Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

The dual space of ℓ^∞

Thesis submitted to
Delft Institute of Applied Mathematics
as partial fulfilment of the requirements

for the degree of

BACHELOR OF SCIENCE
in
APPLIED MATHEMATICS

by

Rick Schermerhorn

Delft, The Netherlands
August 2023

BSc THESIS APPLIED MATHEMATICS

"The dual space of ℓ^∞"

Rick Schermerhorn

Delft University of Technology

Supervisor

Dr. K.P. Hart

Other thesis committee members

Dr. J.A.M. de Groot

August, 2023 Delft

Abstract

In this report we examine the dual space of ℓ^∞. If $p \in [1, \infty)$ and $q \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$, then one can identify the spaces ℓ^q and $(\ell^p)'$ in a natural way via an isometric isomorphism. This identification does not extend to the case $p = \infty$ and $q = 1$. We prove that the obvious candidate for an isometric isomorphism from ℓ^1 into $(\ell^\infty)'$ fails to be surjective, and moreover, that an isometric isomorphism (even a homeomorphism) between these spaces does not exist at all.

We introduce a space that we can identify with $(\ell^\infty)'$ via an isometric isomorphism. This is the space of bounded finitely additive measures on $\mathbb{N}$, denoted by $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Having found this characterization of $(\ell^\infty)'$, we examine what kinds of finitely additive measures on $\mathbb{N}$ exist. These include σ-additive measures that are induced by ℓ^1, diffuse measures, shift-invariant and more general invariant measures, measures that extend the asymptotic density, $0, 1$-valued measures and stretchable, thinnable and elastic measures. Elastic measures can be considered the nicest measures on $\mathbb{N}$, from an intuitive point of view.

We also describe the functionals that correspond to particular types of measures and vice versa. Moreover, we prove that the collection of ultrafilters on $\mathbb{N}$ can be identified with the collection of $0, 1$-valued measures on $\mathbb{N}$, which, in turn, can be identified with the collection of multiplicative functionals on ℓ^∞.

Contents

Preface

The final part of my bachelor Applied Mathematics at the TU Delft was writing a thesis about some mathematical topic. For me it was not a difficult decision to choose a topic in analysis: all three years of the bachelor I enjoyed the analysis courses most. Especially the course Real Analysis was one that I enjoyed a lot. Already in the first semester of the third year I told my supervisor K.P. Hart about my interests and I asked him if he knew a topic in analysis suitable for a bachelor project. He came up with this topic and it sounded very interesting to me, so we agreed to do this project. This was probably the best topic I could have chosen; I found it so interesting that writing this thesis almost did not even feel like an obligation.

This report is partially based on a paper written by Eric Karel van Douwen (1946-1987). Eric was a Dutch mathematician specializing in general topology. He was a very influential figure in topology, known by all topologists around the world. He obtained his master's degree in mathematics at the TU Delft in 1972 and his PhD in mathematics at the Vrije Universiteit in Amsterdam in 1975. Unfortunately he died of a heart attack in 1987. In his short academic career, he published over 70 research papers.

The amazing coincidence is that my dad and Eric were housemates in the time they both studied in Delft. When I told my dad that I was doing a course in topology, he remembered that one of his housemates had obtained a PhD in topology, but he did not know that Eric had become such an influential mathematician. Even if I had not known of Eric before starting this project, I would probably have ended up using his paper, as it is such an important paper in the subject of this thesis.

I would like to thank my supervisor K.P. Hart, whose advice and feedback have been invaluable over the course of this project. I would also like to thank Joost de Groot for being part of my thesis committee. Finally, I would like to thank my family, my friends and my cat for the support they gave me during the writing of this thesis.

Introduction

For $p \in [1, \infty)$ and $q \in [1, \infty]$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, one can identify the spaces ℓ^q and $(\ell^p)'$ in the following way. An element $a \in \ell^q$ induces a bounded linear functional $f_a \in (\ell^p)'$, where $f_a : \ell^p \to \mathbb{R}$ is given by

$$f_a(x) = \sum_{n=1}^{\infty} a_n x_n. \tag{1}$$

Then the map $a \to f_a$ is an isometric isomorphism from ℓ^q onto $(\ell^p)'$. This is explained in more detail in section 2.1. This identification does not extend to the case $p = \infty$ and $q = 1$: the map $a \to f_a$ from ℓ^1 into $(\ell^\infty)'$, with $f_a \in (\ell^\infty)'$ defined in the obvious way (as in (1)), is a linear isometric embedding, but fails to be surjective. This is explained in section 2.2. The question that arises then is whether ℓ^1 and $(\ell^\infty)'$ can be identified in a different way. To answer this question, we show in section 2.3 that the spaces ℓ^1 and $(\ell^\infty)'$ are not even homeomorphic.

In chapter 3 we introduce a space that we can identify with $(\ell^\infty)'$. This is the space of bounded finitely additive measures on $\mathbb{N}$, written $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Finitely additive measures are similar to measures, but instead of *countable* additivity we only require *finite* additivity.[1] This is explained in section 3.1. A functional $\varphi \in (\ell^\infty)'$ induces a finitely additive measure μ_φ on $\mathbb{N}$ by setting $\mu_\varphi(E) = \varphi(\mathbb{1}_E)$. Then the map $\varphi \to \mu_\varphi$ is an isometric isomorphism from $(\ell^\infty)'$ onto $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. This is explained in section 3.2.

In chapter 4 we examine what kinds of finitely additive measures[2] on $\mathbb{N}$ exist.

- In section 4.1 we describe the measures that are induced by ℓ^1, where ℓ^1 is viewed as a (proper) subset of $(\ell^\infty)'$. Such measures are characterized by their values on singletons $\mu(\{n\})$. We prove that the positive measures that arise from ℓ^1 form precisely the collection of σ-additive measures on $\mathbb{N}$.
- In section 4.2 we prove the existence of shift-invariant measures on $\mathbb{N}$. These are measures that satisfy $\mu(E) = \mu(E+1)$ for all $E \subseteq \mathbb{N}$, where $E+1 = \{n+1 : n \in E\}$. We also describe how such measures are related to functionals in $(\ell^\infty)'$. We show that some of these measures correspond to so-called *Banach limits*, which are functionals in $(\ell^\infty)'$ with particularly nice properties.
- In section 4.3 we construct measures that are invariant under a given function. That is, if $f : \mathbb{N} \to \mathbb{N}$ is any function, then one can construct a measure μ on $\mathbb{N}$ satisfying $\mu(f^{-1}[E]) = \mu(E)$ for all $E \subseteq \mathbb{N}$.
- In section 4.4 we prove the existence of measures that extend the asymptotic density. The asymptotic density of a subset $A \subseteq \mathbb{N}$ is an intuitive way to measure how 'large' this set. Not all sets admit such a density. A measure μ is said to extend density if $\mu(A)$ is equal to the density of A whenever the latter exists. Hence measures that extend density have intuitively desirable properties.
- In section 4.5 we introduce ultrafilters and show how an ultrafilter on $\mathbb{N}$ induces in a natural way a measure on $\mathbb{N}$ that assumes only the values 0 and 1. We prove that some of these 0, 1-valued measures were already given by ℓ^1, but not all. We introduce the notion of an ultrafilter limit and show that ultrafilter limits are so-called multiplicative functionals. We show that the collection of ultrafilters on $\mathbb{N}$ can be identified with the collection of 0, 1-valued measures on $\mathbb{N}$, and the latter can in turn be identified with the collection of multiplicative functionals.
- In section 4.6 we prove the existence of so-called stretchable, thinnable and elastic measures. These are some of the nicest measures that one can think of, with intuitively desirable properties. Such measures extend density and moreover respect several forms of scaling in a desirable way.
- In section 4.7 we prove the existence of *unbounded* finitely additive measures on $\mathbb{N}$.

As a final note, let us mention that although this thesis describes a rather large collection of measures on $\mathbb{N}$, there is much more that can be said about these measures. The interested reader is referred to [5].

[1] We also allow finitely additive measures to assume negative values.

[2] In the description below, the term 'measure' means a *finitely* additive measure. A measure in the sense of countable additivity is called a *σ-additive measure* here.

1 Preliminaries

In this chapter we introduce some concepts from functional analysis that are needed in this report. Theorems are stated without proof; these proofs can be found in appendix A. The material presented here is based on [7].

Vector spaces are the central object of study in functional analysis. Often a vector space is endowed with some additional structure, such as a norm. For a linear transformation between two normed vector spaces we have the following characterization of continuity. The proof is easy so it is omitted.

Theorem 1.1 (Continuity of a linear transformation)**.** *Let X and Y be real*[3] *normed vector spaces and $T : X \to Y$ a linear transformation. The following are equivalent:*

(1) T is continuous;
(2) T is uniformly continuous;
(3) T is continuous at 0;
(4) There exists $k \in [0, \infty)$ such that for all $x \in X$ with $\|x\|_X \le 1$, we have $\|T(x)\|_Y \le k$;
(5) There exists $k \in [0, \infty)$ such that for all $x \in X$, we have $\|T(x)\|_Y \le k\|x\|_X$.

In view of property (5), we introduce the following definition.

Definition 1.2 (Boundedness)**.** *Let X and Y be real normed vector spaces and let $T : X \to Y$ be a linear transformation. We say that T is* ***bounded*** *if there exists $k \in [0, \infty)$ such that $\|T(x)\|_Y \le k\|x\|_X$ for all $x \in X$.*

From theorem 1.1, we see that a linear transformation is continuous if and only if it is bounded. We will use the words 'bounded' and 'continuous' interchangeably.

Remark. Note that our definition of boundedness does not agree with the usual definition of boundedness for a function T from X to $\mathbb{R}$, namely that there exists $M \in [0, \infty)$ such that $|T(x)| \le M$ for all $x \in X$. This potential ambiguity is not a really a problem, since apart from the zero map there is no linear transformation that is bounded in the latter sense.

Sometimes one has two normed vector spaces with a bijection between them that preserves all relevant properties. Such spaces can be considered the same. More precisely, we define the following.

Definition 1.3 (Isometry, isometric isomorphism)**.** *Let X and Y be real normed vector spaces. A linear transformation $T : X \to Y$ is called an* ***isometry*** *if $\|T(x)\|_Y = \|x\|_X$ for all $x \in X$. If T is moreover surjective, we call it an* ***isometric isomorphism****. If there exists an isometric isomorphism between X and Y, then X and Y are called* ***isometrically isomorphic****.*

Observe that an isometry is injective. An isometric isomorphism is continuous and bijective, and its inverse is also an isometric isomorphism. Note that being isometrically isomorphic is an equivalence relation (hence the last line in definition 1.3 makes sense). When X and Y are isometrically isomorphic we write this as $X \simeq Y$.

From a functional analytic point of view, spaces that are isometrically isomorphic are essentially identical. This is a useful notion, as a different way of looking at a space can sometimes give more insight into the structure of this space. We will see examples of this in the next chapters.

1.1 The space ℓ^∞

Recall that ℓ^∞ is the space of all bounded real sequences with norm $\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n|$. The following describes a useful dense subset (subspace) of ℓ^∞.

[3] In this thesis we will only be concerned with vector spaces over $\mathbb{R}$. Many results for real vector spaces hold equally well for complex vector spaces. However, sometimes results do not extend as easily as it seems, and one needs to be more careful.

Let us call a sequence $(a_n)_{n=1}^{\infty} \in \ell^{\infty}$ a **simple sequence** if its range is finite. Writing the range as $\{a_n : n \in \mathbb{N}\} = \{\alpha_1, \alpha_2, ..., \alpha_N\}$ (for $\alpha_j \in \mathbb{R}$ all different) and $A_j = a^{-1}(\{\alpha_j\}) \subseteq \mathbb{N}$, we see that we can write

$$a = \sum_{j=1}^{N} \alpha_j \mathbb{1}_{A_j}.$$

Conversely, every sequence of the previous form is a simple sequence. Note that the A_j may be chosen disjoint, as is the case here.

Lemma 1.4 (Density of simple sequences)**.** *The set of simple sequences is a dense subspace of ℓ^{∞}.*[4] *In particular, if $x \in \ell^{\infty}$ satisfies $x_n \geq 0$ for all n, then for every $\varepsilon > 0$, there exists a simple sequence a such that $\|x - a\|_{\infty} < \varepsilon$ and moreover $0 \leq a_n \leq x_n$ for all n.*

Proof. See lemma A.1. □

The next result will be used many times throughout this report. It gives easy conditions to check whether two bounded linear transformations from ℓ^{∞} to $\mathbb{R}$ are equal.

Proposition 1.5. *Let $\varphi, \psi : \ell^{\infty} \to \mathbb{R}$ be bounded linear transformations. If $\varphi(\mathbb{1}_E) = \psi(\mathbb{1}_E)$ for all $E \subseteq \mathbb{N}$, then $\varphi = \psi$.*

Proof. See proposition A.2. □

1.2 Dual spaces

Definition 1.6 (Linear functionals, dual space)**.** *Let X be a real normed vector space. Linear transformations from X to $\mathbb{R}$ are called* ***linear functionals****. The space of continuous linear functionals from X to $\mathbb{R}$ is called the* ***dual space of*** *X and is denoted X'. Endowed with pointwise addition and (real) scalar multiplication, this is a real vector space. The norm on X' is defined as*

$$\|f\|_{X'} = \sup\{|f(x)| : x \in X, \|x\|_X \leq 1\}.$$

Dual spaces are an important concept in functional analysis as they contain a lot of information from the original space. We will see this later in this report.

Remark. An important observation is that $|f(x)| \leq \|f\|_{X'} \|x\|_X$ for all $f \in X'$ and $x \in X$.

Theorem 1.7 (Duals are Banach spaces)**.** *Let X be a real normed vector space. Then X' is a Banach space.*

Proof. See theorem A.3. □

Consider a real vector space X. A situation that arises often is that we have a linear subspace $W \subseteq X$ and a linear map $f_W : W \to \mathbb{R}$, and we want to have a linear extension $f_X : X \to \mathbb{R}$ of f_W. In particular, if X is normed then we would like to have that $\|f_W\|_{W'} = \|f_X\|_{X'}$, i.e. the norm of the functional does not increase.

This is indeed possible, as we will see in theorem 1.9 and 1.10. First we introduce the following definition, which is useful to describe the 'size' of a linear functional.

Definition 1.8 (Sublinear functional)**.** *Let X be a real vector space. A* ***sublinear functional*** *on X is a map $p : X \to \mathbb{R}$ that satisfies*

(1) $p(x + y) \leq p(x) + p(y) \quad \forall x, y \in X$ (subadditivity);

[4] One can even show that for $x \in \ell^{\infty}$, there exists a sequence $(y^{(n)})_{n=1}^{\infty}$ of simple sequences $y^{(n)} \in \ell^{\infty}$ such that $y^{(n)} \uparrow x$ uniformly on $\mathbb{N}$; that is, $y_m^{(n)} \uparrow x_m \quad (n \to \infty)$ for every $m \in \mathbb{N}$ and $\|y^{(n)} - x\|_{\infty} \to 0$. This can be proven in a similar manner as here; then one uses a clever partition of the range in dyadic intervals. A proof can be found in [8, thm. 4.12].

(2) $p(cx) = cp(x) \quad \forall x \in X, \forall c \geq 0$ *(positive homogeneity).*

Remark. The following properties of a sublinear functional are easily checked: $p(0) = 0$, $-p(-x) \leq p(x)$ and $-p(y - x) \leq p(x) - p(y) \leq p(x - y)$.

The following is the main result of this section. It is one of the cornerstones of functional analysis.

Theorem 1.9 (Hahn-Banach theorem)**.** *Let X be a real vector space. Let $p : X \to \mathbb{R}$ be a sublinear functional on X. Suppose $W \subseteq X$ is a linear subspace of X and $f_W : W \to \mathbb{R}$ is a linear functional on W satisfying*

$$f_W(w) \leq p(w) \quad \forall w \in W.$$

Then f_W has a linear extension $f_X : X \to \mathbb{R}$ satisfying

$$f_X(x) \leq p(x) \quad \forall x \in X.$$

Proof. See lemma A.4 and theorem A.5. □

The following[5] is a direct consequence of theorem 1.9.

Theorem 1.10 (Hahn-Banach theorem)**.** *Let X be a real normed vector space, $W \subseteq X$ a linear subspace. Suppose $f_W \in W'$. Then there exists an extension $f_X \in X'$ of f_W such that $\|f_X\|_{X'} = \|f_W\|_{W'}$.*

Proof. See theorem A.6. □

1.3 Zorn's lemma

Definition 1.11 (Partial/total order, partially/totally ordered set)**.** *Let $\mathcal{M}$ be a nonempty set. A binary relation $\preceq$ on $\mathcal{M}$ is called a **partial order on** $\mathcal{M}$ if the following conditions are satisfied:*
(1) $x \preceq x$, for all $x \in \mathcal{M}$ (reflexivity);
(2) $x \preceq y$ and $y \preceq x$ implies $x = y$, for all $x \in \mathcal{M}$ (antisymmetry);
(3) $x \preceq y$ and $y \preceq z$ implies $x \preceq z$, for all $x, y, z \in \mathcal{M}$ (transitivity).
*In that case, the pair $(\mathcal{M}, \preceq)$ is called a **partially ordered set**. If $\preceq$ moreover satisfies*
(4) $x \preceq y$ or $y \preceq x$, for all $x, y \in \mathcal{M}$ (any two elements are comparable),
*then $\preceq$ is called a **total order on** $\mathcal{M}$ and $(M, \preceq)$ is called a **totally ordered set**.*

If $(\mathcal{M}, \preceq)$ is a partially (totally) ordered set and $\mathcal{N} \subseteq \mathcal{M}$ is a nonempty subset, the one can restrict the order $\preceq$ to $\mathcal{N}$ (we still write this restriction $\preceq_{\mathcal{N}}$ as $\preceq$), and then $(\mathcal{N}, \preceq)$ is a partially (totally) ordered set. That is, $\mathcal{N}$ inherits the order of $\mathcal{M}$.

Definition 1.12 (Maximal element)**.** *Let $(\mathcal{M}, \preceq)$ be a partially ordered set. An element $y \in \mathcal{M}$ is called a **maximal element** of $(\mathcal{M}, \preceq)$ if $y \preceq x \implies y = x$ holds for all $x \in \mathcal{M}$.*

It is very important to note that a *maximal* element y need not be a *maximum* in the sense that it does not necessarily hold that $x \preceq y$ for all $x \in \mathcal{M}$. A maximal element need not always exist, and if it exists, it may not be unique. However, if $(\mathcal{M}, \preceq)$ is totally ordered, then there exists at most one maximal element. Zorn's lemma will give us sufficient conditions for the existence of a maximal element. To formulate this theorem we first need to introduce the following definitions.

Definition 1.13 (Chain)**.** *Let $(\mathcal{M}, \preceq)$ be a partially ordered set and let $\mathcal{N} \subseteq \mathcal{M}$ be a nonempty subset. Then $\mathcal{N}$ is called a **chain** in $(\mathcal{M}, \preceq)$ if $\mathcal{N}$ is totally ordered (under the inherited order).*

Definition 1.14 (Upper bound)**.** *Let $(\mathcal{M}, \preceq)$ be a partially ordered set and let $\mathcal{N} \subseteq \mathcal{M}$ be nonempty. An element $y \in \mathcal{M}$ is called an **upper bound** for $\mathcal{N}$ in $(\mathcal{M}, \preceq)$ if $x \preceq y$ holds for all $x \in \mathcal{N}$.*

[5] The name 'Hahn-Banach theorem' is used to describe a collection of closely related results.

Note that for y to be an upper bound for $\mathcal{N}$, in particular y needs to be comparable to every element x of $\mathcal{N}$. But y need not belong to $\mathcal{N}$.

Theorem 1.15 (Zorn's lemma)**.** *Let $(\mathcal{M}, \preceq)$ be a nonempty partially ordered set with the property that every nonempty chain has an upper bound in $\mathcal{M}$.*[6] *Then $\mathcal{M}$ contains at least one maximal element.*

Zorn's lemma is equivalent to the axiom of choice. A proof of this result can be found in [3, thm. 5.10.3].

[6]Note that it is not required that the upper bound belongs to the chain itself.

2 The dual of ℓ^p, $1 \leq p < \infty$

In this chapter we investigate the dual space of ℓ^p for $1 \leq p < \infty$. For such p, we can find very elegant characterizations of these dual spaces. For $p = \infty$, the situation is more complicated, as we will see later. This chapter uses material from [6] and [7].

2.1 $\ell^q \simeq (\ell^p)'$ for $1 \leq p < \infty$

Let $p \in (1, \infty)$ and let $q = \frac{p}{p-1} \in (1, \infty)$, so that $\frac{1}{p} + \frac{1}{q} = 1$. We establish an isometric isomorphism $T : \ell^q \to (\ell^p)'$. First we assign to each element $a \in \ell^q$ a suitable functional $f_a : \ell^p \to \mathbb{R}$. Consider a fixed sequence $a = (a_n)_{n=1}^{\infty} \in \ell^q$. For all $x = (x_n)_{n=1}^{\infty} \in \ell^p$, we have by Hölder's inequality that

$$\sum_{n=1}^{\infty} |a_n x_n| \leq \|x\|_p \|a\|_q < \infty, \tag{2}$$

thus $\sum_{n=1}^{\infty} a_n x_n$ converges (absolutely). Therefore we may define for all $a \in \ell^q$,

$$f_a : \ell^p \to \mathbb{R} \quad f_a(x) = \sum_{n=1}^{\infty} a_n x_n.$$

Note that $f_a \in (\ell^p)'$: linearity of f_a is clear, and boundedness follows from equation (2). Thefore we can define the linear transformation

$$T : \ell^q \to (\ell^p)' \quad T(a) = f_a.$$

We claim that T is an isometric isomorphism. First we show that T is an isometry. Let $a \in \ell^q$, then we have that $\|f_a\|_{(\ell^p)'} \leq \|a\|_q$: for $x \in \ell^p$ with $\|x\|_p \leq 1$, we have

$$|f_a(x)| = \left| \sum_{n=1}^{\infty} a_n x_n \right| \leq \sum_{n=1}^{\infty} |a_n x_n| \leq \|a\|_q \|x\|_p \leq \|a\|_q.$$

Thus it remains to show that

$$\|a\|_q = \left(\sum_{n=1}^{\infty} |a_n|^q \right)^{1/q} \leq \|f_a\|_{(\ell^p)'}.$$

Fix $k \in \mathbb{N}$, it suffices to show that

$$\left(\sum_{n=1}^{k} |a_n|^q \right)^{1/q} \leq \|f_a\|_{(\ell^p)'}.$$

Define for all $1 \leq n \leq k$,

$$\gamma_n = \begin{cases} \frac{|a_n|^q}{a_n}, & a_n \neq 0, \\ 0, & a_n = 0, \end{cases}$$

so that $\gamma_n a_n = |a_n|^q$ for $1 \leq n \leq k$. Define $\gamma = (\gamma_1, \gamma_2, ..., \gamma_k, 0, 0, ...) \in \ell^p$. Writing $e^{(n)}$ for the vector in ℓ^p with a one in n-th position and zero elsewhere, we have $a_n = f_a(e^{(n)})$ and due to linearity of f_a we find

$$\sum_{n=1}^{k} |a_n|^q = \sum_{n=1}^{k} \gamma_n a_n = \sum_{n=1}^{k} \gamma_n f_a(e^{(n)}) = f_a \left(\sum_{n=1}^{k} \gamma_n e^{(n)} \right) = f_a(\gamma) = |f_a(\gamma)| \leq \|f_a\|_{(\ell^p)'} \|\gamma\|_p. \tag{3}$$

As p and q are Hölder conjugates we have that $pq - p = q$. It follows that for all $1 \leq n \leq k$,

$$|\gamma_n|^p = (|a_n|^{q-1})^p = |a_n|^{pq-p} = |a_n|^q.$$

Therefore equation (3) becomes

$$\sum_{n=1}^{k}|a_n|^q \le \|f_a\|_{(\ell^p)'}\|\gamma\|_p = \|f_a\|_{(\ell^p)'}\left(\sum_{n=1}^{k}|a_n|^q\right)^{1/p} \quad \text{hence} \quad \left(\sum_{n=1}^{k}|a_n|^q\right)^{1-1/p} = \left(\sum_{n=1}^{k}|a_n|^q\right)^{1/q} \le \|f_a\|_{(\ell^p)'}.$$

Letting $k \to \infty$, we find $\|a\|_q \le \|f_a\|_{(\ell^p)'}$. Combined with the converse estimate, it follows that T is an isometry.

Lastly we need to show that T is surjective. Let $f \in (\ell^p)'$. Define $a_n = f(e^{(n)})$ and $a = (a_1, a_2, ...)$. We claim that $f_a = f$. We first check that $a \in \ell^q$. Fix $k \in \mathbb{N}$; it suffices to have that

$$\sum_{n=1}^{k}|a_n|^q \le \|f\|^q_{(\ell^p)'} \tag{4}$$

Let $a' = (a_1, a_2, ..., a_k, 0, 0, ...) \in \ell^q$, so that

$$\sum_{n=1}^{k}|a_n|^q = \|a'\|^q_q = \|f_{a'}\|^q_{(\ell^p)'}. \tag{5}$$

Here we used that T is an isometry. Now observe that $\|f_{a'}\|_{(\ell^p)'} \le \|f\|_{(\ell^p)'}$: let $x \in \ell^p$ with $\|x\|_p \le 1$, consider $x' := (x_1, x_2, ..., x_k, 0, 0, ...) \in \ell^p$; then $\|x'\|_p \le 1$ and

$$|f_{a'}(x)| = \left|\sum_{n=1}^{k} a_n x_n\right| = |f(x')| \le \|f\|_{(\ell^p)'},$$

thus indeed $\|f_{a'}\|_{(\ell^p)'} \le \|f\|_{(\ell^p)'}$. Combining this with (5), we see that (4) follows, hence $a \in \ell^q$.

Finally we must check that $f = f_a$. Let $x \in \ell^p$. The crucial part in the proof of surjectivity (which does not hold for $p = \infty$) is that

$$x = \sum_{n=1}^{\infty} x_n e^{(n)}.$$

Then using continuity and linearity of f, we find

$$f(x) = \sum_{n=1}^{\infty} f(x_n e^{(n)}) = \sum_{n=1}^{\infty} a_n x_n = f_a(x).$$

Altogether, we conclude that T is an isometric isomorphism.

The dual of ℓ^1

With exactly the same procedure one can construct an isometric isomorphism $T : \ell^\infty \to (\ell^1)'$ with $T(a) = f_a$, where T and f_a are defined in the obvious way. Showing that T is an isometric isomorphism requires even less effort than before as the norm on ℓ^∞ is very simple to work with.

2.2 The problem with ℓ^∞

In this section we show that the characterization of $(\ell^p)'$ discussed before does not extend to the case $p = \infty$. For all $a \in \ell^1$, let

$$f_a : \ell^\infty \to \mathbb{R} \quad f_a(x) = \sum_{n=1}^{\infty} a_n x_n$$

As before, f_a is well defined (using in this case the simple estimate $|a_n x_n| \leq |a_n| \cdot \|x\|_\infty$) and satisfies $f_a \in (\ell^\infty)'$. Define $T : \ell^1 \to (\ell^\infty)'$ by $T(a) = f_a$. In the same way as before, one checks that T is an isometry.

Now let us investigate the surjectivity of T. Consider any functional $f \in (\ell^\infty)'$. Suppose $a \in \ell^1$ is to satisfy $f = f_a$, then in particular for $x = e^{(n)} \in \ell^\infty$, we must have $f(e^{(n)}) = f_a(e^{(n)}) = a_n$. Thus define $a_n = f(e^{(n)})$ and $a = (a_1, a_2, ...)$. First we check that $a \in \ell^1$. Let $k \in \mathbb{N}$ and consider $a' = (a_1, a_2, ..., a_k, 0, 0, ...) \in \ell^1$; then $\sum_{n=1}^k |a_n| = \|a'\|_1 = \|f_{a'}\|_{(\ell^\infty)'} \leq \|f\|_{(\ell^\infty)'}$. Letting $k \to \infty$, we find $a \in \ell^1$.

The crucial part is now to show that $f = f_a$. In the case where $p < \infty$, we used that for all $x \in \ell^p$,

$$x = \sum_{n=1}^{\infty} x_n e^{(n)} \tag{6}$$

to interchange the limit and the continuous function f. However, such an approach is not possible for ℓ^∞, as we do not have (6) for general $x \in \ell^\infty$. Consider for instance $x = (1, 1, 1, ...) \in \ell^\infty$; then (6) does not even converge.

