1 The limit of a sequence of probability spaces, from scratch.

For each $n \in \mathbb{N}$ let

- X_n be a nonempty set.
- $\mathcal{B}_n \subseteq \mathcal{P}(X)$ an algebra.
- $\mu_n : \mathcal{B}_n \to [0, 1]$ a finitely additive probability measure.
- $T_n : X_n \to X_n$ measure-preserving, that is, for all $A \in \mathcal{B}_n$, $T_n^{-1}(A) \in \mathcal{B}_n$ and $\mu_n(T_n^{-1}(A)) = \mu_n(A)$.

In the following we explain a reasonable way to construct limits of the sequences $X_n, \mathcal{B}_n, \mu_n$ and T_n, which we suggestively denote as $X_\infty, \mathcal{B}_\infty, \mu_\infty$ and T_∞. As a plus, the measure space $(X_\infty, \mathcal{B}_\infty, \mu_\infty)$ will be countably additive. Thus, for any finitely additive probability measure space we can construct a (countably additive) probability measure space with similar properties, see Theorem 1.12.

The results presented here are well-known; this document aims to be a self-contained, elementary exposition of them, without requiring previous knowledge of non-standard analysis. The construction of the 'limit measure' we use below is known as a Loeb measure, and was discovered by Loeb in [L]. For other mentions of Loeb measures in the literature, see for example [DW], [AB, Definition 7.5], or [AEHL, Section 3.1].

The main ingredient in our constructions is a non-principal ultrafilter $\mathcal{F} \subseteq \mathcal{P}(\mathbb{N})$ in $\mathbb{N}$, that is, a family of sets such that:

1. $\varnothing \notin \mathcal{F}$.
2. If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$. If $A, B \notin \mathcal{F}$, then $A \cup B \notin \mathcal{F}$.
3. If $A \in \mathcal{F}$ and $A \subseteq B \subseteq \mathbb{N}$, then $B \in \mathcal{F}$.
4. For every $A \subseteq \mathbb{N}$, exactly one of A or $\mathbb{N} \setminus A$ belongs to $\mathcal{F}$.
5. $\mathcal{F}$ is non-principal, that is, it contains no finite set.

One can prove the existence of non-principal ultrafilters using the axiom of choice, see e.g. The existence of a non-principal ultrafilter needs the axiom of choice to be proved, or some equivalent form of it. One can check the following using the definition of ultrafilter:

Proposition 1.1. *Let $\mu_\mathcal{F} : \mathcal{P}(\mathbb{N}) \to \{0, 1\}$; $\mu(A) = 1$ if $A \in \mathcal{F}$ and $\mu(A) = 0$ if $A \notin \mathcal{F}$. Then $\mu_\mathcal{F}$ is a finitely additive probability measure in $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.* $\square$

So ultrafilters give rise to $\{0, 1\}$-valued, finitely additive probability measures. When we say 'almost all n satisfies some property P', we mean that the set of numbers $n \in \mathbb{N}$ that satisfy P is in $\mathcal{F}$.

Definition 1.2 (Ultraproduct of sets)**.** The *ultraproduct* $X_\infty := \lim_{n\to\mathcal{F}} X_n$ is the quotient set $\frac{\prod_{n\in\mathbb{N}} X_n}{\sim}$, where $(x_n)_n \sim (y_n)_n$ iff $x_n = y_n$ for almost all n. We denote by $[y_n]_n \in \lim_{n\to\mathcal{F}} X_n$ the class of an element $(y_n)_n$.

Definition 1.3 (Internal subsets)**.** Given a sequence of sets $A_n \subseteq X_n$, we denote $\lim_{n\to\mathcal{F}} A_n = \{[x_n]_n \in X_\infty; x_n \in A_n \text{ for almost all } n\}$. Subsets of X_∞ which can be obtained this way are called *internal*.

The following properties follow from the definitions:

1. $\lim_{n\to\mathcal{F}} A_n \cup B_n = (\lim_{n\to\mathcal{F}} A_n) \cup (\lim_{n\to\mathcal{F}} B_n)$.
2. $\lim_{n\to\mathcal{F}} A_n \cap B_n = (\lim_{n\to\mathcal{F}} A_n) \cap (\lim_{n\to\mathcal{F}} B_n)$.
3. $\lim_{n\to\mathcal{F}} A_n \setminus B_n = (\lim_{n\to\mathcal{F}} A_n) \setminus (\lim_{n\to\mathcal{F}} B_n)$.
4. $\lim_{n\to\mathcal{F}} A_n = \lim_{n\to\mathcal{F}} B_n$ iff $A_n = B_n$ for $\mathcal{F}$-almost all n.
5. $\lim_{n\to\mathcal{F}} A_n = \varnothing$ iff $A_n = \varnothing$ for almost all n.

