1

Preliminaries

1.1 Littlewood's Three Principles

We shall be dealing extensively with measurable sets and functions, and begin by recalling Littlewood's three principles (J.E. Littlewood (1885–1977) from 1944; see Lit1944, §4), according to which a general situation is 'nearly' an easy situation:

(i) any measurable set is nearly a finite union of intervals;
(ii) any measurable function is nearly continuous;
(iii) any convergent sequence of measurable functions is nearly uniformly convergent.

These statements can be made precise, as follows.

Littlewood's first principle is essentially the regularity of Lebesgue measure. That is, with $|.|$ denoting Lebesgue measure, for A (Lebesgue) measurable $|A|$ is the infimum of $|U|$ over open sets $U \supseteq A$ (open supersets of A) and the supremum of $|K|$ over compact subsets K of A. So one can approximate to within any $\epsilon > 0$ from without by open sets and from within by compact sets; taking $\epsilon = 1/n$ (with $n = 1, 2, \ldots$), one can find a $\mathcal{G}_\delta$ set $G \supseteq A$ and a $\mathcal{F}_\sigma$ set $F \subseteq A$ with $|G \setminus A| = 0$ and $|A \setminus F| = 0$. (For the place of $\mathcal{G}_\delta$ and $\mathcal{F}_\sigma$ sets in the Borel hierarchy, see pages 32 and 51.) As each open set on the line is a countable disjoint union of open intervals, for each $\epsilon > 0$ one can find a finite (disjoint) union U of open intervals whose symmetric difference with A has measure $|U \Delta A| < \epsilon$. See, e.g., Bog2007a, §1.5 or Roy1988, §3.3.

Littlewood's second principle is essentially Lusin's Theorem (N.N. Lusin, or Luzin (1883–1950), in 1912; see, e.g., Bog2007a, Th. 2.2.10 or Hal1950, §55).

Theorem 1.1.1 (Lusin's Continuity Theorem, or Almost-Continuity Theorem) *Given a regular measure and a finite measure space (e.g. Lebesgue measure on* *a compact interval), for f measurable and a.e. finite, f is* almost continuous*: for $\epsilon > 0$, there exists a closed set F on which f is continuous and whose complement has measure $|F^c| < \epsilon$.*

This property of almost continuity in fact characterizes measurability.

Littlewood's third principle is essentially Egorov's Theorem (D.F. Egorov (1869–1931) in 1911; see, e.g., Bog2007a, Th. 2.2.1; Hal1950, Th. 21A).

Theorem 1.1.2 (Egorov's Theorem) *For A measurable of finite measure, and f_n a sequence of measurable functions convergent a.e. to f on A, f_n converges almost uniformly: for each $\epsilon > 0$ there exists $B \subseteq A$ with $|A \setminus B| < \epsilon$ and $f_n \to f$ uniformly on B.*

Thus almost everywhere convergence (in particular, pointwise convergence) implies almost uniform convergence.

For textbook accounts of Littlewood's three principles, see SteS2005, §4.3; Roy1988, §3.6.

Our standard reference texts for measure theory will be Bogachev (Bog2007a; Bog2007b) and Fremlin (Fre2000b; Fre2001; Fre2002; Fre2003; Fre2008).

1.2 Topology: Preliminaries and Notation

We gather here a variety of results which we will need later. These are all known to the experts; others may prefer to return to this for reference as may be needed.

Our standard references for general topology, as already mentioned in the Preface, will be Engelking (Eng1989) and also the *Handbook of Set-Theoretic Topology* (KunV1984).

In the earlier parts of the book we will, for the most part, work in metric spaces. But we may need to use alternative metrics, for instance to take advantage of completeness. So it will be convenient to work from the start with topological spaces (Hausdorff, by assumption), and these may often turn out to be *metrizable*. Under such circumstances there may also often be a second stronger, or as we shall say *finer*, topology in play (i.e. one with more open sets) – which we call *submetrizable*. A helpful analogy is the interplay of weak and strong topologies in function spaces.

In general we develop topological machinery when needed. But here we briefly review basic concepts to establish conventions and notation – for details and proofs of results mentioned below we refer to Eng1989. So below we are concerned with:

(i) separation properties (regular, completely regular, normal, etc.);
(ii) covering and refinement properties (compactness, countable compactness, local finiteness and paracompactness, etc.);
(iii) base properties (second countability and σ-local finiteness, etc.);
(iv) neighbourhood base properties (first countability, regular bases).

