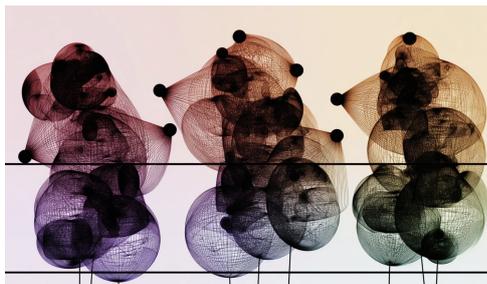# Finitely Additive Measures and Filters

Stephen J. Mildenhall

2026-03-15

1 Background: filters, ideals, and ultrafilters

Filters and ideals formalize the distinction between "large" and "small" sets. A filter records which sets are regarded as large, and an ideal those regarded as small. They are dual notions[1]. Large objects are what remains after filtering.

Throughout, let Ω be a nonempty set.

1.1 Filters: the large sets

A **filter** $\mathcal{F}$ on Ω is a collection of subsets of Ω such that

1. $\emptyset \notin \mathcal{F}$,
2. if $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$,
3. if $A \in \mathcal{F}$ and $A \subseteq B \subseteq \Omega$, then $B \in \mathcal{F}$.

A filter is a family of sets closed upward under inclusion and closed under finite intersections. If a set is large, then any larger set is also large. If two sets are large, then their overlap is still large.

The simplest example is a principal filter. Fix a point $\omega_0 \in \Omega$. Then

$$\mathcal{F}_{\omega_0} = \{\, A \subseteq \Omega : \omega_0 \in A \,\}$$

is a filter. Here "large" simply means "contains ω_0". More generally, if $E \subseteq \Omega$ is nonempty, then

$$\mathcal{F}_E = \{\, A \subseteq \Omega : E \subseteq A \,\}$$

is the filter generated by E. A filter is called **principal** if it is of this form for some nonempty E.

In a probability space, sets of measure 1 form a filter[2]. The dual is the idea of null sets.

---

[1] Filters and ideals are dual via complementation:

$$A \leftrightarrow A^c.$$

A filter identifies the large sets; the corresponding ideal identifies the small sets, namely those whose complements are large:

$$\mathcal{I}_{\mathcal{F}} = A : A^c \in \mathcal{F}.$$

De Morgan's laws turn finite intersections into finite unions, and upward closure into downward closure, so the filter axioms become the ideal axioms, and conversely. This is a Boolean algebra duality, not a ring-theoretic one in the sense of classical Herstein-style algebra: the key ingredient is the complement operation in the Boolean algebra of sets.

[2] If $\lambda(A) = \lambda(B) = 1$, then

$$\lambda(A^c) = \lambda(B^c) = 0.$$

1.2 Ideals: the small sets

Dually, an **ideal** $\mathcal{I}$ on Ω is a collection of subsets of Ω such that

1. $\emptyset \in \mathcal{I}$,
2. if $A, B \in \mathcal{I}$ then $A \cup B \in \mathcal{I}$,
3. if $A \in \mathcal{I}$ and $B \subseteq A$, then $B \in \mathcal{I}$.

An ideal is closed downward under inclusion and closed under finite unions. If a set is small, then any smaller set is also small. If two sets are small, then their union is still small.

The archetypal example is the ideal of finite subsets of $\mathbb{N}$:

$$\mathrm{Fin} = \{ A \subseteq \mathbb{N} : A \text{ is finite} \}.$$

Filters and ideals are complementary. Given a filter $\mathcal{F}$, one may think of the sets outside $\mathcal{F}$ as "not definitely large", but the cleaner dual object is

$$\mathcal{I}_{\mathcal{F}} = \{ A \subseteq \Omega : A^c \in \mathcal{F} \}.$$

If $\mathcal{F}$ is an ultrafilter, this becomes an ideal of negligible sets.

1.3 The Fréchet filter on $\mathbb{N}$

A basic non-principal example is the Fréchet filter on $\mathbb{N}$:

$$\mathcal{F}^{\mathrm{cof}} = \{ A \subseteq \mathbb{N} : A^c \text{ is finite} \}.$$

These are the cofinite sets. A set is large if it contains all but finitely many integers. This filter is not principal. No finite set of indices determines it, and indeed no single point plays a distinguished role. This is the first hint that non-principal filters capture an asymptotic notion of largeness.

1.4 Generated filters

If $\mathcal{A}$ is a family of subsets of Ω with the finite intersection property, meaning that every finite intersection of members of $\mathcal{A}$ is nonempty, then $\mathcal{A}$ generates a filter:

$$\langle \mathcal{A} \rangle = \{ B \subseteq \Omega : \exists A_1, \dots, A_m \in \mathcal{A} \text{ with } A_1 \cap \cdots \cap A_m \subseteq B \}.$$

In words, a set is large if it contains some finite intersection of the prescribed large sets. If the family $\mathcal{A}$ is nested decreasing this simplifies: the generated filter consists of all sets that contain some A_n.

1.5 Ultrafilters: maximal notions of largeness

An **ultrafilter** $\mathcal{U}$ on Ω is a filter that is maximal among proper filters. Equivalently, $\mathcal{U}$ is an ultrafilter if for every subset $A \subseteq \Omega$,

$$A \in \mathcal{U} \quad \text{or} \quad A^c \in \mathcal{U},$$

and exactly one of these holds[3]. This dichotomy is what makes ultrafilters powerful. Every set is either large or its complement is large; there is no middle ground.

Hence

$$(A \cap B)^c = A^c \cup B^c$$

has measure 0, so

$$\lambda(A \cap B) = 1.$$

[3] A proper (non-trivial) filter $\mathcal{U}$ on Ω is an ultrafilter if and only if for every $A \subseteq \Omega$, exactly one of A and A^c belongs to $\mathcal{U}$. To see that this condition implies maximality, suppose $\mathcal{U} \subseteq \mathcal{F}$ where $\mathcal{F}$ is a proper filter. If $A \in \mathcal{F}$ and $A \notin \mathcal{U}$, then by assumption $A^c \in \mathcal{U} \subseteq \mathcal{F}$. Hence $A \cap A^c = \emptyset \in \mathcal{F}$, impossible. So every $A \in \mathcal{F}$ already lies in $\mathcal{U}$, and therefore $\mathcal{F} = \mathcal{U}$. Conversely, suppose $\mathcal{U}$ is an ultrafilter, and let $A \subseteq \Omega$. If neither A nor A^c belongs to $\mathcal{U}$, then one may adjoin A to $\mathcal{U}$ and generate a larger proper filter, contradicting maximality. Indeed, if adding A made the filter improper, there would exist $U \in \mathcal{U}$ with

$$U \cap A = \emptyset,$$

so $U \subseteq A^c$. Since filters are upward closed, this would imply $A^c \in \mathcal{U}$, contrary to assumption. Thus at least one of A or A^c lies in $\mathcal{U}$. They cannot both lie in $\mathcal{U}$, because then $\emptyset = A \cap A^c$ would lie in $\mathcal{U}$. So exactly one does.

