R}$.

Clearly, a singleton is nowhere dense; therefore Cantor’s uncountability theorem follows from Baire category theorem.
1b From Cantor’s uncountability theorem to null sets

Here is another proof of Cantor’s uncountability theorem. Let \( a_1, a_2, \cdots \in \mathbb{R} \); we need \( x \in \mathbb{R} \) such that \( \forall n \ x \neq a_n \). To this end we take \( \varepsilon_1, \varepsilon_2, \cdots > 0 \) such that \( \sum_n \varepsilon_n < 1/2 \) and consider open intervals \((a_n - \varepsilon_n, a_n + \varepsilon_n)\). A finite number of these intervals cannot cover \([0, 1]\) since their total length is less than 1. (Take the Riemann integral of the sum of indicators...) By the Heine-Borel theorem, the infinite sequence of these intervals still does not cover \([0, 1]\).

1b1 Definition. A set \( A \subset \mathbb{R} \) is a null set if for every \( \varepsilon > 0 \) there exist \( \varepsilon_1, \varepsilon_2, \cdots > 0 \) and \( a_1, a_2, \cdots \in \mathbb{R} \) such that \( A \subset \bigcup_n (a_n - \varepsilon_n, a_n + \varepsilon_n) \) and \( 2 \sum_n \varepsilon_n \leq \varepsilon \).

1b2 Theorem. If \( A_1, A_2, \cdots \subset \mathbb{R} \) are null sets then \( \text{Int}(\bigcup_n A_n) = \emptyset \).

1b3 Exercise. (a) Prove that \( \bigcup_n A_n \) is also a null set.
(b) Prove the theorem.

1c Two approaches to small sets and typical objects

1c1 Definition. Given a set \( X \), a set \( \mathcal{N} \) of subsets of \( X \) is called
(a) an ideal\(^1\) (on \( X \)), if

\[
(A \subset B \land B \in \mathcal{N}) \implies A \in \mathcal{N}; \\
A, B \in \mathcal{N} \implies A \cup B \in \mathcal{N}; \\
\emptyset \in \mathcal{N}.
\]

(b) a \( \sigma \)-ideal (on \( X \)), if it is an ideal and

\[
A_1, A_2, \cdots \in \mathcal{N} \implies \bigcup_n A_n \in \mathcal{N}.
\]

An ideal (or \( \sigma \)-ideal) \( \mathcal{N} \) on \( X \) is proper if \( X / \notin \mathcal{N} \).

Clearly, null sets are a proper \( \sigma \)-ideal on \( \mathbb{R} \).

The complement of a null set is called a set of full measure.

1c2 Definition. A set \( A \subset \mathbb{R} \) is meager\(^2\) if \( A \subset \bigcup_n A_n \) for some nowhere dense sets \( A_1, A_2, \cdots \subset \mathbb{R} \).

\(^1\)This notion of set theory is different from (but related to) ideals in ring theory, order theory etc.

\(^2\)Or of the first category.
Clearly, meager sets are a proper $\sigma$-ideal on $\mathbb{R}$.
The complement of a meager set is called comeager.$^1$

When a property holds off a null set (in other words, on a set of full
measure), one says that it holds almost everywhere or for almost all elements.
Dealing with a probability measure one also says almost sure(ly).

When a property holds off a meager set (in other words, on a comea-
ger set), one says that it holds quasi-everywhere or for quasi all elements.$^2$
One also says that this property holds generically, for a generic element,
or for most of elements. Sometimes the word “typical” is used rather than
“generic”.

1d Compact metrizable spaces; sequence spaces

1d1 Definition. (a) A metric space is a pair $(X, \rho)$ of a set $X$ and a metric $\rho$
on $X$, that is, a function $\rho: X \times X \to [0, \infty)$ such that $\rho(x, y) = 0 \iff x = y$,$\rho(x, y) = \rho(y, x)$, $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for all $x, y, z \in X$.

(b) Let $\rho_1, \rho_2$ be two metrics on $X$; $\rho_2$ is stronger than $\rho_1$ if $\rho_2(x_n, x) \to 0 \implies \rho_1(x_n, x) \to 0$ for all $x, x_1, x_2, \cdots \in X$;$^3$ further, $\rho_1, \rho_2$ are equivalent, if $\rho_1(x_n, x) \to 0 \iff \rho_2(x_n, x) \to 0$ for all $x, x_1, x_2, \cdots \in X$.

(c) A metrizable space$^4$ is a pair $(X, R)$ where $X$ is a set and $R$ is an
equivalence class of metrics on $X$ (metrizable topology; metrics of $R$ are
called compatible).

(d) A metrizable space (as well as its metrizable topology) is compact$^5$ if
every sequence has a convergent subsequence.