Notation. Inclusion will be denoted by $A \subseteq B$, proper inclusion by $A \subset B$, complement by A^{c}, symmetric difference by $A\Delta B := (A \setminus B) \cup (B \setminus A)$. By $\bar{A}$ we will denote the closure of A when there is only one topology in play. Otherwise the closure will be written $\mathrm{cl}_{\mathcal{T}}(A)$ or just $\mathrm{cl}_{\mathcal{T}} A$ to imply the relevant topology $\mathcal{T}$.

By $\omega := \mathbb{N} \cup \{0\}$ we denote the set of *finite* ordinals.

A subset A of natural numbers ω will often be identified with a real number in (0,1). Indeed, A is identified uniquely by its *indicator function* (on the natural numbers) with $1_A(n) = 1$ or 0, according as $n \in A$ or not. In turn, 1_A, as a binary sequence, determines a real number in (0,1) with binary expansion of that sequence.

Separation Properties. A topological space X is *regular* if for each neighbourhood U of any point x there is a neighbourhood V with $x \in V \subseteq \bar{V} \subseteq U$. Further X is *completely regular* (or Tychonoff) if for each neighbourhood U of any point x there is a continuous function f with $f(x) = 1$ and $f = 0$ outside U and it is *normal* (or Urysohn) if for any disjoint pair of closed sets A, B there is a continuous function f with $f = 1$ on A and $f = 0$ on B.

A space is *pseudonormal* if matters are as in the last definition but one of the two closed sets is countable (e.g. a convergent sequence). Thus a pseudonormal space is completely regular.

Covering and Refinement Properties. A space X is *compact* if every *open* covering, i.e. open family $\mathcal{U}$ covering X (family of open sets with union X), contains a finite subcovering (finite subfamily $\mathcal{U}'$ covering X). The space is *countably compact* if any countable open covering has a finite subcovering. This is to be contrasted with *sequential compactness*, which demands that every sequence $\langle x_n \rangle$ must have a convergent subsequence; such a space is countably compact (see Eng1989, Th. 3.10.30). A space X is *pseudocompact* if every real-valued continuous function is bounded. A countably compact completely regular space is pseudocompact. Every normal pseudocompact space is countably compact. In a metrizable space these three concepts are equivalent.

A space is *Lindelöf* if any open covering contains (has) a countable subcovering.

A family $\mathcal{F}$ is *locally finite* if each point x has a neighbourhood U meeting (intersecting) only finitely many members of $\mathcal{F}$. If each x has a neighbourhood U meeting at most one member of $\mathcal{F}$, then $\mathcal{F}$ is *discrete*.

The family is σ-locally finite (or σ-discrete) if $\mathcal{F} = \bigcup_{i\in\omega} \mathcal{F}_i$ and each $\mathcal{F}_i$ is locally finite (resp. discrete). Any open σ-locally finite covering has a locally finite refinement (by sets not necessarily open) – see, e.g., Eng1989, Lemma 5.1.10.

A family $\mathcal{V}$ *refines* the family $\mathcal{U}$ if each member of $\mathcal{V}$ is included in a member of $\mathcal{U}$.

A space is *paracompact* if every open covering $\mathcal{U}$ has a locally finite open refinement. Examples are Lindelöf (and in particular compact) spaces and metrizable spaces. By Stone's Theorem (Eng1989, Th. 4.4.1) every open cover of a metrizable space has an open refinement that is both locally finite and σ-discrete.

Every paracompact space is normal. A regular space is paracompact if and only if every open cover has an open σ-locally finite refinement (Eng1989, Th. 5.1.11).

By analogy with countable compactness, X is *countably paracompact* if every countable open covering $\mathcal{U}$ has a locally finite open refinement.

Dowker's Theorem (Eng1989, Th. 5.2.8) asserts that X is normal and countably paracompact if and only if $X \times [0, 1]$ is normal. This links normal spaces to Borsuk's Homotopy Extension Theorem (see, e.g., Eng1989, §5.5.21 and note in particular Starbird's Theorem).

Base Properties. The simplest base property is its countability: a topology is called second countable (i.e. satisfies the second axiom of countability) if it has a countable base. The topology of a metric space is second countable if and only if the space is separable (has a countable dense subset). Generalizations of this simplest of all topological countability properties include σ-discrete bases and σ-locally finite bases.

