Appendix A. Some Topology and Measure Theory

(i) Topology

The concept of a topological space is so fundamental in modern mathematics that we don't feel obliged to recall its definitions or basic properties. Therefore we refer to Dugundji 1966 for everything concerning topology, nevertheless we shall briefly quote some results on compact and metric spaces which we use frequently.

A.1. Compactness:
A topological space $(X, \mathscr{O})$, $\mathscr{O}$ the family of open sets in X, is called *compact* if it is Hausdorff and if every open cover of X has a finite subcover. The second property is equivalent to the *finite intersection property*: every family of closed subsets of X, every finite subfamily of which has non-empty intersection, has itself non-empty intersection.

A.2. The continuous image of a compact space is compact if it is Hausdorff. Moreover, if X is compact, a mapping $\varphi : X \to X$ is already a homeomorphism if it is continuous and bijective. If X is compact for some topology $\mathscr{O}$ and if $\mathscr{O}'$ is another topology on X, coarser than $\mathscr{O}$ but still Hausdorff, then $\mathscr{O} = \mathscr{O}'$.

A.3. Product spaces:
Let $(X_\alpha)_{\alpha\in A}$ a non-empty family of non-empty topological spaces. The product $X := \prod_{\alpha\in A} X_\alpha$ becomes a topological space if we construct a topology on X starting with the base of open rectangles, i.e. with sets of the form $\{x = (x_\alpha)_{\alpha\in A} : x_{\alpha_i} \in O_{\alpha_i} \text{ for } i = 1, \ldots, n\}$ for $\alpha_1, \ldots, \alpha_n \in A$, $n \in \mathbb{N}$ and O_{α_i} open in X_{α_i}. Then Tychonov's theorem asserts that for this topology, X is compact if and only if each X_α, $\alpha \in A$ is compact.

A.4. Urysohn's lemma:
Let X be compact and A, B disjoint closed subsets of X. Then there exists a continuous function $f : X \to [0, 1]$ with $f(A) \subseteq \{0\}$ and $f(B) \subseteq \{1\}$.

A.5. Lebesgue's covering lemma: If (X, d) is a compact metric space and α is is a finite open cover of X, then there exists a $\delta > 0$ such that every set $A \subseteq X$ with diameter diam $(A) < \delta$ is contained in some element of α.

A.6. Category: A subset A of a topological space X is called *nowhere dense* if the closure of A, denoted by $\overline{A}$, has empty interior: $\mathring{\overline{A}} = \emptyset$. A is called of *first category* in X if A is the union of countably many nowhere dense subsets of X. A is called of *second category* in X if it is not of first category. Now let X be a compact or a complete metric space. Then Baire's category theorem states that every non-empty open set is of second category.

(ii) Measure theory

Somewhat less elementary but even more important for ergodic theory is the concept of an abstract measure space. We shall use the standard approach to measure- and integration theory and refer to Bauer [1972] and Halmos [1950]. The advanced reader is also directed to Jacobs [l978]. Although we again assume that the reader is familiar with the basic results, we present a list of more or less known definitions and results.

A.7. Measure spaces and null sets:
A triple (X, Σ, μ) is a *measure space* if X is a set, Σ σ-algebra of subsets of X and μ a measure on Σ, i.e.

$$\mu : \Sigma \to \mathbb{R}_+ \cup \{\infty\}$$

σ-additive and and $\mu(\varnothing) = 0$.

If $\mu(X) < \infty$ (resp. $\mu(X) = 1$), X, Σ, μ is called a *finite* measure space (resp. a *probability space*); it is called σ*-finite*, if $X = \bigcup_{n \in \mathbb{N}} A_n$ with $\mu(A_n) < \infty$ for all $n \in \mathbb{N}$.

A set $N \subseteq \Sigma$ is a μ*-null set* if $\mu(N) = 0$.