A filter leaves many sets undecided. An ultrafilter is what you get by adjoining as many large sets as possible without ever forcing the empty set to become large.

A **principal ultrafilter** is one of the form

$$\mathcal{U}_{\omega_0} = \{\, A \subseteq \Omega : \omega_0 \in A \,\}.$$

On an infinite set there are also non-principal ultrafilters, but their existence requires a choice principle. A filter on an infinite set is called **free** if it is not principal.

1.6 Existence of ultrafilters

A standard theorem says:

Theorem 1.1. *Every proper filter on a set Ω is contained in an ultrafilter*[4].

Equivalently, every family of sets with the finite intersection property (the intersection of any finite subcollection is non-empty) extends to an ultrafilter. This is the ultrafilter lemma. It is weaker than the full axiom of choice, but it is not provable in ordinary ZF alone. For our purposes, we use it as a standard existence result. It exactly what we need later. Once we identify a natural filter of "near-maximizer" sets, an ultrafilter extension turns that vague asymptotic notion into a sharp yes/no notion of largeness.

If $\mathcal{F}$ is a proper filter, the ultrafilter lemma says there exists at least one ultrafilter $\mathcal{U}$ with

$$\mathcal{F} \subseteq \mathcal{U}.$$

But in general the extension is very far from unique. An ultrafilter extension is a way of making a yes/no decision about every set not already decided by $\mathcal{F}$, and there are usually many such decisions available.

For the cofinite filter on $\mathbb{N}$, this is especially dramatic. The cofinite filter only declares sets with finite complement to be large. But there are infinitely many other sets—evens, odds, primes, unions of blocks, and so on—for which neither the set nor its complement is cofinite. An ultrafilter extending the cofinite filter must choose exactly one of each such pair

$$A, \quad A^c,$$

and must do so consistently with finite intersections. There are therefore many different extensions.

Indeed, on $\mathbb{N}$ there are not just many but enormously many non-principal ultrafilters extending the cofinite filter: in fact

$$2^{2^{\aleph_0}}$$

---

[4] Proof sketch via Zorn's lemma. Let

$$\mathcal{P} := \{\, \mathcal{G} : \mathcal{G} \text{ is a proper filter on } \Omega \text{ and } \mathcal{F} \subseteq \mathcal{G} \,\},$$

ordered by inclusion. We want a maximal element of $\mathcal{P}$. Step 1: $\mathcal{P}$ is nonempty because it contains $\mathcal{F}$ itself. Step 2: every chain has an upper bound. Let $\{\, \mathcal{G}_i \,\}_{i \in I}$ be a chain in $\mathcal{P}$, meaning the filters are totally ordered by inclusion. Define

$$\mathcal{G} := \bigcup_{i \in I} \mathcal{G}_i.$$

Then $\mathcal{G}$ is again a proper filter containing $\mathcal{F}$ since

- It contains $\mathcal{F}$, since every $\mathcal{G}_i$ does.
- It is upward closed: if $A \in \mathcal{G}$, then $A \in \mathcal{G}_i$ for some i, and any superset of A lies in $\mathcal{G}_i$, hence in $\mathcal{G}$.
- It is closed under finite intersections: if $A, B \in \mathcal{G}$, then $A \in \mathcal{G}_i$ and $B \in \mathcal{G}_j$ for some i, j. Since the family is a chain, one of these filters contains the other; say $\mathcal{G}_i \subseteq \mathcal{G}_j$. Then both A and B lie in $\mathcal{G}_j$, so

$$A \cap B \in \mathcal{G}_j \subseteq \mathcal{G}.$$

- It is proper: no $\mathcal{G}_i$ contains $\emptyset$, so neither does the union.

Thus $\mathcal{G}$ is an upper bound for the chain. Step 3: apply Zorn's lemma: every chain in $\mathcal{P}$ has an upper bound. By Zorn's lemma, $\mathcal{P}$ has a maximal element, say $\mathcal{U}$. Step 4: maximal proper filter is an ultrafilter: By definition, $\mathcal{U}$ is maximal among proper filters containing $\mathcal{F}$, so it is an ultrafilter.

The only slightly delicate step is taking the union of a chain of filters. The union of arbitrary filters need not be a filter, but for a chain it is, because when you take two sets from the union, they already lie together in one member of the chain.

You can also state the result as:

If a family $\mathcal{A}$ of subsets of Ω has the finite intersection property, then there exists an ultrafilter containing $\mathcal{A}$.

This follows because the family generated by $\mathcal{A}$ is a proper filter, and then the theorem above applies.

The proof above uses Zorn's lemma, hence the axiom of choice. In fact, the ultrafilter theorem is weaker than full choice, so Zorn is more than one strictly needs. But for ordinary mathematical use, Zorn's lemma is the standard proof.

of them. So the cofinite filter is a very small amount of asymptotic information, and an ultrafilter extension is a huge refinement of it. That is why the notation

$$\lim_{\mathcal{U}} x_n$$

depends strongly on the choice of $\mathcal{U}$: different ultrafilters extending the same cofinite filter can produce different generalized limits for the same bounded sequence.

1.7 Ultrafilters as $\{0,1\}$-valued finitely additive measures

Given an ultrafilter $\mathcal{U}$ on Ω, define

$$\mu_{\mathcal{U}}(A) = \begin{cases} 1, & A \in \mathcal{U}, \\ 0, & A \notin \mathcal{U}. \end{cases}$$

Because $\mathcal{U}$ is an ultrafilter, this is equivalently

$$\mu_{\mathcal{U}}(A) = 1_{A \in \mathcal{U}}.$$

Then $\mu_{\mathcal{U}}$ is a finitely additive probability on $\mathcal{P}(\Omega)$. Indeed,

$$\mu_{\mathcal{U}}(\Omega) = 1, \qquad \mu_{\mathcal{U}}(\emptyset) = 0,$$

and if $A \cap B = \emptyset$, then

$$\mu_{\mathcal{U}}(A \cup B) = \mu_{\mathcal{U}}(A) + \mu_{\mathcal{U}}(B).$$

For principal ultrafilters this is just a Dirac mass. For non-principal ultrafilters it is a purely finitely additive measure: it behaves like a probability measure with respect to finite unions, but not with respect to countable unions.

That is the bridge to finitely additive measures. Filters and ultrafilters are set-theoretic objects; finitely additive measures are analytic objects. The passage from one to the other is immediate once the ultrafilter is in hand.

---

2 Ultrafilters, generalized limits, and finitely additive mass on $\mathbb{N}$

To see why ultrafilters matter analytically, consider a bounded sequence. An ordinary limit looks at what happens eventually. An ultrafilter lets us replace "eventually" with a more flexible notion of "largeness". That produces a generalized limit, and behind that generalized limit sits a finitely additive measure on $\mathbb{N}$.