For instance, the Nagata–Smirnov Theorem (Eng1989, Th. 4.4.7) asserts that a space is metrizable if and only if it is regular and has a σ-locally finite base. In the same spirit is Bing's Theorem (Eng1989, Th. 4.4.8) that a space is metrizable if and only if it is regular and has a σ-discrete base.

A further generalization is provided through regular neighbourhood bases below.

Neighbourhood Base Properties. A *neighbourhood base* is an assignment to each point x in space of a family $\mathcal{B}(x)$ of neighbourhoods of x such that for every neighbourhood U of any point x there is a smaller neighbourhood

B in $\mathcal{B}(x)$. The simplest neighbourhood base property is again countability (for each x): a topology is called *first countable* (i.e. satisfies the first axiom of countability) if it has a neighbourhood base assigning to each point x a countable family $\mathcal{B}(x)$ as above.

One obtains a neighbourhood base from a base $\mathcal{B}$ by setting $\mathcal{B}(x) := \{B \in \mathcal{B} : x \in B\}$. This leads to the notion of a base $\mathcal{B}$ that is *point-regular* by requiring that every point x has a neighbourhood U such that all but finitely many members of $\mathcal{B}(x)$ lie in U (i.e. all but finitely many of those members B of $\mathcal{B}$ that contain x lie in U). This notion motivates a 'localized' version.

A base $\mathcal{B}$ is *regular* if for every neighbourhood U of any point x there is a smaller neighbourhood V such that all but finitely many members B of $\mathcal{B}$ meeting V lie in U. Arhangelskii's Theorem (Eng1989, Th. 5.4.6) asserts in particular that a Hausdorff space is metrizable if and only if it has a regular base.

1.3 Convergence Properties

We shall need various modes of convergence, which we now discuss briefly. Let $\langle \Omega, \mathcal{S}, m\rangle$ be a measure space. For Φ a property of subsets of Ω we write $m\{\Phi\}$ as an abbreviation for $m(\{\omega \in \Omega : \Phi(\omega)\})$. In particular, if m is finite we can divide by $m(\Omega)$ to make m a probability, so without loss of generality m is a probability if finite. For $\langle X, \mathcal{T}\rangle$ a topological space, recall that in this latter context a random variable with values in X is an $\mathcal{S}$-measurable map $Y : \Omega \to X$; we write $L^0(\Omega, X)$ for the set of random variables.

Modes of Convergence. For $\langle X, d\rangle$ a metric space and $\langle Y_n : n \in \omega\rangle$, a sequence of random variables with values in X the sequence converges to Y_0:

(i) m-a.e. or almost surely (a.s.) if $Y_n(\omega) \to Y_0(\omega)$ almost everywhere;
(ii) in measure/in probability if, for every $\epsilon > 0$, $m\{d(Y_n, Y_0) > \epsilon\} \to 0$ as $n \to \infty$.

Convergence a.e. implies convergence in measure/in probability (Bog2007a, 2.2.3), but not conversely. The standard example here is constructed from the subintervals $I_j := [j/2^{n-1} - 1, (j+1)/2^{n-1} - 1]$ for $2^{n-1} \le j < 2^n$ of $[0, 1]$. As these have lengths shrinking to zero, their indicator functions $f_j(\omega)$ regarded as random variables on $\Omega = [0, 1]$, equipped with Lebesgue measure, converge to zero in probability. They do not converge almost surely: for any irrational ω there is an infinite sequence $n_j(\omega)$ where the n_jth function is 1 and an infinite sequence $m_j(\omega)$ where the n_jth function is 0. So we do not have a.e. convergence, but do have a.e. convergence along a subsequence. This example is canonical, as a result below (the 'subsequence theorem', due to F. Riesz in 1909) shows.

Thus convergence a.e. (or a.s. in the probability case) is a strong mode of convergence and convergence in measure (or in probability) a weaker one. This is reflected in the terminology of the basic limit theorems of probability theory, the strong law of large numbers (SLLN) and the weak law of large numbers (WLLN) (see, e.g., Dud1989, Ch. 8). Other strong modes of convergence are convergence in L_p (or in pth mean; we shall only need $p = 1$ – convergence in mean – and $p = 2$ – convergence in mean square). These are not comparable to convergence a.e. At the other extreme is *convergence in distribution*, or *in law*, as in the central limit theorem (CLT) of probability theory; this is implied by convergence in measure/probability, but not conversely unless the limit is constant. See, e.g., Dud1989, Ch. 9.