Thus the previous method of proving surjectivity cannot be used here. In fact, T is not surjective. We construct a functional $f \in (\ell^\infty)'$ that is not of the form f_a. Consider the subspace c of ℓ^∞ consisting of convergent sequences. Define

$$\tilde{f} : c \to \mathbb{R} \quad \tilde{f}(x) = \lim_{n\to\infty} x_n.$$

This is a bounded linear functional, hence the Hahn-Banach theorem (theorem 1.10) gives us an extension $f \in (\ell^\infty)'$ of $\tilde{f}$. We claim that f is not of the form f_a. Suppose $a \in \ell^1$ is such that $f = f_a$. As $\sum_{n=1}^\infty |a_n| < \infty$, there exists $N \in \mathbb{N}$ such that $\sum_{n=N}^\infty |a_n| < 1$. Let $x = (0, 0, ..., 0, 1, 1, ...) \in c$ (the first one in N-th position). Then we have $f(x) = \tilde{f}(x) = 1$, but

$$f_a(x) = \sum_{n=1}^{\infty} a_n x_n = \sum_{n=N}^{\infty} a_n \leq \sum_{n=N}^{\infty} |a_n| < 1,$$

which is a contradiction. Hence T is not surjective.

Remark. The proof of surjectivity of the mapping $T : \ell^1 \to (\ell^\infty)'$ that we used earlier for finite p did not work because we could not write (6) for general $x \in \ell^\infty$. Consider the subspace c_0 of ℓ^∞ and note that (6) holds for all $x \in c_0$. Hence if we define $T : \ell^1 \to (c_0)'$ in the obvious way then we do have surjectivity of the map T. This proves that ℓ^1 and $(c_0)'$ are isometrically isomorphic.

In fact, if we define $T : \ell^1 \to c'$ in the obvious way, then we can also show that this map is surjective. Equation (6) may not hold, but this is resolved by writing $x \in c$ as $x = \alpha \mathbb{1} + y$ where $\alpha = \lim_{n\to\infty} x_n$, $\mathbb{1} = (1, 1, ...)$ and $y \in c_0$ is defined by $y_n = x_n - \alpha$ and then applying (6) to y. Thus ℓ^1 and c' are isometrically isomorphic.

We saw in this section that the obvious candidate T for an isometric isomorphism between ℓ^1 and $(\ell^\infty)'$ did not work. But this does not rule out the possibility that other choices for T could work. In the next section we prove that this is not the case.

2.3 No homeomorphism between ℓ^1 and $(\ell^\infty)'$

In this section we prove that the spaces ℓ^1 and $(\ell^\infty)'$ are not isometrically isomorphic at all. In fact, we prove the stronger claim that they are not even homeomorphic. For the proof we first need some auxiliary results, which are interesting in themselves.

We start by proving that homeomorphisms preserve separability.

Theorem 2.1. *Let X and Y be real normed vector spaces, and suppose X and Y are homeomorphic. Then X is separable if and only if Y is separable.*

Proof. By symmetry it suffices to prove only one direction. Suppose X is separable and let $A \subseteq X$ be a countable dense subset. Let $f : X \to Y$ be a homeomorphism. The $f[A]$ is the desired countable dense set in Y. Countability is clear, and note that

$$Y = f[X] = f[\overline{A}] \overset{(*)}{\subseteq} \overline{f[A]},$$

so that $Y = \overline{f[A]}$. Note that $(*)$ follows from the fact that $\overline{A} \subseteq f^{-1}[\overline{f[A]}]$ as $f^{-1}[\overline{f[A]}]$ is a closed set containing A. This finishes the proof. □

Next we need the following results about the separability of ℓ^1 and ℓ^∞.

Lemma 2.2. *ℓ^1 is separable.*

Proof. Let for all $k \in \mathbb{N}$, $A_k := \{(x_1, x_2, ..., x_k, 0, 0, ...) : x_j \in \mathbb{Q}, 1 \leq j \leq k\}$ and define $A := \bigcup_{k=1}^\infty A_k \subseteq \ell^1$. Then A is countable, being a countable union of countable sets. We claim that A is dense in ℓ^1. Let $z \in \ell^1$, and let $\varepsilon > 0$. We need $x \in A$ such that $\|z - x\|_1 < \varepsilon$. As $\sum_{n=1}^\infty |z_n| < \infty$, there exists $N \in \mathbb{N}$ such that $\sum_{n=N+1}^\infty |z_n| < \varepsilon/2$. Define $x' := (z_1, z_2, ..., z_N, 0, 0, ...) \in \ell^1$. Then $\|z - x'\|_1 = \sum_{n=N+1}^\infty |z_n| < \varepsilon/2$. For every $1 \leq n \leq N$, we may choose $x_n \in \mathbb{Q}$ such that $|x_n - x'_n| < \frac{\varepsilon}{2N}$. Let $x := (x_1, x_2, ..., x_N, 0, 0, ...) \in A$, then we see

$$\|x - x'\|_1 = \sum_{n=1}^N |x_n - x'_n| < \sum_{n=1}^N \frac{\varepsilon}{2N} = \varepsilon/2.$$

Altogether, we find

$$\|z - x\|_1 = \|z - x' + x' - x\|_1 \leq \|z - x'\|_1 + \|x - x'\|_1 < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

which concludes the proof. □

Remark. With a small modification of the previous proof one shows that ℓ^p is separable for each $1 \leq p < \infty$.

Lemma 2.3. *ℓ^∞ is not separable.*

Proof. Let $A \subseteq \ell^\infty$ be a countable subset. We show that A cannot be dense. Write $A = \{a^{(n)} : n \in \mathbb{N}\}$. Define for all $j \in \mathbb{N}$

$$x_j = \begin{cases} a_j^{(j)} + 1, & |a_j^{(j)}| \leq 1, \\ 0, & |a_j^{(j)}| > 1, \end{cases}$$

and set $x = (x_1, x_2, ...)$. Then $x \in \ell^\infty$, as $|x_j| \leq 2$ for every $j \in \mathbb{N}$. Then for every $n \in \mathbb{N}$, we have $|a_n^{(n)} - x_n| \geq 1$ hence $\left\|a^{(n)} - x\right\|_\infty \geq 1$. Hence A is not dense in ℓ^∞. As A was arbitrary, we conclude that ℓ^∞ is not separable. □

Our next aim is to show that separability of the dual space X' implies separability of the original space X. For this the next theorem will be used.

Theorem 2.4. *Let X be a real normed vector space and $W \subseteq X$ a linear subspace. Suppose $x \in X$ is such that*

$$\delta := \inf_{w \in W} \|x - w\|_X > 0.$$

Then there exists $f \in X'$ such that $\|f\|_{X'} = 1$ and $f(w) = 0$ for all $w \in W$.

Proof. Define $Y = \text{span}\{x\} \oplus W = \{cx + w : c \in \mathbb{R}, w \in W\}$ and let $g_Y : Y \to \mathbb{R}$ be defined by $g_Y(cx+w) = c\delta$. Clearly g_Y is linear. We claim that $g_Y \in Y'$. Let $c \in \mathbb{R}$ and $w \in W$, then (w.l.o.g. $c \neq 0$)

$$|g_Y(cx+w)| = |c|\delta \leq |c|\left\|x - \frac{-w}{c}\right\|_X = \|cx+w\|_X.$$

From this we see that $g_Y \in Y'$. Let $g \in X'$ be a Hahn-Banach extension of g_Y (theorem 1.10). Note that $\|g\|_{X'} \neq 0$ as $g(x) = g_Y(x) = \delta$. Finally, define $f = \frac{1}{\|g\|_{X'}} g \in X'$. Then f has the desired properties.[7] ☐

Theorem 2.5. *Let X be a real normed vector space. If X' is separable, then so is X.*

Proof. We may without loss of generality assume $X \neq \{0\}$.[8] Define $\mathcal{B}_{X'} = \{f \in X' : \|f\|_{X'} = 1\}$. As X' is separable, so is $\mathcal{B}_{X'}$.[9] Thus we may choose a collection $\{f_1, f_2, ...\} \subseteq \mathcal{B}_{X'}$ that is dense in $\mathcal{B}_{X'}$. Then for every $n \in \mathbb{N}$, we may choose $x_n \in X$ such that $\|x_n\|_X = 1$ and $|f_n(x_n)| \geq \frac{1}{2}$. Define $W = \text{span}_{\mathbb{Q}}\{x_1, x_2, ...\}$, by which we mean the set of finite *rational* linear combinations of elements from $\{x_1, x_2, ...\}$. We claim that W is the desired countable dense set in X. Countability is clear. We also have that $\overline{W} = X$. Suppose namely that $\overline{W} \subsetneq X$, then there is some $x^* \in X \setminus \overline{W}$. Then we have that $\inf_{w \in \overline{W}} \|x^* - w\|_X > 0$ thus by theorem 2.4 there exists $f \in X'$ such that $\|f\|_{X'} = 1$ and $f(w) = 0$ for all $w \in \overline{W}$. Then we obtain for all $n \in \mathbb{N}$ that

$$\frac{1}{2} \leq |f_n(x_n)| = |f_n(x_n) - f(x_n)| \leq \|f_n - f\|_{X'}\|x_n\|_X = \|f_n - f\|_{X'}.$$

But this contradicts density of $\{f_1, f_2, ...\}$ in $\mathcal{B}_{X'}$. Hence $\overline{W} = X$. This finishes the proof. ☐

Finally we have the main theorem of this section.

Theorem 2.6. *ℓ^1 and $(\ell^\infty)'$ are not homeomorphic.*

Proof. This follows immediately by combining the results from this section. ☐

[7] In fact one can show that $\|g_Y\|_{Y'} = 1$, so that (by the Hahn-Banach theorem) the function g can be chosen to satisfy $\|g\|_{X'} = 1$. Then we have that $f = g$ hence $f(x) = \delta$. This has very useful implications. For instance, this improvement of theorem 2.4 implies the following, with X a real normed vector space:

- For all $x \in X$, we have $\|x\|_X = \sup\{|f(x)| : f \in X', \|f\|_{X'} \leq 1\}$;
- For all $x, y \in X$ with $x \neq y$, there exists $f \in X'$ such that $f(x) \neq f(y)$.

[8] We exclude this trivial case because then we would have $\mathcal{B}_{X'} = \varnothing$.

[9] Here we use that for a separable metric space (M, d) and a subset $A \subseteq M$, we have that A is separable as well (under the inherited metric). This is not as trivial as it might look. To see why this is true, let $(x_n)_{n=1}^\infty$ be a dense sequence in M and set

$$\Theta = \{(n, m) \in \mathbb{N} \times \mathbb{N} : B_{1/m}(x_n) \cap A \neq \varnothing\}.$$

For all $(n, m) \in \Theta$, choose $y_{n,m} \in B_{1/m}(x_n) \cap A$. Then $\{y_{n,m} : (n, m) \in \Theta\}$ is a countable dense set in A.

3 A description of $(\ell^\infty)'$

So far we have not found a way to characterize $(\ell^\infty)'$. In this chapter we introduce a space that we can identify with $(\ell^\infty)'$. This is a space of measures, but in a more general sense. This chapter uses material from [2].

3.1 Finitely additive measures

Definition 3.1 (Measure). *A* ***measure*** *on a measurable space $(S, \mathcal{A})$ is a function $\mu : \mathcal{A} \to [0, \infty]$ that satisfies $\mu(\varnothing) = 0$ and is σ-additive, i.e. for all pairwise disjoint sequences $A_1, A_2, ...$ of sets in $\mathcal{A}$, we have that*

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j). \tag{7}$$

Sometimes it is impossible to construct a measure with desired properties on a given measurable space, due to the σ-additivity requirement. For instance, it is impossible to construct a probability measure on $\mathbb{N}$ (that is, $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ being the measurable space) that assigns to each singleton the same probability. One suggestion is to relax σ-additivity to the weaker condition of *finite additivity.* In that case we speak of a *finitely additive measure.*

Definition 3.2 (Finitely additive measure). *Let $(S, \mathcal{A})$ be a measurable space.*

(1) A map $\mu : \mathcal{A} \to \mathbb{R}$ is called a ***finitely additive measure*** *on $\mathcal{A}$ if the following conditions are satisfied:*
  - *$\mu(\varnothing) = 0$;*[10]
  - *If $A, B \in \mathcal{A}$ are disjoint, then $\mu(A \cup B) = \mu(A) + \mu(B)$ (**finite additivity**).*

(2) A finitely additive measure μ on $\mathcal{A}$ is called ***bounded*** *if $\sup\{|\mu(F)| : F \in \mathcal{A}\} < \infty$; that is, if the range of μ is a bounded subset of $\mathbb{R}$.*

(3) A finitely additive measure μ on $\mathcal{A}$ is called ***positive*** *if $\mu(F) \geq 0$ for all $F \in \mathcal{A}$. Similarly, μ is called* ***negative*** *if $\mu(F) \leq 0$ for all $F \in \mathcal{A}$.*[11]

Remark. One should note that a measure (in the sense of definition 3.1) is a finitely additive measure if and only if it does not assume value ∞. Moreover, a finitely additive measure need not be a measure (as in definition 3.1), for two reasons:

- It may assume negative values.
- Even if it is positive, it may not be σ-additive. We will see examples of this in chapter 4.

To prevent confusion, a measure as in definition 3.1 will from now on explicitly be called a **σ-additive measure**. A finitely additive measure will sometimes (sloppily) be referred to as a measure.

We continue with some basic properties of finitely additive measures.

Theorem 3.3. *Let $(S, \mathcal{A})$ be a measurable space and let $\mu : \mathcal{A} \to \mathbb{R}$ be a finitely additive measure. Then the following hold:*

(1) If $F_1, F_2, ..., F_n \in \mathcal{A}$ are disjoint, then

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) = \sum_{i=1}^{n} \mu(F_i).$$

(2) If $E, F \in \mathcal{A}$ satisfy $E \subseteq F$, then $\mu(F \setminus E) = \mu(F) - \mu(E)$.

(3) For all $E, F \in \mathcal{A}$, we have $\mu(E \cup F) = \mu(E) + \mu(F) - \mu(E \cap F)$.

(4) If μ is positive and $E, F \in \mathcal{A}$ satisfy $E \subseteq F$, then $\mu(E) \leq \mu(F)$.

[10] Actually we already get this from the second requirement: $\mu(\varnothing) = \mu(\varnothing) + \mu(\varnothing)$. This does not generally follow if μ is allowed to take value $\pm\infty$.

[11] It is important to note that 'negative' is not the negation of 'positive', and vice versa.

(5) If μ is positive, then μ is bounded.
(6) If μ is positive and $F_1, F_2, ..., F_n \in \mathcal{A}$, then

$$\mu\left(\bigcup_{i=1}^{n} F_i\right) \leq \sum_{i=1}^{n} \mu(F_i).$$

(7) If μ is positive and $E_1, E_2, ... \in \mathcal{A}$ are pairwise disjoint, then

$$\sum_{n=1}^{\infty} \mu(E_n) \leq \mu\left(\bigcup_{n=1}^{\infty} E_n\right).$$

Proof. The proofs are not difficult so they are left to the reader. In (6), one can write an arbitrary union $\bigcup_{i=1}^{n} F_i$ as a disjoint union $\bigcup_{i=1}^{n} E_i$ where $E_1 = F_1$, $E_2 = F_2 \setminus F_2$, $E_3 = F_3 \setminus (F_1 \cup F_2)$ et cetera. $\square$

In the context of this thesis we are mainly interested in *bounded* finitely additive measures.[12] The space of bounded finitely additive measures on the measurable space $(S, \mathcal{A})$ will be denoted $\mathrm{ba}(S, \mathcal{A})$ (ba meaning 'bounded additive'). For $\mu, \nu \in \mathrm{ba}(S, \mathcal{A})$ we define for $F \in \mathcal{A}$, $(\mu + \nu)(F) = \mu(F) + \nu(F)$, and for scalars $a \in \mathbb{R}$ we set $(a\mu)(F) = a\mu(F)$. The space $\mathrm{ba}(S, \mathcal{A})$ is closed under addition and scalar multiplication, thus it is a real vector space. For $\mu, \nu \in \mathrm{ba}(S, \mathcal{A})$, we say that $\mu \leq \nu$ if $\mu(F) \leq \nu(F)$ for all $F \in \mathcal{A}$.

Definition 3.4 (Positive, negative and total variation)**.** *Let $(S, \mathcal{A})$ be a measurable space and $\mu \in \mathrm{ba}(S, \mathcal{A})$. Define*
*(1) $\mu^+ : \mathcal{A} \to \mathbb{R}$ $\quad \mu^+(F) = \sup\{\mu(E) : E \subseteq F, E \in \mathcal{A}\}$. Then μ^+ is called the **positive variation** of μ.*
*(2) $\mu^- : \mathcal{A} \to \mathbb{R}$ $\quad \mu^-(F) = -\inf\{\mu(E) : E \subseteq F, E \in \mathcal{A}\}$. Then μ^- is called the **negative variation** of μ.*
*(3) $|\mu| : \mathcal{A} \to \mathbb{R}$ $\quad |\mu| = \mu^+ + \mu^-$. Then $|\mu|$ is called the **total variation** of μ.*

From boundedness of μ it follows that $\mu^\pm$ indeed map into $\mathbb{R}$, so they are well-defined.

Proposition 3.5. *Let $(S, \mathcal{A})$ be a measurable space, let $\mu \in \mathrm{ba}(S, \mathcal{A})$. Then the following hold:*
(1) μ^+, $\mu^-, |\mu| \in \mathrm{ba}(S, \mathcal{A})$ and they are positive.
(2) $\mu = \mu^+ - \mu^-$.
(3) For any $F \in \mathcal{A}$, we have

$$|\mu|(F) = \sup\{\mu(E) - \mu(F \setminus E) : E \subseteq F, E \in \mathcal{A}\}.$$

(4) For any $F \in \mathcal{A}$, we have

$$|\mu|(F) = \sup\Big\{\sum_{i=1}^{n} |\mu(E_i)| : (E_i)_{i=1}^{n} \text{ disjoint in } \mathcal{A},\ \bigcup_{i=1}^{n} E_i \subseteq F\Big\}.$$

Proof. (1). Let $A, B \in \mathcal{A}$ be disjoint. We show that $\mu^+(A) + \mu^+(B) = \mu^+(A \cup B)$. Let $E, F \in \mathcal{A}$ be such that $E \subseteq A$ and $F \subseteq B$. Then E and F are disjoint and $E \cup F \subseteq A \cup B$, hence

$$\mu(E) + \mu(F) = \mu(E \cup F) \leq \mu^+(A \cup B).$$

Taking the supremum over all such E and F, we find $\mu^+(A) + \mu^+(B) \leq \mu^+(A \cup B)$. For the converse estimate, let $G \in \mathcal{A}$ be such that $G \subseteq A \cup B$. Then $G \cap A$ and $G \cap B$ are in $\mathcal{A}$ and

$$\mu(G) = \mu(G \cap A) + \mu(G \cap B) \leq \mu^+(A) + \mu^+(B).$$

[12] In section 4.7 we give necessary and sufficient conditions on a measurable space $(S, \mathcal{A})$ for the existence of an unbounded finitely additive measure on this space.

Taking the supremum over all such G, the converse estimate $\mu^+(A \cup B) \leq \mu^+(A) + \mu^+(B)$ follows, hence this must be an equality. It is clear that μ^+ is positive, and it is bounded as μ is bounded. Thus we have $\mu^+ \in \mathrm{ba}(S, \mathcal{A})$. Now observe that $\mu^- = (-\mu)^+$ hence the claim for μ^- follows immediately. Finally, we have that $|\mu| = \mu^+ + \mu^- \in \mathrm{ba}(S, \mathcal{A})$ and it is positive.

(2). Let $F \in \mathcal{A}$. Observe that $\{\mu(E) : E \subseteq F, E \in \mathcal{A}\} = \{\mu(F \setminus G) : G \subseteq F, G \in \mathcal{A}\}$. Thus we have

$$\begin{aligned}\mu^+(F) &= \sup\{\mu(E) : E \subseteq F, E \in \mathcal{A}\} = \sup\{\mu(F \setminus G) : G \subseteq F, G \in \mathcal{A}\} = \sup\{\mu(F) - \mu(G) : G \subseteq F, G \in \mathcal{A}\} \\ &= \mu(F) + \sup\{-\mu(G) : G \subseteq F, G \in \mathcal{A}\} = \mu(F) - \inf\{\mu(G) : G \subseteq F, G \in \mathcal{A}\} = \mu(F) + \mu^-(F).\end{aligned}$$

From this the desired result follows.

(3). Let $F \in \mathcal{A}$ and define

$$\xi = \sup\{\mu(E) - \mu(F \setminus E) : E \subseteq F, E \in \mathcal{A}\}.$$

Note that this is equal to $\xi = \sup\{\mu(F \setminus E) - \mu(E) : E \subseteq F, E \in \mathcal{A}\}$ as these sets are equal. We first show that $\mu^+(F) + \mu^-(F) \leq \xi$. Let $A, B \in \mathcal{A}$ be such that $A \subseteq F$ and $B \subseteq F$. It suffices to show that

$$\mu(A) - \mu(B) \leq \xi. \tag{8}$$

Note that $\mu(A) - \mu(B) = \mu(A \setminus B) - \mu(B \setminus A)$. Define[13] $C = F \setminus (A \triangle B)$. Observation:

$$C \cup (A \setminus B) = F \setminus (B \setminus A) \quad \text{and} \quad C \cup (B \setminus A) = F \setminus (A \setminus B),$$

and these are disjoint unions. Now if $\mu(C) \geq 0$, we have

$$\begin{aligned}\mu(A) - \mu(B) &= \mu(A \setminus B) - \mu(B \setminus A) \leq \mu(A \setminus B) + \mu(C) - \mu(B \setminus A) = \mu((A \setminus B) \cup C) - \mu(B \setminus A) \\ &= \mu(F \setminus (B \setminus A)) - \mu(B \setminus A) \leq \xi,\end{aligned}$$

so that (8) holds. Similarly, if $\mu(C) < 0$, we have

$$\begin{aligned}\mu(A) - \mu(B) &= \mu(A \setminus B) - \mu(B \setminus A) \leq \mu(A \setminus B) - \mu(B \setminus A) - \mu(C) = \mu(A \setminus B) - \mu((B \setminus A) \cup C) \\ &= \mu(A \setminus B) - \mu(F \setminus (A \setminus B)) \leq \xi,\end{aligned}$$

which proves (8). The converse estimate $\xi \leq \mu^+(F) + \mu^-(F)$ is clear by the definition of $\mu^\pm$. This proves the desired result.

(4). Let $F \in \mathcal{A}$ and write

$$\gamma = \sup\Big\{\sum_{i=1}^n |\mu(E_i)| : (E_i)_{i=1}^n \text{ disjoint in } \mathcal{A}, \ \bigcup_{i=1}^n E_i \subseteq F\Big\}.$$

We first show that $|\mu|(F) \leq \gamma$. Let $E \in \mathcal{A}$ be such that $E \subseteq F$. By property (3) in proposition 3.5, it suffices to have that $\mu(E) - \mu(F \setminus E) \leq \gamma$. Indeed, as E and $F \setminus E$ are disjoint sets in $\mathcal{A}$ contained in F, we have that

$$\mu(E) - \mu(F \setminus E) \leq |\mu(E)| + |\mu(F \setminus E)| \leq \gamma.$$

For the converse estimate, let $E_1, E_2, ..., E_n \in \mathcal{A}$ be pairwise disjoint sets such that $\bigcup_{i=1}^n E_i \subseteq F$. Define

[13] $A \triangle B := (A \setminus B) \cup (B \setminus A)$ is called the **symmetric difference** of A and B.

$E_{n+1} = F \setminus \bigcup_{i=1}^n E_i$. Write $I = \{1 \leq i \leq n+1 : \mu(E_i) \geq 0\}$ and $J = \{1 \leq i \leq n+1 : \mu(E_i) < 0\}$. Then

$$\sum_{i=1}^n |\mu(E_i)| \leq \sum_{i=1}^{n+1} |\mu(E_i)| = \sum_{i\in I} \mu(E_i) - \sum_{i \in J} \mu(E_i) = \mu\left(\bigcup_{i\in I} E_i\right) - \mu\left(\bigcup_{i\in J} E_i\right)$$
$$= \mu\left(\bigcup_{i\in I} E_i\right) - \mu\left(F \setminus \bigcup_{i\in I} E_i\right) \leq |\mu|(F),$$

using property (3) in the final step. Taking the supremum over all admissible collections $(E_i)_{i=1}^n$, we find $\gamma \leq |\mu|(F)$. Combining this with the converse estimate, we derive property (4). ☐

The decomposition of a bounded measure μ into two positive measures $\mu = \mu^+ - \mu^-$ as in proposition 3.5 is called a **Jordan decomposition**.

We introduce the following norm on $\mathrm{ba}(S, \mathcal{A})$.

Definition 3.6. *For $\mu \in \mathrm{ba}(S, \mathcal{A})$, we define the **total variation norm** of μ by $\|\mu\|_{\mathrm{ba}} = |\mu|(S)$.*

Proposition 3.7. *The total variation norm $\|\cdot\|_{\mathrm{ba}}$ is a norm on $\mathrm{ba}(S, \mathcal{A})$.*

Proof. The defining properties of a norm are easily derived from (4) in proposition 3.5. ☐

3.2 $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ as the dual space of ℓ^∞

From now on the measurable space that we consider is $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. In this section we construct an isometric isomorphism from $(\ell^\infty)'$ onto $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

With every functional $\varphi \in (\ell^\infty)'$ we associate the following finitely additive measure $\mu_\varphi \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$:

$$\mu_\varphi : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \mu_\varphi(E) = \varphi(\mathbb{1}_E).$$

From linearity of φ it follows that μ_φ is indeed finitely additive. For boundedness of this measure, observe that for all $E \subseteq \mathbb{N}$, we have

$$|\mu_\varphi(E)| = |\varphi(\mathbb{1}_E)| \leq \|\varphi\|_{(\ell^\infty)'} \|\mathbb{1}_E\|_\infty \leq \|\varphi\|_{(\ell^\infty)'}.$$

Define the linear map $R : (\ell^\infty)' \to \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ by $R(\varphi) = \mu_\varphi$. We claim that R is an isometric isomorphism.

Proposition 3.8. *R is an isometry.*

Proof. Let $\varphi \in (\ell^\infty)'$. We first show that $\|\varphi\|_{(\ell^\infty)'} \leq \|\mu_\varphi\|_{\mathrm{ba}}$. We may assume without loss of generality that $\varphi \neq 0$. Let $x \in \ell^\infty$ be such that $\|x\|_\infty \leq 1$ and let $\varepsilon > 0$. We show that $|\varphi(x)| \leq \|\mu_\varphi\|_{\mathrm{ba}} + \varepsilon$. Let $a \in \ell^\infty$ be a simple sequence such that

$$\|x - a\|_\infty < \frac{\varepsilon}{\|\varphi\|_{(\ell^\infty)'}},$$

and moreover $|a_n| \leq 1$ for all n (lemma 1.4). Then we have that

$$|\varphi(x)| < |\varphi(a)| + \varepsilon \tag{9}$$

Now write $a = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ with $|\alpha_i| \leq 1$ and $A_1, A_2, ..., A_n \subseteq \mathbb{N}$ pairwise disjoint. Then

$$|\varphi(a)| = \left|\varphi\left(\sum_{i=1}^n \alpha_i \mathbb{1} A_i\right)\right| \leq \sum_{i=1}^n |\alpha_i| \cdot |\varphi(\mathbb{1}_{A_i})| \leq \sum_{i=1}^n |\mu_\varphi(A_i)| \overset{(*)}{\leq} \|\mu_\varphi\|_{\mathrm{ba}},$$

using proposition 3.5 in $(*)$. Combined with (9), this proves that $|\varphi(x)| < \|\mu_\varphi\|_{\mathrm{ba}} + \varepsilon$. Letting $\varepsilon \downarrow 0$ and taking the supremum over all admissible x, the desired inequality follows.

Now for the converse estimate, let $A_1, A_2, ..., A_n \subseteq \mathbb{N}$ be pairwise disjoint. For $1 \leq i \leq n$, take $\gamma_i \in \{-1, 1\}$ such that $|\mu(A_i)| = \gamma_i \mu(A_i)$. Define $x = \sum_{i=1}^n \gamma_i \mathbb{1}_{A_i}$, then we have that $\|x\|_\infty \leq 1$. Then

$$\sum_{i=1}^n |\mu(A_i)| = \sum_{i=1}^n \gamma_i \mu(A_i) = \sum_{i=1}^n \gamma_i \varphi(\mathbb{1}_{A_i}) = \varphi(x) \leq |\varphi(x)| \leq \|\varphi\|_{(\ell^\infty)'}. \tag{10}$$

Taking the supremum over all disjoint collections $(A_i)_{i=1}^n$ of subsets of $\mathbb{N}$ and using proposition 3.5, we derive that $\|\mu_\varphi\|_{\text{ba}} \leq \|\varphi\|_{(\ell^\infty)'}$. Combined with the converse estimate, this proves that R is an isometry. □

The proof of surjectivity of the map R involves integrating with respect to a finitely additive measure. We briefly indicate how one can integrate a function $x \in \ell^\infty$ with respect to a finitely additive measure μ.

Integration with respect to a finitely additive measure

Let $\mu \in \text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ be a finitely additive measure. First suppose μ is positive.