2.1 The usual limit and the cofinite filter

Let $x = (x_n)_{n \geq 1} \in \ell_\infty$, so x is a bounded real sequence. To say that $x_n \to L$ in the ordinary sense means that for every $\varepsilon > 0$, the set

$$\{\, n \in \mathbb{N} : \; |x_n - L| < \varepsilon \,\}$$

contains all sufficiently large n. Equivalently, this set is cofinite. Thus ordinary convergence can be expressed in terms of the Fréchet filter

$$\mathcal{F}^{\mathrm{cof}} = \{\, A \subseteq \mathbb{N} : A^c \text{ is finite} \,\}.$$

Namely,

$$x_n \to L \quad \Longleftrightarrow \quad \{\, n : \; |x_n - L| < \varepsilon \,\} \in \mathcal{F}^{\mathrm{cof}} \quad \forall \varepsilon > 0.$$

Thus, the ordinary limit is already filter-based. The phrase "for all sufficiently large n" just means "on a set in the cofinite filter".

2.2 Extending "eventually" to an ultrafilter

The cofinite filter is not an ultrafilter. For example, the even integers are not cofinite, and neither are the odd integers, so the cofinite filter does not decide which of these two sets is large.

By the ultrafilter extension theorem, there exists an ultrafilter $\mathcal{U}$ on $\mathbb{N}$ such that

$$\mathcal{F}^{\mathrm{cof}} \subseteq \mathcal{U}.$$

If $\mathcal{U}$ is chosen non-principal, then it contains every cofinite set, no finite set, and for every $A \subseteq \mathbb{N}$ exactly one of A and A^c belongs to $\mathcal{U}$. The ultrafilter $\mathcal{U}$ is a sharpened notion of asymptotic largeness. A set of indices is either large or its complement is large. There is no ambiguity.

2.3 The ultrafilter limit

Given a bounded sequence $x = (x_n)$ and an ultrafilter $\mathcal{U}$ on $\mathbb{N}$, one defines the $\mathcal{U}$-limit of x to be the real number L such that

$$\{\, n : \ |x_n - L| < \varepsilon \,\} \in \mathcal{U} \quad \text{for every } \varepsilon > 0.$$

One writes

$$L = \lim_{\mathcal{U}} x_n.$$

For bounded real sequences this limit exists and is unique. Obviously, L must be a limit point of the sequence, in the usual sense.

Uniqueness is easy. If both L and M satisfied the definition, with $L \neq M$, choose $\varepsilon < |L - M|/2$. Then the sets

$$\{\, n : |x_n - L| < \varepsilon \,\} \quad \text{and} \quad \{\, n : |x_n - M| < \varepsilon \,\}$$

are disjoint. They cannot both belong to a filter, since filters are closed under intersections and do not contain the empty set.

Existence uses boundedness and compactness. Since x is bounded, its range lies in a compact interval $[-M, M]$. Repeatedly bisect the interval. Because $\mathcal{U}$ is an ultrafilter, at each stage one of the two half-intervals captures a $\mathcal{U}$-large set of indices:

$$\{\, n : x_n \in I_1 \,\} \in \mathcal{U} \quad \text{or} \quad \{\, n : x_n \in I_2 \,\} \in \mathcal{U}.$$

Choosing a large half at each step gives a nested sequence of closed intervals whose lengths tend to 0. Their intersection is a single point L, and this L is the $\mathcal{U}$-limit. Thus every bounded sequence has an ultrafilter limit along every ultrafilter.

2.4 Agreement with the ordinary limit

If $x_n \to L$ in the ordinary sense, then

$$\{\, n : \ |x_n - L| < \varepsilon \,\}$$

is cofinite for every $\varepsilon > 0$, hence belongs to every ultrafilter extending the cofinite filter. Therefore

$$\lim_{\mathcal{U}} x_n = L.$$

So the ultrafilter limit extends the usual limit. It does not replace ordinary convergence; it continues it.

2.5 A simple example

Consider the sequence

$$x_n = (-1)^n.$$

This sequence does not converge ordinarily. But for any ultrafilter $\mathcal{U}$ on $\mathbb{N}$, exactly one of the sets

$$E = \{\, n : x_n = 1 \,\} = \{\, 2, 4, 6, \ldots \,\}, \qquad O = \{\, n : x_n = -1 \,\} = \{\, 1, 3, 5, \ldots \,\}$$

belongs to $\mathcal{U}$.

If $E \in \mathcal{U}$, then
$$\lim_{\mathcal{U}} x_n = 1.$$

If $O \in \mathcal{U}$, then
$$\lim_{\mathcal{U}} x_n = -1.$$

In this way the ultrafilter decides one of the two subsequences to be large, and that becomes the generalized limit.

This is the key point: an ultrafilter gives a binary decision about which index sets count as large, and that turns oscillation into a well-defined asymptotic value.

2.6 The associated linear functional on ℓ_∞

For a fixed ultrafilter $\mathcal{U}$, define
$$\Lambda_{\mathcal{U}}(x) := \lim_{\mathcal{U}} x_n, \qquad x \in \ell_\infty.$$

Then $\Lambda_{\mathcal{U}}$ is a positive linear functional on ℓ_∞ with norm 1.

The essential properties are:

1. linearity:
$$\Lambda_{\mathcal{U}}(ax + by) = a\Lambda_{\mathcal{U}}(x) + b\Lambda_{\mathcal{U}}(y),$$

2. positivity: if $x_n \geq 0$ for all n, then
$$\Lambda_{\mathcal{U}}(x) \geq 0,$$

3. normalization: for the constant sequence $\mathbf{1} = (1, 1, \dots)$,
$$\Lambda_{\mathcal{U}}(\mathbf{1}) = 1.$$

It follows that
$$\|\Lambda_{\mathcal{U}}\| = 1.$$

This is already a finitely additive integration functional in disguise.

2.7 The corresponding finitely additive probability on $\mathbb{N}$

To each subset $A \subseteq \mathbb{N}$ associate its indicator sequence $1_A \in \ell_\infty$ (i.e., the sequence $1_A(n) = 1$ if $n \in A$ and 0 otherwise). Define
$$\mu_{\mathcal{U}}(A) := \Lambda_{\mathcal{U}}(1_A).$$

Since 1_A only takes the values 0 and 1, its $\mathcal{U}$-limit must also be 0 or 1. In fact,
$$\mu_{\mathcal{U}}(A) = \begin{cases} 1, & A \in \mathcal{U}, \\ 0, & A \notin \mathcal{U}. \end{cases}$$

Thus $\mu_{\mathcal{U}}$ is exactly the $\{0, 1\}$-valued set function attached to the ultrafilter.

The measure $\mu_{\mathcal{U}}$ is finitely additive. If A and B are disjoint, then
$$1_{A \cup B} = 1_A + 1_B,$$
so
$$\mu_{\mathcal{U}}(A \cup B) = \mu_{\mathcal{U}}(A) + \mu_{\mathcal{U}}(B).$$

Also,
$$\mu_{\mathcal{U}}(\mathbb{N}) = 1.$$

So $\mu_{\mathcal{U}}$ is a finitely additive probability on $\mathbb{N}$.