Convergence in pth mean is metric and generated by the norm of L_p. Convergence in measure/probability is metric and generated by the *Ky Fan metric*, α say. Convergence in distribution is metric and generated by the *Prohorov metric*, ρ say. That convergence in probability implies convergence in distribution follows from $\rho \leq \alpha$; see, e.g., Dud1989, Th. 11.3.5. These results also show that convergence a.e. is metrizable only when it coincides with convergence in measure. This does not happen in general, as the example below shows – indeed, in general convergence a.e. is not even topological (see, e.g., Dud1989, Problem 9.2.2). But it does happen if the measure space is purely atomic, as then there are no non-trivial null sets; examples such as the one above, which take place on [0, 1] under the Lebesgue-measurable sets and Lebesgue measure, do not then apply.

Say that the sequence $\langle Y_n : n \in \omega \rangle$ has the *sub-subsequence property* relative to the null sets if for every subsequence $\langle Y_{n(k)} \rangle$ there is a sub-subsequence $\langle Y_{n(k(j))} \rangle$ converging a.e. to Y_0.

Theorem 1.3.1 (Subsequence Theorem; Dud1989, Th. 9.2.1; Bog2007a, 2.2.5(i)) *For $\langle Y_n : n \in \omega \rangle$ a sequence of random variables with values in a separable metric space $\langle X, d \rangle$, the sequence $\langle Y_n : n \in \omega \rangle$ converges to Y_0 in probability if and only if $\langle Y_n : n \in \omega \rangle$ has the sub-subsequence property relative to the null sets.*

Proof Suppose that $\langle Y_n : n \in \omega \rangle$ does not converge to Y_0 in probability. Then for some $\varepsilon > 0$ there is a subsequence $n(k)$ such that

$$P\{d(Y_{n(k)}, Y_0) > \varepsilon\} > \varepsilon, \text{ for all } k.$$

Thus for any sub-subsequence $\langle Y_{n(k(j))}\rangle$ we have also

$$P\{d(Y_{n(k(j))}, Y_0) > \varepsilon\} > \varepsilon, \text{ for all } j,$$

and so $Y_{n(k(j))}(\omega)$ does not converge to $Y_0(\omega)$ for a non-null set of ω, i.e. $\langle Y_n : n \in \omega\rangle$ does not have the sub-subsequence property.

Now suppose that $\langle Y_n : n \in \omega\rangle$ converges in probability to Y_0. Then so does $\langle Y_{n(k)} : k \in \omega\rangle$ for any given subsequence $n(k)$. Choose for each j an integer $k(j) > j$ such that

$$P\{d(Y_{n(k(j))}, Y_0) > 1/j\} < 1/j^2.$$

Hence, by summability of $1/j^2$, the set

$$\bigcap\nolimits_k \bigcup\nolimits_{j>k} \{\omega : d(Y_{n(k(j))}(\omega), Y_0(\omega)) > 1/j\}$$

has P-measure 0, i.e.

$$\bigcup\nolimits_k \bigcap\nolimits_{j>k} \{\omega : d(Y_{n(k(j))}(\omega), Y_0(\omega)) \leq 1/j\}$$

has P-measure 1. So for ω in this latter set and any $\varepsilon > 0$ there is some $k = k(\omega)$ such that for $j > \max\{k(\omega), 1/\varepsilon\}$

$$d(Y_{n(k(j))}(\omega), Y_0(\omega)) \leq 1/j < \varepsilon.$$

That is, $\langle Y_{n(k(j))}\rangle$ converges almost surely to Y_0. □

Definition The sequence $\langle Y_n\rangle$ is *Cauchy* or *fundamental* in measure if

$$P\{\sup_{k\geq n} d(Y_n, Y_k) \geq \varepsilon\} \to 0 \text{ as } n \to \infty.$$

Lemma 1.3.2 (Almost Sure Convergence Criterion; Dud1989, 9.2.4; cf. Bog2007a, 2.2.5 (ii)) *If $\langle Y_n\rangle$ is a Cauchy/fundamental sequence, then Y_n a.s. converges to some Y_0 a.s. (and so also in measure).*

Remark Wagner and Wilczyński (WagW2000) proved that the category version also holds.