- *Step 1: Positive simple sequences.* Let $a = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ be a positive simple sequence ($\alpha_i \geq 0$) with $A_1, A_2, ..., A_n \subseteq \mathbb{N}$ pairwise disjoint. Define

$$\int_{\mathbb{N}} a \, d\mu = \sum_{i=1}^n \alpha_i \mu(A_i).$$

  As in the construction of the Lebesgue integral, one can show that this is well-defined (i.e. does not depend on the representation) and this integral is linear on positive simple sequences.
- *Step 2: General positive sequences.* Let $x \in \ell^\infty$ be a positive sequence. Define

$$\int_{\mathbb{N}} x \, d\mu = \sup\Big\{ \int_{\mathbb{N}} a \, d\mu : 0 \leq a \leq x, \ a \text{ simple} \Big\}. \tag{11}$$

  Again, this integral is well-defined (i.e., coincides with step 1 whenever x is simple) and linear on positive sequences. To prove this, one uses the fact that simple sequences are dense in ℓ^∞. Note that (11) is finite as x is bounded.
- *Step 3: General sequences.* Let $x \in \ell^\infty$ and write $x = x^+ - x^-$, where $x_n^\pm := \max\{\pm x_n, 0\} \geq 0$. Define

$$\int_{\mathbb{N}} x \, d\mu = \int_{\mathbb{N}} x^+ \, d\mu - \int_{\mathbb{N}} x^- \, d\mu$$

Now suppose μ is a general finitely additive measure in $\text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

- *Step 4: General finitely additive measure.* Write $\mu = \mu^+ - \mu^-$. Define for all $x \in \ell^\infty$,

$$\int_{\mathbb{N}} x \, d\mu = \int_{\mathbb{N}} x \, d\mu^+ - \int_{\mathbb{N}} x \, d\mu^-.$$

Proposition 3.9. *R is surjective.*

Proof. Let $\mu \in \text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Define

$$\varphi : \ell^\infty \to \mathbb{R} \quad \varphi(x) = \int_{\mathbb{N}} x \, d\mu.$$

From linearity of the integral it follows that φ is linear. For boundedness, note that for all $x \in \ell^\infty$, we have[14]

$$\begin{aligned}|\varphi(x)| &= \left|\int_{\mathbb{N}} x\,d\mu^+ - \int_{\mathbb{N}} x\,d\mu^-\right| \leq \left|\int_{\mathbb{N}} x\,d\mu^+\right| + \left|\int_{\mathbb{N}} x\,d\mu^-\right| \leq \int_{\mathbb{N}} |x|\,d\mu^+ + \int_{\mathbb{N}} |x|\,d\mu^- \\ &\leq \int_{\mathbb{N}} \|x\|_\infty\,d\mu^+ + \int_{\mathbb{N}} \|x\|_\infty\,d\mu^- = \|x\|_\infty \left(\mu^+(\mathbb{N}) + \mu^-(\mathbb{N})\right) = \|x\|_\infty \|\mu\|_{\text{ba}}.\end{aligned}$$

Thus indeed $\varphi \in (\ell^\infty)'$. Finally we have that $\mu_\varphi = \mu$: for all $E \subseteq \mathbb{N}$,

$$\mu_\varphi(E) = \varphi(\mathbb{1}_E) = \int_{\mathbb{N}} \mathbb{1}_E\,d\mu = \mu(E).$$

This shows that R is surjective. □

Combining propositions 3.8 and 3.9, we conclude that the map $R : (\ell^\infty)' \to \text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ is an isometric isomorphism. Hence we have found a way to characterize $(\ell^\infty)'$ as the space $\text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. However, we are not entirely satisfied with this answer. What does the space $\text{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ look like? What kinds of measures on $\mathbb{N}$ do there exist? We investigate this in the next chapter.

[14] One can easily check the obvious properties of the integral $\int \cdot\, d\mu^\pm$ that are used here.

4 Finitely additive measures on $\mathbb{N}$

In this chapter we investigate the space $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. We describe how one can construct (finitely additive) measures on $\mathbb{N}$ with certain desirable properties. We also describe how certain measures are related to certain types of functionals and vice versa. In the final section we discuss the existence of *unbounded* finitely additive measures. This chapter uses material from [1], [2] and [5].

It is convenient to introduce the following notation and terminology. Let $T : \ell^1 \to (\ell^\infty)'$ be the linear isometric embedding from ℓ^1 into $(\ell^\infty)'$ constructed before, i.e., $T(a) = f_a$, where $f_a(x) = \sum_{n=1}^\infty a_n x_n$. Furthermore, let $R : (\ell^\infty)' \to \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ be the isometric isomorphism from $(\ell^\infty)'$ onto $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ constructed before, i.e., $R(\varphi) = \mu_\varphi$, where $\mu_\varphi(E) = \varphi(\mathbb{1}_E)$.

- For $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ and $\varphi \in (\ell^\infty)'$, we say that μ *is induced by* φ, or that φ *induces* μ, if $\mu = \mu_\varphi$.
- Similarly, for $a \in \ell^1$ and $\varphi \in (\ell^\infty)'$, we say that φ *is induced by* a, or that a *induces* φ, if $\varphi = f_a$.
- Finally, for $a \in \ell^1$ and $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, we say that a *induces* μ, or that μ *is induced by* a, if $\mu = \mu_{f_a}$.

The notation μ will denote either a measure in $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ or the function $\mu_{(\cdot)}$ that sends $\varphi \in (\ell^\infty)'$ to μ_φ. It will be clear from the context which of these meanings we use.

We say that a functional $\varphi \in (\ell^\infty)'$ **preserves positivity** if $\varphi(x) \geq 0$ for all $x \in \ell^\infty$ with $x_n \geq 0$ for all n. Similarly, a measure $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ is said to preserve positivity if $\mu(E) \geq 0$ for all $E \subseteq \mathbb{N}$. The next theorem states that our identification of $(\ell^\infty)'$ with ba preserves positivity in both directions.

Theorem 4.1. *Let $\varphi \in (\ell^\infty)'$. Then φ preserves positivity if and only if μ_φ preserves positivity.*

Proof. It is trivial that a positive functional induces a positive measure. Conversely, suppose $\mu := \mu_\varphi$ is a positive measure. Let $x \in \ell^\infty$ with $x_n \geq 0$ for all n; we must show that $\varphi(x) \geq 0$. We may assume without loss of generality that $\varphi \neq 0$.

Let $\varepsilon > 0$. We show $\varphi(x) > -\varepsilon$. Let $a \in \ell^\infty$ be a *positive* simple sequence (lemma 1.4) such that $\|x - a\|_\infty < \frac{\varepsilon}{\|\varphi\|_{(\ell^\infty)'}}$. Then

$$\varphi(a) - \varphi(x) \leq |\varphi(x) - \varphi(a)| = |\varphi(x - a)| \leq \|\varphi\|_{(\ell^\infty)'} \|x - a\|_\infty < \varepsilon,$$

so that $\varphi(a) - \varepsilon < \varphi(x)$. Now write

$$a = \sum_{j=1}^N \alpha_j \mathbb{1}_{A_j},$$

where $\alpha_j \geq 0$. Then

$$\varphi(a) = \sum_{j=1}^N \alpha_j \varphi(\mathbb{1}_{A_j}) = \sum_{j=1}^N \alpha_j \mu(A_j) \geq 0,$$

where we use that μ is a positive measure. Thus we see that $-\varepsilon < \varphi(x)$. Letting $\varepsilon \downarrow 0$, the result follows. ☐

4.1 Measures induced by ℓ^1

We have identified ℓ^1 as a proper subspace of $(\ell^\infty)'$. The identification $(\ell^\infty)' \simeq \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ allows us to view ℓ^1 as a proper subspace of $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, just by composing the maps T and R: an element $a \in \ell^1$ induces a finitely additive measure $\mu := R(T(a))$, i.e.,[15]

$$\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \mu(E) = \sum_{n=1}^\infty a_n \mathbb{1}_E(n) = \sum_{n \in E} a_n.$$

[15] One has to be careful when writing expressions such as $\sum_{n \in E} a_n$. Such expressions may be defined as follows. For a function $f : \mathbb{N} \to \mathbb{R}$ and a subset $E \subseteq \mathbb{N}$, the expression $\sum_{n \in E} f(n)$ may be defined as $\sum_{n \in E} f(n) := \sum_{n=1}^\infty f(n) \mathbb{1}_E(n)$, whenever the latter series either converges absolutely or contains only positive terms. If this series converges conditionally, then by Riemann's rearrangement theorem, the terms in this series can be rearranged so that the new series converges to any real number or diverges to $\pm\infty$.

As $T[\ell^1] \subsetneq (\ell^\infty)'$, we also have $(R \circ T)[\ell^1] \subsetneq \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, hence the set of measures induced by ℓ^1 is not all of $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Later sections in this chapter deal with measures that do not arise from ℓ^1. First we describe some properties of the (set of) measures induced by ℓ^1.

Theorem 4.2. *The set of measures induced by ℓ^1 is a closed subspace of* $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$*.*

Proof. $(R \circ T)[\ell^1]$ is a subspace of $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, being the image of a linear space under a linear map. We show closedness. Let $(\mu_n)_{n=1}^\infty \subseteq (R \circ T)[\ell^1]$ be a sequence, $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ and suppose $\mu_n \to \mu$ in $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. We show $\mu \in (R \circ T)[\ell^1]$. Now $\mu_n = R(\varphi_n)$ for $\varphi_n \in T[\ell^1]$, and as $(\mu_n)_{n=1}^\infty$ is Cauchy, so is $(\varphi_n)_{n=1}^\infty$:

$$\|\varphi_n - \varphi_m\|_{(\ell^\infty)'} = \|R(\varphi_n - \varphi_m)\|_{\mathrm{ba}} = \|\mu_n - \mu_m\|_{\mathrm{ba}}.$$

As $(\ell^\infty)'$ is a Banach space (theorem 1.7), we have that $\varphi_n \to \varphi$ in $(\ell^\infty)'$ for some $\varphi \in (\ell^\infty)'$. Now $\varphi_n = T(a^{(n)})$ for $a^{(n)} \in \ell^1$, and in the same way we see that $(a^{(n)})_{n=1}^\infty$ is Cauchy, hence $a^{(n)} \to a$ in ℓ^1 for some $a \in \ell^1$. Then by continuity,

$$\varphi_n = T(a^{(n)}) \to T(a)$$

hence $\varphi = T(a)$. Then finally,

$$\mu_n = R(\varphi_n) \to R(\varphi) = R(T(a)).$$

Hence $\mu = R(T(a)) \in (R \circ T)[\ell^1]$, as desired. □

Proposition 4.3. *Let $a \in \ell^1$. Then $a_n \geq 0$ for all n if and only if μ_{f_a} preserves positivity.*

Proof. It is clear that $a_n \geq 0$ for all n implies that $\mu := \mu_{f_a}$ is positive. Conversely, if μ preserves positivity, just note that $0 \leq \mu(\{n\}) = a_n$ for all n. □

Next we discuss σ-additivity of the measures induced by ℓ^1.[16]

Theorem 4.4. *Let $a \in \ell^1$ be a sequence such that $a_n \geq 0$ for all n. Then the measure induced by a, i.e., $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ given by*

$$\mu(E) = \sum_{n=1}^\infty a_n \mathbb{1}_E(n)$$

is σ-additive. Conversely, any positive σ-additive measure $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ is induced by some $a \in \ell^1$.

Proof. Let $E_1, E_2, ... \in \mathcal{P}(\mathbb{N})$ be pairwise disjoint. Then using the monotone convergence theorem to interchange the limit and summation, we obtain

$$\mu\left(\bigcup_{m=1}^\infty E_m\right) = \sum_{n=1}^\infty a_n \mathbb{1}_{\bigcup_{m=1}^\infty E_m}(n) = \sum_{n=1}^\infty \sum_{m=1}^\infty a_n \mathbb{1}_{E_m}(n) \overset{\text{(MCT)}}{=} \sum_{m=1}^\infty \sum_{n=1}^\infty a_n \mathbb{1}_{E_m}(n) = \sum_{m=1}^\infty \mu(E_m).$$

This shows that μ is σ-additive.

Conversely, suppose $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ is a positive σ-additive measure on $\mathbb{N}$. Define $a_n = \mu(\{n\})$ and set $a = (a_1, a_2, ...)$. Then $a \in \ell^1$, as $\sum_{n=1}^\infty |a_n| = \sum_{n=1}^\infty \mu(\{n\}) = \mu(\mathbb{N}) < \infty$. Now let $E \subseteq \mathbb{N}$. We show that $\mu(E) = \sum_{n=1}^\infty a_n \mathbb{1}_E(n)$. First suppose that E is an infinite set, then we may let $m_{(\cdot)} : \mathbb{N} \to E$ be a bijection. Then

$$\mu(E) = \sum_{k=1}^\infty \mu(\{m_k\}) = \sum_{k=1}^\infty a_{m_k} = \sum_{k=1}^\infty a_{m_k} \mathbb{1}_E(m_k) = \sum_{n=1}^\infty a_n \mathbb{1}_E(n),$$

as desired. A similar reasoning applies when E is finite. This finishes the proof. □

[16] For positive finitely additive measures, σ-additivity is defined in the usual way. We do not define σ-additivity for measures that can take negative values. One has to be careful with convergence statements when both positive and negative terms are involved.

Remark. We have seen that any absolutely convergent series $\sum_{n=1}^{\infty} a_n$ (that is, $a \in \ell^1$) induces a measure in a natural way. One can wonder whether this idea can be extended to conditionally convergent series. This is not possible. Consider any conditionally convergent series $\sum_{n=1}^{\infty} a_n$, then $\sum_{n=1}^{\infty} a_n^+ = \infty$ (where $a_n^+ = \max\{a_n, 0\}$). Then by choosing $E = \{n \in \mathbb{N} : a_n \geq 0\}$, we find

$$\sum_{n=1}^{\infty} a_n \mathbb{1}_E(n) = \sum_{n=1}^{\infty} a_n^+ \mathbb{1}_E(n) = \infty,$$

so this does not define a finitely additive measure.[17]

As said, the measures induced by ℓ^1 form a *proper* subspace of $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. In the next sections we describe some measures that do not arise from ℓ^1.

4.2 Shift-invariant measures

In this section we prove the existence of so-called shift-invariant measures on $\mathbb{N}$. For subsets $E \subseteq \mathbb{N}$, we define $E + 1 = \{n + 1 : n \in E\}$, i.e., E shifted by one to the right.

Definition 4.5 (Shift-invariance). *We call a finitely additive measure μ on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ **shift-invariant** if $\mu(E) = \mu(E + 1)$ for all $E \subseteq \mathbb{N}$.*

Our construction of shift-invariant measures uses so-called *Banach limits.* Let S denote the shift operator on ℓ^∞, defined by

$$S : \ell^\infty \to \ell^\infty \quad S((x_1, x_2, ...)) = (x_2, x_3, ...).$$

Theorem 4.6 (Banach limits). *There exists a function $\Phi : \ell^\infty \to \mathbb{R}$ that satisfies*

(1) $\Phi(ax + by) = a\Phi(x) + b\Phi(y)$ for all $x, y \in \ell^\infty$ and $a, b \in \mathbb{R}$ (linearity);

(2) $\Phi(x) \geq 0$ whenever $x \in \ell^\infty$ with $x_n \geq 0 \quad \forall n \in \mathbb{N}$ (positivity);

(3) $\Phi(\mathbb{1}) = 1$, where $\mathbb{1} := (1, 1, ...)$ (Φ is normalized);

(4) $\Phi(S(x)) = \Phi(x)$ for all $x \in \ell^\infty$ (shift-invariance).

Moreover, any function $\Phi : \ell^\infty \to \mathbb{R}$ satisfying the above also satisfies

(5) $|\Phi(x)| \leq \|x\|_\infty$ for all $x \in \ell^\infty$ ($\Phi \in (\ell^\infty)'$ with $\|\Phi\|_{(\ell^\infty)'} = 1$);

(6) $\Phi(x) = \lim_{n\to\infty} x_n$ whenever $(x_n)_{n=1}^{\infty}$ is convergent (Φ extends limits);

(7) $\liminf_{n\to\infty} x_n \leq \Phi(x) \leq \limsup_{n\to\infty} x_n$ for all $(x_n)_{n=1}^{\infty} \in \ell^\infty$.

Proof. Define the linear subspace $W = \{S(x) - x : x \in \ell^\infty\}$ of ℓ^∞. We would like to define a linear functional on W that takes value 0 on W and use Hahn-Banach to extend this to a linear functional on ℓ^∞; this will guarantee property (4) (and (1) trivially). But since we also want to satisfy property (3) it is useful to start with a functional on a slightly bigger subspace than W, namely $W \oplus \mathrm{span}\{\mathbb{1}\}$.

First note that $W \cap \mathrm{span}\{\mathbb{1}\} = \{0\}$. This follows from the fact that $\mathbb{1} \notin W$: if $\mathbb{1} = S(x) - x$ for $x \in \ell^\infty$, then it must hold that $x_n = (n - 1) + x_1$, which is impossible. We may therefore define a linear functional $\phi : W \oplus \mathrm{span}\{\mathbb{1}\} \to \mathbb{R}$ by $\phi(w + c\mathbb{1}) = c$.

[17] One can even show that for a conditionally convergent series $\sum_{n=1}^{\infty} a_n$, there exists $E \subseteq \mathbb{N}$ such that $\sum_{n=1}^{\infty} a_n \mathbb{1}_E(n)$ does not even converge or diverge to $\pm\infty$. The construction of such a set uses similar techniques as the proof of Riemann's rearrangement theorem.

It is now useful to show that $|\phi(y)| \le \|y\|_\infty$ for all $y \in W \oplus \operatorname{span}\{\mathbb{1}\}$. Let $w = S(x) - x \in W$ ($x \in \ell^\infty$) and $c \in \mathbb{R}$; then for every $N \in \mathbb{N}$ we have the inequalities

$$\begin{aligned} \|w + c\mathbb{1}\|_\infty &= \sup_{n\in\mathbb{N}} |x_{n+1} - x_n + c| \ge \sup_{n \le N} |x_{n+1} - x_n + c| \ge \frac{1}{N}\sum_{n=1}^{N} |x_{n+1} - x_n + c| \\ &\ge \frac{1}{N}\left|\sum_{n=1}^{N}(x_{n+1} - x_n) + Nc\right| = \frac{1}{N}|x_{N+1} - x_1 + Nc|. \end{aligned} \tag{12}$$

Now let N tend to ∞, then as $(x_n)_{n=1}^\infty$ is bounded, the last term in (12) tends to $|c|$. This proves that $\|w + c\mathbb{1}\|_\infty \ge |c|$, as desired.

As $\|\cdot\|_\infty$ is a sublinear functional on ℓ^∞, we can apply the Hahn-Banach theorem (theorem 1.9) to find a linear extension $\Phi : \ell^\infty \to \mathbb{R}$ of ϕ satisfying $|\Phi(x)| \le \|x\|_\infty$ for all $x \in \ell^\infty$. By construction, Φ has properties (1), (3) and (4): for (4), just note that $\Phi(S(x)) - \Phi(x) = \Phi(S(x) - x) = \phi(S(x) - x) = 0$. Now for (2), let $x \in \ell^\infty$ with $x_n \ge 0$; write $x = cy$ with $c = \|x\|_\infty$ and $y_n \in [0, 1]$. Then we obtain

$$1 - \Phi(y) = \Phi(\mathbb{1} - y) \le \|\mathbb{1} - y\|_\infty \le 1,$$

as $1 - y_n \in [0, 1]$. Therefore we find $\Phi(y) \ge 0$ and $\Phi(x) = c\Phi(y) \ge 0$, proving (2).

Now suppose $\Phi : \ell^\infty \to \mathbb{R}$ is any function satisfying (1), (2), (3) and (4) (not necessarily the one constructed here). Let $x \in \ell^\infty$. We show that $|\Phi(x)| \le \|x\|_\infty$. The case $x = 0$ is trivial by linearity of Φ, so we assume $x \ne 0$. First suppose $\|x\|_\infty = 1$. Let $y := \mathbb{1} - x$; then $y_n \ge 0$. Then we find

$$1 = \Phi(\mathbb{1}) = \Phi(x + y) = \Phi(x) + \Phi(y),$$

and therefore $1 - \Phi(x) = \Phi(y) \ge 0$, and we obtain $\Phi(x) \le 1$. From this reasoning applied to $-x$ we also derive that $\Phi(-x) \le 1$, so that $|\Phi(x)| \le 1$. Finally, for general $x \ne 0$ we write $x = cy$ where $c = \|x\|_\infty$ and $\|y\|_\infty = 1$ and the result follows from the previous and homogeneity, proving (5).

Now we prove (7). Let $x \in \ell^\infty$. We first show $\Phi(x) \le \limsup_{n\to\infty} x_n$. Let $\xi > \limsup_{n\to\infty} x_n$. Then we can find $n \in \mathbb{N}$ such that $\sup_{k\ge n} x_k < \xi$. Then for $y := S^n(x)$, we find $y_m = x_{n+m} < \xi$ for all $m \in \mathbb{N}$ thus $\|y\|_\infty \le \xi$. Then note that $\Phi(x) = \Phi(S^n(x)) = \Phi(y) \le \xi$. Letting $\xi \downarrow \limsup_{n\to\infty} x_n$, we obtain the desired estimate. The estimate $\liminf_{n\to\infty} x_n \le \Phi(x)$ follows from the previous and the observation that

$$\liminf_{n\to\infty} x_n = -\limsup_{n\to\infty} -x_n.$$

Finally, (6) follows directly from (7). □

A map $\Phi : \ell^\infty \to \mathbb{R}$ that satisfies the first four properties (hence all) in theorem 4.6 is called a **Banach limit.**[18] When we know that a functional is a Banach limit, we will often denote it by Φ or Ψ instead of φ or ψ to emphasize this.

Theorem 4.7 (Existence of shift-invariant measures)**.** *There exists a positive, shift-invariant finitely additive measure μ on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ satisfying $\mu(\mathbb{N}) = 1$.*

Proof. Let $\Phi : \ell^\infty \to \mathbb{R}$ be a Banach limit and define $\mu = \mu_\Phi$. Positivity of Φ (see (2) in theorem 4.6) implies that μ is positive, from (3) we find that $\mu(\mathbb{N}) = 1$ and shift-invariance of Φ implies shift-invariance of μ. □

[18] Banach limits are not unique (corollary 4.14). However, for some (non-convergent) sequences the value of a Banach limit is uniquely determined. Consider for instance the sequence $x = (1, 0, 1, 0, ...)$. Then $x + S(x) = (1, 1, 1, ...)$ thus for any Banach limit Φ, $1 = \Phi(x + S(x)) = 2\Phi(x)$ hence $\Phi(x) = \frac{1}{2}$. In the same way, one shows that for any Banach limit Φ, it holds that $\Phi(\mathbb{1}_{a\mathbb{N}+b}) = \frac{1}{a}$ (for all $a, b \in \mathbb{N}$). Sequences $(x_n)_{n=1}^\infty \in \ell^\infty$ for which the value of a Banach limit is uniquely determined are called **almost convergent**.

We see that the measure constructed in theorem 4.7 is precisely the measure induced by the Banach limit Φ under the identification $(\ell^\infty)' \simeq \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Thus we see that Banach limits induce positive, shift-invariant measures on $\mathbb{N}$ satisfying $\mu(\mathbb{N}) = 1$. In theorem 4.17 we will see that the converse also holds, i.e. that any such measure is induced by a Banach limit.

Remark. If μ is a (bounded or unbounded) shift-invariant finitely additive measure on $\mathbb{N}$, then $\mu(\{n\}) = 0$ for all $n \in \mathbb{N}$.[19] To see this, note that $\mu(\mathbb{N}) = \mu(\mathbb{N}+1) + \mu(\{1\}) = \mu(\mathbb{N}) + \mu(\{1\})$. Thus it follows that $\mu(\{1\}) = 0$, and inductively we see that $\mu(\{n\}) = 0$ for all n. From this we also see that for every finite subset $E \subseteq \mathbb{N}$ we must have $\mu(E) = 0$. Moreover, it follows from this observation that a nontrivial positive, shift-invariant measure cannot be σ-additive.

Theorem 4.8. *The set of bounded, shift-invariant measures on* $\mathbb{N}$ *is a closed subspace of* $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Proof. It is clear that this set forms a subspace. Now let $(\mu_n)_{n=1}^\infty$ be a sequence of bounded, shift-invariant measures on $\mathbb{N}$ and suppose $\mu_n \to \mu$ for $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. We show that μ is shift-invariant. First note that for every $A \subseteq \mathbb{N}$, we have $\mu(A) = \lim_{n\to\infty} \mu_n(A)$:

$$|\mu_n(A) - \mu(A)| = |(\mu_n - \mu)(A)| \leq |\mu_n - \mu|(A) \leq |\mu_n - \mu|(\mathbb{N}) = \|\mu_n - \mu\|_{\mathrm{ba}} \to 0 \quad (n \to \infty).$$

Thus for every $E \subseteq \mathbb{N}$, we obtain

$$\mu(E+1) = \lim_{n\to\infty} \mu_n(E+1) = \lim_{n\to\infty} \mu_n(E) = \mu(E),$$

as desired. ☐

Remark. The subspace of measures induced by ℓ^1 has trivial intersection with the subspace of bounded shift-invariant measures. To see this, suppose μ is a bounded shift-invariant measure induced by $(a_n)_{n=1}^\infty \in \ell^1$. Then we have for all $n \in \mathbb{N}$ that

$$0 = \mu(\{n\}) = \sum_{j=1}^{\infty} a_j \mathbb{1}_{\{n\}}(j) = a_n,$$

hence $(a_n)_{n=1}^\infty$ is zero and μ is the trivial measure.

In theorem 4.7 we have constructed a positive, shift-invariant finitely additive measure. By scaling, we can also construct negative bounded, shift-invariant measures (i.e. taking values in $(-\infty, 0]$). It is also possible to construct shift-invariant measures that take on both (strictly) positive and negative values. We will prove this in proposition 4.15.

Our first step towards the proof of proposition 4.15 is to show that Banach limits are not unique. To prove this we use the so-called *Cesàro*[20] *operator*, a bounded linear operator on ℓ^∞ with some very useful properties. This operator will also be used in sections 4.4 and 4.5.

The Cesàro operator

Definition 4.9 (Cesàro operator)**.** *The* ***Cesàro operator*** C *is defined by*

$$C : \ell^\infty \to \ell^\infty \quad C(x) = \left(\frac{1}{n} \sum_{k=1}^{n} x_k \right)_{n=1}^{\infty}.$$

[19] A question that arises naturally from this observation is whether every positive (or negative) finitely additive measure that vanishes on singletons is shift-invariant. In corollary 4.21 we show that counterexamples exist.

[20] Ernesto Cesàro (1859-1906) was an Italian mathematician who worked in differential geometry. He is mainly known for his method of Cesàro summation, which is a method of assigning a 'generalized limit' to some series that may not be convergent in the usual sense. Lemma 4.10 demonstrates how this can be done: instead of considering the convergence of the sequence of partial sums, one considers the convergence of the sequence of *averages* of the partial sums. The method of Cesàro summation has important applications in Fourier analysis, where it is used to describe convergence of Fourier series. Details can be found in [1].

One can immediately see that C is a bounded linear operator, is moreover positive (i.e. sends positive sequences to positive sequences) and satisfies $C(\mathbb{1}) = \mathbb{1}$. The Cesàro operator is the sequence of averages of $(x_n)_{n=1}^\infty$. Therefore it is reasonable to expect that whenever $(x_n)_{n=1}^\infty$ is convergent, the sequence of averages also converges and to the same limit. This is the case.

Lemma 4.10. *Suppose* $\lim_{n\to\infty} x_n = \xi$. *Then*[21]

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} x_n = \xi.$$

Proof. We may without loss of generality assume $\xi = 0$ (otherwise we replace x_n by $y_n = x_n - \xi$). Let $\varepsilon > 0$. Let $N_0 \in \mathbb{N}$ be such that $|x_n| \le \varepsilon$ for all $n \ge N_0$. Then for all $N \ge N_0$, we obtain

$$\left|\frac{1}{N}\sum_{n=1}^{N} x_n\right| \le \frac{1}{N}\sum_{n=1}^{N_0-1}|x_n| + \frac{1}{N}\sum_{n=N_0}^{N}|x_n| \le \frac{1}{N}\sum_{n=1}^{N_0-1}|x_n| + \frac{(N-N_0)\varepsilon}{N} \le \frac{1}{N}\sum_{n=1}^{N_0-1}|x_n| + \varepsilon.$$

There exists $N_1 \in \mathbb{N}$ such that for all $N \ge N_1$,

$$\frac{1}{N}\sum_{n=1}^{N_0-1}|x_n| \le \varepsilon.$$

Thus for all $N \ge \max\{N_0, N_1\}$, we obtain

$$\left|\frac{1}{N}\sum_{n=1}^{N} x_n\right| \le 2\varepsilon,$$

as desired. □

We say that a functional $\varphi \in (\ell^\infty)'$ *extends limits*, or *extends the* lim *operator*, if $\varphi(x) = \lim_{n\to\infty} x_n$ whenever $(x_n)_{n=1}^\infty$ is convergent. If a functional φ extends limits and has norm 1, then one can build a Banach limit by taking the composition with the Cesàro operator.