Every subset of $\mathbb{R}$ is a metric space with the metric $\rho(x, y) = |x - y|$.
This space is compact if and only if the set is closed and bounded.

The Cantor set $C \subset [0, 1]$ may be defined as consisting of all numbers of
the form

$$\varphi(x) = \sum_{k=1}^{\infty} \frac{2x(k)}{3^k}$$

for $x \in \{0, 1\}^\infty$, that is $x: \{1, 2, \ldots\} \to \{0, 1\}$.

1d2 Exercise. (a) $\varphi: \{0, 1\}^\infty \to C$ is a bijection;

---

1 Or residual.
2 However, “quasi” is also used in potential theory (in relation to capacity).
3 However, a Cauchy sequence in $(X, \rho_2)$ need not be Cauchy in $(X, \rho_1)$.
4 Equivalently, and usually, a metrizable space is defined as a special case of a topological
space; but here we do not need the notion of general (not just metrizable) topological space.
5 Equivalently (for metrizable spaces), and usually, a compact space is defined by the
Heine-Borel property: every open cover has a finite subcover.
(b) if \( x, x_1, x_2, \ldots \in \{0, 1\}^\infty \) then
\[
\varphi(x_n) \xrightarrow{n \to \infty} \varphi(x) \iff \forall k \left( x_n(k) \xrightarrow{n \to \infty} x(k) \right).
\]
Prove it.

The metric \( \rho(x, y) = |\varphi(x) - \varphi(y)| \) is not invariant under permutations of coordinates on \( \{0, 1\}^\infty \), but its equivalence class \( R \) is (see 1d2(b)). Thus, we have a compact metrizable space \( \{0, 1\}^\infty \), and moreover, the compact metrizable space \( \{0, 1\}^S \) is well-defined for an arbitrary countable set \( S \) (irrespective of its enumeration). The space \( \{0, 1\}^S \) may also be thought of as the space of all subsets of \( S \).

1d3 Definition. A set \( A \) in a metrizable space \( X \) is nowhere dense if every nonempty open set contains some nonempty open subset that does not intersect \( A \).

Still, \( A \) is nowhere dense if and only if \( \text{Int}(\text{Cl}(A)) = \emptyset \).

1d4 Exercise. (a) Prove that nowhere dense sets are an ideal (on a metrizable space).

(b) On \( \mathbb{R} \), prove that they are not a \( \sigma \)-ideal.

1d5 Exercise. A set \( A \subset \{0, 1\}^\infty \) is nowhere dense if and only if for all \( m \) and \( t_1, \ldots, t_m \in \{0, 1\} \) there exist \( n > m \) and \( t_{m+1}, \ldots, t_n \in \{0, 1\} \) such that all sequences that start with \( t_1, \ldots, t_n \) do not belong to \( A \).

Prove it.

1d6 Theorem (Baire). Let \( X \) be a compact metrizable space. If \( A_1, A_2, \ldots \subset X \) are nowhere dense then \( \text{Int}(\bigcup_n A_n) = \emptyset \).

1d7 Exercise. (a) Prove the theorem.

(b) Find an example of a non-compact metrizable space such that the \( \sigma \)-ideal of meager sets is not proper.

Thus, the proper \( \sigma \)-ideal of meager sets is well-defined on every compact metrizable space, in particular, on \( \{0, 1\}^\infty \), and we may speak about generic elements, quasi-everywhere etc. Now, what about null sets? Can we transfer Lebesgue measure from \( \mathbb{R} \) to \( \{0, 1\}^\infty \) by \( \varphi^{-1} \)? No, we cannot, since the Cantor set is itself a null set. But on the other hand, endless coin tossing should provide a useful probability measure on \( \{0, 1\}^\infty \); and binary digits can be thought of as endless coin tossing over Lebesgue measure!

We consider the map \( \psi : [0, 1) \to \{0, 1\}^\infty \),
\[
\psi(u) = (b_1(u), b_2(u), \ldots),
\]
where $b_k(u)$ are the binary digits of $u$, that is,

$$b_k(u) \in \{0, 1\}, \quad \sum_{k=1}^{\infty} \frac{b_k(u)}{2^k} = u, \quad \liminf_k b_k(u) = 0.$$ 

True, $\psi$ is not a bijection, but do not bother: the countable set $\{x : \liminf_k x(k) = 1\}$ is anyway a null set, and outside it $\psi$ is a bijection,

$$\psi^{-1}(x) = \sum_{k=1}^{\infty} \frac{x(k)}{2^k}.$$ 

We transfer Lebesgue measure to $\{0, 1\}^\infty$ by $\psi$. That is, a set $A \subset \{0, 1\}^\infty$ is measurable if $\psi^{-1}(A)$ is Lebesgue measurable, and then $\mu(A)$ is equal to the Lebesgue measure of $\psi^{-1}(A)$. This probability measure $\mu$ is sometimes called Lebesgue measure on $\{0, 1\}^\infty$. It is invariant under permutations of coordinates on $\{0, 1\}^\infty$. Thus, we have a probability space $\{0, 1\}^\infty$, and moreover, the probability space $\{0, 1\}^S$ is well-defined for an arbitrary countable set $S$ (irrespective of its enumeration). It gives us the proper $\sigma$-ideal of null sets on such space, and we may speak about almost all elements etc.