1.4 Miscellaneous

We will use $\mathcal{F}$, $\mathcal{G}$ for the families of closed and open sets ('f for *fermé*, g for *geöffnet*'), $\mathcal{H}$ as our usual letter for a family of sets in general, $\mathcal{K}$ for the family of compact sets ('k for *kompakt*'). A family $\mathcal{H}$ is *multiplicative* if it is closed under finite intersection, and so analogously a set-valued map $F : \mathcal{H} \to \mathcal{H}$ is *multiplicative* if $F(A \cap B) = F(A) \cap F(B)$. Dually, say that the family $\mathcal{H}$ is *additive* if it is closed under finite unions, and so also a set-valued map F is *additive* if $F(A \cup B) = F(A) \cup F(B)$. A family $\mathcal{H}$ is *maximally additive* if whenever $A \cup B$ is in $\mathcal{H}$, then at least one of A or B is in $\mathcal{H}$; this is motivated by a context in which a family $\mathcal{H}$ with some property P may be extended to a *maximal* such family, and the extension process permits the inclusion for each $A \cup B$ one of A or B, without violating P. We write $\sigma(\mathcal{H})$ for the σ-algebra generated by (the smallest σ-algebra containing) $\mathcal{H}$. Given a σ-algebra, we will often need to consider a sub-σ-algebra of 'small sets' (prototypical examples: null sets or meagre sets). Such a sub-σ-algebra will be closed under intersection with any set in the σ-algebra, and so will have the structure of an *ideal* (using the viewpoint of Boolean algebra, with intersection as product). We will use $\mathcal{I}$ to denote an ideal generically, qualifying this to distinguish one ideal from another; we write $\mathcal{M}$ for the ideal of meagre sets (see Chapter 2), and $\mathcal{N}$ for the ideal of null sets. Here the measure μ will be Lebesgue measure, n-dimensional Lebesgue or Haar measure. Given a sequence $\sigma := \langle \sigma_1, \sigma_2, \ldots \rangle$, write $\sigma \mid n$ for the first n terms. For $\mathcal{H}$ a family of sets, write $\mathbf{S}(\mathcal{H})$ for the class of sets of the form

$$\bigcup_{\sigma} H(\sigma), \quad \text{where } H(\sigma) := \bigcap_{n=1}^{\infty} H(\sigma \mid n),$$

with each $H(\sigma \mid n) \in \mathcal{H}$. Here $\mathbf{S}$ is the *Souslin operation* (M. Ya. Souslin, or Suslin (1894–1919), in 1917), and the sets in $\mathbf{S}(\mathcal{H})$ are the Souslin-$\mathcal{H}$ sets (the term analytic and notation A are also used; see, e.g., Bog2007a, §1.10). The Souslin operation is important in descriptive set theory, and we shall return to it later in §3.1. Meanwhile we note the following:

(i) The Souslin operation is idempotent: $\mathbf{S}(\mathbf{S}(\mathcal{H})) = \mathbf{S}(\mathcal{H})$.

(ii) The class of measurable sets is closed under the Souslin operation (Lusin and Sierpiński in 1918, Nikodym in 1925, Marczewski (as Szpilrajn) in 1929 and 1933 – see, e.g., RogJ1980, Cor. 2.9.3). We will meet the sets with the Baire property in Chapter 2; for them we have the dual result, also due to Marczewski (see, e.g., RogJ1980, Cor. 2.9.4):

(iii) the class of sets with the Baire property is closed under the Souslin operation.

Evidently $\mathbf{S}(\mathcal{H})$ includes the family of all countable intersections of members of $\mathcal{H}$, i.e. $\mathbf{S}(\mathcal{H}) \supseteq \mathcal{H}_\delta$; with Hausdorff's δ for Durchschnitt notation, writing $\mathbf{i} \mid n = (i_1, \ldots, i_n)$, and noting that

$$\bigcup_{\mathbf{i} \in \mathbb{N}^{\mathbb{N}}} \bigcap_n H(i_1, \ldots, i_n) = \bigcup_{i_i=1}^{\infty} \left(\bigcup_{\mathbf{i} \in \mathbb{N}^{\mathbb{N}}} \bigcap_{n=2}^{\infty} H(i_2, \ldots, i_n) \right),$$

we see also the inclusion of countable unions, i.e. $\mathbf{S}(\mathcal{H}) \supseteq \mathcal{H}_\sigma$, so that

$$\mathcal{H}_\sigma \subseteq \mathcal{H}_{\sigma\delta} \subseteq \mathcal{H}_{\sigma\delta\sigma} \subseteq \cdots \subseteq \mathbf{S}(\mathcal{H}).$$

Beyond the initial levels displayed, the Borelian-$\mathcal{H}$ hierarchy is thus obtained, by transfinite induction through the countable ordinals, by alternating the σ and δ operations at successor ordinals and at limit ordinals 'coalescing' (taking unions of) the preceding families. The hierarchy is thus included in $\mathbf{S}(\mathcal{H})$. We shall refer to the case $\mathcal{H} = \mathcal{K}$ where $\mathcal{K} = \mathcal{K}(X)$ denotes the family of compact sets in X.