Proposition 4.11. *Suppose* $\varphi \in (\ell^\infty)'$ *extends limits and satisfies*[22] $\|\varphi\|_{(\ell^\infty)'} = 1$. *Then* $\varphi \circ C$ *is a Banach limit.*

Proof. We check the four defining properties of Banach limits mentioned in theorem 4.6. It is clear that $\varphi \circ C$ is linear and we have $(\varphi \circ C)(\mathbb{1}) = \varphi(\mathbb{1}) = 1$. It remains to show positivity and shift-invariance.

Observe that C sends positive sequences to positive sequences. Hence to show that $\varphi \circ C$ is positive, it suffices to show that $\varphi(x) \ge 0$ whenever $x \in \ell^\infty$ satisfies $x_n \ge 0$ for all n. Let $x \in \ell^\infty$ be such that $x_n \ge 0$ for all n. First assume that $x_n \in [0,1]$ for all n. Then we have that

$$1 - \varphi(x) = \varphi(\mathbb{1} - x) \le |\varphi(\mathbb{1} - x)| \le \|1 - x\|_\infty \le 1, \tag{13}$$

so that $\varphi(x) \ge 0$. The result for general positive x follows from equation (13) and homogeneity. Thus we derive that $\varphi \circ C$ is positive.

Finally we show shift-invariance. Let $x \in \ell^\infty$. Note that we have

$$\frac{1}{n}\sum_{k=1}^{n} x_k - \frac{1}{n}\sum_{k=1}^{n} x_{k+1} = \frac{1}{n}\sum_{k=1}^{n} x_k - x_{k+1} = \frac{1}{n}(x_1 - x_{n+1}) \to 0 \quad (n \to \infty),$$

[21] The converse does not generally hold. For instance, consider the sequence $x_n = \sum_{k=1}^{n}(-1)^k$.
[22] This condition $\|\varphi\|_{(\ell^\infty)'} = 1$ is to ensure that φ preserves positivity.

where we used that $(x_n)_{n=1}^{\infty}$ is bounded. This shows that $C(x) - C(S(x))$ is a convergent sequence with limit 0, hence $\varphi(C(x) - C(S(x))) = 0$, which proves shift-invariance of $\varphi \circ C$. We conclude that $\varphi \circ C$ is a Banach limit. □

The next lemma will be used in the construction of two different Banach limits.

Lemma 4.12. *There exists a subset $E \subseteq \mathbb{N}$ such that $C(\mathbb{1}_E)$ has a subsequence $(C(\mathbb{1}_E)_{n_k})_{k=1}^{\infty}$ that satisfies $C(\mathbb{1}_E)_{n_k} = \frac{1}{3}$ for all $k \in \mathbb{N}$ and another subsequence $(C(\mathbb{1}_E)_{m_k})_{k=1}^{\infty}$ that satisfies $C(\mathbb{1}_E)_{m_k} = \frac{2}{3}$ for all $k \in \mathbb{N}$.*

Proof. Consider

$$x = (0, 0, \underset{x_3}{1}, \overbrace{1, 1, \underset{x_6}{1}}^{3 \text{ times}}, \overbrace{0, 0, 0, 0, 0, \underset{x_{12}}{0}}^{6 \text{ times}}, \overbrace{1, 1, ..., 1, \underset{x_{24}}{1}}^{12 \text{ times}}, \overbrace{0, 0, ..., 0, \underset{x_{48}}{0}}^{24 \text{ times}}, \overbrace{1, 1, ..., 1, \underset{x_{96}}{1}}^{48 \text{ times}}, ...).$$

The idea is as follows. We start with the terms $(0, 0, 1)$; the average of these is $\frac{1}{3}$. Next, we add 1's until the average of the terms defined so far is $\frac{2}{3}$. Then we add 0's until the average of the terms defined so far is $\frac{1}{3}$. It is clear that we can repeat this indefinitely: if the first $3 \cdot 2^{n-1}$ terms are defined ($n \in \mathbb{N}$) then we add $3 \cdot 2^{n-1}$ times a 0 or a 1 (depending on the $3 \cdot 2^{n-1}$-th term). Now consider the sequence of averages $y := C(x)$, i.e.,

$$y_n = \frac{1}{n} \sum_{k=1}^{n} x_k.$$

Let $n_k = 3 \cdot 2^{2(k-1)}$ and $m_k = 3 \cdot 2^{2k-1}$ for all $k \in \mathbb{N}$. By construction, we have that $y_{n_k} = \frac{1}{3}$ and $y_{m_k} = \frac{2}{3}$ for all $k \in \mathbb{N}$. Now let $E = \{n \in \mathbb{N} : x_n = 1\}$, so that $x = \mathbb{1}_E$. Then E is the desired set and $(C(\mathbb{1}_E)_{n_k})_{k=1}^{\infty}$ and $(C(\mathbb{1}_E)_{m_k})_{k=1}^{\infty}$ are the desired subsequences. □

Proposition 4.13. *There exist functionals $\varphi, \vartheta \in (\ell^\infty)'$ that extend limits, satisfy $\|\varphi\|_{(\ell^\infty)'} = \|\vartheta\|_{(\ell^\infty)'} = 1$ and satisfy $(\varphi \circ C)(\mathbb{1}_E) \neq (\vartheta \circ C)(\mathbb{1}_E)$ for some $E \subseteq \mathbb{N}$.*

Proof. Let $E \subseteq \mathbb{N}$, $(n_k)_{k=1}^{\infty}$ and $(m_k)_{k=1}^{\infty}$ be as in lemma 4.12. Define the linear subspaces of ℓ^∞

$$c^{(n)} = \{x \in \ell^\infty : \lim_{k\to\infty} x_{n_k} \text{ exists}\} \quad \text{and} \quad c^{(m)} = \{x \in \ell^\infty : \lim_{k\to\infty} x_{m_k} \text{ exists}\}.$$

Note that $c \subseteq c^{(n)} \cap c^{(m)}$. Define

$$\phi : c^{(n)} \to \mathbb{R} \quad \phi(x) = \lim_{k\to\infty} x_{n_k},$$
$$\theta : c^{(m)} \to \mathbb{R} \quad \theta(x) = \lim_{k\to\infty} x_{m_k}.$$

These are bounded linear functionals satisfying $\|\phi\|_{(c^{(n)})'} = \|\theta\|_{(c^{(m)})'} = 1$. The Hahn-Banach theorem gives us extensions $\varphi, \vartheta \in (\ell^\infty)'$ of ϕ and θ, respectively, satisfying $\|\varphi\|_{(\ell^\infty)'} = \|\vartheta\|_{(\ell^\infty)'} = 1$. Note that φ and ϑ extend limits. Moreover, we have that $\varphi \circ C$ and $\vartheta \circ C$ are Banach limits, in view of proposition 4.11. For the final assertion, note that

$$\varphi(C(\mathbb{1}_E)) = \phi(C(\mathbb{1}_E)) = \lim_{k\to\infty} C(\mathbb{1}_E)_{n_k} = \frac{1}{3},$$
$$\vartheta(C(\mathbb{1}_E)) = \theta(C(\mathbb{1}_E)) = \lim_{k\to\infty} C(\mathbb{1}_E)_{m_k} = \frac{2}{3}.$$

This finishes the proof. □

The next result is immediate from (the proof of) proposition 4.13.

Corollary 4.14. *Banach limits are not unique.*

We can now prove the existence of a shift-invariant measure that assumes both (strictly) positive and negative values.

Proposition 4.15. *For any $\gamma \in \mathbb{R}$, there exists a bounded, shift-invariant measure μ on $\mathbb{N}$ and $A \subseteq \mathbb{N}$ such that $\mu(A) < 0$ and $\mu(\mathbb{N} \setminus A) > 0$, and moreover $\mu(\mathbb{N}) = \gamma$.*

Proof. Let Φ and Ψ be two different Banach limits. Then Φ and Ψ differ on an indicator function (see[23] proposition 1.5), say $\Phi(\mathbb{1}_A) > \Psi(\mathbb{1}_A)$. Then $\Phi(\mathbb{1}_{\mathbb{N}\setminus A}) = \Phi(\mathbb{1}_{\mathbb{N}}) - \Phi(\mathbb{1}_A) = 1 - \Phi(\mathbb{1}_A)$ and also $\Psi(\mathbb{1}_{\mathbb{N}\setminus A}) = 1 - \Psi(\mathbb{1}_A)$. Thus we see that $\Phi(\mathbb{1}_{\mathbb{N}\setminus A}) < \Psi(\mathbb{1}_{\mathbb{N}\setminus A})$. Now let $\delta > 1$ be such that $\delta\Phi(\mathbb{1}_{\mathbb{N}\setminus A}) < \Psi(\mathbb{1}_{\mathbb{N}\setminus A})$. Define $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by $\mu(E) = \delta\Phi(\mathbb{1}_E) - \Psi(\mathbb{1}_E)$. Then μ is a bounded, shift-invariant finitely additive measure with $\mu(\mathbb{N}) = \delta - 1 \neq 0$, and note that $\mu(A) > 0$ and $\mu(\mathbb{N} \setminus A) < 0$. The result for $\gamma \neq 0$ follows by appropriate scaling of μ, and the result for $\gamma = 0$ follows by a modification of this proof, namely choosing $\delta = 1$ instead. □

We now consider again the dual space of ℓ^∞. We have seen in the proof of theorem 4.7 that a Banach limit Φ induces a positive, finitely additive shift-invariant measure μ satisfying $\mu(\mathbb{N}) = 1$. In fact, any linear combination of Banach limits induces a bounded shift-invariant measure. We can also go in the converse direction: what are the properties of functionals that induce such measures? We start with the following observation.

Theorem 4.16. *Let $\varphi \in (\ell^\infty)'$. Then φ is shift-invariant if and only if μ_φ is shift-invariant.*

Proof. We have already seen that a shift-invariant functional φ induces a shift-invariant measure. Now suppose $\mu := \mu_\varphi$ is shift-invariant. We show that φ is shift-invariant. We may without loss of generality assume that $\varphi \neq 0$.

We introduce the following notation: for $A \subseteq \mathbb{N}$ with $1 \notin A$, we define $A - 1 = \{n - 1 : n \in A\} \subseteq \mathbb{N}$.

Let $x \in \ell^\infty$; we must show that $\varphi(x) = \varphi(S(x))$. Note that we may assume that $x_1 = 0$: consider $y = (0, x_2, x_3, ...)$, then $S(x) = S(y)$ and assuming that $\varphi(y) = \varphi(S(y))$, we get

$$\varphi(x) = \varphi((x_1, 0, 0, ...)) + \varphi(y) = x_1\varphi(\mathbb{1}_{\{1\}}) + \varphi(S(y)) = x_1\mu(\{1\}) + \varphi(S(x)) = \varphi(S(x)).$$

Let $\varepsilon > 0$; we show that $|\varphi(x - S(x))| < \varepsilon$. Let $a \in \ell^\infty$ be a simple sequence with $\|x - a\|_\infty \leq \frac{\varepsilon}{2\|\varphi\|_{(\ell^\infty)'}}$. As $x_1 = 0$, we may assume $a_1 = 0$ (if a_1 is nonzero, we replace it by zero; this does not increase $\|x - a\|_\infty$). Then we have that

$$\begin{aligned} |\varphi(x - S(x))| &= |\varphi(x - a + a - S(x) - S(a) + S(a))| = |\varphi((x - a) + (S(a) - S(x)) + (a - S(a))| \\ &\leq |\varphi(x - a)| + |\varphi(S(a) - S(x))| + |\varphi(a - S(a))| \\ &\leq \|\varphi\|_{(\ell^\infty)'}\|x - a\|_\infty + \|\varphi\|_{(\ell^\infty)'}\|S(a) - S(x)\|_\infty + |\varphi(a - S(a))| < \varepsilon + |\varphi(a - S(a))|. \end{aligned} \tag{14}$$

Now write

$$a = \sum_{j=1}^{N} \alpha_j \mathbb{1}_{A_j},$$

where $A_1, A_2, ..., A_N \subseteq \mathbb{N}$ are disjoint and $\alpha_1, \alpha_2, ..., \alpha_N \in \mathbb{R}$ are *nonzero*[24], so that $1 \notin A_j$ for all j. Then we have that

$$a - S(a) = \sum_{j=1}^{N} \alpha_j(\mathbb{1}_{A_j} - \mathbb{1}_{A_j - 1}). \tag{15}$$

By shift-invariance of μ, we have that for all j,

$$\varphi(\mathbb{1}_{A_j - 1}) = \mu(A_j - 1) = \mu(A_j) = \varphi(\mathbb{1}_{A_j}).$$

[23] Alternatively; this is immediately clear if we take the Banach limits constructed in proposition 4.13.

[24] We just omit $\alpha_k\mathbb{1}_{A_k}$ from the sum if α_k happens to be zero. If this enforces that $N = 0$, then a is the zero sequence and we certainly have $a - S(a) = 0$, so we implicitly avoid this triviality here.

Combined with equation (15), this gives

$$\varphi(a - S(a)) = \sum_{j=1}^{N} \alpha_j \varphi(\mathbb{1}_{A_j} - \mathbb{1}_{A_j - 1}) = 0.$$

Combining this with equation (14), we find $|\varphi(x - S(x))| < \varepsilon$. Letting $\varepsilon \downarrow 0$, we obtain $\varphi(x) = \varphi(S(x))$, as desired. □

Theorem 4.17. *Let $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ be a positive (hence bounded), shift-invariant finitely additive measure satisfying $\mu(\mathbb{N}) = 1$. Let $\varphi \in (\ell^\infty)'$ be the functional that induces μ, i.e. $\mu = \mu_\varphi$. Then φ is a Banach limit.*

Proof. We check the four defining properties in of Banach limits mentioned in theorem 4.6. We have that φ is linear, it preserves positivity (theorem 4.1), we have $\varphi(\mathbb{1}) = \varphi(\mathbb{1}_{\mathbb{N}}) = \mu(\mathbb{N}) = 1$ and finally, shift-invariance of φ follows from theorem 4.16. We conclude that φ is a Banach limit. □

Remark. If μ is a positive shift-invariant finitely additive measure with $\mu(\mathbb{N}) = 1$, then $\mu(a\mathbb{N}) = \frac{1}{a}$ for all $a \in \mathbb{N}$. This follows from theorem 4.17 and the discussion in footnote 18.

If we have a bounded, shift-invariant measure, we can write it as the difference of two *positive* shift-invariant measures.

Proposition 4.18. *Let $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ be a bounded, shift-invariant finitely additive measure, and write $\mu = \mu^+ - \mu^-$. Then $\mu^\pm$ are shift-invariant.*

Proof. We first show that μ^+ is shift-invariant. Let $E \subseteq \mathbb{N}$ and define $\xi = \sup\{\mu(F) : F \subseteq E\}$ and $\gamma = \sup\{\mu(G) : G \subseteq E + 1\}$. We must show that $\xi = \gamma$. Let $F \subseteq E$. Then $F + 1 \subseteq E + 1$ thus $\mu(F) = \mu(F + 1) \leq \gamma$, hence taking the supremum over F, we derive $\xi \leq \gamma$. The converse estimate follows similarly. Thus we have that $\gamma = \xi$ hence μ^+ is shift-invariant. To see that μ^- is shift-invariant, just note that

$$\mu^-(E + 1) = \mu^+(E + 1) - \mu(E + 1) = \mu^+(E) - \mu(E) = \mu^-(E),$$

which concludes the proof. □

We can extend the result from theorem 4.17 in the following way. Consider a (nontrivial) positive shift-invariant finitely additive measure μ on $\mathbb{N}$, say $\alpha := \mu(\mathbb{N}) \neq 0$. Then $\frac{1}{\alpha}\mu$ is a positive, shift-invariant finitely additive measure satisfying $\frac{1}{\alpha}\mu(\mathbb{N}) = 1$, hence by theorem 4.17 we have that $\frac{1}{\alpha}\mu$ is induced by a Banach limit Φ: $\frac{1}{\alpha}\mu = \mu_\Phi$. By linearity of the identification $(\ell^\infty)' \simeq \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, we have that $\mu = \mu_{\alpha\Phi}$.

Now suppose μ is any bounded shift-invariant measure. Writing $\mu = \mu^+ - \mu^-$, we have that $\mu^\pm$ are induced by scaled Banach limits, say $\mu^+ = \mu_{\alpha\Phi}$ and $\mu^- = \mu_{\beta\Psi}$. Then $\mu = \mu_{\alpha\Phi - \beta\Psi}$. Thus we conclude that set of bounded, shift-invariant measures corresponds to the set of linear combinations of Banach limits.

Diffuse measures that are not shift-invariant

Now we address a question raised in footnote 19. We show that there exists a positive, finitely additive measure that vanishes on singletons but is not shift-invariant. A (finitely additive) measure that vanishes on singletons (hence on all finite sets) is called **diffuse**. We start by examining what properties the inducing functional of a diffuse measure has.

Proposition 4.19. *Let $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ be a bounded finitely additive diffuse measure. Let $\varphi \in (\ell^\infty)'$ be the functional that induces, μ, i.e. $\mu = \mu_\varphi$. Then for all $x \in c$, we have*

$$\varphi(x) = \varphi(\mathbb{1}) \cdot \lim_{n\to\infty} x_n.$$

In particular, if $\varphi(\mathbb{1}) = 1$, then φ extends the lim *operator.*

Proof. Step 1. Let $x \in c_{00}$, the set of sequences with only finitely many nonzero terms. Then writing $x = (x_1, x_2, ..., x_n, 0, 0, ...)$, we see that

$$\varphi(x) = \varphi\left(\sum_{k=1}^{n} x_k \mathbb{1}_{\{k\}}\right) = \sum_{k=1}^{n} x_k \varphi(\mathbb{1}_{\{k\}}) = \sum_{k=1}^{n} x_k \mu(\{k\}) = 0.$$

Step 2. Let $x \in c_0$. Let $\varepsilon > 0$ and let $N \in \mathbb{N}$ be such that $\sup_{n \geq N} |x_n| < \varepsilon$. Set $x^* = (x_1, x_2, ..., x_N, 0, 0, ...)$. Then using step 1,

$$|\varphi(x)| = |\varphi((x - x^*) + x^*)| = |\varphi(x - x^*)| \leq \|\varphi\|_{(\ell^\infty)'} \|x - x^*\|_\infty < \varepsilon \|\varphi\|_{(\ell^\infty)'}.$$

Letting $\varepsilon \downarrow 0$, we see that $\varphi(x) = \lim_{n\to\infty} x_n = 0$.

Step 3. Let $x \in c$. Define $\alpha = \lim_{n\to\infty} x_n$ and let $y = x - \alpha\mathbb{1} = (x_1 - \alpha, x_2 - \alpha, ...)$ and note $y \in c_0$. Then using step 2, we obtain

$$\varphi(x) = \varphi(y + \alpha\mathbb{1}) = \alpha\varphi(\mathbb{1}) = \varphi(\mathbb{1}) \cdot \lim_{n\to\infty} x_n,$$

as desired. □

The following converse to proposition 4.19 holds: if $\varphi \in (\ell^\infty)'$ extends the lim operator, then the induced functional μ_φ is diffuse.

Theorem 4.20. *There exists a functional $\varphi \in (\ell^\infty)'$ preserves positivity, satisfies $\varphi(\mathbb{1}) = 1$, extends limits but is not shift-invariant.*

Proof. We define a functional on a subspace of ℓ^∞ that extends limits but is not shift-invariant (on its domain); then any Hahn-Banach extension cannot be shift-invariant. Let $c^{(*)} := \{x \in \ell^\infty : \lim_{n\to\infty} x_{2n} \text{ exists}\}$; a linear subspace of ℓ^∞ with $c \subsetneq c^{(*)}$. Let

$$\phi : c^{(*)} \to \mathbb{R} \quad \phi(x) = \lim_{n\to\infty} x_{2n}.$$

Then ϕ is a bounded linear functional with $\|\phi\|_{(c^{(*)})'} = 1$. The Hahn-Banach theorem (theorem 1.10) gives us an extension $\varphi \in (\ell^\infty)'$ of ϕ satisfying $\|\varphi\|_{(\ell^\infty)'} = \|\phi\|_{(c^{(*)})'} = 1$. Then note that φ extends limits, and that for all $x \in \ell^\infty$, we have $|\varphi(x)| \leq \|x\|_\infty$. We show that φ is positive. Let $x \in \ell^\infty$ with $x_n \geq 0$ for all n. By scaling and homogeneity we may assume $x_n \in [0, 1]$ for all n. Then we have

$$1 - \varphi(x) = \varphi(\mathbb{1} - x) \leq |\varphi(\mathbb{1} - x)| \leq \|1 - x\|_\infty \leq 1,$$

so that $\varphi(x) \geq 0$. Finally note that φ is not shift-invariant: this follows by considering the sequence $((-1)^n)_{n=1}^\infty \in c^{(*)}$. □

Remark. Note that there exist infinitely many such functionals as in theorem 4.20: any convex combination $t\varphi + (1 - t)\Psi$ (for $t \in (0, 1]$) satisfies the hypotheses, where φ is as in theorem 4.20 and Ψ is any Banach limit.

The next result describes a diffuse measure that is not shift-invariant.

Corollary 4.21. *There exists a positive, finitely additive diffuse measure $\mu \in$ ba that satisfies $\mu(\mathbb{N}) = 1$, assumes some value in $(0, 1)$ and is not shift-invariant.*

Proof. Let φ be as in theorem 4.20. If the induced measure μ_φ assumes some value in $(0, 1)$ then $\mu := \mu_\varphi$ has the desired properties and we are done.

So assume that μ_φ assumes only values 0 and 1. Let Ψ be any Banach limit. Then note that the functional $\psi := \frac{1}{3}\varphi + \frac{2}{3}\Psi$ preserves positivity, extends limits, satisfies $\psi(\mathbb{1}) = 1$ and is not shift-invariant.

We claim that $\mu := \mu_\psi$ satisfies the hypotheses in the theorem. To see that μ assumes some value in $(0,1)$, note that $\Psi(\mathbb{1}_{2\mathbb{N}}) = \frac{1}{2}$ thus we have $\psi(\mathbb{1}_{2\mathbb{N}}) = \frac{1}{3}\varphi(\mathbb{1}_{2\mathbb{N}}) + \frac{1}{3}$, hence $\psi(\mathbb{1}_{2\mathbb{N}}) \in \{\frac{1}{3}, \frac{2}{3}\}$. By construction, μ also has the other properties. This finishes the proof. □

In fact, we will see in section 4.5 an alternative method of constructing diffuse measures that are not shift-invariant. There we construct measures that assume only the values 0 and 1 (and value 1 is assumed).

If we know all *diffuse* measures on $\mathbb{N}$ then we know the whole space $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, in the following sense.

Theorem 4.22. *Let μ be any bounded, finitely additive measure on $\mathbb{N}$. Then there exists a unique measure η induced by ℓ^1 and a unique diffuse measure ν such that $\mu = \eta + \nu$.*

Proof. Define $a_n = \mu(\{n\})$ and $a = (a_1, a_2, ...)$. Then $a \in \ell^1$ as

$$\sum_{n=1}^{\infty} |a_n| \leq \sum_{n=1}^{\infty} \mu^+(\{n\}) + \mu^-(\{n\}) \leq \mu^+(\mathbb{N}) + \mu^-(\mathbb{N}) < \infty.$$

Define $\eta : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by $\eta(E) = \sum_{n=1}^{\infty} a_n \mathbb{1}_E(n)$, the measure induced by a. Furthermore, set $\nu := \mu - \eta$. Then $\mu = \eta + \nu$, and ν is diffuse. Finally, uniqueness is clear. □

4.3 More general invariant measures

In the previous section we have constructed shift-invariant measures. We now consider a more general notion of invariance.

Let $f : \mathbb{N} \to \mathbb{N}$ be any function. A finitely additive measure μ is called **f-invariant** if $\mu(A) = \mu(f^{-1}[A])$ for all $A \subseteq \mathbb{N}$.

Theorem 4.23. *Let $f : \mathbb{N} \to \mathbb{N}$ be any function. Then there exists an f-invariant, positive finitely additive measure μ on $\mathbb{N}$ satisfying $\mu(\mathbb{N}) = 1$.*

Proof. Let us define for all $n \in \mathbb{N}$ and $E \subseteq \mathbb{N}$ the notation $f^{-(n+1)}[E] := f^{-1}[f^{-n}[E]]$. Let $\Phi : \ell^\infty \to \mathbb{R}$ be a Banach limit and let ν be any positive finitely additive measure with $\nu(\mathbb{N}) = 1$. Define

$$\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \mu(E) = \Phi(\nu(E), \nu(f^{-1}[E]), \nu(f^{-2}[E]), ...).$$

Then by observing that for disjoint $E, F \subseteq \mathbb{N}$ and $n \in \mathbb{N}$, we have $\nu(f^{-n}[E \cup F]) = \nu(f^{-n}[E] \cup f^{-n}[F]) = \nu(f^{-n}[E]) + \nu(f^{-n}[F])$ and the fact that Φ is linear, we see that μ is finitely additive. Moreover, it is positive and satisfies $\mu(\mathbb{N}) = 1$. Finally, f-invariance of μ follows from shift-invariance of Φ. □

With a restriction on f we have the following similar result.[25]

Theorem 4.24. *Let $f : \mathbb{N} \to \mathbb{N}$ be a bijection. Then there exists a positive finitely additive measure μ satisfying $\mu(E) = \mu(f[E])$ for all $E \subseteq \mathbb{N}$ and $\mu(\mathbb{N}) = 1$.*

Proof. Let Φ be any Banach limit and let ν be a positive finitely additive measure satisfying $\nu(\mathbb{N}) = 1$. Define

$$\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \mu(E) = \Phi(\nu(E), \nu(f[E]), \nu(f^2[E]), ...).$$

Note that for disjoint sets E and F, we have that $f[E \cup F] = f[E] \cup f[F]$ is a disjoint union (due to injectivity). Combined with linearity of Φ, this proves that μ is finitely additive. Clearly μ is positive and satisfies $\mu(\mathbb{N}) = 1$, and from shift-invariance of Φ we see that $\mu(E) = \mu(f[E])$ for $E \subseteq \mathbb{N}$. □

[25] In section 4.6 we will construct measures with much nicer properties than described here.

4.4 Measures that extend density

Definition 4.25. *For a subset $E \subseteq \mathbb{N}$, we define*

$$d(E) = \lim_{n\to\infty} \frac{|E \cap \{1, 2, ..., n\}|}{n},$$

whenever this limit exists. We call $d(E)$ the ***(asymptotic) density*** *of E.*

Asymptotic density is an intuitive way to measure how 'large' a subset of $\mathbb{N}$ is. Let $E \subseteq \mathbb{N}$ and consider the probability that a randomly chosen number from $\{1, 2, ..., n\}$ lies in E. This probability is equal to the number of elements in E that are in $\{1, 2, ..., n\}$, divided by the number of elements in $\{1, 2, ..., n\}$. If this probability tends to a limit for n tending to infinity, then this number is called the density of E and it can be interpreted as a kind of probability that a randomly chosen natural number lies in E.

The next theorem describes some basic properties of density.

Theorem 4.26. *Let $E, F \subseteq \mathbb{N}$.*

(1) If E is finite, then $d(E) = 0$.
(2) If $d(E)$ exists, then $d(\mathbb{N} \setminus E)$ exists and $d(\mathbb{N} \setminus E) = 1 - d(E)$.
(3) If E is finite, then $d(\mathbb{N} \setminus E) = 1$. In particular, $d(\mathbb{N}) = 1$.
(4) If E is finite and $d(F)$ exists, then $d(E \cup F)$ exists and $d(E \cup F) = d(F)$.
(5) If $E \subseteq F$ and $d(F) = 0$, then $d(E) = 0$.
(6) If $d(E)$, $d(F)$ and $d(E \cup F)$ all exist, then

$$\max\{d(E), d(F)\} \leq d(E \cup F) \leq \min\{d(E) + d(F), 1\}.$$

(7) If $E = \{an + b : n \in \mathbb{N}\}$ (for some $a, b \in \mathbb{N}$), then $d(E) = \frac{1}{a}$. In particular, the set of even numbers has density $\frac{1}{2}$.
(8) If $E = \{2^n : n \in \mathbb{N}\}$, then $d(E) = 0$.
(9) If $E = \{n^2 : n \in \mathbb{N}\}$, then $d(E) = 0$.
(10) The set of prime numbers has density 0.
(11) There exists sets which do not admit a density.

Proof. We only prove (8), (9), (10) and (11); the other properties are trivial.

(8). Suppose $E = \{2^n : n \in \mathbb{N}\}$. Note that for all $n \in \mathbb{N}$, we have

$$|E \cap \{1, 2, ..., n\}| \leq |\{1, 2, ..., \lceil \log_2(n) \rceil\}|,$$

as an injection is given by $2^m \to m$. Thus we have

$$\frac{|E \cap \{1, 2, ..., n\}|}{n} \leq \frac{\lceil \log_2(n) \rceil}{n} \leq \frac{\log_2(n) + 1}{n} \to 0 \quad (n \to \infty),$$

which shows that $d(E) = 0$.