In particular, internal subsets are an algebra of subsets of X_∞. In order to define our measure μ_∞ we will need the following result.

Proposition 1.4. *For each $k \in \mathbb{N}$ let $A^k = \lim_{n\to\mathcal{F}} A_n^k$ be an internal subset of X_∞. If the sets A^k are nonempty and pairwise disjoint, then $A := \cup_{k\in\mathbb{N}} A_k$ is not internal.*

Proof. Suppose that $A = \lim_n A_n$ for some sets $A_n \subseteq X_n$. For each $k, n \in \mathbb{N}$ let

$$B_n^k := A_n \cap \left(A_n^k \setminus (A_n^1 \cup \cdots \cup A_n^{k-1})\right).$$

Then for each fixed k, the fact that $\varnothing \neq A \cap \left(A^k \setminus (A^1 \cup \cdots \cup A^{k-1})\right)$ implies that $B_n^k \neq \varnothing$ for almost all n. We define a point $[x_n]_n \in A$ by letting $x_n \in X_n$ be some point of $B_n^{k_n}$, where k_n is given by:

1. If $B_n^j \neq \varnothing$ for finitely many values of j, let k_n be the maximum such j.
2. If $B_n^j \neq \varnothing$ for infinitely many j, let k_n satisfy $k_n > n$ and $B_n^{k_n} \neq \varnothing$.

There may be some values of n such that $B_n^k = \varnothing$ for all k. But the set of such values n is $\mathcal{F}$-small, so we can choose x_n however we want in that case.

We then have $[x_n]_n \in A$, because $x_n \in B_n^{k_n} \subseteq A_n$ for almost all n.

But for each fixed value of k, $[x_n]_n \notin A_k$. Indeed, for almost all n we have $k_n > k$ (this is obvious in Item 2, and in Item 1 it follows from the fact that $B_n^{k+1} \neq \varnothing$ for almost all n), so $x_n \in B_n^{k_n} \subseteq A_n \setminus A_n^k$, so $x_n \notin A_n^k$. $\square$

We also need to define what $\lim_{n\to\omega} x_n$ means, when $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence of real/complex numbers.

Proposition 1.5. *For any sequence $(x_n)_{n\in\mathbb{N}}$ in a compact Hausdorff space X, there exists a unique $x \in X$, which we denote*

$$\lim_{n\to\mathcal{F}} x_n,$$

such that for all neighborhoods U of x, the set $\{n \in \mathbb{N}; x_n \in U\}$ is in $\mathcal{F}$.

Proof. Existence of the limit: Suppose for contradiction that for all $x \in X$ there is a nhood U_x such that $\mu_{\mathcal{F}}(\{n \in \mathbb{N}; x_n \in U\}) = 0$. Take a finite cover $X = U_1 \cup \cdots \cup U_n$ of X by such neighborhoods. Then, we have $\mathbb{N} = \cup_{i=1}^n \{n \in \mathbb{N}; x_n \in U_i\}$, a contradiction as $\mu_{\mathcal{F}}(\mathbb{N}) > 0$.

Uniqueness of the limit: Suppose there are two 'limit points' $x \neq y$ in X satisfying the property above. Let U_x, U_y be disjoint neighborhoods of x, y. Then $\{n \in \mathbb{N}; x_n \in U_x\}$ and $\{n \in \mathbb{N}; x_n \in U_y\}$ have $\mu_{\mathcal{F}}$-measure 1, a contradiction as they are disjoint. $\square$

If a sequence (x_n) of complex numbers is bounded, then we can interpret it as a sequence in a compact subset of $\mathbb{C}$ and define $\lim_{n\to\omega} x_n$ according to Theorem 1.5; the limit will only depend on the sequence, not on the compact set we choose.

Definition 1.6. Let $\mathcal{A}_\infty$ be the algebra of all internal sets of the form $\lim_{n\to\mathcal{F}} A_n$, where $A_n \in \mathcal{B}_n$ for all n. Let $\mathcal{B}_\infty \subseteq \mathcal{P}(X_\infty)$ be the σ-algebra generated by $\mathcal{A}_\infty$.