If $\mathcal{U}$ is non-principal, then every finite set has measure 0. Indeed, if $\{n\} \in \mathcal{U}$ for some n, then $\mathcal{U}$ would be principal. Hence $\{n\} \notin \mathcal{U}$ and so by definition
$$\mu_{\mathcal{U}}(n) = 0 \quad \text{for every } n.$$

6

Therefore

$$\mu_{\mathcal{U}}(\mathbb{N}) = 1 \qquad \text{but} \qquad \sum_{n=1}^{\infty} \mu_{\mathcal{U}}(n) = 0.$$

This shows concretely that $\mu_{\mathcal{U}}$ is not countably additive. That failure is a feature, not a bug. Finite additivity allows mass to live on an asymptotic notion of largeness rather than on individual points.

This simple sequence example contains the whole mechanism we need later.

- Start with a family of sets that are regarded as large.
- Extend that family to an ultrafilter.
- Use the ultrafilter to define a $\{0, 1\}$-valued finitely additive measure.
- Interpret that measure as concentrating all mass on the chosen notion of largeness.

---

3 Example of a finitely additive measure concentrated near the essential supremum

We now turn to a motivating example. The point is to build a positive finitely additive measure μ that places full mass on every near-maximizer set

$$\{ X > \operatorname{ess\,sup} X - \varepsilon \},$$

even though these sets shrink down to a Lebesgue-null limit set. This is exactly the kind of concentration that countable additivity forbids and finite additivity permits.

3.1 The function and its essential supremum

Define

$$X(x) = \begin{cases} \sin(1/x), & x \in (0, 1], \\ 0, & x = 0. \end{cases}$$

We view X as a bounded measurable function on $\Omega = [0, 1]$ with its Lebesgue σ-algebra.

Since $\sin(1/x) \leq 1$ for all $x > 0$, we have

$$\operatorname{ess\,sup} X \leq 1.$$

On the other hand, for every $\varepsilon > 0$, the set

$$\{ x \in (0, 1] : X(x) > 1 - \varepsilon \}$$

has positive Lebesgue measure, because near each point where $\sin(1/x) = 1$ the function stays above $1 - \varepsilon$ on a nontrivial interval. Hence

$$\operatorname{ess\,sup} X = 1.$$

The exact maximizer set is

$$\{ X = 1 \} = \left\{ \frac{2}{(4k + 1)\pi} : k = 0, 1, 2, \dots \right\}.$$

This set is countable and hence has Lebesgue measure 0. Thus, X attains its essential supremum only on a null set, but it comes arbitrarily close to that supremum on many small intervals accumulating at 0.

3.2 The near-maximizer sets

For each $n \geq 2$, define

$$A_n := \{ x \in [0, 1] : X(x) > 1 - 1/n \}.$$

These are the sets on which X lies within $1/n$ of its essential supremum. Because $1 - 1/(n+1) > 1 - 1/n$, the sets are nested decreasing:

$$A_{n+1} \subseteq A_n.$$

Their intersection is exactly

$$\bigcap_{n=2}^{\infty} A_n = \{\, x : X(x) = 1 \,\}.$$

Indeed, a point belongs to every A_n if and only if

$$X(x) > 1 - \frac{1}{n} \quad \text{for all } n,$$

which is equivalent to $X(x) = 1$.

Thus (A_n) is a decreasing sequence of measurable sets with

$$\lambda(A_n) > 0 \quad \text{for every } n,$$

but

$$\lambda\left(\bigcap_{n=2}^{\infty} A_n\right) = \lambda(\{\, X = 1 \,\}) = 0.$$

This is the setup we want. The sets A_n say "arbitrarily close to the essential supremum", while their limit is too small to carry any countably additive mass.

3.3 The filter generated by the sets A_n

Because the sets A_n are nested and nonempty, they have the finite intersection property. In fact, for any finite collection,

$$A_{n_1} \cap \cdots \cap A_{n_k} = A_{\max(n_1,\ldots,n_k)} \neq \emptyset.$$

As a result, they generate a proper filter $\mathcal{F}$ on $[0,1]$, namely

$$\mathcal{F} = \{\, B \subseteq [0,1] : A_n \subseteq B \text{ for some } n \,\}.$$

In words, a set is large if it contains one of the near-maximizer sets. This is the natural filter attached to the family (A_n). It says that any set containing a sufficiently strong near-maximizer region should count as large.

By the ultrafilter extension theorem, there exists an ultrafilter $\mathcal{U}$ on $[0,1]$ such that

$$\mathcal{F} \subseteq \mathcal{U}.$$

Since each A_n belongs to $\mathcal{F}$, we have

$$A_n \in \mathcal{U} \qquad \text{for every } n.$$

This is the key step. The ultrafilter takes the decreasing family (A_n) and declares every one of them to be large.

3.4 The induced finitely additive measure

Now define

$$\mu(B) := \begin{cases} 1, & B \in \mathcal{U}, \\ 0, & B \notin \mathcal{U}, \end{cases} \qquad B \subseteq [0,1].$$

As in section 2, this is a $\{\, 0, 1 \,\}$-valued finitely additive probability on the power set $\mathcal{P}([0,1])$. In particular,

1. $\mu([0,1]) = 1$,
2. $\mu(\emptyset) = 0$,
3. if B and C are disjoint, then

$$\mu(B \cup C) = \mu(B) + \mu(C).$$

Because $A_n \in \mathcal{U}$ for every n, we get

$$\mu(A_n) = 1 \qquad \text{for every } n.$$

Since

$$A_n = \{\, X > 1 - 1/n \,\},$$

this becomes

$$\mu\{\, X > 1 - 1/n \,\} = 1 \qquad \text{for every } n \geq 2.$$

More generally, for every $\varepsilon > 0$, choose n so large that $1/n < \varepsilon$. Then

$$A_n = \{\, X > 1 - 1/n \,\} \subseteq \{\, X > 1 - \varepsilon \,\}.$$

Since $A_n \in \mathcal{U}$ and ultrafilters are upward closed, it follows that

$$\{\, X > 1 - \varepsilon \,\} \in \mathcal{U}.$$

Therefore

$$\mu\{\, X > 1 - \varepsilon \,\} = 1 \qquad \text{for every } \varepsilon > 0.$$

Since $\operatorname{ess\,sup} X = 1$, this is exactly

$$\mu\{\, X > \operatorname{ess\,sup} X - \varepsilon \,\} = 1 \qquad \text{for every } \varepsilon > 0.$$

Why does this not contradict the nullity of $X = 1$? At first sight this may look paradoxical. The sets A_n decrease to the set $\{\, X = 1 \,\}$, and that limit set is countable, hence Lebesgue-null. How can every A_n have μ-mass 1? The answer is that μ is not countably additive. Indeed,

$$\bigcap_{n=2}^{\infty} A_n = \{\, X = 1 \,\},$$

but finite additivity alone does not imply continuity from above[5].