(9). The proof is similar to the proof of (8), in this case by bounding $|E \cap \{1, 2, ..., n\}| \leq \lceil \sqrt{n} \rceil$.

(10). Let $\pi(n)$ denote the number of prime numbers less than or equal to n. The claim now follows from the prime number theorem, which states that

$$\lim_{n\to\infty} \frac{\pi(n)}{\left(\frac{n}{\log(n)}\right)} = 1.$$

(11). The idea is illustrated in figure 1. For all $n \geq 0$, define $I_n = [2^n, 2^{n+1})$, $E_n = I_n \cap \mathbb{N}$ and set

$$A = \bigcup_{n=0}^{\infty} E_{2n}.$$

Define $\alpha(m) = |A \cap \{1, 2, ..., m\}|$ for all $m \geq 1$. We show that $(\frac{\alpha(m)}{m})_{m=1}^{\infty}$ has two convergent subsequences with different limits, so that $d(A)$ does not exist. Indeed, if we define $m_k := 2^{2k-1}$, then one can readily check that $\frac{\alpha(m_k)}{m_k} \to \frac{2}{3}$ as $k \to \infty$. On the other hand, if we consider $n_k := 2^{2k}$, then we see that $\frac{\alpha(n_k)}{n_k} \to \frac{1}{3}$ for $k \to \infty$. Thus $d(A)$ does not exist.[26] □

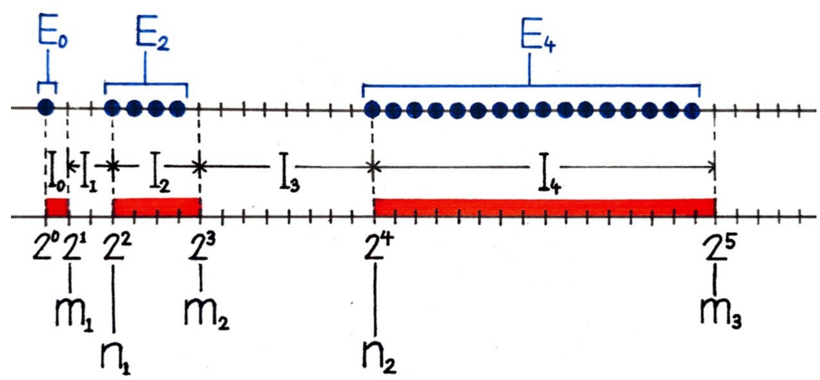

Figure 1: Proof of (11) in theorem 4.26.

We will prove shortly the existence of a finitely additive measure μ on $\mathbb{N}$ that extends density. That is, μ satisfies $\mu(E) = d(E)$ whenever $d(E)$ exists. First we prove that this is a stronger condition than shift-invariance.

Theorem 4.27. *If $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ extends density, then it is shift-invariant.*

Proof. Let $E \subseteq \mathbb{N}$. First suppose that $E \subseteq A$, where A is either $2\mathbb{N}$ or $2\mathbb{N} - 1$. Then we have

$$\mu((1+E) \cup (A \setminus E)) = \mu(1+E) + \mu(A \setminus E) = \mu(1+E) + \mu(A) - \mu(E) = \mu(1+E) - \mu(E) + \frac{1}{2}, \quad (16)$$

where we used disjointness in the first step and $\mu(A) = d(A) = \frac{1}{2}$ in the last step. Also, we have for every n that

$$\frac{|\{1,2,...,n\} \cap (E \cup (A \setminus E))| - 1}{n} \overset{(*)}{\leq} \frac{|\{1,2,...,n\} \cap ((1+E) \cup (A \setminus E))|}{n} \overset{(**)}{\leq} \frac{|\{1,2,...,n\} \cap (E \cup (A \setminus E))| + 1}{n}.$$

We will shortly prove $(*)$ and $(**)$. Letting $n \to \infty$, the leftmost and rightmost term tend to $d(A) = \frac{1}{2}$ hence we conclude $\mu((1+E) \cup (A \setminus E)) = d((1+E) \cup (A \setminus E)) = \frac{1}{2}$. Combining this with equation (16), we see that $\mu(E) = \mu(1+E)$.

Now we show why $(*)$ holds. The idea is illustrated in figure 2. We may assume without loss of generality that $\{1, 2, ..., n\} \cap E \neq \varnothing$. Let $\xi = \max(\{1, 2, ..., n\} \cap E)$ and consider the map

$$f : \{1,2,...,n\} \cap (E \cup (A \setminus E)) \to (\{1,2,...,n\} \cap ((1+E) \cup (A \setminus E))) \cup \{\xi + 1\},$$
$$f(m) = \begin{cases} m+1 & \text{if } m \in E, \\ m & \text{if } m \in A \setminus E. \end{cases}$$

[26] Observe the similarities between this proof and the proof of lemma 4.12.

It is easily checked that f is injective, and therefore $(*)$ follows. The proof of $(**)$ is similar.

Now consider a general $E \subseteq \mathbb{N}$ and write $E = E_1 \cup E_2$, where $E_1 = E \cap (2\mathbb{N} - 1)$ and $E_2 = E \cap (2\mathbb{N})$. Then by the previous we have

$$\mu(E+1) = \mu((E_1+1) \cup (E_2+1)) = \mu(E_1+1) + \mu(E_2+1) = \mu(E_1) + \mu(E_2) = \mu(E),$$

as desired. □

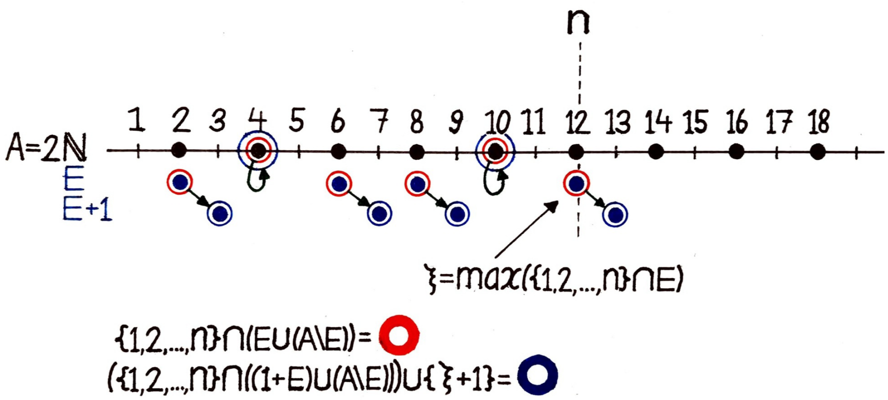

Figure 2: Proof of $(*)$ in theorem 4.27.

Theorem 4.28. *There exists a positive finitely additive measure μ on $\mathbb{N}$ that extends density.*

Proof. Define for all $n \in \mathbb{N}$, $\mu_n : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by

$$\mu_n(E) = \frac{|E \cap \{1, 2, ..., n\}|}{n}.$$

Then each μ_n is a finitely additive measure and $0 \leq \mu_n(E) \leq 1$. Now let $\Phi : \ell^\infty \to \mathbb{R}$ be a Banach limit. Define $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by

$$\mu(E) = \Phi(\mu_1(E), \mu_2(E), \mu_3(E), ...).$$

Then μ is a finitely additive measure on $\mathbb{N}$. Moreover, from the fact that Banach limits extend the usual limit, it follows that μ extends density. Positivity of μ follows from positivity of Φ. □

One might wonder whether a measure that extends density is necessarily positive. In corollary 4.30 we show that this is not the case. The following proposition is the main ingredient for the proof.

Proposition 4.29. *There exist two different positive finitely additive measures that extend density.*

Proof. Let $\varphi, \vartheta \in (\ell^\infty)'$ and $E \subseteq \mathbb{N}$ be as in proposition 4.13. Then $\varphi \circ C$ and $\vartheta \circ C$ are Banach limits by proposition 4.11. Define

$$\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \mu(A) = (\varphi \circ C)(\mathbb{1}_A),$$
$$\nu : \mathcal{P}(\mathbb{N}) \to \mathbb{R} \quad \nu(A) = (\vartheta \circ C)(\mathbb{1}_A).$$

These are finitely additive measures, and they are positive as $\varphi \circ C$ and $\vartheta \circ C$ are Banach limits. To see that μ extends density, let $A \subseteq \mathbb{N}$ be such that $d(A)$ exists. Then

$$d(A) = \lim_{n\to\infty} \frac{|A \cap \{1, 2, ..., n\}|}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_A(k) = \lim_{n\to\infty} C(\mathbb{1}_A)_n.$$

As φ extends limits, we have

$$\mu(A) = \varphi(C(\mathbb{1}_A)) = \lim_{n\to\infty} C(\mathbb{1}_A)_n = d(A).$$

Hence μ extends density. The same reasoning with μ replaced by ν and φ replaced by ϑ shows that ν extends density. Finally, we have that $\mu \neq \nu$ as $\mu(E) = (\varphi \circ C)(\mathbb{1}_E) \neq (\vartheta \circ C)(\mathbb{1}_E) = \nu(E)$. □

Corollary 4.30. *There exists a bounded finitely additive measure that extends density and assumes a negative value.*

Proof. Let μ and ν be two different positive measures that extend density, as in proposition 4.29. Let $E \subseteq \mathbb{N}$ be such that $\mu(E) \neq \nu(E)$. Let $t \in \mathbb{R}$ be such that $t\mu(E) + (1-t)\nu(E) < 0$. Then the measure $t\mu + (1-t)\nu$ extends density and assumes a negative value. □

Remark. In fact we see from the proof of corollary 4.30 that for any choice of $\gamma \in \mathbb{R}$, we can construct a measure that extends density and satisfies $\mu(E) = \gamma$ for some $E \subseteq \mathbb{N}$. Indeed, just solve for t in the equation $t\mu(E) + (1-t)\nu(E) = \gamma$ in the proof of corollary 4.30.

The set of bounded, density-extending measures does not form a subspace of $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. However, we do have the following result.

Theorem 4.31. *The set of bounded, density-extending measures is a closed convex*[27] *subset of* $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Proof. Convexity is clear. Let $(\mu_n)_{n=1}^{\infty}$ be a sequence of bounded, density-extending measures and suppose $\mu_n \to \mu$ for $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. We show that μ extends density. Let $E \subseteq \mathbb{N}$ be such that $d(E)$ exists. Then we have that

$$|\mu_n(E) - \mu(E)| = |(\mu_n - \mu)(E)| \leq |\mu_n - \mu|(E) \leq |\mu_n - \mu|(\mathbb{N}) = \|\mu_n - \mu\|_{\mathrm{ba}} \to 0 \quad (n \to \infty),$$

hence $\lim_{n\to\infty} \mu_n(E) = \mu(E)$. But we have $\mu_n(E) = d(E)$ for all n, hence $\mu(E) = d(E)$. □

4.5 Ultrafilters and $0, 1$-valued measures

In this section we construct finitely additive measures that assume only the values 0 and 1. To this end we introduce the following definition.

Definition 4.32. *A measure* $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ *is called* $0, 1$***-valued*** *if* $\mu(E) \in \{0, 1\}$ *for all* $E \subseteq \mathbb{N}$ *and there exists* $F \subseteq \mathbb{N}$ *such that* $\mu(F) = 1$

Remark. Suppose μ is a $0, 1$-valued measure.

(1) By monotonicity, we have $\mu(\mathbb{N}) = 1$.

(2) Then μ is not shift-invariant: if it were shift-invariant, then it would assume all values of the form $\frac{1}{n}$ for $n \in \mathbb{N}$.

(3) For all $A \subseteq \mathbb{N}$, we have precisely one of $\mu(A) = 1$ or $\mu(A \setminus \mathbb{N}) = 1$.

We will see that some $0, 1$-values measures are induced by ℓ^1, but in this section we construct many more that are not of a form we have seen earlier. For this we need some theory regarding filters and ultrafilters.

Filters and ultrafilters

Definition 4.33 (Filter)**.** *Let X be a nonempty*[28] *set. A **filter** on X is a nonempty collection of subsets $\mathcal{F} \subseteq \mathcal{P}(X)$ satisfying*

(1) $\varnothing \notin \mathcal{F}$;

[27] Recall that a subset A of a real vector space is called **convex** if $(1-t)x + ty \in A$ for all $x, y \in A$ and $t \in [0, 1]$.

[28] In this section we will always assume that X is nonempty.

(2) If $F, G \in \mathcal{F}$, then $F \cap G \in \mathcal{F}$ (closed under finite intersections);
(3) If $F \in \mathcal{F}$ and $F \subseteq G$, then $G \in \mathcal{F}$ (closed under supersets).

By induction, the second property implies that $\bigcap_{i=1}^{n} F_i \in \mathcal{F}$ for any finite collection $F_1, F_2, ..., F_n \in \mathcal{F}$. Note that a filter must contain the full set X.

Intuitively, a filter can be thought of as a collection of subsets of X that are considered 'almost everything'. The following example explains this interpretation.

Example. Let $(S, \mathcal{A}, \mu)$ be a measure space with μ not the zero measure. For a subset $A \subseteq S$, let us write $A \approx S$ if and only if there exists $\Delta \in \mathcal{A}$ such that $\mu(\Delta^C) = 0$ and $\Delta \subseteq A$.[29] Thus intuitively, $A \approx S$ if and only if A differs from S only on a (subset of a) null set, i.e. A is 'almost everything'. Then one can easily check that the following intuitive properties hold:

- $\varnothing \not\approx S$ and $S \approx S$ ($\varnothing$ is not almost everything and S is almost everything);
- If $A \approx S$ and $B \approx S$ then $A \cap B \approx S$ (if A and B are almost everything, then $A \cap B$ is almost everything);
- If $A \approx S$ and $A \subseteq B$, then $B \approx S$ (if A is almost everything and B is even larger, then B is almost everything).

Then the collection $\mathcal{F} := \{A \subseteq S : A \approx S\}$ is a filter, and it is precisely the collection of sets that are 'almost everything'.[30]

Example. The collection $\{X\}$ is a filter, called the *trivial filter*. The power set $\mathcal{P}(X)$ is not a filter as it contains the empty set, and the collection $\mathcal{P}(X) \setminus \{\varnothing\}$ is not a filter if X contains at least two different elements.

Example. Let (M, d) be a metric space and fix some $x \in M$. A subset $U \subseteq M$ is called a *neighborhood* of x if there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$. The set of all neighborhoods of x is denoted by $\mathcal{U}(x)$. Then $\mathcal{U}(x)$ is a filter, called the *neighborhood filter of x*.[31]

Definition 4.34 (Principal filter)**.** *Consider a nonempty subset $A \subseteq X$. The collection*

$$\mathcal{F}_A = \{F \subseteq X : A \subseteq F\}$$

is a filter, called the ***principal filter generated by*** *A. A filter $\mathcal{F}$ is called* ***principal*** *if it is of the form $\mathcal{F}_A$ for some (nonempty) $A \subseteq X$, and non-principal otherwise.*

In the special case where $A = \{x\}$ is a singleton, the collection $\mathcal{F}_x := \mathcal{F}_{\{x\}} = \{F \subseteq X : x \in F\}$ is called the ***principal filter generated by*** *x.*

A question that arises naturally in view of definition 4.34 is whether there exist any non-principal filters on a set X. We will see that this is the case if and only X is an infinite set. First we give a simple criterion to decide whether a filter is principal.

Proposition 4.35 (Criterion for principal filters)**.** *Let $\mathcal{F}$ be a filter. Then $\mathcal{F}$ is principal if and only if $\bigcap_{F \in \mathcal{F}} F \in \mathcal{F}$.*

Proof. First suppose $\mathcal{F}$ is principal, so $\mathcal{F} = \mathcal{F}_A$ for some nonempty $A \subseteq X$. Then trivially we have that $\bigcap_{F \in \mathcal{F}} F = \{A\} \in \mathcal{F}$, as desired.

[29] Loosely speaking, $A \approx S$ means that $\mu(A^C) = 0$, but since A need not be measurable we introduce the set Δ. Note that $A \approx S$ is precisely the statement '$x \in A$ for μ-almost every $x \in S$'.

[30] There is an important caveat here. Let $\mathcal{F}$ be a filter on X and let us say $E \subseteq X$ is 'almost everything' if $E \in \mathcal{F}$ and 'almost nothing' if $X \setminus E \in \mathcal{F}$. It does *not* hold in general that if $E \subseteq X$ is not in $\mathcal{F}$, then E is 'almost nothing' (indeed, think of the filter that we associated with the measure space, for instance the subset $[0, 1] \subseteq [0, 2]$ where we consider $[0, 2]$ with the Lebesgue measure). However, for some types of filters called *ultrafilters* we will see that this does hold.

[31] One can define a neighborhood filter in more generality for topological spaces. Consider a topological space (X, τ) and a fixed $x \in X$. A subset $U \subseteq X$ is called a *neighborhood* of x if there exists $O \in \tau$ such that $x \in O \subseteq U$. The set of all neighborhoods of x is denoted $\mathcal{U}(x)$ and is called the *neighborhood filter of x*.

Conversely, suppose $A := \bigcap_{F\in\mathcal{F}} F \in \mathcal{F}$. Then A is nonempty as $\mathcal{F}$ is a filter. We show that $\mathcal{F} = \mathcal{F}_A$. Indeed, if $G \in \mathcal{F}$, then $A \subseteq G$ hence $G \in \mathcal{F}_A$. And for $G \in \mathcal{F}_A$, we have $A \subseteq G$ hence $G \in \mathcal{F}$. Thus we see that $\mathcal{F} = \mathcal{F}_A$, hence $\mathcal{F}$ is principal. □

From proposition 4.35 and the fact that filters are closed under finite intersections, we immediately derive the following.

Corollary 4.36. *If X is a finite set, then any filter on X is principal.*

Thus we cannot construct non-principal filters over finite sets. But over infinite sets this is possible; the following provides an example.

Definition 4.37. *Let X be an infinite set. The collection*

$$\mathcal{F}r := \{F \subseteq X : X \setminus F \text{ is finite}\}$$

*is called the **Fréchet filter**.*

It is readily checked that $\mathcal{F}r$ is indeed a filter. The Fréchet filter is a non-principal filter. To see this, we use proposition 4.35. Define $F_x = X \setminus \{x\}$ for all $x \in X$, then $F_x \in \mathcal{F}r$ and $\bigcap_{x\in X} F_x = \varnothing$. Now note that $\bigcap_{F\in\mathcal{F}r} F \subseteq \bigcap_{x\in X} F_x$, hence the former set must be empty and from proposition 4.35 we conclude that $\mathcal{F}r$ is non-principal.

If we have a collection of filters, then we can build a new filter in the following way.

Proposition 4.38 (Intersection and union of filters)**.** *Let $\mathcal{I}$ be a nonempty index set and let $(\mathcal{F}^{(\alpha)})_{\alpha\in\mathcal{I}}$ be a collection of filters on X.*

(1) Then $\bigcap_{\alpha\in\mathcal{I}} \mathcal{F}^{(\alpha)}$ is a filter.

(2) If additionally $(\mathcal{F}^{(\alpha)})_{\alpha\in\mathcal{I}}$ is a chain, in the sense that for all $\alpha, \beta \in \mathcal{I}$ we have either $\mathcal{F}^{(\alpha)} \subseteq \mathcal{F}^{(\beta)}$ or $\mathcal{F}^{(\beta)} \subseteq \mathcal{F}^{(\alpha)}$, then the union $\bigcup_{\alpha\in\mathcal{I}} \mathcal{F}^{(\alpha)}$ is a filter.

Proof. The proof of (1) is straightforward. For (2), denote $\mathcal{F} := \bigcup_{\alpha\in\mathcal{I}} \mathcal{F}^{(\alpha)}$. It is clear that $\mathcal{F}$ is nonempty and satisfies (1) and (3) in the definition of a filter. For (2), take $F, G \in \mathcal{F}$ and let $\alpha, \beta \in \mathcal{I}$ be such that $F \in \mathcal{F}^{(\alpha)}$ and $G \in \mathcal{F}^{(\beta)}$. Then one of $\mathcal{F}^{(\alpha)}$ and $\mathcal{F}^{(\beta)}$ is contained in the other one, say without loss of generality $\mathcal{F}^{(\alpha)} \subseteq \mathcal{F}^{(\beta)}$. Then $F, G \in \mathcal{F}^{(\beta)}$ hence $F \cap G \in \mathcal{F}^{(\beta)} \subseteq \mathcal{F}$, as desired. □

Remark. The requirement that $(\mathcal{F}^{(\alpha)})_{\alpha\in\mathcal{I}}$ is a chain is essential. For instance, take $x, y \in X$ with $x \neq y$, then $\mathcal{F}_x \cup \mathcal{F}_y$ is not a filter as $\{x\} \cap \{y\} = \varnothing$.

There is a natural way to construct filters with desired properties. Given a collection $\mathcal{C}$ of subsets of X, one can wonder whether there exists a filter $\mathcal{F}$ containing $\mathcal{C}$. The next definition introduces a necessary and sufficient condition on $\mathcal{C}$ for this to hold.

Definition 4.39. *A nonempty family of subsets $\mathcal{C} \subseteq \mathcal{P}(X)$ is said to have the **finite intersection property** if the intersection of any collection of finitely many sets from $\mathcal{C}$ is nonempty; that is,*

$$\text{for all } E_1, E_2, ..., E_n \in \mathcal{C}, \text{ we have that } \bigcap_{i=1}^{n} E_i \neq \varnothing.$$

Note that filters have the finite intersection property, and so does every (nonempty) subset of a filter.

Lemma 4.40 (Constructing filters)**.** *Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be a nonempty collection of sets with the finite intersection property. Then there exists a filter $\mathcal{F}$ on X that contains $\mathcal{C}$.*

Proof. We want to construct a filter $\mathcal{F}$ that contains $\mathcal{C}$. If this is to hold, then in particular we must have for all $E_1, E_2, ..., E_n \in \mathcal{C}$ that $\bigcap_{i=1}^n E_i \in \mathcal{F}$. But then $\mathcal{F}$ must also contain all supersets of the latter. This leads us to define

$$\mathcal{F} = \{F \in \mathcal{P}(X) : \text{ there exist } E_1, E_2, ..., E_n \in \mathcal{C} \text{ such that } \bigcap_{i=1}^n E_i \subseteq F\}.$$

It is readily checked that $\mathcal{F}$ is filter, where we use the finite intersection property to derive that $\varnothing \notin \mathcal{F}$. By construction, $\mathcal{F}$ contains $\mathcal{C}$.

We can actually show that the $\mathcal{F}$ constructed here is the *minimal* filter that contains $\mathcal{C}$, i.e. if $\mathcal{G}$ is any filter that contains $\mathcal{C}$, then $\mathcal{F} \subseteq \mathcal{G}$. Let $\mathcal{G}$ be any filter that contains $\mathcal{C}$. Let $F \in \mathcal{F}$ and let $E_1, E_2, ..., E_n \in \mathcal{C}$ be such that $\bigcap_{i=1}^n E_i \subseteq F$. Then as $\mathcal{G}$ is a filter and contains, $E_1, E_2, ..., E_n$, we have that $\bigcap_{i=1}^n E_i \in \mathcal{G}$ hence $A \in \mathcal{G}$. This proves the desired minimality. ☐

If $\mathcal{C}$ has the finite intersection property, then the (unique) minimal filter $\mathcal{F}$ that contains $\mathcal{C}$ is called the **filter generated by** $\mathcal{C}$ and is usually denoted $\langle \mathcal{C} \rangle$. Note that a 'principal filter generated by A' as in definition 4.34 is just a special case of this, with $\mathcal{C} = \{A\}$.

There is always a smallest filter, namely the trivial filter $\{X\}$. In general, there is not something like a largest filter (in the sense that it contains every filter). However, we can define the following notion of maximality for a filter.

Definition 4.41 (Ultrafilters)**.** *A filter $\mathcal{U}$ on X is called an* ***ultrafilter*** *if it is not properly contained in another filter. That is, if $\mathcal{F}$ is any filter on X with $\mathcal{U} \subseteq \mathcal{F}$, then $\mathcal{U} = \mathcal{F}$.*

We have already seen examples of ultrafilters: for fixed $x \in X$, the principal filter $\mathcal{F}_x$ is an ultrafilter. To see this, suppose $\mathcal{G}$ is any filter on X satisfying $\mathcal{F}_x \subseteq \mathcal{G}$. Then we have $\mathcal{F}_x = \mathcal{G}$: let $G \in \mathcal{G}$, then we must have $x \in G$: if $x \notin G$, then $G \cap \{x\} = \varnothing$, which is impossible as $\{x\} \in \mathcal{G}$. Thus $x \in G$, hence $G \in \mathcal{F}_x$ and so $\mathcal{F}_x = \mathcal{G}$.[32]

If $A \subseteq X$ contains at least two different elements, then the principal filter $\mathcal{F}_A$ generated by A is *not* an ultrafilter. Indeed, suppose this is an ultrafilter and fix $a \in A$, then $\mathcal{F}_{A \setminus \{a\}}$ is a filter satisfying $\mathcal{F}_A \subseteq \mathcal{F}_{A \setminus \{a\}}$ hence by maximality of $\mathcal{F}_A$, we have $\mathcal{F}_A = \mathcal{F}_{A \setminus \{a\}}$. But then we have $A \setminus \{a\} \in \mathcal{F}_A$ hence $A \subseteq A \setminus \{a\}$, which is impossible.

Proposition 4.42. *The principal ultrafilters on X are precisely the principal filters generated by some $x \in X$.*

A natural question is whether there exist non-principal *ultrafilters*. We will prove in corollary 4.44 that this is the case. The idea is to *extend* a non-principal filter (such as the Fréchet filter[33]) to an ultrafilter. For this we need the following theorem.

Theorem 4.43 (Ultrafilter lemma)**.** *Let $\mathcal{F}$ be a filter. Then there exists an ultrafilter $\mathcal{U}$ that contains $\mathcal{F}$.*

Proof. Let $\mathcal{M}$ denote the collection of all filters that contain $\mathcal{F}$. Then $\mathcal{M}$ is nonempty as $\mathcal{F} \in \mathcal{M}$. Consider the partial order $\subseteq$ on $\mathcal{M}$. We check the conditions in Zorn's lemma. Let $\mathcal{N} \subseteq \mathcal{M}$ be a nonempty chain. Then an upper bound for $\mathcal{N}$ is given by $\bigcup_{\mathcal{G} \in \mathcal{N}} \mathcal{G}$. The fact that this lies in $\mathcal{M}$ follows from proposition 4.38.

Thus by Zorn's lemma, $(\mathcal{M}, \subseteq)$ has a maximal element $\mathcal{U}$. This is an ultrafilter that contains $\mathcal{F}$. ☐

[32] This example proves that while there exist 'maximal' filters in the sense of definition 4.41, a 'maximum' does not exist (i.e. a filter that contains every other filter) if X contains at least two different elements. Indeed, if such a maximum exists then it must be $\mathcal{F}_x$, but for $x, y \in X$ with $x \neq y$ we have $\mathcal{F}_x \subsetneq \mathcal{F}_y$ (and $\mathcal{F}_y \subsetneq \mathcal{F}_x$).

[33] The Fréchet filter (on an infinite set) is not an ultrafilter: one can write the infinite set X as a disjoint union $X = E \cup F$ where E and F are both infinite (and thus so are their complements), so E and F do not belong to $\mathcal{F}r$ hence $\mathcal{F}r$ is not an ultrafilter (see proposition 4.45). For the construction of such sets E and F one can use the axiom of choice, but we will not make this precise here. In the case $X = \mathbb{N}$ one can take $E = 2\mathbb{N} - 1$ and $F = 2\mathbb{N}$ for instance.

Note that the ultrafilter lemma is trivial when we want to extend a filter $\mathcal{F}$ that satisfies $\bigcap_{F \in \mathcal{F}} F \neq \varnothing$. (Filters for which this intersection is empty will later be called *free*.) Indeed, then we may choose any point x in this intersection, and then $\mathcal{F}_x$ is an ultrafilter that contains $\mathcal{F}$.

The ultrafilter lemma is very useful: it allows one to construct an ultrafilter containing certain desired sets. We have seen that any collection of sets with the finite intersection property is contained in a filter, and we now know that such a filter is in turn contained in an ultrafilter.

Corollary 4.44. *Let X be an infinite set. Then there exists at least one non-principal ultrafilter on X.*

Proof. Let $\mathcal{F}$ be a non-principal filter on X, for instance the Fréchet filter. Let $\mathcal{U}$ be an ultrafilter that contains $\mathcal{F}$. We claim that $\mathcal{U}$ is non-principal. Note that $\bigcap_{F \in \mathcal{F}} F \notin \mathcal{F}$, so $\bigcap_{F \in \mathcal{F}} F \notin \mathcal{U}$. Now if $\mathcal{U}$ were principal, we would have $\bigcap_{F \in \mathcal{U}} F \in \mathcal{U}$, but then we have $\bigcap_{F \in \mathcal{U}} F \subseteq \bigcap_{F \in \mathcal{F}} F$ hence $\bigcap_{F \in \mathcal{F}} F \in \mathcal{U}$, which is a contradiction. We conclude that $\mathcal{U}$ is non-principal. □

We now give some characterizations of ultrafilters that will be used throughout this section.