1e “Almost all” versus “quasi all”: first examples

1e1 Example. The famous strong law of large numbers (SLLN) states that

(a) $$\lim_{n} \frac{1}{n} \sum_{k=1}^{n} x(k) = \frac{1}{2} \quad \text{for almost all } x \in \{0, 1\}^\infty.$$ 

In contrast,

(b) $$\liminf_{n} \frac{1}{n} \sum_{k=1}^{n} x(k) = 0, \quad \limsup_{n} \frac{1}{n} \sum_{k=1}^{n} x(k) = 1 \quad \text{for quasi all } x \in \{0, 1\}^\infty,$$

as we will see soon.

1e2 Example. Consider sets

$$A_n = \{x : x(1) = x(n+1), x(2) = x(n+2), \ldots, x(n) = x(2n)\} \subset \{0, 1\}^\infty.$$ 

Clearly, $\mu(A_n) = 2^{-n}$, thus $\sum_n \mu(A_n) < \infty$; by the first Borel-Cantelli lemma,

(a1) $$\mu(\limsup_n A_n) = 0.$$ 

\footnote{It is in fact the Haar measure on the topological group $(\mathbb{Z}_2)^\infty$.}
In other words, almost every \( x \) belongs to \( A_n \) only for finitely many \( n \). Equivalently,\(^1\)

\[
\sum_n \mathbb{1}_{A_n}(x) < \infty \quad \text{for almost all } x \in \{0,1\}^\infty
\]

(\( \mathbb{1}_A \) being the indicator of \( A \)). In contrast,

\[
\sum_n \mathbb{1}_{A_n}(x) = \infty \quad \text{for quasi all } x \in \{0,1\}^\infty,
\]

as we will see soon. That is, quasi every \( x \) belongs to \( A_n \) for infinitely many \( n \). (Of course, the infinite set of \( n \) depends on \( x \).)

**1e3 Exercise.** Denote by \( B_n \) the complement of \( A_n \), and by \( C_n \) the set \( B_n \cap B_{n+1} \cap \ldots \) Prove that

(a) \( C_n \) is closed;

(b) \( C_n \) is nowhere dense.

Thus, \( C = \cup_n C_n \) is meager, and its complement \( \cap_n (A_n \cup A_{n+1} \cup \ldots) = \limsup_n A_n \) is comeager, which proves **1e2(b)**.

**1e4 Exercise.** Now consider sets \( A_n = \{ x : x(n) = x(n+1) = \cdots = x(n^2) = 0 \} \). Prove that

(a) the set \( \limsup_n A_n \) is comeager;

(b) \( \liminf_n \frac{1}{n} \sum_{k=1}^n x(k) = 0 \) for all \( x \in \limsup_n A_n \).

A half of **1e1(b)** is thus proved; the other half is similar.

**1f Digits of a typical number**

We return to the map \( \psi : [0,1) \to \{0,1\}^\infty, \psi(u) = (b_1(u), b_2(u), \ldots) \) where \( b_k(u) \) are the binary digits of \( u \). Of course, \( \psi \) is discontinuous; and nevertheless...

**1f1 Exercise.** Prove that

(a) If \( A \subset \{0,1\}^\infty \) is nowhere dense then \( \psi^{-1}(A) \subset [0,1) \) is nowhere dense.

(b) If \( A \subset \{0,1\}^\infty \) is meager then \( \psi^{-1}(A) \subset [0,1) \) is meager.

(c) If \( A \subset \{0,1\}^\infty \) is comeager then \( \psi^{-1}(A) \subset [0,1) \) is comeager.

---

\(^1\)The sum of the indicators is integrable, therefore, finite almost everywhere. (This is the proof of the first Borel-Cantelli lemma.)
Exercise. Let $A \subset \{0,1\}^\infty$. Prove or disprove:
(a) If $\psi^{-1}(A)$ is nowhere dense then $A$ is nowhere dense.
(b) If $\psi^{-1}(A)$ is meager then $A$ is meager.

Remark. A map satisfying the equivalent conditions $1f1(b,c)$ (but not necessarily (a)) may be called genericity preserving. Informally, such map transforms a generic element of the first space into a generic element of the second space.