Properties, implications, conclusions etc. are valid "μ-almost everywhere" or for "almost all $x \in X$" if they are valid for all $x \in X \backslash N$ where N is some μ-null set. If no confusion seems possible we sometimes write "... is valid for all x" meaning "...is valid for almost all $x \in X$".

A.8. Equivalent measures:
Let (X, Σ, μ) be a σ-finite measure space and ν another measure on Σ. ν is called *absolutely continuous* with respect to μ if every μ-null set is ν-null set. ν is equivalent to μ iff ν is absolutely continuous with respect to μ and conversely. The measures which are absolutely continuous with respect to μ can be characterized by the Radon-Nikodỳm theorem (see Halmos [1950], §31).

A.9. The measure algebra:
In a measure space (X, Σ, μ) the μ-null sets form a σ-ideal $\mathscr{N}$. The Boolean algebra

$$\check{\Sigma} := \Sigma / \mathscr{N}$$

is called the corresponding *measure algebra*. We remark that $\check{\Sigma}$ is isomorphic to the algebra of characteristic functions in $L^\infty(X, \Sigma, \mu)$ (see App.B.20) and therefore is a *complete Boolean algebra*.

For two subsets A, B of X,

$$A \triangle B := (A \cup B) \backslash (A \cap B) = (A \backslash B) \cup (B \backslash A)$$

denotes the symmetric difference of A and B, and

$$d(A, B) := \mu(A \triangle B)$$

defines a semi-metric on X vanishing on Σ the elements of $\mathscr{N}$ (if $\mu(X) < \infty$). Therefore we obtain a metric on $\check{\Sigma}$ still denoted by d.

A.10 Proposition: The measure algebra $(\check{Z}, d)$ of a finite measure space X, Σ, μ is a complete metric space.

Proof. It suffices to show that (Σ, d) is complete. For a Cauchy sequence $(A_n)_{n \in \mathbb{N}} \subseteq \Sigma$, choose a subsequence $(A_{n_i})_{i \in \mathbb{N}}$ such that $d(A_k, A_l) < 2^{-i}$ for $k, l > n_i$. Then $A := \bigcap_{m=1}^\infty \bigcup_{j=m}^\infty A_{n_j}$ is the limit of (A_n). Indeed, with $B_m := \bigcup_{j=m}^\infty A_{n_j}$ we have

$$d(B_m, A_{n_m}) \leqslant \sum_{j=m}^\infty \mu(A_{n_{j+1}} \backslash A_{n_j}) \leqslant \sum_{j=m}^\infty 2^{-j} = 2 \cdot 2^{-m}$$

and

$$
\begin{aligned}
d(A, B_m) &\leqslant \sum_{j=m}^{\infty} \mu(B_j \backslash B_{j+1}) \leqslant \sum_{j=m}^{\infty} \big(d(B_j, A_{n_j}) + d(A_{n_j}, A_{n_{j+1}}) + d(A_{n_{j+1}}, B_{j+1})\big) \\
&\leqslant \sum_{j=m}^{\infty} \big(2 \cdot 2^{-j} + 2^{-j} + 2 \cdot 2^{-(j+1)}\big) \leqslant 8 \cdot 2^{-m}.
\end{aligned}
$$

Therefore

$$d(A, A_k) \leqslant d(A, B_m) + d(B_m, A_{n_m}) + d(A_{n_m}, A_k) \leqslant 11 \cdot 2^{-m}$$

for $k \geqslant n_m$. ∎

A.11. For a subset $\breve{W}$ of $\check{\Sigma}$ we denote by $a(\breve{W})$ the *Boolean algebra generated* by $\breve{W}$, by $\sigma(\breve{W})$ the *Boolean σ-algebra generated* by $\breve{W}$.
$\check{\Sigma}$ is called countably generated, if there exists a countable subset $\breve{W} \subseteq \check{\Sigma}$ such that $\sigma(\breve{W}) = \check{\Sigma}$.
The metric d relates $a(\breve{W})$ and $\sigma(\breve{W})$. More precisely, using an argument as in (A.10) one can prove that in a finite measure space

$$\sigma(\breve{W}) = \overline{a(\breve{W})}^{d} \quad \text{for every } \breve{W} \subseteq \check{\Sigma}$$

.