Proposition 4.45 (Characterization of ultrafilters)**.** *Let $\mathcal{F}$ be a filter on X. The following are equivalent:*

(1) $\mathcal{F}$ is an ultrafilter;
(2) For every $A \subseteq X$, either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$;
(3) For any finite collection $E_1, E_2, ..., E_n \in \mathcal{P}(X)$, $\bigcup_{i=1}^n E_i \in \mathcal{F}$ implies that $E_i \in \mathcal{F}$ for some $i \in \{1, ..., n\}$.

Proof. (1) $\implies$ (2). Suppose $\mathcal{F}$ is an ultrafilter and let $A \subseteq X$. Suppose $X \setminus A \notin \mathcal{F}$; we prove that $A \in \mathcal{F}$. We check that the collection $\mathcal{F} \cup \{A\}$ has the finite intersection property: let $E_1, E_2, ..., E_n \in \mathcal{F}$ and write $E := \bigcap_{i=1}^n E_i \in \mathcal{F}$. Then we have $E \cap A \neq \varnothing$: if $E \cap A = \varnothing$ then $E \subseteq X \setminus A$ hence $X \setminus A \in \mathcal{F}$, which is a contradiction. Thus $\mathcal{F} \cup \{A\}$ has the finite intersection property hence there exists a filter $\mathcal{G}$ that contains $\mathcal{F} \cup \{A\}$. By maximality of $\mathcal{F}$, we must have $\mathcal{F} = \mathcal{G}$ hence $A \in \mathcal{F}$, as desired.

(2) $\implies$ (3). Suppose (2); by induction it suffices to consider the case $n = 2$ in (3). Let $A, B \subseteq X$ be such that $A \cup B \in \mathcal{F}$. Now if both $A \notin \mathcal{F}$ and $B \notin \mathcal{F}$ then we must have $X \setminus A \in \mathcal{F}$ and $X \setminus B \in \mathcal{F}$ hence $(X \setminus A) \cap (X \setminus B) = X \setminus (A \cup B) \in \mathcal{F}$, which is impossible. The result follows.

(3) $\implies$ (1). Assume (3) and suppose $\mathcal{F}$ is not maximal: there exists a filter $\mathcal{G}$ such that $\mathcal{F} \subsetneq \mathcal{G}$. Fix some $A \in \mathcal{G} \setminus \mathcal{F}$, then as $X = A \cup (X \setminus A) \in \mathcal{F}$ we must have $X \setminus A \in \mathcal{F}$. But then $X \setminus A \in \mathcal{G}$, hence $A \cap (X \setminus A) = \varnothing \in \mathcal{G}$, which is a contradiction. The claim follows. □

Intuitively, property (2) states that an ultrafilter decides for every $A \subseteq \mathbb{N}$ whether either A is 'almost everything' ($A \in \mathcal{F}$) or 'almost nothing' (in the sense that $X \setminus A \in \mathcal{F}$), and (3) states that if $A \cup B$ is 'almost everything' then at least one of A and B must be 'almost everything'.

We introduce one more notion to classify filters.

Definition 4.46. *A filter $\mathcal{F}$ is called **free** if $\bigcap_{F \in \mathcal{F}} F = \varnothing$.*

An example of a free filter is the Fréchet filter on an infinite set X: in the discussion following definition 4.37 we have seen that $\bigcap_{F \in \mathcal{F}r} F = \varnothing$.

Note that a free filter is necessarily non-principal, in view of proposition 4.35. The converse does not hold in general. Let (M, d) be a metric space and fix an accumulation point $x \in M$ (for instance one can take $M = \mathbb{R}$ and any $x \in \mathbb{R}$). Consider the neighborhood filter $\mathcal{U}(x)$, then we have $\bigcap_{F \in \mathcal{U}(x)} F = \{x\} \notin \mathcal{U}(x)$. Thus $\mathcal{U}(x)$ is neither principal nor free.

We can characterize free filters with the following proposition.

Proposition 4.47. *Let X be an infinite set.*[34] *A filter $\mathcal{F}$ on X is free if and only if it contains the Fréchet filter.*

[34] Note that on a finite set X every filter is principal, so not free.

Proof. Suppose $\mathcal{F}r \subseteq \mathcal{F}$. Then as $\mathcal{F}r$ is free, we have $\bigcap_{F\in\mathcal{F}} F \subseteq \bigcap_{F\in\mathcal{F}r} F = \varnothing$, thus $\mathcal{F}$ is free.

Conversely, suppose that $\mathcal{F}$ is free and let $G \in \mathcal{F}r$. We show that $G \in \mathcal{F}$; we may assume without loss of generality that $G \subsetneq X$. For all $x \in X \setminus G$, choose $F_x \in \mathcal{F}$ such that $x \notin F_x$. Then as $X \setminus G$ is finite, the intersection $F := \bigcap_{x\in X\setminus G} F_x$ lies in $\mathcal{F}$ and we have that $F \subseteq G$ by construction. Hence $G \in \mathcal{F}$, as desired. □

Corollary 4.48. *Let X be an infinite set and let $\mathcal{F}$ be a free ultrafilter on X. Then $\mathcal{F}$ does not contain a finite set.*

Proof. If $\mathcal{F}$ contains a finite set E, then it also contains its complement $X\setminus E$ as the Fréchet filter is contained in $\mathcal{F}$. But then $E \cap (X \setminus E) = \varnothing \in \mathcal{F}$, which is a contradiction. □

For ultrafilters, 'free' and 'non-principal' are the same.

Theorem 4.49. *An ultrafilter is either free or principal (and not both).*

Proof. It is clear that at most one of these can hold. Let $\mathcal{U}$ be an ultrafilter on X. If X is finite then every (ultra)filter is principal, so we may assume that X is infinite. Suppose that $\mathcal{U}$ is non-principal. We show that $\mathcal{U}$ contains the Fréchet filter.

First consider fixed $x \in X$, then by proposition 4.45 we have that either $\{x\} \in \mathcal{U}$ or $X \setminus \{x\} \in \mathcal{U}$. As $\mathcal{U}$ is non-principal, the former cannot hold, so we have $X \setminus \{x\} \in \mathcal{U}$.

Now take $F \in \mathcal{F}r$ (w.l.o.g. $F \subsetneq X$); then $X \setminus F$ is finite and $F = X \setminus (\bigcup_{x\in X\setminus F}\{x\}) = \bigcap_{x\in X\setminus F}(X \setminus \{x\})$ is a finite intersection of sets in $\mathcal{U}$ hence lies in $\mathcal{U}$. This proves that $\mathcal{F}r \subseteq \mathcal{U}$. Now by proposition 4.47 it follows that $\mathcal{U}$ is free. □

Measures induced by ultrafilters

From now on we consider (ultra)filters on the set $X = \mathbb{N}$. The set of ultrafilters on $\mathbb{N}$ is denoted $\beta\mathbb{N}$. Ultrafilters induce a $0,1$-valued measure on $\mathbb{N}$ in a natural way, by assigning measure 1 to a set if it is 'almost everything' ($E \in \mathcal{U}$) and 0 if it is 'almost nothing' ($X \setminus E \in \mathcal{U}$, or equivalently $E \notin \mathcal{U}$).

Let $\mathcal{U} \subseteq \mathcal{P}(\mathbb{N})$ be an ultrafilter on $\mathbb{N}$. Define $\mu_{\mathcal{U}} : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by

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

We check that $\mu_{\mathcal{U}}$ is a finitely additive measure. Let $E, F \subseteq \mathbb{N}$ be disjoint. If $E \cup F \notin \mathcal{U}$ then by properties of filters we have $E, F \notin \mathcal{U}$ and we have the desired additivity. Suppose $E\cup F \in \mathcal{U}$, then as $\mathcal{U}$ is an ultrafilter we have either $E \in \mathcal{U}$ or $F \in \mathcal{U}$, say without loss of generality $E \in \mathcal{U}$. Then $F \notin \mathcal{U}$ as $E \cap F = \varnothing$, and again we have the desired additivity. We call $\mu_{\mathcal{U}}$ the measure *associated with* (or *induced by*) the ultrafilter $\mathcal{U}$.

Remark. This way of constructing measures only works for *ultrafilters*. Suppose $\mathcal{F}$ is an arbitrary filter on $\mathbb{N}$ and suppose that the map $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ defined by

$$\mu(E) = \begin{cases} 0 & \text{if } E \notin \mathcal{F}, \\ 1 & \text{if } E \in \mathcal{F}. \end{cases}$$

defines a finitely additive, $0,1$-valued measure on $\mathbb{N}$. We claim that $\mathcal{F}$ is an ultrafilter. Indeed let $E \subseteq \mathbb{N}$, then $\mu(E) + \mu(\mathbb{N} \setminus E) = \mu(\mathbb{N}) = 1$ hence (precisely) one of $\mu(E)$ and $\mu(\mathbb{N} \setminus E)$ equals 1. That is, (precisely) one of E and $\mathbb{N} \setminus E$ lies in $\mathcal{F}$, and by proposition 4.45 it follows that $\mathcal{F}$ is an ultrafilter.

Before we present any applications of our new idea of constructing measures, let us check that the map $\mathcal{U} \to \mu_{\mathcal{U}}$ is a bijection from the set of ultrafilters on $\mathbb{N}$ onto the set of $0,1$-valued measures. Trivially, it is injective. Now consider any $0,1$-valued measure μ. Define $\mathcal{U} = \{E \subseteq \mathbb{N} : \mu(E) = 1\}$. Then $\mathcal{U}$ is an ultrafilter:

- $\mathcal{U} \neq \varnothing$ as $\mathbb{N} \in \mathcal{U}$, and $\varnothing \notin \mathcal{U}$;
- From the fact that μ is $0,1$-valued one can easily check that $\mu(E \cap F) = \mu(E)\mu(F)$ for all $E, F \subseteq \mathbb{N}$, so that $\mathcal{U}$ is closed under finite intersections;
- By monotonicity of μ, $\mathcal{U}$ is closed under supersets.

Finally, we have that $\mu = \mu_{\mathcal{U}}$. Thus our identification $\mathcal{U} \leftrightarrow \mu_{\mathcal{U}}$ is a bijective correspondence.

One can construct $0,1$-valued measures on $\mathbb{N}$ that take value 1 on specifically chosen sets: take a collection $\mathcal{C}$ of subsets of $\mathbb{N}$ with the finite intersection property, let $\mathcal{U}$ be an ultrafilter that contains $\mathcal{C}$; then the associated measure $\mu_{\mathcal{U}}$ takes value 1 on the elements of $\mathcal{C}$. For instance, letting $\mathcal{C} = \{a\mathbb{N} : a \in \mathbb{N}\}$, we get a $0,1$-valued measure μ that satisfies $\mu(a\mathbb{N}) = 1$ for all $a \in \mathbb{N}$.

Let us now investigate whether the $0,1$-valued measures induced by ultrafilters are of a form we have seen before. Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. First suppose $\mathcal{U}$ is a *principal* ultrafilter, then there exists $n \in \mathbb{N}$ such that $\mathcal{U} = \mathcal{U}_n$, the (ultra)filter generated by n. Then the associated measure $\mu = \mu_{\mathcal{U}_n}$ is given by

$$\mu(E) = \begin{cases} 0 & \text{if } n \notin E, \\ 1 & \text{if } n \in E. \end{cases}$$

This is precisely the Dirac measure δ_n centered at n. Conversely, any Dirac measure δ_n on $\mathbb{N}$ is induced by a principal ultrafilter, namely $\mathcal{U} = \mathcal{U}_n = \{E \subseteq \mathbb{N} : n \in E\}$. Hence we have a bijective correspondence $\mathcal{U}_n \leftrightarrow \mu_{\mathcal{U}_n} = \delta_n$ between principal ultrafilters and Dirac measures. Also note that the collection of 'Dirac functionals', i.e., functionals $\varphi_n \in (\ell^\infty)'$ of the form $\varphi_n(x) = x_n$, corresponds bijectively to the collection of Dirac measures via the identification $\varphi_n \leftrightarrow \mu_{\varphi_n} = \delta_n$. Finally, the collection of standard (Schauder) basis vectors $e^{(n)} = (0, 0, ..., 0, 1, 0, ...) \in \ell^1$ corresponds bijectively to the collection of 'Dirac functionals' via the identification $e^{(n)} \leftrightarrow \varphi_{e^{(n)}} = \varphi_n$. Summarizing, we have a bijective correspondence

$$e^{(n)} \leftrightarrow \varphi_{e^{(n)}} = \varphi_n \leftrightarrow \mu_{\varphi_n} = \delta_n = \mu_{\mathcal{U}_n} \leftrightarrow \mathcal{U}_n.$$

Now suppose $\mathcal{U}$ is a non-principal (hence free) ultrafilter and let μ be the associated measure. Then μ vanishes on finite sets as $\mathcal{U}$ does not contain finite sets, hence μ is not induced by ℓ^1. Note that $\mu_{\mathcal{U}}$ is also not shift-invariant as it assumes only the values 0 and 1. The relation between free ultrafilters and functionals will be discussed later in this section.

The idea of a measure induced by an ultrafilter can be extended in the following sense. Let $\mathcal{U}^{(1)}, \mathcal{U}^{(2)}, ...$ be (not necessarily distinct) ultrafilters on $\mathbb{N}$ and let $a \in \ell^1$. Define $\mu : \mathcal{P}(\mathbb{N}) \to \mathbb{R}$ by

$$\mu(E) = \sum_{n=1}^{\infty} a_n \mathbb{1}_{\mathcal{U}^{(n)}}(E) = \sum_{n=1}^{\infty} a_n \mu_{\mathcal{U}^{(n)}}(E). \tag{17}$$

Then μ is a finitely additive measure. If every $\mathcal{U}^{(n)}$ is principal, say $\mathcal{U}^{(n)}$ is generated by $N_n \in \mathbb{N}$, then one can check that μ is the measure induced by $b = (b_1, b_2, ...) \in \ell^1$ defined by $b_m = \sum_{n=1}^{\infty} a_n \mathbb{1}_{\{m\}}(N_n)$.

Conversely, every measure μ induced by some $a \in \ell^1$ is of the form (17) just by taking $\mathcal{U}^{(n)}$ to be the principal ultrafilter generated by n.

Limits along an ultrafilter

Ultrafilters provide a way to describe convergence of a sequence. The idea is as follows. Suppose we have a sequence $(x_n)_{n=1}^{\infty}$ in a compact metric space; for us this will be a compact subset of $\mathbb{R}$. Most likely this sequence does not have a limit $\lim_{n\to\infty} x_n$: it can oscillate wildly or show other strange behaviour. However, as the sequence lies in a compact set, we know that it has a *convergent subsequence*. In general it will have many different convergent subsequences, whose limits we call *subsequential limits*. An ultrafilter can be used to choose in a consistent way precisely one of these subsequential limits and define this to be the (generalized) 'limit' of the sequence. For instance, consider the sequence $(-1, 1, -1, 1, ...)$. If an ultrafilter $\mathcal{U}$ contains the coefficients corresponding to the subsequential limit 1 (that is, $2\mathbb{N} \in \mathcal{U}$), then the limit along this ultrafilter will be 1; if it contains the coefficients corresponding to -1 then limit will be -1. This intuition will become more clear in the proof of theorem 4.52.

We first define what it means for a sequence to *converge along an ultrafilter.*

Definition 4.50 (Limits along an ultrafilter)**.** *Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$ and let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. We say that a real number α is a **$\mathcal{U}$-limit of** $(x_n)_{n=1}^{\infty}$ if for every $\varepsilon > 0$, we have*

$$\{n \in \mathbb{N} : |x_n - \alpha| < \varepsilon\} \in \mathcal{U}.$$

In that case, α is unique and we define $\mathcal{U}\text{-}\lim x_n := \alpha$.

Intuitively, this describes that the set $\{n \in \mathbb{N} : |x_n - \alpha| < \varepsilon\}$ is 'almost everything': $|x_n - \alpha| < \varepsilon$ for $\mu_{\mathcal{U}}$-almost every $n \in \mathbb{N}$.[35]

Limits along an ultrafilter offer a very useful way to extend the usual limit operator lim from the set of convergent sequences to an operator $\mathcal{U}\text{-}\lim$ defined on all of ℓ^{∞}, similar to a Banach limit. However, whereas Banach limits were *shift-invariant* functionals, we will see that limits along an ultrafilter give a *multiplicative* functional, a concept that will be introduced later.

Before we continue, we have to show that a so-called *ultralimit* is indeed unique whenever it exists.

Proposition 4.51. *Let $\mathcal{U}$ be any ultrafilter on $\mathbb{N}$ and $(x_n)_{n=1}^{\infty}$ a sequence of real numbers. If a $\mathcal{U}$-limit of $(x_n)_{n=1}^{\infty}$ exists, then it is unique.*

Proof. Suppose $\alpha, \beta \in \mathbb{R}$ are both $\mathcal{U}$-limits of $(x_n)_{n=1}^{\infty}$. Let $\varepsilon > 0$. Then the sets $\{n \in \mathbb{N} : |x_n - \alpha| < \varepsilon/2\}$ and $\{n \in \mathbb{N} : |x_n - \beta| < \varepsilon/2\}$ both belong to $\mathcal{U}$, thus so does their intersection, which is then nonempty. Take some n in this intersection, then it follows that $|\alpha - \beta| \leq |\alpha - x_n| + |x_n - \beta| < \varepsilon$. Letting $\varepsilon \downarrow 0$, we obtain $\alpha = \beta$, as desired. □

In general, existence of an ultralimit is not guaranteed.

Example. Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$ and let $(x_n)_{n=1}^{\infty}$ be defined by $x_n = n$. Then $\mathcal{U}\text{-}\lim x_n$ does not exists: for every $\alpha \in \mathbb{R}$, the set $\{n \in \mathbb{N} : |x_n - \alpha| < 1\}$ (or any other choice of $\varepsilon > 0$) is finite hence does not belong to $\mathcal{U}$.

The case that will be interesting to us is when we have a *free* ultrafilter and a *bounded* sequence.

Theorem 4.52 ($\mathcal{U}$ is free and $(x_n)_{n=1}^{\infty}$ is bounded)**.** *Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$ and suppose $(x_n)_{n=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. Then $\mathcal{U}\text{-}\lim x_n$ exists. Moreover, it is a subsequential limit of $(x_n)_{n=1}^{\infty}$.*

Proof. We need only show existence and the final assertion. We first prove the result for the special case when $(x_n)_{n=1}^{\infty}$ has only *finitely* many subsequential limits, to explain the intuition described earlier in this section. The idea is illustrated in figure 3.

[35] For a *finitely* additive measure, we define 'μ-almost everywhere' in the same way as this is done for σ-additive measures. More concretely, suppose $(S, \mathcal{A})$ is a measurable space and μ is a finitely additive measure on this space. Let $P(x)$ be a statement for every $x \in S$. We say that $P(x)$ *for μ-almost every $x \in S$* if there exists $\Delta \in \mathcal{A}$ such that $P(x)$ holds for all $x \in \Delta$, and moreover $\mu(\Delta^C) = 0$.

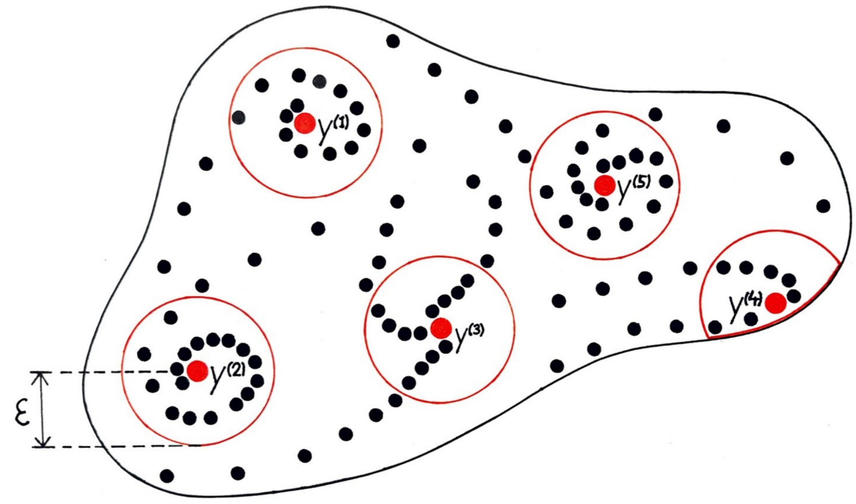

Figure 3: Proof of part 1 of theorem 4.52. The red points represent the subsequential limits $y^{(1)}, y^{(2)}, ..., y^{(N)}$. The points inside the red circles correspond to a set of coordinates $\{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\}$. The finite collection of points outside the red circles corresponds to the set of coordinates E_ε.

Part 1: Special case. Let $D(x)$ denote the set of subsequential limits of $(x_n)_{n=1}^\infty$ and assume this set is finite. As $(x_n)_{n=1}^\infty$ is bounded, $D(x)$ is nonempty. We may write $D(x) = \{y^{(1)}, y^{(2)}, ..., y^{(N)}\}$ (for $y^{(m)}$ all different). For every $\varepsilon > 0$, we have

$$\mathbb{N} = \bigcup_{m=1}^{N} \{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\} \cup E_\varepsilon,$$

where we define $E_\varepsilon = \mathbb{N} \setminus \bigcup_{m=1}^N \{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\}$. Then E_ε is a finite set (we will prove this), hence $E_\varepsilon \notin \mathcal{U}$.[36] Thus by properties of ultrafilters, $\mathcal{U}$ contains one of the sets $\{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\}$. Intuitively, this means that $\mathcal{U}$ contains the coefficients corresponding to a subsequence that converges to $y^{(m)}$. To see that E_ε is indeed finite, suppose to the contrary that it is infinite. Then one may construct a convergent subsequence $(x_{n_k})_{k=1}^\infty$ with coefficients $n_k \in E_\varepsilon$. This sequence converges to some $y \in D(x)$, say $y = y^{(m)}$. Then there exists $k \in \mathbb{N}$ such that $|x_{n_k} - y^{(m)}| < \varepsilon$. But then $n_k \in \{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\}$, which contradicts the definition of E_ε.

Now define $\delta = \min\{|y^{(m)} - y^{(k)}| : 1 \le k, m \le N, k \ne m\}$[37] and take $m \in \{1, ..., N\}$ such that $\{n \in \mathbb{N} : |x_n - y^{(m)}| < \delta/2\} \in \mathcal{U}$. We claim that $\{n \in \mathbb{N} : |x_n - y^{(m)}| < \varepsilon\} \in \mathcal{U}$ for all $\varepsilon > 0$, so that $y^{(m)}$ is the desired $\mathcal{U}$-limit. Indeed let $\varepsilon > 0$, without loss of generality $\varepsilon < \delta/2$. Then there is $1 \le k \le N$ such that $\{n \in \mathbb{N} : |x_n - y^{(k)}| < \varepsilon\} \in \mathcal{U}$. Then we have $k = m$. Indeed, the intersection $\{n \in \mathbb{N} : |x_n - y^{(m)}| < \delta/2, |x_n - y^{(k)}| < \varepsilon\}$ lies in $\mathcal{U}$ hence is nonempty; take n in this intersection, then $|y^{(m)} - y^{(k)}| \le |x_n - y^{(m)}| + |x_n - y^{(k)}| < \delta$, hence $m = k$. Thus $y^{(m)}$ is the (unique) $\mathcal{U}$-limit, and it is a subsequential limit of $(x_n)_{n=1}^\infty$.

Part 2: General case. We first prove existence of a $\mathcal{U}$-limit and then show that it must be a subsequential limit. Note that the sequence lies in the compact set $B := [-\|x\|_\infty, \|x\|_\infty]$. We claim that there exists $y \in B$ such that y is a $\mathcal{U}$-limit of $(x_n)_{n=1}^\infty$. Suppose not, then for every $y \in B$ there exists $\varepsilon_y > 0$ such that $\{n \in \mathbb{N} : |x_n - y| < \varepsilon_y\} \notin \mathcal{U}$. Then the collection $((y - \varepsilon_y, y + \varepsilon_y) \cap B)_{y \in B}$ is an open cover of B. By

[36] Here we use that $\mathcal{U}$ is free.

[37] This is only well-defined if $N \ge 2$. We may for instance define $\delta = 1$ when $N = 1$, but one can check that what follows is trivial when $N = 1$.

compactness of B, this has a finite subcover, say

$$B = \bigcup_{m=1}^{N} (y_m - \varepsilon_{y_m}, y_m + \varepsilon_{y_m}) \cap B.$$

Then we have that $\mathbb{N} = \{n \in \mathbb{N} : x_n \in B\}$, i.e.,

$$\mathbb{N} = \bigcup_{m=1}^{N} \{n \in \mathbb{N} : |x_n - y_m| < \varepsilon_{y_m}\}.$$

By properties of ultrafilters, one of these sets is contained in $\mathcal{U}$. But this is a contradiction with the definition of ε_{y_m}. We conclude that some $y \in B$ is a $\mathcal{U}$-limit of $(x_n)_{n=1}^{\infty}$. For every $\varepsilon > 0$, the set $\{n \in \mathbb{N} : |x_n - y| < \varepsilon\}$ lies in $\mathcal{U}$ thus it is an infinite set ($\mathcal{U}$ is free). Then one may construct a subsequence $(x_{n_k})_{k=1}^{\infty}$ that converges to y, which shows that y is a subsequential limit. □

Remark. From the final part of the proof we can actually see the following. Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$ and let $(x_n)_{n=1}^{\infty}$ be any sequence (not necessarily bounded) such that $\mathcal{U}\text{-}\lim x_n$ exists. Then $(x_n)_{n=1}^{\infty}$ has at least one convergent subsequence, and $\mathcal{U}\text{-}\lim x_n$ is equal to a subsequential limit.

The 'converse' to this does not hold generally, in the following sense. Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$ containing $2\mathbb{N}$[38] and consider $(x_n)_{n=1}^{\infty} = (0, 2, 0, 4, 0, 6, ...)$. This has a convergent subsequence, but $\mathcal{U}\text{-}\lim x_n$ does not exist. Indeed, suppose it exists, then it must be equal to a subsequential limit, hence $\mathcal{U}\text{-}\lim x_n = 0$. But for small $\varepsilon > 0$, the set $\{n \in \mathbb{N} : |x_n| < \varepsilon\} = 2\mathbb{N} - 1$ is not in $\mathcal{U}$, a contradiction.

We have seen that free ultrafilters allow us to attribute a 'generalized' limit to any bounded sequence by consistenly choosing a subsequential limit. The 'converse' is also possible, in the following sense.

Proposition 4.53. *Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers and suppose there exists a convergent subsequence $(x_{n_k})_{k=1}^{\infty}$ with limit α. Then there exists a free ultrafilter $\mathcal{U}$ on $\mathbb{N}$ such that*

$$\mathcal{U}\text{-}\lim x_n = \alpha.$$

Proof. Let $A = \{n_k : k \in \mathbb{N}\}$ and note that this is an infinite set as $n_{(\cdot)}$ is strictly increasing. The collection $\mathcal{F}r \cup \{A\}$ (with $\mathcal{F}r$ the Fréchet filter) has the finite intersection property. To see this, it suffices to show that for all $E \in \mathcal{F}r$, we have $E \cap A \neq \varnothing$. Now if $E \cap A = \varnothing$, then $A \subseteq \mathbb{N} \setminus E$ and the latter is a finite set, which is impossible.

Thus we may extend $\mathcal{F}r \cup \{A\}$ to an ultrafilter $\mathcal{U}$. As $\mathcal{F}r$ is free, so is $\mathcal{U}$. We claim that $\mathcal{U}\text{-}\lim x_n = \alpha$. Let $\varepsilon > 0$ and take $N \in \mathbb{N}$ such that $|x_{n_k} - \alpha| < \varepsilon$ for all $k \geq N$. Then $\{n_k : k \geq N\} \subseteq \{n \in \mathbb{N} : |x_n - \alpha| < \varepsilon\}$. The first set lies in $\mathcal{U}$ as it can be written as $\{n_k : k \in \mathbb{N}\} \cap \{n_N, n_N + 1, n_N + 1, ...\}$, the intersection of two sets in $\mathcal{U}$. Therefore $\{n \in \mathbb{N} : |x_n - \alpha| < \varepsilon\} \in \mathcal{U}$, as desired. □

If a sequence is convergent, then any subsequence converges to the same limit, so by theorem 4.52 we have the following result.

Corollary 4.54. *Let $\mathcal{U}$ be a free ultrafilter on $\mathbb{N}$ and let $(x_n)_{n=1}^{\infty}$ be a convergent sequence. Then*

$$\mathcal{U}\text{-}\lim x_n = \lim_{n \to \infty} x_n.$$

Thus a free ultrafilter $\mathcal{U}$ induces a functional $\mathcal{U}\text{-}\lim$ defined on ℓ^{∞} that extends the usual limit.