Proposition 1.7. *The map $\mu : \mathcal{A}_\infty \to [0,1]$; $\mu(\lim_{n\to\mathcal{F}} A_n) = \lim_{n\to\mathcal{F}} \mu_n(A_n)$ is a pre-measure, thus it extends to a probability measure $\mu_\infty : \mathcal{B}_\infty \to [0,1]$.*

Proof. We omit checking that μ is well defined. Now suppose we have disjoint sets $A^1, A^2, \ldots$ in $\mathcal{A}_\infty$, with $A_i = \lim_{n\to\mathcal{F}} A_n^i$. Note that for $i \neq j$ we have $A_n^i \cap A_n^j = \varnothing$ for almost all n. And suppose that $A := \cup_{k\in\mathbb{N}} A^k \in \mathcal{A}_\infty$ (say, $A = \lim_{n\to\mathcal{F}} A_n$), then by Theorem 1.4, A must be a finite union of the sets A^k, say $A = \cup_{k=1}^K A^k$, and in particular $A^n = \varnothing$ for $n > K$). So

$$\begin{aligned}\mu(A) = \mu\left(\cup_{k=1}^K A^k\right) = \mu\left(\lim_{n\to\mathcal{F}} (A_n^1 \cup \cdots \cup A_n^K)\right) &= \lim_{n\to\mathcal{F}} \mu_n(A_n^1 \cup \cdots \cup A_n^K) \\ &= \sum_{k=1}^K \lim_{n\to\mathcal{F}} \mu_n(A_n^k) = \sum_{k=1}^K \mu(A^k) = \sum_{k=1}^\infty \mu(A^k).\end{aligned}$$

So as we wanted, μ is a premeasure. So by Caratheodory's extension theorem, it extends to a probability measure μ_∞ in $(X_\infty, \mathcal{B}_\infty)$. $\square$

The following again follows from the definitions.

Proposition 1.8 (Limits of measure preserving actions)**.** *If $T_n : X_n \to X_n$ is a measure preserving map $\forall n$, then the map $T_\infty : X_\infty \to X_\infty; [x_n]_n \mapsto [Tx_n]_n$ satisfies $T_\infty^{-1}(\lim_{n\to\mathcal{F}} A_n) = \lim_{n\to\mathcal{F}} T_n^{-1}(A_n)$, so it is measurable and measure preserving.* $\square$

Remark 1.9. The limit of a sequence of ergodic measure-preserving systems need not be ergodic. E.g. let $\mathbb{T}$ be the torus, i.e. the quotient $\frac{\mathbb{R}}{\mathbb{Z}}$ with Lebesgue measure, and $T_n : \mathbb{T} \to \mathbb{T}$ be given by $T_n(x) = x + \frac{1}{n} \bmod 1$. Consider the map $T_\infty : \mathbb{T}_\infty \to \mathbb{T}_\infty$, and $A = \lim_{n\to\mathcal{F}}[0, 1/2]$. Then $\mu_\infty(A) = 0$, and $\mu_\infty(A\Delta T_\infty^{-1}(A)) = 0$, so T_∞ is not ergodic. T_∞ is not trivial either, in fact there exist internal sets B such that $\mu_\infty(B\Delta T_\infty^{-1}(B)) = 1$.

Definition 1.10. Given a sequence $(X_n, \mathcal{B}_n, \mu_n, T_n)$ of finitely additive measure preserving systems, we denote by $\lim_{n\to\mathcal{F}}(X_n, \mathcal{B}_n, \mu_n, T_n)$ the measure preserving system $(X_\infty, \mathcal{B}_\infty, \mu_\infty, T_\infty)$, where $X_\infty, \mathcal{B}_\infty, \mu_\infty, T_\infty$ are constructed as in Theorem 1.4, Theorem 1.6,Theorem 1.7 and Theorem 1.8.

We can consider ultralimits of L^∞ functions:

Proposition 1.11. *Let $f_n : X_n \to \mathbb{D}$ be measurable for all n. Then the limit function $f : \lim_{n\to\mathcal{F}} X_n \to \mathbb{D}$ given by $f([x_n]_n) = \lim_{n\to\mathcal{F}} f_n(x_n)$ is measurable, and satisfies*

$$\int_{X_\infty} f d\mu_\infty = \lim_{n\to\mathcal{F}} \int_{X_n} f_n d\mu_n. \tag{1}$$

Note that the preimage of a measurable set need not be internal! But it is still measurable.