A.12. The Borel algebra:
In many applications a set X bears a topological structure and a measure space structure simultaneously. In particular, if X is a compact space, we *always* take the σ-algebra $\mathscr{B}$ generated by the open sets, called the *Borel algebra* on X. The elements of $\mathcal{B}$ are called Borel sets, and a measure defined on $\mathcal{B}$ is a Borel measure. Further, we only consider *regular Borel measures*: here, μ is called regular if for every $A \in \mathcal{B}$ and $\varepsilon > 0$ there is a compact set $K \subseteq A$ and an open set $U \supseteq A$ such that $\mu(A \backslash K) < \varepsilon$ and $\mu(U \backslash A) < \varepsilon$.

A.13 Example:

Let $X = [0, 1][$ be endowed with the usual topology. Then the Borel algebra $\mathcal{B}$ is generated by the set of all dyadic intervals

$$\mathscr{D} := \big\{[k \cdot 2^{-i}, (k+1) \cdot 2^{-i}] : i \in \mathbb{N}; k = 0, \ldots, 2^i - 1\big\}.$$

$\mathscr{D}$ is called a *separating base* because it generates $\mathcal{B}$ and for any $x, y \in X$, $x \neq y$, there is $D \in \mathscr{D}$ such that $x \in D$ and $y \notin D$, or $x \notin D$ and $y \in D$.

A.14. Measurable mappings:
Consider two measure spaces (X, Σ, μ) and (Y, T, ν). A mapping $\varphi : X \to Y$ is called *measurable*, if $\varphi^{-1}(A) \in \Sigma$ for every $A \in T$, and called *measure-preserving*, if, in addition, $\mu(\varphi^{-1}(A)) = \nu(A)$ for all $A \in T$ for all $A \in T$ (abbreviated: $\mu \circ \varphi^{-1} = \nu$).

For real-valued measurable functions f and g on (X, Σ, μ), where $\mathbb{R}$ is endowed with the Borel algebra, we use the following notation:

$$
\begin{aligned}
[f \in B] &:= f^{-1}(B) \quad \text{for } B \in \mathcal{B}, \\
[f = g] &= \{x \in X : f(x) = g(x)\}, \\
[f \leqslant g] &:= \{x \in X : f(x) \leqslant g(x)\}.
\end{aligned}
$$

Finally,

$$\mathbf{1}_A : x \mapsto \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad \text{denotes the characteristic}$$

function of $A \subseteq X$. If $A = X$, we often write $\mathbf{1}$ instead of $\mathbf{1}_X$.

A.15. Continuous vs. measurable functions:
Let X be compact, $\mathcal{B}$ the Borel algebra on X and μ a regular Borel measure. Clearly, every continuous function $f : X \to \mathbb{C}$ is measurable for the corresponding Borel algebras. On the other hand there is a partial converse:

Theorem (Lusin): Let $f : X \to \mathbb{C}$ be measurable and $\varepsilon > 0$. Then there exists a compact set $A \subseteq X$ such that $\mu(X \backslash A) < \varepsilon$ and f is continuous on A.

Proof (Feldman [1981]): Let $\{U_j\}_{j\in\mathbb{N}}$ be a countable base of open subsets of $\mathbb{C}$. Let V_j be open such that $f^{-1}(U_j) \subseteq V_+j$ and $\mu(VV_j \backslash f^{-1}(U_j)) < \frac{\varepsilon}{2} 2^{-j}$. If we take $B := \bigcup_{j=1}^{\infty} (V_j \backslash f^{-1}(U_j))$, we obtain $\mu(B) < \frac{\varepsilon}{2}$, and we show that $g := f|_{B^c}$ is continuous. To this end observe that

$$\begin{aligned} V_j \cap B^c &= V_j \cap (V_j \backslash f^{-1}(U_j))^c \cap B^c = V_j \cap (V_j^c \cup f^{-1}(U_j)) \cap B^c \\ &= V_j \cap f^{-1}(U_j) \cap B^c = f^{-1}(U_j) \cap B^c = g^{-1}(U_j). \end{aligned}$$