In the proof of theorem 4.52 we used the fact that $\mathcal{U}$ is free to conclude that $\mathcal{U}\text{-}\lim x_n$ is equal to some subsequential limit of $(x_n)_{n=1}^{\infty}$. This suggests that this might fail for $\mathcal{U}$-limits over *principal* ultrafilters. Indeed, principal filters do not give us something useful, as the next proposition shows.

[38] One can easily construct such an ultrafilter by extending the collection $\mathcal{F}r \cup \{2\mathbb{N}\}$ (with the finite intersection property) to an ultrafilter.

Proposition 4.55 ($\mathcal{U}$ is principal)**.** *Let $\mathcal{U}$ be a principal ultrafilter on $\mathbb{N}$. Take $m \in \mathbb{N}$ such that $\mathcal{U} = \mathcal{U}_m$, the principal filter generated by m. Then for any sequence $(x_n)_{n=1}^{\infty}$ of real numbers, $\mathcal{U}\text{-}\lim x_n$ exists and $\mathcal{U}\text{-}\lim x_n = x_m$.*

Proof. To see that x_m is the (unique) $\mathcal{U}$-limit of $(x_n)_{n=1}^{\infty}$, just note that for all $\varepsilon > 0$, we have that $x_m \in \{n \in \mathbb{N} : |x_n - x_m| < \varepsilon\}$ hence this set lies in $\mathcal{U}$. □

Before we discuss more properties of $\mathcal{U}\text{-}\lim$, let us mention that the measures $\mu_{\mathcal{U}}$ associated with an ultrafilter $\mathcal{U}$ (either free or principal) are really just limits along this ultrafilter. More precisely, we have that $\mu_{\mathcal{U}}(E) = \mathcal{U}\text{-}\lim \mathbb{1}_E(n)$. Indeed, first suppose that $E \in \mathcal{U}$, then for all $0 < \varepsilon < 1$, we have that $\{n \in \mathbb{N} : |\mathbb{1}_E(n) - 1| < \varepsilon\} = E \in \mathcal{U}$, thus $\mu_{\mathcal{U}}(E) = \mathcal{U}\text{-}\lim \mathbb{1}_E(n)$. If $E \notin \mathcal{U}$, then $\mathbb{N} \setminus E \in \mathcal{U}$ and a similar reasoning applies.

An ultrafilter (either free or principal) $\mathcal{U}$ gives rise to a map $\mathcal{U}\text{-}\lim : \ell^\infty \to \mathbb{R}$. This function has some special properties: it is a so-called *multiplicative functional.*

Multiplicative functionals

Definition 4.56 (Multiplicative functional)**.** *A function $\Phi : \ell^\infty \to \mathbb{R}$ is called a* ***multiplicative functional*** *if it satisfies*

(1) $\Phi(ax + by) = a\Phi(x) + b\Phi(y)$ for all $x, y \in \ell^\infty$ and $a, b \in \mathbb{R}$ (linearity);
(2) $\Phi(x) \geq 0$ whenever $x \in \ell^\infty$ with $x_n \geq 0 \quad \forall n \in \mathbb{N}$ (positivity);
(3) $\Phi(\mathbb{1}) = 1$, where $\mathbb{1} := (1, 1, ...)$ (Φ is normalized);
(4) $\Phi(xy) = \Phi(x)\Phi(y)$ for all $x, y \in \ell^\infty$, where $xy = (x_n y_n)_{n=1}^{\infty}$ (Φ is multiplicative).

Remark. Any map $\Phi : \ell^\infty \to \mathbb{R}$ with properties (1), (2) and (3) satisfies $\Phi \in (\ell^\infty)'$ and $\|\Phi\|_{(\ell^\infty)'} = 1$. Indeed, let $x \in \ell^\infty$, first consider the case $\|x\|_\infty = 1$. Let $y := \mathbb{1} - x$, so that $y_n \geq 0$. Then $1 = \Phi(\mathbb{1}) = \Phi(x + y) = \Phi(x) + \Phi(y)$, so that $1 - \Phi(x) = \Phi(y) \geq 0$ hence $\Phi(x) \leq 1$. This reasoning applied to $-x$ shows that $\Phi(-x) \leq 1$, hence $|\Phi(x)| \leq 1 = \|x\|_\infty$. For general nonzero $x \in \ell^\infty$, the claim $|\Phi(x)| \leq \|x\|_\infty$ follows from the previous by homogeneity, and the case $x = 0$ is trivial. Thus we see that $\Phi \in (\ell^\infty)'$ and $\|\Phi\|_{(\ell^\infty)'} \leq 1$; combined with (3) this gives $\|\Phi\|_{(\ell^\infty)'} = 1$

We sometimes say that a functional $\varphi \in (\ell^\infty)'$ is *multiplicative on a subset $W \subseteq \ell^\infty$* if $\varphi(xy) = \varphi(x)\varphi(y)$ holds for all $x, y \in W$. Note that a functional being multiplicative on ℓ^∞ does not necessarily mean that it is a *multiplicative functional.*

When we know that a function is a multiplicative functional, we will often write it as Φ or Ψ instead of φ or ψ to emphasize that it has these additional properties. This should not be confused with Banach limits. A Banach limit cannot be multiplicative: this follows by considering the sequence $(1, 0, 1, 0, ...)$.

The next lemma gives sufficient (and necessary) conditions for a functional to be multiplicative on ℓ^∞.

Lemma 4.57. *Suppose $\varphi \in (\ell^\infty)'$ satisfies*

$$\varphi(\mathbb{1}_E \mathbb{1}_F) = \varphi(\mathbb{1}_E)\varphi(\mathbb{1}_F)$$

for all $E, F \subseteq \mathbb{N}$. Then $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in \ell^\infty$.

Proof. First we prove the claim for simple sequences. Indeed, if $a = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ and $b = \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j}$ are simple sequences, then

$$ab = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \mathbb{1}_{A_i} \mathbb{1}_{B_j}.$$

Then using linearity of φ and the assumption, we obtain

$$\varphi(ab) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \varphi(\mathbb{1}_{A_i} \mathbb{1}_{B_j}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j \varphi(\mathbb{1}_{A_i})\varphi(\mathbb{1}_{B_j}) = \varphi(a)\varphi(b).$$

Now consider general $x, y \in \ell^\infty$. By splitting $x = x^+ - x^-$ and $y = y^+ - y^-$ ($x_n^\pm = \max\{\pm x_n, 0\} \geq 0$, similar for y) and using linearity of φ, we may assume without loss of generality that $x_n, y_n \geq 0$ for all n. Let $\varepsilon > 0$. Let a be a positive simple sequence such that $\|x - a\|_\infty \leq \varepsilon$ and $\|a\|_\infty \leq \|x\|_\infty$, and let b be a positive simple sequence such that $\|y - b\|_\infty \leq \varepsilon$ and $\|b\|_\infty \leq \|y\|_\infty$. Then for all n, we have

$$|x_n y_n - a_n b_n| = |x_n(y_n - b_n) + b_n(x_n - a_n)| \leq \varepsilon\|x\|_\infty + \varepsilon\|y\|_\infty,$$

so that $\|xy - ab\|_\infty \leq \varepsilon(\|x\|_\infty + \|y\|_\infty)$. Then

$$\begin{aligned}|\varphi(xy) - \varphi(x)\varphi(y)| &= |\varphi(xy) - \varphi(ab) + \varphi(ab) - \varphi(x)\varphi(y) + \varphi(a)\varphi(b) - \varphi(a)\varphi(b)| \\ &= |(\varphi(xy) - \varphi(ab)) + (\varphi(a)\varphi(b) - \varphi(x)\varphi(y))| \\ &\leq |\varphi(xy) - \varphi(ab)| + |\varphi(a)\varphi(b) - \varphi(x)\varphi(y)|.\end{aligned}$$

We can estimate the first term by

$$|\varphi(xy) - \varphi(ab)| = |\varphi(xy - ab)| \leq \|\varphi\|_{(\ell^\infty)'}\|xy - ab\|_\infty \leq \varepsilon(\|x\|_\infty + \|y\|_\infty)\|\varphi\|_{(\ell^\infty)'}.$$

The second term can be estimated by

$$\begin{aligned}|\varphi(a)\varphi(b) - \varphi(x)\varphi(y)| &= |\varphi(a)(\varphi(b) - \varphi(y)) + \varphi(y)(\varphi(a) - \varphi(x))| \leq |\varphi(a)| \cdot |\varphi(b - y)| + |\varphi(y)| \cdot |\varphi(a - x)| \\ &\leq \|\varphi\|_{(\ell^\infty)'}\|a\|_\infty\|\varphi\|_{(\ell^\infty)'}\|b - y\|_\infty + \|\varphi\|_{(\ell^\infty)'}\|y\|_\infty\|\varphi\|_{(\ell^\infty)'}\|x - a\|_\infty \\ &\leq \|\varphi\|^2_{(\ell^\infty)'}(\|x\|_\infty\|b - y\|_\infty + \|y\|_\infty\|a - x\|_\infty) \leq \varepsilon\|\varphi\|^2_{(\ell^\infty)'}(\|x\|_\infty + \|y\|_\infty).\end{aligned}$$

Altogether, we find

$$|\varphi(xy) - \varphi(x)\varphi(y)| \leq \varepsilon(\|\varphi\|_{(\ell^\infty)'} + \|\varphi\|^2_{(\ell^\infty)'})(\|x\|_\infty + \|y\|_\infty).$$

Letting $\varepsilon \downarrow 0$, the desired result follows. □

The usual identification of $(\ell^\infty)'$ with $\mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, i.e., the map $\varphi \to \mu_\varphi$, is an injection from the set of multiplicative functionals into the set of $0,1$-valued measures. To see that this maps indeed into the set of $0,1$-valued measures, suppose $\varphi \in (\ell^\infty)'$ is a multiplicative functional, then for all $E \subseteq \mathbb{N}$ we have

$$\mu_\varphi(E) = \varphi(\mathbb{1}_E) = \varphi(\mathbb{1}_E\mathbb{1}_E) = \varphi(\mathbb{1}_E)\varphi(\mathbb{1}_E) = \mu_\varphi(E)^2.$$

This can only hold if $\mu_\varphi(E) \in \{0, 1\}$. Finally, note that the value 1 is attained in $\mu_\varphi(\mathbb{N}) = 1$. The next proposition shows that the map $\varphi \to \mu_\varphi$ from the set of multiplicative functionals to the set of $0,1$-valued measures is also surjective.

Proposition 4.58. *Suppose $\mu \in \mathrm{ba}(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ is a $0,1$-valued measure. Then there exists a multiplicative functional $\varphi \in (\ell^\infty)'$ that satisfies $\mu_\varphi = \mu$. That is, the inducing functional of μ is a multiplicative functional.*

Proof. Let $\varphi \in (\ell^\infty)'$ be the inducing functional of μ. We check the defining properties of multiplicative functionals. Trivially φ is linear, and $\varphi(\mathbb{1}) = \mu(\mathbb{N}) = 1$. It remains to show that φ is positive and multiplicative on ℓ^∞. Let $x \in \ell^\infty$ be a sequence with $x_n \geq 0$ for all n. Let $a \in \ell^\infty$ be a simple sequence such that $0 \leq a_n \leq x_n$ for all n (lemma 1.4). Write $a = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}$ with $\alpha_i \geq 0$ and $(A_i)_{i=1}^n$ pairwise disjoint. Then

$$0 \overset{(*)}{\leq} \sum_{i=1}^n \alpha_i\mu(A_i) = \int_{\mathbb{N}} a\,d\mu \overset{(**)}{\leq} \int_{\mathbb{N}} x\,d\mu = \varphi(x).$$

In $(*)$ we used the fact that μ assumes only positive values and in $(**)$ we used the definition of the integral. This proves that φ is positive.

Now we show that φ is multiplicative on ℓ^∞. We use lemma 4.57. Let $E, F \subseteq \mathbb{N}$. From the fact that μ is $0,1$-valued, one can easily check that $\mu(E \cap F) = \mu(E)\mu(F)$. We have

$$\varphi(\mathbb{1}_E \mathbb{1}_F) = \varphi(\mathbb{1}_{E\cap F}) = \mu(E \cap F) = \mu(E)\mu(F) = \varphi(\mathbb{1}_E)\varphi(\mathbb{1}_F).$$

As E and F were arbitrary, lemma 4.57 shows that φ is multiplicative on ℓ^∞. We conclude that φ is a multiplicative functional. ☐

Combining proposition 4.58 and the discussion preceding it, we conclude that the map $\varphi \leftrightarrow \mu_\varphi$ is a bijective correspondence between multiplicative functionals and $0,1$-valued measures. Going in the reverse direction, this states that the map $\mu \leftrightarrow \int_{\mathbb{N}} \cdot \, d\mu$ is a bijective correspondence between $0,1$-valued measures and multiplicative functionals.

Earlier we have seen that the bijective correspondence $\mathcal{U} \leftrightarrow \mu_{\mathcal{U}}$ between ultrafilters on $\mathbb{N}$ and $0,1$-valued measures. By composing the maps $\mathcal{U} \to \mu_{\mathcal{U}}$ and $\mu \to \int_{\mathbb{N}} \cdot \, d\mu$, we obtain a bijective correspondence $\mathcal{U} \leftrightarrow \int_{\mathbb{N}} \cdot \, d\mu_{\mathcal{U}}$ between ultrafilters and multiplicative functionals. In fact, this is precisely the map $\mathcal{U} \to \mathcal{U}\text{-}\lim$. To see this, let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. We first show that $\mathcal{U}\text{-}\lim \in (\ell^\infty)'$. By the remark after definition 4.56, we need only show that it is linear, positive and normalized (as in definition 4.56). This is easy to see if $\mathcal{U}$ is principal, so we may assume it is free.

- *Linearity.* Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be sequences in ℓ^∞ and set $x = \mathcal{U}\text{-}\lim x_n$, $y = \mathcal{U}\text{-}\lim y_n$. Fix $\varepsilon > 0$, then we have

  $$\{n \in \mathbb{N} : |x_n - x| < \varepsilon/2\} \cap \{n \in \mathbb{N} : |y_n - y| < \varepsilon/2\} \subseteq \{n \in \mathbb{N} : |(x_n + y_n) - (x + y)| < \varepsilon\}.$$

  The first intersection is contained in $\mathcal{U}$, thus so is the right set. Thus $\mathcal{U}\text{-}\lim$ preserves addition. Now let $\alpha \in \mathbb{R}$ be a scalar. For $\alpha = 0$ we see that $\mathcal{U}\text{-}\lim \alpha x_n = \alpha x$, so suppose $\alpha \neq 0$. Then for $\varepsilon > 0$, we have

  $$\{n \in \mathbb{N} : |\alpha x_n - \alpha x| < \varepsilon\} = \{n \in \mathbb{N} : |x_n - x| < \frac{\varepsilon}{|\alpha|}\} \in \mathcal{U},$$

  which shows scalar homogeneity.
- *Positivity.* If $(x_n)_{n=1}^\infty$ has only positive terms, then every subsequential limit is positive, hence the result follows.
- $\mathcal{U}\text{-}\lim$ *is normalized.* The sequence $(1, 1, 1, ...)$ is convergent with limit 1, and the result follows.

By the remark after definition 4.56, we conclude that $\mathcal{U}\text{-}\lim \in (\ell^\infty)'$. To show that $\mathcal{U}\text{-}\lim = \int_{\mathbb{N}} \cdot \, d\mu_{\mathcal{U}}$, it suffices to prove that they agree on all indicator functions. Indeed, let $E \subseteq \mathbb{N}$, then we have that

$$\mathcal{U}\text{-}\lim \mathbb{1}_E(n) = \mu_{\mathcal{U}}(E) = \int_{\mathbb{N}} \mathbb{1}_E \, d\mu_{\mathcal{U}},$$

as desired.

We can summarize these identifications in the following way:

$$\mathcal{U} \leftrightarrow \mu_{\mathcal{U}} \leftrightarrow \int_{\mathbb{N}} \cdot \, d\mu_{\mathcal{U}} = \mathcal{U}\text{-}\lim.$$

If $\varphi \in (\ell^\infty)'$ is a multiplicative functional, then the ultrafilter $\mathcal{U}$ that induces φ via the above identification is $\mathcal{U} = \{E \subseteq \mathbb{N} : \mu_\varphi(E) = 1\} = \{E \subseteq \mathbb{N} : \varphi(\mathbb{1}_E) = 1\}$.

From the previous discussion and proposition 4.58 we derive the following.

Theorem 4.59. *Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. Then $\mathcal{U}\text{-}\lim : \ell^\infty \to \mathbb{R}$ is a multiplicative functional.*

Proof. Note that $\mathcal{U}\text{-}\lim = \int_{\mathbb{N}} \cdot \, d\mu_{\mathcal{U}}$ is the inducing functional of $\mu_{\mathcal{U}}$, and the latter is a $0,1$-valued measure. The result now follows from proposition 4.58. ☐

The functional $\mathcal{U}$- lim is not very interesting when $\mathcal{U}$ is principal, but when $\mathcal{U}$ is free it has very useful properties. In this case we have a multiplicative functional $\mathcal{U}$- lim that extends the usual limit. Using this functional $\mathcal{U}$- lim, we can build a Banach limit with some additional properties by taking the composition with the Cesàro operator, introduced in definition 4.9.

Theorem 4.60. *Let $\mathcal{U}$ be a free ultrafilter; for notational convenience write $\Phi := \mathcal{U}\text{-}\lim$. Then $\Phi \circ C$ is a Banach limit. Moreover, $\Phi \circ C$ is multiplicative on convergent sequences; that is, if $x, y \in \ell^\infty$ are convergent, then $(\Phi \circ C)(xy) = (\Phi \circ C)(x)(\Phi \circ C)(y)$.*

Proof. As Φ is a multiplicative functional (theorem 4.59), we have that $\|\Phi\|_{(\ell^\infty)'} = 1$ (see discussion after definition 4.56). Moreover, Φ extends limits (corollary 4.52). Hence by proposition 4.11, the functional $\Phi \circ C$ is a Banach limit. For the final assertion, note that whenever $x, y \in \ell^\infty$ are convergent, we have by lemma 4.10 that

$$C(xy) = \lim_{n\to\infty} x_n y_n = \lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n = C(x)C(y),$$

and as Φ is multiplicative the claim follows. This finishes the proof. $\square$

Remark. Note that the functional $\Phi \circ C$ cannot be multiplicative on ℓ^∞, as it would then be a multiplicative Banach limit, which does not exist.

As an application of theorem 4.60, one can construct a positive measure μ on $\mathbb{N}$ that extends density by taking a free ultrafilter $\mathcal{U}$ and defining

$$\mu(E) = \mathcal{U}\text{-}\lim \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_E(k) = \mathcal{U}\text{-}\lim \frac{|E \cap \{1, 2, ..., n\}|}{n}.$$

4.6 Stretchable, thinnable and elastic measures

In this section we examine in greater detail how 'nice' a measure on $\mathbb{N}$ can be, in the sense that it has intuitively desirable properties. We will restrict ourselves here to finitely additive measures that map into $[0, 1]$. The material presented in this section is based on [5].

Besides extending density, we would like measures to have some additional properties.

- Suppose we have a set $K \subseteq \mathbb{N}$. Consider the set $3K$; this is the set K 'stretched out' by a factor 3. Intuitively, $3K$ is 3 times less dense than K. Therefore we would like to have $\mu(3K) = \frac{1}{3}\mu(K)$.
- Somewhat similar to 'stretching out'; we can do something that can be described as 'thinning out'. Let $K \subseteq \mathbb{N}$ be a set and let F be the set obtained by taking every third element of K. For instance, if

  $$K = \{2, 3, 7, 9, 10, 12, 15, 17, 23, 27\},$$

  then F is given by $F = \{7, 12, 23\}$. As F is intuitively 3 times less dense than K, we would like to have $\mu(F) = \frac{1}{3}\mu(K)$.

We can consider 'stretching out' and 'thinning out' in a more general way. Suppose $f : \mathbb{N} \to \mathbb{N}$ is an injection[39] and suppose

$$\alpha := \lim_{n\to\infty} \frac{n}{f(n)}$$

exists. Then loosely speaking,[40] $f(n) \approx \beta n$ for large n, where we define $\beta = \frac{1}{\alpha}$.

- Now consider any subset $K \subseteq \mathbb{N}$, then $f[K] \approx \beta K$, i.e. $f[K]$ is approximately the set K 'stretched out' by a factor β. Hence we would like to have $\mu(f[K]) \approx \frac{1}{\beta}\mu(K)$, i.e. $\mu(f[K]) \approx \alpha\mu(K)$.

[39] We will later explain why we require that f be injective.

[40] We sacrifice some rigour in this reasoning, for the sake of intuition. For instance, to define β we assume that $\alpha \neq 0$, the notion '$\approx$' is not properly defined, and 'taking every β-th element of K' only makes sense if $\beta \in \mathbb{N}$.

- 'Thinning out' has an analogous interpretation. Suppose F is obtained from K by taking the $f(n)$-th element of K, for all $n \in \mathbb{N}$ for which an $f(n)$-th element of K exists. Then intuitively $F \approx F'$, where F is the set obtained from K by taking every β-th element. Thus we would like to have that $\mu(F) \approx \mu(F') = \frac{1}{\beta}\mu(K)$, i.e. $\mu(F) \approx \alpha\mu(K)$.

Observe that the first considerations are just a special case of the latter by taking $f(n) = 3n$.

Before giving a definition of 'stretchability' and 'thinnability', let us make precise the idea of 'taking the m-th element of a set'. Let $K \subseteq \mathbb{N}$ be a set. First suppose K is infinite. Define $c_K : \mathbb{N} \to K$ as the unique strictly increasing surjection from $\mathbb{N}$ onto K, the *counting function* of K. If we then for instance take $f : \mathbb{N} \to \mathbb{N}$ as $f(n) = 3n$, then $(c_K \circ f)[\mathbb{N}]$ contains precisely every third element of K. We use a similar construction for K finite: in that case, let $c_K : \{1, 2, ..., |K|\} \to K$ be the unique strictly increasing surjection.

Definition 4.61. *Let $f : \mathbb{N} \to \mathbb{N}$ be any function and let $K \subseteq \mathbb{N}$.*
- *Define $f \circ K = f[K]$.*
- *Define $K \circ f$ as follows.*
  - *If K is infinite, set $K \circ f = (c_K \circ f)[\mathbb{N}]$.*
  - *If K is finite, set $K \circ f = c_K[f[\mathbb{N}] \cap \{1, 2, ..., |K|\}]$.*

We use the notation $f \circ K$ instead of (the shorter notation) $f[K]$ to emphasize that $f \circ K$ and $K \circ f$ are similar concepts. Note that in fact $f \circ K = (f \circ c_K)[\mathbb{N}]$ or $(f \circ c_K)[\{1, 2, ..., |K|\}]$ (depending on whether K is finite or not).

The notions $f \circ K$ and $K \circ f$ capture the idea of 'stretching out' and 'thinning out', respectively. If we take for instance $f(n) = 3n$ in definition 4.61, then $f \circ K$ corresponds to $3K$, and $K \circ f$ corresponds to the set F that is obtained by taking every third element of K. Definition 4.61 allows more f than just ones that are linear; the idea is that we will consider functions f that are 'almost linear'.

Remark. Observe that $f \circ \mathbb{N} = \mathbb{N} \circ f$ for any f: this follows from the fact that $c_{\mathbb{N}} = \mathrm{id}_{\mathbb{N}}$. It is also intuitively clear if we take for example $f(n) = 3n$: stretching out $\mathbb{N}$ by a factor 3 is the same as taking every third element.

Define $\mathbb{L}$ as the set of all injections $f : \mathbb{N} \to \mathbb{N}$ for which the following limit exists:

$$\lambda(f) := \lim_{n \to \infty} \frac{n}{f(n)}. \tag{18}$$

Thus $\mathbb{L}$ is the set of injections that are 'eventually linear': loosely speaking, $f(n) \approx \frac{1}{\lambda(f)}n$ for large n.[41]

Definition 4.62 (Stretchability, thinnability)**.** *Let $\mu : \mathcal{P}(\mathbb{N}) \to [0, 1]$ be a finitely additive measure.*
*(1) μ is called **stretchable** if $\mu(f \circ K) = \lambda(f)\mu(K)$ for all $f \in \mathbb{L}$ and all $K \subseteq \mathbb{N}$.*
*(2) μ is called **thinnable** if $\mu(K \circ f) = \lambda(f)\mu(K)$ for all $f \in \mathbb{L}$ and all $K \subseteq \mathbb{N}$.*

Remark. There exists a thinnable measure on $\mathbb{N}$ that is not stretchable (see [5, thm. 1.5]). It is not known whether every stretchable measure is thinnable ([5, p. 227]).

We will show that $\mathbb{N}$ has a measure that is both stretchable and thinnable. Such a measure automatically extends density; we will see this in corollary 4.64.

The next proposition is in accordance with our intuition. Statement (1) shows that if K is obtained by stretching out $\mathbb{N}$ by a factor $\frac{1}{\lambda(f)}$ (that is to say, $K := f[\mathbb{N}]$ for some $f \in \mathbb{L}$; recall $f(n) \approx \frac{1}{\lambda(f)}n$), then $d(K)$ exists and $d(K) = \lambda(f) \cdot d(\mathbb{N}) = \lambda(f)$. Note this set K is the same as the set K' obtained by thinning out $\mathbb{N}$ by a factor $\frac{1}{\lambda(f)}$, by which we mean that $K' := \mathbb{N} \circ f$.

[41] As said, this is not very precise, and certainly does not make sense if $\lambda(f) = 0$. If $\lambda(f) = 0$, then the interpretation 'almost linear' does not work very well, as for instance the function $f(n) = n^2$ is by no means close to being linear.

Proposition 4.63. *Let $f : \mathbb{N} \to \mathbb{N}$ be an injection. Define $K = f \circ \mathbb{N} = \mathbb{N} \circ f$.*
(1) If $f \in \mathbb{L}$, then $d(K) = \lambda(f)$.
(2) If f is strictly increasing and $d(K)$ exists, then $f \in \mathbb{L}$ and $\lambda(f) = d(K)$.

Proof. We first prove (1). Figure 4 illustrates the intuition of the proof.

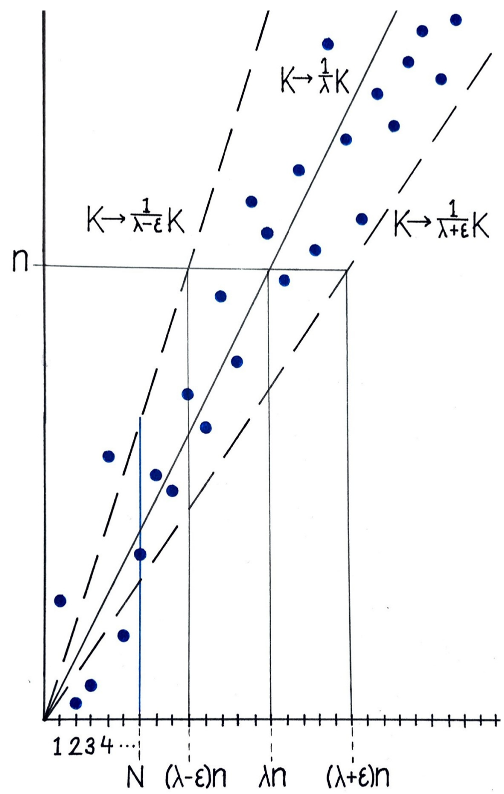

Figure 4: Proof of case 1 of theorem 4.63.

Define $\lambda := \lambda(f) = \lim_{n\to\infty} \frac{n}{f(n)}$.

Case 1: $\lambda > 0$. First assume that $\lambda > 0$. Let $\varepsilon \in (0, \lambda)$. Due to injectivity of f, we have for all $n \in \mathbb{N}$

$$|\{K \cap \{1, 2, ..., n\}| = |\{k \in \mathbb{N} : f(k) \leq n\}|.$$

Let $N \in \mathbb{N}$ be such that for all $k \geq N$, we have $|\frac{k}{f(k)} - \lambda| < \varepsilon$. Thus for all $k \geq N$, we have

$$\frac{k}{\lambda + \varepsilon} < f(k) < \frac{k}{\lambda - \varepsilon}.$$

Now let

$$n \geq \max\Big\{N, \frac{N}{\lambda + \varepsilon}, \max_{1 \leq k \leq N} f(k)\Big\}$$

be arbitrary.[42] We claim that

$$\{1, 2, ..., \lfloor(\lambda - \varepsilon)n\rfloor\} \overset{(i)}{\subseteq} \{k \in \mathbb{N} : f(k) \leq n\} \overset{(ii)}{\subseteq} \{1, 2, ..., \lceil(\lambda + \varepsilon)n\rceil\}. \tag{19}$$

[42] In fact $n \geq \max\{f(k) : 1 \leq k \leq N\}$ already enforces $n \geq N$ due to injectivity. For the interpretation in the picture this is added.