Taking $\mathcal{H} = \mathcal{G}$, the open sets of a space (recall 'G for geöffnet'), and $\mathcal{H} = \mathcal{F}$, the closed sets ('F for fermé'), we see that the Borelian-$\mathcal{H}$ and Borelian-$\mathcal{F}$ hierarchies are also included in the corresponding $\mathbf{S}(\mathcal{H})$:

$$\mathcal{G}_\sigma = \mathcal{G} \subseteq \mathcal{G}_{\delta\sigma} \subseteq \mathcal{G}_{\delta\sigma\delta} \subseteq \cdots \subseteq \mathbf{S}(\mathcal{G}),$$
$$\mathcal{F}_\sigma \subseteq \mathcal{F}_{\sigma\delta} \subseteq \mathcal{F}_{\sigma\delta\sigma} \subseteq \cdots \subseteq \mathbf{S}(\mathcal{F}).$$

Evidently, in a metric space $\mathcal{G} \subseteq \mathcal{F}_\sigma$, and similarly $\mathcal{F} \subseteq \mathcal{G}_\delta$. Continuing in this way, $\mathcal{F}_\sigma \subseteq \mathcal{G}_{\delta\sigma}$, $\mathcal{G}_\delta \subseteq \mathcal{F}_{\sigma\delta}$ and so on. So taking the union the two hierarchies amalgamate, creating the *Borel hierarchy*, which is closed under complementarity. We note, for future comparison, the capitalized Greek notation, a compromise between the two languages, emanating from mathematical logic using bold-face Σ for sum and $\mathbf{\Pi}$ for product:

$$\mathbf{\Pi}^0_0 \subseteq \mathbf{\Pi}^0_1 \subseteq \mathbf{\Pi}^0_2 \subseteq \cdots$$
$$\mathbf{\Sigma}^0_0 \subseteq \mathbf{\Sigma}^0_1 \subseteq \mathbf{\Sigma}^0_2 \subseteq \cdots$$

with $\mathbf{\Pi}^0_0$ for $\mathcal{F}$ and $\mathbf{\Sigma}^0_0$ for $\mathcal{G}$ and later levels beyond these initial ones indexed by ordinals. The notation stresses the implied reference to the existential quantifier $\exists n$ use of the integers (0-level objects) in the countable union operation over a sequence of sets and the complementary universal quantifier $\forall n$ in the countable intersection. The bold-face signals the implied coding of the sequences involved which are in general unconstructively enumerated. When constructive (more accurately 'effectively enumerated') one drops down to light-face notation.

The sets in $\mathbf{S}(\mathcal{F})$ are the Souslin-$\mathcal{F}$ sets; in the context of the closed subsets of a complete metric space, these are called *absolutely analytic* (subsets) since such subsets when embedded in any 'enveloping' metric space are Souslin-$\mathcal{F}$ in any enveloping metric space. We address their properties in later chapters.

We use 'measure' to mean 'countably additive measure'. If a measure is defined on the power set $\wp(X)$ of all subsets of a countable set X, and vanishes on singletons, it vanishes identically by countable additivity (note that a finitely additive measure may very well vanish on singletons but not identically).

The same statement holds for X of cardinality that of the first uncountable ordinal (Ulam's Theorem of 1930 – see Bog2007a, Th. 1.12.40).

A cardinal κ is called *non-measurable* if whenever a measure is defined on all subsets of a set of cardinality κ, and vanishes on singletons, it vanishes identically. Other cardinals are called *measurable*. Whereas one thinks of measurable sets as being 'nice', here it is non-measurable cardinals that are 'nice'. For background, see, e.g., Bog2007a; Bog2007b; Fre2008. We will meet non-measurable cardinality in connection with results of Pol, 2.1.6 and §12.5*, and in §16.3.

We will study category–measure duality in Chapter 9, and a certain non-metric topology, the *density topology* $\mathcal{D}$, in Chapter 7. This is convenient for our purposes because using it one can bring the category and measure aspects together. We note here that this can only happen because the topology is non-metric. For decompositions of metric measure spaces into two parts, one meagre (small in category) and one null (small in measure), see MarS1949. See also Oxt1980, Ch. 16.