Proof. For the first part it is enough to check that the preimage of any closed set is measurable. And indeed, for any closed $C \subseteq \mathbb{D}$ and letting C_k be the $\frac{1}{k}$-nhood of C, we have

$$f^{-1}(C) = \bigcap_{k\in\mathbb{N}} \lim_{n\to\mathcal{F}} f_n^{-1}(C_k).$$

To prove Equation (1) it is enough to prove the case where all the functions f_n take values in some finite set $D_0 \subseteq \mathbb{D}$ (and then we approximate arbitrary functions in the L^∞ norm).

But if $f_n : X_n \to D_0$ for all n, then f also takes values in D_0, and for each $d \in D_0$ we have $f^{-1}(d) = \lim_{n\to\infty} f_n^{-1}(d)$. Thus,

$$\begin{aligned}\int_{X_\infty} fd \cdot \mu_\infty &= \sum_{d\in D_0} d \cdot \mu_\infty(f^{-1}(d)) = \sum_{d\in D_0} d \cdot \lim_{n\to\mathcal{F}} \mu_n(f_n^{-1}(d)) \\ &= \lim_{n\to\mathcal{F}} \sum_{d\in D_0} d \cdot \mu_n(f_n^{-1}(d)) = \lim_{n\to\mathcal{F}} \int_{X_n} f_n d\mu_n.\end{aligned}$$

□

1.1 From finitely additive to countably additive measures

By letting $(X_n, \mathcal{B}_n, \mu_n)$ be a fixed space $(X, \mathcal{B}, \mu)$ for all n, the following follows from the results in the previous section:

Proposition 1.12. *Let $(X, \mathcal{B}, \mu)$ be a finitely additive probability measure space. Then we have a (countably additive) probability space $(\overline{X}, \overline{\mathcal{B}}, \overline{\mu}) := \lim_{n\to\mathcal{F}}(X, \mathcal{B}, \mu)$, and an injective map $\mathcal{B} \to \overline{\mathcal{B}}; A \mapsto \overline{A} := \lim_{n\to\mathcal{F}} A$, which satisfies for all $A, B \in \mathcal{B}$ that*

1. $\overline{\mu}(\overline{A}) = \mu(A)$.

2. $\overline{A \cup B} = \overline{A} \cup \overline{B}$, $\overline{A \cap B} = \overline{A} \cap \overline{B}$.

3. $\overline{\varnothing} = \varnothing$.

4. $\overline{X \setminus A} = \overline{X} \setminus \overline{A}$.

If $(X, \mathcal{B}, \mu)$ *and* $(Y, \mathcal{C}, \nu)$ *are finitely additive probability spaces and* $T : X \to Y$ *is measure preserving, then it has an associated map* $\overline{T} : \overline{X} \to \overline{Y}$*, given by* $\overline{T}([x_n]_n) = [Tx_n]_n$*, such that:*

1. $\overline{T}^{-1}\left(\overline{A}\right) = \overline{T^{-1}(A)}$ *for all* $A \in \mathcal{C}$.

2. If $T : X \to Y$ *and* $S : Y \to Z$ *are measure-preserving, then* $\overline{S \circ T} = \overline{S} \circ \overline{T}$.

In particular, the assignment $(X, \mathcal{B}, \mu) \to (\overline{X}, \overline{\mathcal{B}}, \overline{\mu})$ *can be seen a functor.* □

References

[AB] Ackelsberg, Bergelson. *Furstenberg-Sarkozy theorem and partition regularity of polynomial equations over finite fields.* arXiv preprint, 2023.

[AEHL] Albeverio, Høegh-Krohn, Fenstad, Lindstrøm. *Nonstandard methods in stochastic analysis and mathematical physics.* Courier Dover Publications, 2009.

[L] Loeb. *Conversion from nonstandard to standard measure spaces and applications in probability theory.* Trans. Amer. Math. Soc. 211 (1975), 113–122.

[DW] Duanmu, Weiss. *Finitely-additive, countably-additive and internal probability measures.* Comment. Math. Univ. Carolin. 59,4 (2018) 467–485.