---

[5]Continuity from above for countably additive measures. If μ is countably additive and $A_n \downarrow A$, with $\mu(A_1) < \infty$, then

$$\mu(A_n) \downarrow \mu(A).$$

A standard proof reduces this to continuity from below. Define

$$B_n := A_1 \setminus A_n.$$

Since A_n decreases, the sets B_n increase, and

$$\bigcup_{n=1}^{\infty} B_n = A_1 \setminus A.$$

By countable additivity, measures are continuous from below, so

$$\mu(B_n) \uparrow \mu(A_1 \setminus A).$$

But

$$\mu(B_n) = \mu(A_1) - \mu(A_n)$$

and

$$\mu(A_1 \setminus A) = \mu(A_1) - \mu(A),$$

where finiteness of $\mu(A_1)$ ensures these differences are legitimate. Hence

$$\mu(A_1) - \mu(A_n) \uparrow \mu(A_1) - \mu(A),$$

which is equivalent to

$$\mu(A_n) \downarrow \mu(A).$$

Continuity from below. Continuity from below. If μ is countably additive and $A_n \uparrow A$, then

$$\mu(A_n) \uparrow \mu(A).$$

Indeed, let

$$B_1 := A_1, \qquad B_n := A_n \setminus A_{n-1} \quad (n \geq 2).$$

Then the sets B_n are pairwise disjoint and

$$A_n = \bigcup_{k=1}^{n} B_k, \qquad A = \bigcup_{k=1}^{\infty} B_k.$$

By countable additivity,

$$\mu(A_n) = \sum_{k=1}^{n} \mu(B_k) \uparrow \sum_{k=1}^{\infty} \mu(B_k) = \mu(A).$$

Countably additive measures satisfy

$$A_n \downarrow A \quad \implies \quad \mu(A_n) \downarrow \mu(A),$$

provided $\mu(A_1) < \infty$. Our set function μ need not satisfy this. In fact, for a suitable choice of ultrafilter[6], one has

$$\mu(\{\, x \,\}) = 0 \qquad \text{for every } x \in [0,1],$$

and

$$\mu(\{\, X = 1 \,\}) = 0,$$

while still

$$\mu(A_n) = 1 \qquad \text{for every } n.$$

So mass is not concentrated on the countable set of exact maximizers. It is concentrated on the filter of near-maximizer sets. The measure does not see individual points; it sees the asymptotic pattern of sets that keep X arbitrarily close to its top value.

3.5 Interpreting μ as a positive element of ba

The set function μ defines a positive bounded finitely additive measure on $[0,1]$, that is, a positive element of ba$([0,1])$. Because μ is positive, its total variation norm is simply its total mass:

$$\|\mu\|_{\mathrm{ba}} = \mu([0,1]) = 1.$$

Therefore the previous identity can be written as

$$\mu\{\, X > \operatorname{ess\,sup} X - \varepsilon \,\} = \|\mu\|_{\mathrm{ba}} \qquad \text{for every } \varepsilon > 0.$$

The positive finitely additive measure μ places all of its mass in every neighborhood of the essential supremum, no matter how tight that neighborhood becomes.

This example isolates a basic phenomenon made possible by finite addivitity. If one insists on countable additivity, then a decreasing sequence of sets with null intersection cannot all carry full mass. Continuity from above forces the masses to collapse to the mass of the intersection. Finite additivity removes that constraint. One may have

$$\mu(A_n) = 1 \quad \text{for every } n, \qquad \text{but} \qquad \mu\left(\bigcap_{n=2}^{\infty} A_n \right) = 0.$$

That is precisely why ultrafilters are useful here. They convert a qualitative notion of "arbitrarily close to the supremum" into an honest positive finitely additive measure that charges all such neighborhoods fully.

---

[6]The statement

$$\mu(\{\, x \,\}) = 0 \quad \forall x$$

does not by itself imply

$$\mu(\{\, X = 1 \,\}) = 0,$$

because $\{\, X = 1 \,\}$ is countable and we only have finite additivity. To see the latter, choose the ultrafilter so that it explicitly avoids the whole set $\{\, X = 1 \,\}$, which we can do as follows. Let

$$C := \{\, X = 1 \,\}.$$

We know C is countable, and for every n,

$$A_n = \{\, X > 1 - 1/n \,\}$$

contains infinitely many intervals around the peak points, so in particular

$$A_n \setminus C \neq \emptyset.$$

Indeed, removing the countable set of exact peak points does not remove those intervals. Now consider the family

$$\mathcal{A} := \{\, A_n : n \geq 2 \,\} \cup \{\, C^c \,\}.$$

I claim $\mathcal{A}$ has the finite intersection property. That is easy: any finite intersection has the form

$$A_{n_1} \cap \cdots \cap A_{n_k} \cap C^c = A_m \cap C^c, \qquad m = \max(n_1, \dots, n_k),$$

and this is nonempty because $A_m \setminus C \neq \emptyset$. Therefore $\mathcal{A}$ generates a proper filter, and by the ultrafilter lemma there exists an ultrafilter $\mathcal{U}$ containing $\mathcal{A}$. From this construction we get each $A_n \in \mathcal{U}$, so $\mu(A_n) = 1$ for all n, and $C^c \in \mathcal{U}$, so necessarily $C \notin \mathcal{U}$ and hence $\mu(C) = \mu(\{\, X = 1 \,\}) = 0$ as required.

The construction does not depend on the particular X. The same template works whenever one has a bounded measurable function X and a decreasing family of nonempty near-maximizer sets

$$A_n = \{\, X > \operatorname{ess\,sup} X - \varepsilon_n \,\}, \qquad \varepsilon_n \downarrow 0,$$

with the finite intersection property.

One forms the filter generated by the sets A_n, extends it to an ultrafilter, and takes the associated $\{\, 0, 1 \,\}$-valued finitely additive measure. That measure gives mass 1 to every A_n, hence to every set that contains one of them.

The specific choice

$$X(x) = \sin(1/x)$$

is appealing because the geometry is vivid. The function keeps returning to values near 1 on infinitely many shrinking intervals, while the exact maximizer set remains countable and negligible in the Lebesgue sense. This makes the distinction between countable and finite additivity easy to see.

---

4 Ultrafilter limits of functions

A sequence is just a function on $\mathbb{N}$. Once one has an ultrafilter on an arbitrary set, one can speak of the ultrafilter limit of a function.

Let Ω be a set, let $\mathcal{U}$ be an ultrafilter on Ω, and let

$$f : \Omega \to K$$

be a function taking values in a compact Hausdorff space K. A point $L \in K$ is called the $\mathcal{U}$-**limit of** f if for every open neighborhood V of L,

$$f^{-1}(V) \in \mathcal{U}.$$

One writes

$$L = \lim_{\mathcal{U}} f.$$

This is the exact analog of the definition for bounded sequences. There, $\Omega = \mathbb{N}$ and K is a compact interval containing the range of the sequence.