Since any open subset U of $\mathbb{C}$ can be written as $U = \bigcup_{j\in M} U_j$, we have $G^{-1}(U) = \bigcup_{j\in M} g^{-1}(U_j) = \bigcup_{j\in M} V_j \cap B^c$, which is open in B^c. Now we choose a compact set $A \subseteq B^c$ with $\mu(B^c \backslash A) < \frac{\varepsilon}{2}$, and conclude that f is continuous on A and that $\mu(X \backslash A) = \mu(B) + \mu(B^c \backslash A) < \varepsilon$. ∎

A.16. Convergence of integrable functions:
Let (X, Σ, μ) be a finite measure space and $1 \leqslant p < \infty$. A measurable (real) function f on X is called *p-integrable*, if $\int |f|^p \, d\mu < \infty$ (see Bauer [1972], 2.6.3).

For sequences $(f_n)_{n\in\mathbb{N}}$ of p-integrable functions we have three important types of convergence:

1. $(f_n)_{n\in\mathbb{N}}$ converges to f *μ-almost everywhere* if
$$\lim_{n\to\infty} (f_n(x) - f(x)) = 0 \quad \text{for almost all } x \in X.$$
2. $(f_n)_{n\in\mathbb{N}}$ converges to f *in the p-norm* if
$$\lim_{n\to\infty} \int |f_n - f|^p \, d\mu = 0 \quad \text{see (B.20).}$$
3. $(f_n)_{n\in\mathbb{N}}$ converges to f *μ-stochastically* if
$$\lim_{n\to\infty} \mu[|f_n - f| \geqslant \varepsilon] = 0 \quad \text{for every } \varepsilon > 0.$$

Proposition: Let $(f_n)_{n\in\mathbb{N}}$ be p-integrable functions and f be measurable.

(i) If $f_n \to f$ μ-almost everywhere or in the p-norm, then $f_n \to f$ μ-stochastically (see Bauer [1972], 2.11.3 and 2.11.4).
(ii) If $(f_n)_{n\in\mathbb{N}}$ converges to f in the p-norm, then there exists a subsequence (f_{n_k}) converging to f μ-a.e. (see Bauer [1972], 2.7.5).
(iii) If $(f_n)_{n\in\mathbb{N}}$ converges to f μ-a.e. and if there is a p-integrable function g such that $|f_n(x)| \leqslant g(x)$ μ-a.e., then $f_n \to f$ in the p-norm and f is p-integrable (Lebesgue's dominated convergence theorem, see Bauer [1972], 2.7.4).

Simple examples show that in general no other implications are valid.

A.17. Product spaces:
Given a countable family $(X_\alpha, \Sigma_\alpha, \mu_\alpha)_{\alpha\in A}$ of probability spaces, we can consider the cartesian product $X = \prod_{\alpha\in A}$ and the so-called product σ-algebra $\Sigma = \bigotimes_{\alpha\in A}\Sigma_\alpha$ which is generated by the set of all *measurable rectangles*, i.e. sets of the form

$$R_{\alpha_1,\ldots,\alpha_n}(A_{\alpha_1},\ldots,A_{\alpha_n}) := \left\{x = (x_\alpha)_{\alpha\in A} : x_{\alpha_i} \in A_{\alpha_i} \text{ for } i = 1,\ldots,n\right\}$$

for $\alpha_1,\ldots,\alpha_n \in A$, $n \in \mathbb{N}$, $A_{\alpha_i} \in \Sigma_{\alpha_i}$.