Combining $1f1$ with $1e1(b)$ and $1e2(b)$ we see that quasi all $u \in [0,1)$ satisfy
\[ \liminf_n \frac{1}{n} \sum_{k=1}^n b_k(u) = 0, \quad \limsup_n \frac{1}{n} \sum_{k=1}^n b_k(u) = 1, \]
and the relation
\[ b_1(u) = b_{n+1}(u), \ldots, b_n(u) = b_{2n}(u) \]
holds for infinitely many $n$.

All said about $\{0,1\}^\infty$ and binary digits generalizes readily to $\{0,1,\ldots,9\}^\infty$ and decimal digits, as well as any other basis. Given comeager sets $A_p \subset \{0,\ldots,p-1\}^\infty$, we observe for a generic number $u \in [0,1)$ the following property: for every basis $p = 2,3,\ldots$ the corresponding digits of $u$ are a sequence that belongs to $A_p$.

Hints to exercises

1a4: $[b_1,c_1] \supset [b_2,c_2] \supset \ldots$
1d2: if $x(1) = y(1), \ldots, x(n) = y(n)$ then $|\varphi(x) - \varphi(y)| \leq \frac{2}{3^n+1} + \frac{2}{3^n+2} + \ldots$
otherwise $|\varphi(x) - \varphi(y)| \geq \frac{2}{3^n} - \frac{2}{3^n+1} - \frac{2}{3^n+2} - \ldots$
1d4: (a) $[b_1,c_1] \supset [b_2,c_2] \supset [b_3,c_3]$; (b) the union can be dense.
1d7: (a) similar to 1a4 with balls rather than intervals; (b) try a dense countable set.
1e3 (b) use 1d5
1e4 (b) try $n \in \{1,4,9,16,\ldots\}$
1f1 (a) by 1d5 every binary interval $[\frac{k}{2^n}, \frac{k+1}{2^n})$ contains a binary subinterval such that... (b), (c) follow from (a).
1f2 consider $\{0,1\}^\infty \setminus \psi([0,1))$.

---

1According to Melleray and Tsankov, a continuous map with this property is called category-preserving; see arXiv:1201.4447, Def. 2.7.
<table>
<thead>
<tr>
<th>Term</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ-ideal</td>
<td>2</td>
</tr>
<tr>
<td>almost</td>
<td>3, 5</td>
</tr>
<tr>
<td>Baire theorem</td>
<td>1, 4</td>
</tr>
<tr>
<td>binary digits</td>
<td>5</td>
</tr>
<tr>
<td>Cantor set</td>
<td>3</td>
</tr>
<tr>
<td>comeager</td>
<td>3</td>
</tr>
<tr>
<td>compact</td>
<td>3</td>
</tr>
<tr>
<td>compatible metric</td>
<td>3</td>
</tr>
<tr>
<td>equivalent metrics</td>
<td>3</td>
</tr>
<tr>
<td>full measure</td>
<td>2</td>
</tr>
<tr>
<td>generic</td>
<td>3, 4</td>
</tr>
<tr>
<td>genericity preserving</td>
<td>7</td>
</tr>
<tr>
<td>ideal</td>
<td>2</td>
</tr>
<tr>
<td>Lebesgue measure on {0,1}^\infty</td>
<td>5</td>
</tr>
<tr>
<td>meager</td>
<td>2</td>
</tr>
<tr>
<td>metric</td>
<td>3</td>
</tr>
<tr>
<td>metric space</td>
<td>3</td>
</tr>
<tr>
<td>metrizable space</td>
<td>3</td>
</tr>
<tr>
<td>metrizable topology</td>
<td>3</td>
</tr>
<tr>
<td>nowhere dense</td>
<td>1, 4</td>
</tr>
<tr>
<td>null set</td>
<td>2</td>
</tr>
<tr>
<td>proper</td>
<td>2</td>
</tr>
<tr>
<td>quasi</td>
<td>3, 4</td>
</tr>
<tr>
<td>stronger metric</td>
<td>3</td>
</tr>
<tr>
<td>typical</td>
<td>3</td>
</tr>
<tr>
<td>{0,1}^S</td>
<td>1</td>
</tr>
<tr>
<td>{0,1}^\infty</td>
<td>3</td>
</tr>
<tr>
<td>I_A</td>
<td>9</td>
</tr>
<tr>
<td>b_k(·)</td>
<td>5</td>
</tr>
<tr>
<td>Cl</td>
<td>1</td>
</tr>
<tr>
<td>Int</td>
<td>1</td>
</tr>
<tr>
<td>ϕ(·)</td>
<td>3</td>
</tr>
<tr>
<td>ψ(·)</td>
<td>4</td>
</tr>
</tbody>
</table>