4.1 Existence and uniqueness

Because K is compact Hausdorff, every function $f : \Omega \to K$ has a unique ultrafilter limit along $\mathcal{U}$.

Uniqueness follows from the Hausdorff property. If $L \neq M$, choose disjoint neighborhoods V_L of L and V_M of M. Then

$$f^{-1}(V_L) \in \mathcal{U} \quad \text{and} \quad f^{-1}(V_M) \in \mathcal{U}$$

cannot both hold, because those two preimages are disjoint and an ultrafilter cannot contain two disjoint sets.

Existence is a compactness argument. The family of closed sets

$$\overline{f(A)}, \qquad A \in \mathcal{U},$$

has the finite intersection property, so compactness implies that their total intersection is nonempty. Any point in that intersection is the desired ultrafilter limit.

Thus ultrafilters turn arbitrary functions into convergent objects, provided the range is compact.

4.2 The case of bounded real-valued functions

For the present blog post, we only need the case of bounded real-valued functions. If

$$f : \Omega \to \mathbb{R}$$

is bounded, then its range is contained in some compact interval $[-M, M]$, and so

$$\lim_{\mathcal{U}} f$$

exists and is uniquely defined as a real number. Concretely then, $L = \lim_{\mathcal{U}} f$ means that for every $\varepsilon > 0$,

$$\{\, \omega \in \Omega : |f(\omega) - L| < \varepsilon \,\} \in \mathcal{U}.$$

This is the form most closely parallel to ordinary ε-N convergence.

4.3 The ultrafilter limit of $X(x) = \sin(1/x)$

Now return to the function

$$X(x) = \begin{cases} \sin(1/x), & x \in (0, 1], \\ 0, & x = 0, \end{cases}$$

and to the ultrafilter $\mathcal{U}$ from section 3, chosen so that

$$A_n = \{\, X > 1 - 1/n \,\} \in \mathcal{U} \qquad \text{for every } n.$$

We claim that

$$\lim_{\mathcal{U}} X = 1.$$

Indeed, let $\varepsilon > 0$. Choose n so large that $1/n < \varepsilon$. Then

$$A_n = \{\, X > 1 - 1/n \,\} \subseteq \{\, X > 1 - \varepsilon \,\}.$$

Since $A_n \in \mathcal{U}$ and ultrafilters are upward closed,

$$\{\, X > 1 - \varepsilon \,\} \in \mathcal{U}.$$

But

$$\{\, X > 1 - \varepsilon \,\} \subseteq |X - 1| < \varepsilon,$$

so again by upward closure,

$$\{\, |X - 1| < \varepsilon \,\} \in \mathcal{U}.$$

This is exactly the definition of

$$\lim_{\mathcal{U}} X = 1.$$

So the ultrafilter sees the function X as converging to 1, even though in the ordinary topological sense there is no limit behavior of $X(x)$ as $x \downarrow 0$.

This is a very compact way of expressing what the construction is doing: the ultrafilter chooses the near-maximizer regions as the large sets, and relative to that notion of largeness the function converges to its essential supremum.

---

5 Integration with respect to the finitely additive measure

The ultrafilter from section 3 also gives a positive finitely additive measure

$$\mu(B) = 1_{B \in \mathcal{U}}, \qquad B \subseteq [0, 1].$$

We now connect this measure to integration. The main message is simple: since μ assigns full mass to every near-maximizer set of X, integration against μ behaves as though all mass were concentrated on sets where X is arbitrarily close to 1.

5.1 From finitely additive measures to linear functionals

For countably additive measures, integration is built from measure theory on a σ-algebra. For finitely additive measures, the cleanest viewpoint is functional-analytic: a bounded finitely additive measure defines a bounded linear functional on bounded measurable functions.

A positive finitely additive measure μ on $[0, 1]$ defines a positive linear functional

$$I_\mu : L^\infty([0,1]) \to \mathbb{R}, \qquad I_\mu(f) = \int f\, d\mu.$$

For indicator functions this agrees with the set function:

$$\int 1_B\, d\mu = \mu(B).$$

Linearity then extends the definition to simple functions, and boundedness extends it further to bounded measurable functions.

For a positive finitely additive probability, the norm of this functional is

$$\|I_\mu\| = \mu([0,1]) = 1.$$

For the present example, we do not need the full machinery. We only need the basic monotonicity and normalization properties of integration.

5.2 Basic inequalities

Because μ is positive, integration is monotone: if

$$f \le g,$$

then

$$\int f\, d\mu \le \int g\, d\mu.$$

Also, since μ is a probability,

$$\int 1\, d\mu = 1.$$

Now fix $\varepsilon > 0$ and consider the set

$$B_\varepsilon := \{\, X > 1 - \varepsilon \,\}.$$

From section 3,

$$\mu(B_\varepsilon) = 1.$$

Since $\mu(B_\varepsilon) = 1$, we have

$$\mu(B_\varepsilon^c) = 0.$$

For a positive finitely additive measure, any bounded function supported on a μ-null set has integral 0. Thus

$$\int X\, d\mu = \int_{B_\varepsilon} X\, d\mu.$$

But on B_ε one has

$$1 - \varepsilon < X \le 1.$$

Therefore

$$(1-\varepsilon)\mu(B_\varepsilon) \le \int X\, d\mu \le 1 \cdot \mu(B_\varepsilon).$$

Since $\mu(B_\varepsilon) = 1$, this yields again

$$1 - \varepsilon \le \int X, d\mu \le 1.$$

Letting $\varepsilon \downarrow 0$ gives

$$\int X\, d\mu = 1,$$

as expected. Each near-maximizer set has full mass, so outside it there is no mass, and inside it the function is nearly equal to 1.

5.3 Relation to the ultrafilter limit

The identities

$$\lim_{\mathcal{U}} X = 1 \qquad \text{and} \qquad \int X \, d\mu = 1$$

express the same underlying phenomenon in two different languages.

- The ultrafilter-limit language says that relative to the chosen notion of largeness, the values of X converge to 1.
- The finitely additive integration language says that the associated measure places all mass on regions where X is arbitrarily close to 1.

For this particular $\{0, 1\}$-valued measure, the two viewpoints line up perfectly.

5.4 What this says about essential suprema

For an ordinary countably additive probability P, the integral of a bounded function usually reflects behavior across the whole space. In contrast, the finitely additive measure μ constructed here is entirely selective: it only cares about those sets declared large by the ultrafilter. As a result,

$$\int X \, d\mu = \operatorname{ess\,sup} X,$$

not because $X = 1$ on a set of positive Lebesgue measure, but because every neighborhood of the essential supremum has full μ-mass.