The well known extension theorem of Hahn-Kolmogorov implies that there exists a unique probability measure $\mu : \bigotimes_{\alpha\in A}\mu_\alpha$ on Σ such that

$$\mu(R_{\alpha_1,\ldots,\alpha_n}(A_{\alpha_1},\ldots,A_{\alpha_n})) = \prod_{i=1}^{n}\mu_{\alpha_i}(A_{\alpha_i})$$

for every measurable rectangle (see Halmos [1950], §383 Theorem B).

Then X, Σ, μ is called the *product (measure) space* defined by $(X_\alpha, \Sigma_\alpha, \mu_\alpha)_{\alpha\in A}$

Finally, we mention an extension theorem dealing with a different situation (see also Ash [1972], Theorem 5.11.2).

Theorem: Let $(X_n)_{n\in\mathbb{Z}}$ be a sequence of compact spaces, $\mathcal{B}_n$ the Borel algebra on X_n. Further, we denote by Σ the product σ-algebra on $X = \prod_{n\in\mathbb{Z}} X_n$, by $\mathscr{F}_m$ the set of all measurable sets in X whose elements depend only on the coordinates $-m,\ldots,0,\ldots,m$. Finally we put $\mathscr{F} = \bigcup_{m\in\mathbb{N}}\mathscr{F}_m$. If μ is a function on $\mathscr{F}$ such that it is a regular probability measure on $\mathscr{F}_m$ for each $m \in \mathbb{N}$, then μ has a unique extension to a probability measure on Σ.

Remark: Let $\varphi_n : X \to Y_n := \prod_{-n}^{n} X_i;\ (x_j)_{j\in\mathbb{Z}} \mapsto (x_{-n},\ldots,x_n)$. Then we assume above that $\nu_n(A) := \mu(\varphi_n^{-1}(A))$, A measurable in Y_n, defines a regular Borel probability measure on Y_n for every $n \in \mathbb{N}$.

Proof. The set function μ has to be extended from $\mathscr{F}$ to $\sigma(\mathscr{F}) = \Sigma$. By the classical Carathèodory extension theorem (see Bauer [1972], 1.5) it suffices to show that $\lim_{i\to\infty}\mu(C_i) = 0$ for any decreasing sequence $(C_i)_{i\in\mathbb{N}}$ of sets in $\mathscr{F}$ satisfying $\bigcap_{i\in\mathbb{N}} C_i = \varnothing$. Assume that $\mu(C_i) \geqslant \varepsilon$ for all $i \in \mathbb{N}$ and some $\varepsilon > 0$. For each C_i there is an $n \in \mathbb{N}$ such that $C_i \in \mathscr{F}_n$ and $A_i \subseteq Y_n$ with $C_i = \varphi_n^{-1}(A_i)$. Let B_i a closed subset of A_i such that $\nu_n(A_i\backslash B_i) \leqslant \frac{\varepsilon}{2}\cdot 2^{-i}$. Then $D_i := \varphi_n^{-1}(B_i)$ is compact in X and $\mu(C_i\backslash D_i) \leqslant \frac{\varepsilon}{2}\cdot 2^{-i}$. Now the sets $G_k := \bigcap_{i=1}^{k} D_i$ form a decreasing sequence of compact subsets of X, and we have

$$\begin{aligned} G_k \subseteq C_k \text{ and } \mu(G_k) &= \mu(C_k) - \mu(C_k\backslash G_k) = \mu(C_k) - \mu\Big(\bigcup_{i=1}^{k}(C_i\backslash D_i)\Big) \\ &\geqslant \mu(C_k) - \sum_{i=1}^{k}\mu(C_i\backslash D_i) \geqslant \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}. \end{aligned}$$

Hence $G_k \neq \varnothing$ and therefore $\bigcap_{i\in\mathbb{N}} C_i$, which contains $\bigcap_{i\in\mathbb{N}} G_i$, is non-empty, a contradiction. ∎