This is the analytic content behind the identity

$$\mu \{ X > \operatorname{ess\,sup} X - \varepsilon \} = \|\mu\|_{\mathrm{ba}}.$$

---

6 Why this matters for $(L^\infty)^*$

This discussion is not just a curiosity about filters and finitely additive measures. It appears naturally when one studies the dual of L^∞.

If one works with actual bounded functions on a space Ω, before quotienting out null sets, then point evaluations are available. For each $x \in \Omega$, the Dirac mass

$$\delta_x$$

defines a positive linear functional by

$$f \mapsto f(x).$$

So if a bounded function f attains its supremum at some point x_0, then

$$\delta_{x_0}(f) = \sup f.$$

At that level, the supremum is easily captured by a Dirac mass. Obviously this is not available if f does not achieve its sup.

But L^∞ is not the space of bounded functions. It is the space of essentially bounded functions modulo equality almost everywhere. Once one passes to equivalence classes modulo null sets, pointwise evaluation is no longer well defined. Changing a function at a single point does not change its class in L^∞, so there is no meaningful functional

$$[f] \mapsto f(x).$$

That removes the obvious maximizers. If a function has

$$\operatorname{ess\,sup} f = M$$

but only reaches the value M on a null set, then Dirac masses are no longer available in L^∞ to detect that top value.

This is where finitely additive measures enter. The dual space $(L^\infty)^*$ is much larger than L^1, and positive elements of $(L^\infty)^*$ correspond to bounded finitely additive measures, not just countably additive ones. These finitely additive functionals can charge every near-maximizer set

$$\{\, f > \operatorname{ess\,sup} f - \varepsilon \,\}$$

with full mass, and so can recover the essential supremum even when no point mass survives the passage to L^∞.

That is exactly what the $\sin(1/x)$ example illustrates. Before quotienting by null sets, one could look at Dirac masses at points where $X = 1$. After passing to L^∞, those point masses disappear as well-defined functionals. The ultrafilter construction replaces them with a positive finitely additive measure that still detects the top value:

$$\int X \, d\mu = \operatorname{ess\,sup} X = 1.$$

---

7 Conclusion

The story begins with a very simple idea: some sets are large, some are small. Filters formalize the large sets, ideals formalize the small ones, and ultrafilters push that distinction to its extreme by deciding every set one way or the other. From there, two analytic constructions follow immediately.

First, ultrafilters produce generalized limits. An ordinary limit uses the cofinite filter, that is, what happens eventually. A non-principal ultrafilter refines "eventually" into a maximal notion of asymptotic largeness, and every bounded sequence acquires a well-defined ultrafilter limit.

Second, ultrafilters produce positive finitely additive measures. The associated $\{\, 0, 1 \,\}$-valued set function is finitely additive, not countably additive, and this failure of countable additivity is exactly what makes the construction useful.

The function

$$X(x) = \sin(1/x)$$

shows the phenomenon vividly. Its essential supremum is 1, but the exact maximizer set $X = 1$ is only countable and therefore Lebesgue-null. Nevertheless, by taking the filter generated by the near-maximizer sets $A_n = \{\, X > 1 - 1/n \,\}$ and extending it to an ultrafilter, one obtains a positive finitely additive probability μ such that

$$\mu \{\, X > \operatorname{ess\,sup} X - \varepsilon \,\} = 1 \qquad \text{for every } \varepsilon > 0.$$

Equivalently,

$$\mu \{\, X > \operatorname{ess,\,sup} X - \varepsilon \,\} = \|\mu\|_{\mathrm{ba}}.$$

The same ultrafilter also sees

$$\lim_{\mathcal{U}} X = 1 \qquad \text{and} \qquad \int X, d\mu = 1.$$

So the generalized limit viewpoint and the finitely additive integration viewpoint are really two versions of the same construction.

That is the central lesson. Finite additivity allows mass to concentrate, not on points, but on a chosen asymptotic pattern of sets. Ultrafilters are the mechanism that makes this precise.

8 Appendix: Boolean algebras and the connection with ring theory

A **Boolean algebra** is a set B equipped with two binary operations, usually written

$$a \wedge b, \qquad a \vee b,$$

a unary operation

$$a \mapsto a',$$

and distinguished elements

$$0, \qquad 1,$$

such that for all $a, b, c \in B$:

1. commutativity:
$$a \wedge b = b \wedge a, \qquad a \vee b = b \vee a,$$

2. associativity:
$$(a \wedge b) \wedge c = a \wedge (b \wedge c), \qquad (a \vee b) \vee c = a \vee (b \vee c),$$

3. distributivity:
$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c),$$
$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c),$$

4. identity laws:
$$a \wedge 1 = a, \qquad a \vee 0 = a,$$

5. complement laws:
$$a \wedge a' = 0, \qquad a \vee a' = 1.$$

The standard example is the power set $\mathcal{P}(\Omega)$ of a set Ω, where

$$a \wedge b = a \cap b, \qquad a \vee b = a \cup b, \qquad a' = a^c, \qquad 0 = \emptyset, \qquad 1 = \Omega.$$

Informally, a Boolean algebra is an abstract version of the algebra of subsets of a set, with intersection, union, and complement.

A Boolean algebra may be viewed either as an abstract algebraic structure or as a ring-like structure built from sets. In the concrete model $\mathcal{P}(\Omega)$, the underlying objects are subsets of Ω, the order is inclusion, the meet and join operations are

$$A \wedge B = A \cap B, \qquad A \vee B = A \cup B,$$

and complementation is

$$\neg A = A^c.$$

The same structure can also be written in Boolean ring language, where addition is symmetric difference

$$A + B := A \triangle B = (A \ B) \cup (B \ A),$$

and multiplication is intersection

$$AB := A \cap B.$$

In this dictionary, $0 = \emptyset$ and $1 = \Omega$, and every element is idempotent:

$$A^2 = A.$$

Thus Boolean algebras and Boolean rings are really two presentations of the same thing. Ideals in the Boolean-ring sense are exactly the same as ideals in the order/set-theoretic sense: they are downward closed and closed under finite unions, equivalently closed under ring addition and absorption by multiplication. Filters are the dual notion, and in Boolean algebra they correspond to complements of ideals. More precisely, if I is an ideal, then

$$\mathcal{F} = \{\, A : A^c \in I \,\}$$

is a filter, and conversely. Relative to the top element 1, one can also write a filter as

$$\mathcal{F} = 1 + I := \{\, 1 + a : a \in I \,\},$$

where in Boolean-ring notation $1 + a$ is just the complement of a, since

$$1 + A = \Omega \triangle A = A^c.$$

So a filter is literally "one plus an ideal". All of this can be realized through indicator functions. Identifying a set $A \subseteq \Omega$ with its indicator 1_A, the Boolean-ring operations become pointwise operations mod 2:

$$1_A 1_B = 1_{A \cap B}, \qquad 1_A + 1_B = 1_{A \triangle B}, \qquad 1 + 1_A = 1_{A^c}.$$

Thus $\mathcal{P}(\Omega)$ is isomorphic to the ring of $\{\, 0, 1 \,\}$-valued functions on Ω, with multiplication given by ordinary pointwise multiplication and addition taken mod 2. This is the clean algebraic setting in which ideals, filters, and complements fit together.

9 Annotated bibliography

- Isaac Goldbring, Ultrafilters Throughout Mathematics, Goldbring (2022). A broader and more advanced source. Not as lightweight, but excellent for perspective. The early material is useful for background, and the later material shows how widely ultrafilters appear. This is a good source to cite when you want the reader to know the subject is part of mainstream mathematics, not an odd corner case.

- W. W. Comfort, Ultrafilters: some old and some new results, W. Wistar Comfort (1977). This is a substantial Bulletin of the AMS survey article, published in 1977 in volume 83, issue 4, pages 417–455. It is not an elementary introduction in the style of lecture notes, but it is a major expository survey by one of the central figures in the area. It gives a broad picture of what ultrafilters do across topology, set theory, and related parts of analysis, and it helps the reader see that ultrafilters are not merely a technical trick for generalized limits, but a large and well-developed subject in their own right. For a reader who already knows the basic definitions and wants a serious map of the territory, this is an excellent next step.

- W. W. Comfort and S. Negrepontis, The Theory of Ultrafilters, William Wistar Comfort and Negrepontis (2012). This is the full-length classic book treatment. It is a serious monograph rather than a quick introduction. The scope is much wider and deeper than what you need for this blog post, but that is exactly its value as a reference: if a reader wants the definitive classical treatment, this is the place to go. It develops the theory systematically and documents the older literature in a way that short surveys cannot.

- Max Garcia, Filters and Ultrafilters in Real Analysis, Garcia (2004). This is especially relevant for the bridge from filters to analysis. It helps connect abstract filter language with convergence, limits, and real-analytic applications. Very useful for the transition into generalized limits.

- Alex Kruckman, Notes on Ultrafilters, Kruckman (2012). A very readable introduction to filters and ultrafilters. Good on the basic intuition that filters represent large sets, and good for the elementary formal definitions. This is probably the best place to start for section 1 of the post. It is concise, clear, and does not bury the reader in logic-heavy applications.

- Burak Kaya, Ultrafilters and How to Use Them, Kaya (n.d.). Another accessible set of lecture notes, somewhat broader in scope than Kruckman. These notes also emphasize the intuitive meaning of ultrafilters and then show how they get used. Good as a second source when polishing exposition or checking terminology.

- Jech, Set Theory, Jech (2003). Chapter 7 discussed filters, ultra filters, and Boolean algebras. Mentions that the dual notion to an ultrafilter is a prime ideal[7]. Tarski proved every filter can be extended to an ultrafilter. Pospíšil proved that for every infinite cardinal κ, there exist $2^{2^{\kappa}}$ ultrafilters on κ.

- K. Yosida and E. Hewitt, Finitely additive measures Yosida and Hewitt (1952). This is one of the foundational papers of the subjec. It is the source of what is now called the Yosida–Hewitt

---

[7]A prime ideal in the ultrafilter story is the same formal notion as a prime ideal in algebraic number theory, but it lives in a very special kind of ring, namely a Boolean ring. For the power set $\mathcal{P}(\Omega)$ one defines

$$A + B := A \triangle B, \qquad AB := A \cap B,$$

so every element is idempotent:

$$A^2 = A.$$

An ideal P is prime if

$$AB \in P \implies A \in P \text{ or } B \in P,$$

exactly as in ordinary commutative ring theory. In this setting, because multiplication is intersection, prime ideals are exactly the complements of ultrafilters: if $\mathcal{U}$ is an ultrafilter, then

$$I = \{ A \subseteq \Omega : A^c \in \mathcal{U} \}$$

is a prime ideal, and conversely.

The important point is that this neat duality is special to Boolean algebras and Boolean rings. In ordinary ring theory, such as algebraic number theory, prime ideals and maximal ideals are different notions in general, and there is no complement operation turning one naturally into a filter-like object. In a Boolean ring, by contrast, every prime ideal is automatically maximal, and complementation gives a built-in passage between "small" sets (ideals) and "large" sets (filters, ultrafilters). So the notion of prime ideal is the same one as in algebraic number theory, but the ambient algebra is much more rigid, and that is why ultrafilters appear so naturally there.

decomposition. It is heavily cited (762 Google Scholar citations) and remains a basic reference point whenever one works with bounded finitely additive measures and the dual of L^∞

It shows that bounded finitely additive measures are not just pathological curiosities. They form a structured space, usually denoted ba, and that every positive finitely additive measure splits uniquely into two parts:

$$\mu = \mu_{ca} + \mu_{pfa},$$

where μ_{ca} is countably additive and μ_{pfa} is purely finitely additive. "Purely finitely additive" means, roughly, that it contains no nonzero countably additive part beneath it. This decomposition is exactly the conceptual backdrop for examples like the ultrafilter measure in this post: such measures live in the purely finitely additive world, not in the ordinary countably additive one.

A second reason the paper matters is its connection to duality. The modern viewpoint that positive functionals on L^∞ correspond to bounded finitely additive measures is closely tied to the Yosida–Hewitt framework, and later literature routinely cites the paper in exactly that connection. So if one wants to understand why finitely additive measures appear naturally in $(L^\infty)^*$ rather than as an artificial construction, this is one of the papers behind that story.

That said, the paper is not an easy first read. It is foundational rather than introductory, and much of its importance lies in the structural decomposition it proves, not in elementary examples. A good way to approach it is after one already has a concrete picture in mind—for example, ultrafilter limits on ℓ_∞ and the $\sin(1/x)$ example in this post. With those examples in hand, the paper becomes easier to read as a general theory explaining where such objects sit inside the larger space of finitely additive measures.

References

Comfort, W. Wistar. (1977). Ultrafilters: some old and some new results. *Bulletin of the American Mathematical Society*, *83*(4), 417–455.

Comfort, William Wistar, & Negrepontis, S. (2012). *The Theory of Ultrafilters*. Springer Science & Business Media.

Garcia, M. (2004). Filters and Ultrafilters. In (pp. 427–432). Springer-Verlag. https://doi.org/10.1007/0-387-21796-7/_39

Goldbring, I. (2022). *Ultrafilters Throughout Mathematics. Graduate Studies in Mathematics*. American Mathematical Society. https://doi.org/10.1090/gsm/220

Jech, T. (2003). *Set Theory* (3rd Millennium.). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/3-540-44761-X

Kaya, B. (n.d.). Ultrafilters and How to Use Them.

Kruckman, A. (2012). *Notes on ultrafilters* (pp. 1–5). Berkeley.

Yosida, K., & Hewitt, E. (1952). Finitely Additive Measures. *Transactions of the American Mathematical Society*, *72*(1), 46. https://doi.org/10.2307/1990654