(i) Let $1 \le k \le \lfloor(\lambda-\varepsilon)n\rfloor$, so that $k \le (\lambda-\varepsilon)n$. If $k \le N$ then $f(k) \le n$, as desired, and if $k > N$ then

$$f(k) < \frac{k}{\lambda-\varepsilon} \le \frac{(\lambda-\varepsilon)n}{\lambda-\varepsilon} = n,$$

as desired.

(ii) Let $k \in \mathbb{N}$ with $f(k) \le n$. Suppose to the contrary that $k > (\lambda+\varepsilon)n$, then $k \ge N$ thus

$$f(k) > \frac{k}{\lambda+\varepsilon} > \frac{(\lambda+\varepsilon)n}{\lambda+\varepsilon} = n,$$

which is a contradiction. Hence $k \le \lceil(\lambda+\varepsilon)n\rceil$, as desired.

From equation (19) we conclude that for all n large enough,

$$\frac{\lfloor(\lambda-\varepsilon)n\rfloor}{n} \le \frac{|\{k \in \mathbb{N} : f(k) \le n\}|}{n} \le \frac{\lceil(\lambda+\varepsilon)n\rceil}{n}.$$

This implies that for all n large enough, we have

$$\left|\frac{|\{k \in \mathbb{N} : f(k) \le n\}|}{n} - \lambda\right| \le 2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this proves that $d(K) = \lambda$.

Case 2: $\lambda = 0$. Now suppose $\lambda = 0$; this proof is similar to the previous. Let $\varepsilon > 0$ and let $N \in \mathbb{N}$ be such that for all $k \ge N$, we have $|\frac{k}{f(k)}| < \varepsilon$. Then we have that $\frac{k}{\varepsilon} < f(k)$ for all $k \ge N$. Now let $n \ge \frac{N}{\varepsilon}$ be arbitrary. We claim that

$$\{k \in \mathbb{N} : f(k) \le n\} \subseteq \{1, 2, ..., \lceil\varepsilon n\rceil\}. \tag{20}$$

Indeed, let $k \in \mathbb{N}$ with $f(k) \le n$ and suppose to the contrary that $k > \varepsilon n$. Then $k > N$ hence

$$f(k) > \frac{k}{\varepsilon} > \frac{\varepsilon n}{\varepsilon} = n,$$

which is a contradiction. This proves (20). Then finally, from (20) we see that for all n large enough, we have

$$\frac{|\{k \in \mathbb{N} : f(k) \le n\}|}{n} \le \frac{\lceil\varepsilon n\rceil}{n} \le 2\varepsilon.$$

As $\varepsilon > 0$ was arbitrary, this shows that $d(K) = 0$, as desired.

Finally, we prove (2). This is very straightforward:

$$\begin{aligned} d(K) &= \lim_{n\to\infty} \frac{|K \cap \{1,2,...,n\}|}{n} \overset{(i)}{=} \lim_{n\to\infty} \frac{|\{k \in \mathbb{N} : f(k) \le n\}|}{n} \overset{(ii)}{=} \lim_{n\to\infty} \frac{|\{k \in \mathbb{N} : f(k) \le f(n)\}|}{f(n)} \\ &\overset{(iii)}{=} \lim_{n\to\infty} \frac{|\{1,2,...,n\}|}{f(n)} = \lambda. \end{aligned}$$

In (*i*) we used an observation from case 1; in (*ii*) we pass to a subsequence and (*iii*) follows from the fact that f is strictly increasing. This finishes the proof. □

A stretchable or thinnable measure automatically extends density.

Corollary 4.64. *Suppose* $\mu : \mathcal{P}(\mathbb{N}) \to [0,1]$ *is stretchable or thinnable. Then* μ *extends density.*

Proof. Let $E \subseteq \mathbb{N}$ be such that $d(E)$ exists. First suppose that E is infinite and let $c_E : \mathbb{N} \to \mathbb{N}$ be the (strictly increasing) counting function of E. Then by proposition 4.63, we have that $c_E \in \mathbb{L}$ and $\lambda(c_E) = d(E)$. As μ is stretchable or thinnable, we obtain

$$\mu(E) = \mu(c_E \circ \mathbb{N}) = \mu(\mathbb{N} \circ c_E) = \lambda(c_E)\mu(\mathbb{N}) = d(E),$$

as desired. Now suppose E is finite and take an infinite set A with $d(A) = 0$ such that $E \cap A = \varnothing$. For instance one can take $A = \{2^n : n \geq N\}$ for N large enough. Then using the previous in $(*)$, we have

$$\mu(E) = \mu(E) + d(A) \overset{(*)}{=} \mu(E) + \mu(A) = \mu(E \cup A) \overset{(*)}{=} d(E \cup A) = d(A) = 0 = d(E),$$

as desired. This finishes the proof. □

Remark. We can now see why it is important that we consider only injections in the definition of stretchability and thinnability. Suppose we require that the conditions in definition 4.62 hold for *all* functions $f : \mathbb{N} \to \mathbb{N}$ for which $\lim_{n\to\infty} \frac{n}{f(n)}$ exists. We claim that a stretchable or thinnable measure (under this new definition) does not exist. Consider $f(2n-1) = f(2n) = 2n$ for $n \in \mathbb{N}$, then $\lambda := \lim_{n\to\infty} \frac{n}{f(n)} = 1$. Suppose μ is a stretchable or thinnable measure in the new sense, then we have that

$$\mu(f \circ \mathbb{N}) = \mu(\mathbb{N} \circ f) = \lambda\mu(\mathbb{N}) = 1.$$

We have $f \circ \mathbb{N} = \mathbb{N} \circ f = 2\mathbb{N}$. As μ is also stretchable or thinnable in the sense of definition 4.62, μ must extend density by corollary 4.64, so we must have $\mu(2\mathbb{N}) = \frac{1}{2}$, a contradiction.

We introduce one more type of measure that we consider 'nice'. For functions $f, g : \mathbb{N} \to \mathbb{N}$, define an ordering by setting $f \leq g$ if and only if $f(n) \leq g(n)$ for all $n \in \mathbb{N}$.

Definition 4.65. *A finitely additive measure $\mu : \mathcal{P}(\mathbb{N}) \to [0,1]$ is called* ***elastic*** *if the following two conditions are satisfied:*

(1) $\mu(mK) = m^{-1}\mu(K)$ for all $m \in \mathbb{N}$.

(2) For all $f, g : \mathbb{N} \to \mathbb{N}$, if $f \leq g$ then $\mu(f \circ \mathbb{N}) \geq \mu(g \circ \mathbb{N})$.

Note that a stretchable measure obviously satisfies (1), just by taking $f(n) = mn$ in the definition of stretchability. Property (2) is desirable in the following sense. Intuitively, the set $3\mathbb{N}$ is less dense than the set $2\mathbb{N}$, so we would like to have $\mu(3\mathbb{N}) \leq \mu(2\mathbb{N})$. (This is the case if μ extends density.) Note that this is really just taking $f(n) = 2n$ and $g(n) = 3n$ in (2). Property (2) requires such a condition on μ in more generality, namely for *all* functions f and g.

We have the following very useful result, which we will not prove here. A proof can be found in [5].

Theorem 4.66. *An elastic measure is both stretchable and thinnable.*

In theorem 4.68 we construct an elastic measure, thereby showing (in particular[43]) the existence of a stretchable and thinnable measure. We use the following lemma, which describes how one can construct a measure satisfying (2) in definition 4.65.

Lemma 4.67. *Let Φ be a Banach limit. Suppose $\alpha : \mathbb{N} \times \mathbb{N} \to [0, \infty)$ is a function satisfying the following conditions:*

(1) $\sum_{j=1}^{\infty} \alpha(i,j) = 1$ for all $i \in \mathbb{N}$.

(2) $\alpha(i,j) \geq \alpha(i,j+1)$ for all $i, j \in \mathbb{N}$.

Define the finitely additive measure $\mu : \mathcal{P}(\mathbb{N}) \to [0,1]$ by

$$\mu(K) = \Phi\left(\sum_{j\in K} \alpha(1,j), \sum_{j\in K} \alpha(2,j), \sum_{j\in K} \alpha(3,j), \ldots\right).$$

Suppose $f, g : \mathbb{N} \to \mathbb{N}$ are functions with $f \leq g$. Then we have that $\mu(f \circ \mathbb{N}) \geq \mu(g \circ \mathbb{N})$.

[43] It is not known whether there exists a nonelastic, stretchable and thinnable measure [5].

Proof. Note that μ is well-defined (i.e. maps into $[0,1]$) as Φ preserves positivity, and clearly it is a finitely additive measure. To prove the claim, it suffices to show that for all $i \in \mathbb{N}$, we have

$$\sum_{j \in g[\mathbb{N}]} \alpha(i,j) \leq \sum_{j \in f[\mathbb{N}]} \alpha(i,j).$$

Fix $i \in \mathbb{N}$. Define $\tilde{f}, \tilde{g} : \mathbb{N} \to \mathbb{N} \cup \{0\}$ by

$$\begin{aligned} \tilde{f}(n) &= \begin{cases} f(n) & \text{if } f(n) \notin \{f(1), f(2), ..., f(n-1)\}, \\ 0 & \text{otherwise.} \end{cases} \\ \tilde{g}(n) &= \begin{cases} g(n) & \text{if } g(n) \notin \{g(1), g(2), ..., g(n-1)\}, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{21}$$

Moreover, define $\alpha(i,0) = 0$. We see that the sequence $(\tilde{f}(n))_{n=1}^{\infty}$ contains every $f(m)$ exactly once, and a term $\alpha(i, \tilde{f}(n))$ is zero whenever $f(n)$ has already appeared in $\{f(1), f(2), ..., f(n-1)\}$. Thus we have that

$$\sum_{j \in f[\mathbb{N}]} \alpha(i,j) = \sum_{n=1}^{\infty} \alpha(i, \tilde{f}(n)). \tag{22}$$

The same claims and the same equality (22) hold with f and $\tilde{f}$ replaced by g and $\tilde{g}$, respectively. Thus the result we have to show amounts to the following:

$$\sum_{n=1}^{\infty} \alpha(i, \tilde{g}(n)) \leq \sum_{n=1}^{\infty} \alpha(i, \tilde{f}(n)).$$

This follows by comparing the terms. Let $n \in \mathbb{N}$. If $\tilde{g}(n) = 0$, then $\alpha(i, \tilde{g}(n)) = 0 \leq \alpha(i, \tilde{f}(n))$. Now suppose $\tilde{g}(n) \neq 0$, then we have

$$\tilde{f}(n) \leq f(n) \leq g(n) = \tilde{g}(n).$$

By assumption (2), we have

$$\alpha(i, \tilde{f}(n)) \geq \alpha(i, \tilde{f}(n)+1) \geq \alpha(i, \tilde{f}(n)+2) \geq ... \geq \alpha(i, \tilde{g}(n)).$$

This finishes the proof. □

Theorem 4.68. *There exists an elastic measure on* $\mathbb{N}$.

Proof. Let ζ be the Riemann zeta function and define $\alpha : \mathbb{N} \times \mathbb{N} \to [0, \infty)$ by

$$\alpha(i,j) = \frac{j^{-(1+1/i)}}{\zeta(1+1/i)}.$$

One can easily verify that α satisfies the conditions in lemma 4.67. Let μ be defined as in lemma 4.67. We need only show that μ has the desired scaling property. Let $m \in \mathbb{N}$ and $K \subseteq \mathbb{N}$. It is useful to observe first that for all $j \in \mathbb{N}$, we have $(*)$: $\lim_{i \to \infty} \alpha(i,j) = 0$. This follows from the fact that

$$\lim_{i \to \infty} \zeta(1+1/i) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

using the monotone convergence theorem.

To prove that $\mu(mK) = m^{-1}\mu(K)$, it suffices to show the following (as a Banach limit sends convergent sequences to their limit):

$$\lim_{i\to\infty}\left(\sum_{j\in mK}\alpha(i,j)-\frac{1}{m}\sum_{j\in K}\alpha(i,j)\right)=0.$$

For all $i\in\mathbb{N}$, we have that

$$\sum_{j\in mK}\alpha(i,j)=\sum_{j\in K}\alpha(i,mj)=m^{-(1+1/i)}\sum_{j\in K}\alpha(i,j).$$

Thus we have that

$$\sum_{j\in mK}\alpha(i,j)-\frac{1}{m}\sum_{j\in K}\alpha(i,j)=\left(\frac{1}{m^{(1+1/i)}}-\frac{1}{m}\right)\left(\sum_{j\in K}\alpha(i,j)\right).$$

The left factor tends to 0 as $i\to\infty$. This is also the case for the right factor; this follows from the dominated convergence theorem and $(*)$. This finishes the proof. ☐

A pathological measure

So far we have constructed some well-behaved measures. We can also do the opposite: find out what pathological properties a measure can have. Here we mention an example.

Let us define the following.

- For $K\subseteq\mathbb{N}$ and $r\in[1,\infty)$, define $r+K=\{r+k:k\in K\}$ and $rK=\{rk:k\in K\}$.
- For $D\subseteq[1,\infty)$, define $\lfloor D\rfloor=\{\lfloor d\rfloor:d\in D\}$.
- Let $A\subseteq[1,\infty)$ be a subset and μ a finitely additive measure. Then μ is called **A-stretchable** if $\mu(\lfloor b+aK\rfloor)=a^{-1}\mu(K)$ for each $a\in A$, $b\in[0,\infty)$ and $K\subseteq\mathbb{N}$.

Note that a stretchable measure is A-stretchable for any $A\subseteq[1,\infty)$ as we can take $f(n)=\lfloor b+an\rfloor$ in the definition of stretchability. The converse need not be true, in the sense that not every not every $[1,\infty)$-stretchable measure is stretchable [5, thm.1.4].

Theorem 4.69 (A pathological measure)**.** *There exists a $\mathbb{Q}_{\geq 1}$-stretchable measure μ on $\mathbb{N}$ and a set $K\subseteq\mathbb{N}$ such that $\mu(K)=1$ but $\mu(\lfloor rK\rfloor)=0$ for each irrational number $r\in[1,\infty)$.*

Proof. A construction of this measure can be found in [5, thm. 1.7]. ☐

4.7 Extension of finitely additive measures

In this section we prove that a finitely additive measure defined on a suitable domain $\mathcal{R}\subseteq\mathcal{P}(S)$ can be extended to a finitely additive measure on the whole power set $\mathcal{P}(S)$. This result can be used to prove the existence of unbounded finitely additive measures. Since the results in this section hold very generally, we do not restrict ourselves to $\mathbb{N}$ here.

In order to state the results we need to consider finitely additive measures defined on more general domains than σ-algebras.

Definition 4.70 (Ring)**.** *Let S be a set. A collection of sets $\mathcal{R}\subseteq\mathcal{P}(S)$ is called a **ring** if the following conditions are satisfied:*

(1) $\varnothing\in\mathcal{R}$;
(2) $A,B\in\mathcal{R}\implies B\setminus A\in\mathcal{R}$;
(3) $A\cup B\in\mathcal{R}\implies A\cup B\in\mathcal{R}$.

Definition 4.71 (Finitely additive measure)**.** *Let S be a set and $\mathcal{R}\subseteq\mathcal{P}(S)$ be a ring. A function $\mu:\mathcal{R}\to\mathbb{R}$ is called a **finitely additive measure** if $\mu(\varnothing)=0$ and $\mu(A\cup B)=\mu(A)+\mu(B)$ for all disjoint $A,B\in\mathcal{R}$.*

The following lemma will be useful for constructing extensions of measures.

Lemma 4.72. *Let V be a real vector space and let $W \subseteq V$ be a linear subspace. Suppose $\varphi : W \to \mathbb{R}$ is a linear map. Then φ has a linear extension $\Phi : V \to \mathbb{R}$.*

Proof. Let us denote the domain of a function f by D_f. Let $\mathcal{M}$ be the collection of all linear functions $f : D_f \to \mathbb{R}$ with the following properties:

- f is defined on a linear subspace $W \subseteq D_f \subseteq V$;
- $f(w) = \varphi(w)$ for all $w \in W$.

Observe that $\mathcal{M} \neq \varnothing$ as $\varphi \in \mathcal{M}$. Now define a partial order on $\mathcal{M}$ by setting $f \preceq g$ if and only if $D_f \subseteq D_g$ and $f(x) = g(x)$ for all $x \in D_f$. We check the conditions in Zorn's lemma. Let $\mathcal{N} \subseteq \mathcal{M}$ be a nonempty chain. We construct an upper bound $f_{\mathcal{N}} \in \mathcal{M}$ for $\mathcal{N}$. Define

$$A = \bigcup_{f \in \mathcal{N}} D_f.$$

Then A is a linear subspace of V: this follows from the fact that $\mathcal{N}$ is a chain. Define $f_{\mathcal{N}} : A \to \mathbb{R}$ as follows. Let $y \in A$, then there exists $f \in \mathcal{N}$ such that $y \in D_f$. Define $f_{\mathcal{N}}(y) := f(y)$. From the fact that $\mathcal{N}$ is a chain it follows that this is well-defined (i.e. does not depend on the choice of f). Finally we check that $f_{\mathcal{N}} \in \mathcal{M}$: linearity of $f_{\mathcal{N}}$ follows again from the fact that $\mathcal{N}$ is a chain, we have that $W \subseteq A \subseteq V$, and $f_{\mathcal{N}}$ extends φ.

We conclude that $f_{\mathcal{N}} \in \mathcal{M}$, and by construction it is an upper bound for $\mathcal{N}$. Thus by Zorn's lemma, there exists a maximal element $\Psi \in \mathcal{M}$. We claim that $D_\Psi = V$. Suppose not, and fix some $x \in V \setminus D_\Psi$. Define $\Psi : \text{span}\{x\} \oplus D_\Psi \to \mathbb{R}$ by $\Psi(\alpha x + y) = \Phi(y)$. Then we have that $\Psi \in \mathcal{M}$ and $\Phi \preceq \Psi$, but we do not have $\Phi = \Psi$. This contradicts maximality of Φ. Hence we conclude that $\Phi : V \to \mathbb{R}$ is the desired linear extension. ☐

Theorem 4.73 (Extension of finitely additive measures)**.** *Let $\mathcal{R} \subseteq \mathcal{P}(S)$ be a ring and let $\mu : \mathcal{R} \to \mathbb{R}$ be a finitely additive measure. Then there exists a finitely additive measure $\overline{\mu} : \mathcal{P}(S) \to \mathbb{R}$ that extends μ.*

Proof. Define

$$\begin{aligned} V_{\mathcal{R}} =& \{f : S \to \mathbb{R} : \quad f = \sum_{i=1}^m a_i \mathbb{1}_{A_i},\ a_i \in \mathbb{R},\ A_i \in \mathcal{R}\}, \\ V_{\mathcal{P}(S)} =& \{g : S \to \mathbb{R} : \quad g = \sum_{j=1}^n b_j \mathbb{1}_{B_j},\ b_j \in \mathbb{R},\ B_j \in \mathcal{P}(S)\}. \end{aligned}$$

These are real vector spaces and $V_{\mathcal{R}}$ is a subspace of $V_{\mathcal{P}(S)}$. From the properties of a ring it follows that every $f \in V_{\mathcal{R}}$ can be written as

$$f = \sum_{i=1}^m a_i \mathbb{1}_{A_i}$$

for real scalars $a_1, ..., a_m$ and *pairwise disjoint* sets $A_1, A_2, ..., A_m \in \mathcal{R}$. Now define

$$\varphi : V_{\mathcal{R}} \to \mathbb{R} \quad \varphi(f) = \varphi\left(\sum_{i=1}^m a_i \mathbb{1}_{A_i}\right) = \sum_{i=1}^m a_i \mu(A_i), \tag{23}$$

where f has a representation with disjoint $A_1, A_2, ..., A_m \in \mathcal{R}$. Then φ is well-defined: if

$$f = \sum_{i=1}^m a_i \mathbb{1}_{A_i} = \sum_{j=1}^n b_j \mathbb{1}_{B_j}$$

with $A_1, A_2, ..., A_m \in \mathcal{R}$ pairwise disjoint and $B_1, B_2 ..., B_n \in \mathcal{R}$ pairwise disjoint, then by using a common refinement $C_{ij} = A_i \cap B_j$ one checks that the value of (23) does not depend on the representation. In a similar manner one checks that φ is linear.

Therefore by lemma 4.72, there exists a linear extension $\Phi : V_{\mathcal{P}(S)} \to \mathbb{R}$ of φ. Now define

$$\overline{\mu} : \mathcal{P}(S) \to \mathbb{R} \quad \overline{\mu}(A) = \Phi(\mathbb{1}_A).$$

This map has the desired properties. ☐

There are much more useful results regarding extensions of measures than proposition 4.73. For instance, Carathéodory's extension theorem states that a σ-additive premeasure defined on a ring $\mathcal{R}$ can be extended to a σ-*additive* measure on $\sigma(\mathcal{R})$ (and possibly a slightly bigger collection). But often a σ-additive extension to the whole power set is not possible.

Proposition 4.74. *Let S be a set. Then there exists a finitely additive measure $\mu : \mathcal{P}(S) \to \mathbb{R}$ such that for every finite set $A \subseteq S$, we have $\mu(A) = |A|$.*

Proof. If S is finite then we can just define $\mu(A) = |A|$ and this has the desired properties. So we may assume that S is infinite. Let $\mathcal{R} = \{A \subseteq S : A \text{ or } S \setminus A \text{ is finite}\}$. It can easily be seen that $\mathcal{R}$ is a ring. Now define

$$\mu : \mathcal{R} \to \mathbb{R} \quad \mu(A) = \begin{cases} |A| & \text{if } A \text{ finite,} \\ -|S \setminus A| & \text{if } S \setminus A \text{ finite.} \end{cases}$$

It is not hard to check that μ is a finitely additive measure on $\mathcal{R}$. Now apply theorem 4.73 to find an extension $\overline{\mu} : \mathcal{P}(S) \to \mathbb{R}$. This extension has the desired properties. ☐

Proposition 4.75 (Unbounded finitely additive measures)**.** *Let S be an infinite set*[44] *and let $\mathcal{A}$ be a σ-algebra on S such that $\mathcal{A}$ contains a sequence of finite sets $(A_n)_{n=1}^{\infty}$ for which $\sup_{n \in \mathbb{N}} |A_n| = \infty$. Then there exists an unbounded finitely additive measure $\mu : \mathcal{A} \to \mathbb{R}$.*

Proof. By proposition 4.74, we may choose a finitely additive measure $\mu : \mathcal{P}(S) \to \mathbb{R}$ satisfying $\mu(A) = |A|$ for every finite set $A \subseteq S$. Then the restriction $\mu|_{\mathcal{A}} : \mathcal{A} \to \mathbb{R}$ is a finitely additive measure on $\mathcal{A}$ with the desired properties. ☐

[44] Note that from the condition on the sequence $(A_n)_{n=1}^{\infty}$ it already follows that S must be infinite. Also note that if S is finite then any finitely additive measure on $(S, \mathcal{A})$ is bounded.

Appendix A Proofs of theorems from preliminaries

In this appendix we give the proofs of the results stated in chapter 1.

Lemma A.1 (Density of simple sequences)**.** *The set of simple sequences is a dense subspace of ℓ^∞. In particular, if $x \in \ell^\infty$ satisfies $x_n \geq 0$ for all n, then for every $\varepsilon > 0$, there exists a simple sequence a such that $\|x-a\|_\infty < \varepsilon$ and moreover $0 \leq a_n \leq x_n$ for all n.*

Proof. Clearly this is a subspace; we show density. Let $x \in \ell^\infty$ and let $\varepsilon > 0$. First suppose $x_n \geq 0$ for all n. Let $N \in \mathbb{N}$ be such that $N\varepsilon > \|x\|_\infty$. Note that $[0, N\varepsilon) = \bigcup_{j=1}^N [(j-1)\varepsilon, j\varepsilon)$. For all $1 \leq j \leq N$, let $A_j = x^{-1}([(j-1)\varepsilon, j\varepsilon))$; then note that $(A_j)_{j=1}^N$ is pairwise disjoint. Define $\alpha_j = (j-1)\varepsilon$ for $1 \leq j \leq N$. Finally, let $a \in \ell^\infty$ be given by

$$a = \sum_{j=1}^N \alpha_j \mathbb{1}_{A_j}.$$

Then a is a simple sequence. We check that $\|x-a\|_\infty \leq \varepsilon$. Let $n \in \mathbb{N}$, then

$$x_n \in [0, \|x\|_\infty] \subseteq [0, N\varepsilon) = \bigcup_{j=1}^N [(j-1)\varepsilon, j\varepsilon).$$

Thus $x_n \in [(j-1)\varepsilon, j\varepsilon)$ for some (unique) $1 \leq j \leq N$. Then $n \in A_j$ and $a_n = \alpha_j = (j-1)\varepsilon$. Thus we see that $|x_n - a_n| = |x_n - (j-1)\varepsilon| \leq \varepsilon$. Taking the supremum over n, we find $\|x-a\|_\infty \leq \varepsilon$. Moreover, we have that $0 \leq a_n \leq x_n$ for all n. This proves that positive sequences can be approximated from below by positive simple sequences.

Now consider general $x \in \ell^\infty$ and write $x = x^+ - x^-$, where $x^\pm \in \ell^\infty$ are given by $x_n^\pm = \max\{\pm x_n, 0\} \geq 0$. Let a and b be (positive) simple sequences such that $\|x^+ - a\|_\infty < \frac{\varepsilon}{2}$ and $\|x^- - b\|_\infty < \frac{\varepsilon}{2}$. Then $a-b$ is a simple sequence, and

$$\|x-(a-b)\|_\infty = \left\|(x^+ - a) + (b - x^-)\right\|_\infty \leq \left\|x^+ - a\right\|_\infty + \left\|x^- - b\right\|_\infty < \varepsilon.$$

This finishes the proof. □

Proposition A.2. *Let $\varphi, \psi : \ell^\infty \to \mathbb{R}$ be bounded linear functionals. If $\varphi(\mathbb{1}_E) = \psi(\mathbb{1}_E)$ for all $E \subseteq \mathbb{N}$, then $\varphi = \psi$.*

Proof. First note that by linearity, φ and ψ agree on all simple sequences. The claim now follows by a general result that two continuous functions that agree on a dense set must agree on the whole space. Indeed, let $x \in \ell^\infty$ and choose a sequence of simple functions $(a^{(n)})_{n=1}^\infty$ such that $a^{(n)} \to x$ in ℓ^∞, then by continuity we find

$$\varphi(x) = \lim_{n\to\infty} \varphi(a^{(n)}) = \lim_{n\to\infty} \psi(a^{(n)}) = \psi(x),$$

which proves the claim. □

Theorem A.3 (Duals are Banach spaces)**.** *Let X be a real normed vector space. Then X' is a Banach space.*

Proof. Let $(f_n)_{n=1}^\infty$ be a Cauchy sequence in X'. We show that this sequence has a limit $f \in X'$.

Step 1. We construct a potential limit f. For all $x \in X$, we have that $(f_n(x))_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{R}$; indeed, note that $|f_n(x) - f_m(x)| = |(f_n - f_m)(x)| \leq \|f_n - f_m\|_{X'} \|x\|_X \to 0$ as $m, n \to \infty$. As $\mathbb{R}$ is complete, this sequence converges and we may define $f : X \to \mathbb{R}$ by $f(x) = \lim_{n\to\infty} f_n(x)$.

Step 2. We show that $f \in X'$. For linearity, observe that

$$f(\alpha x + \beta y) = \lim_{n\to\infty} f_n(\alpha x + \beta y) = \lim_{n\to\infty} \alpha f_n(x) + \beta f_n(y) = \alpha f(x) + \beta f(y).$$

Next we show boundedness. As $(f_n)_{n=1}^{\infty}$ is Cauchy, it is bounded hence there exists $M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$, we have $\|f_n\|_{X'} \le M$. Then we obtain for all $x \in X$,

$$|f(x)| = |\lim_{n\to\infty} f_n(x)| = \lim_{n\to\infty} |f_n(x)| \le M\|x\|_X.$$

This proves boundedness, thus we have $f \in X'$.

Step 3. We show that $f_n \to f$, i.e. $\|f_n - f\|_{X'} \to 0$. Let $\varepsilon > 0$ and let $N_0 \in \mathbb{N}$ be such that for all $n, m \ge N_0$, we have $\|f_n - f_m\|_{X'} < \frac{\varepsilon}{2}$. We claim that for all $n \ge N_0$, it holds that $\|f - f_n\|_{X'} \le \varepsilon$. Indeed let $n \ge N_0$ and $x \in X$ with $\|x\|_X \le 1$, then we see that for all $m \in \mathbb{N}$,

$$|f(x)-f_n(x)| \le |f(x)-f_m(x)|+|f_m(x)-f_n(x)| \le |f(x)-f_m(x)|+|(f_m-f_n)(x)| \le |f(x)-f_m(x)|+\|f_m - f_n\|_{X'}.$$

There exists $N_1 \in \mathbb{N}$ such that for all $m \ge N_1$, we have $|f(x) - f_m(x)| < \frac{\varepsilon}{2}$. Fix some $m \ge \max\{N_0, N_1\}$, then $(n \ge N_0)$ we see that

$$|f(x) - f_n(x)| \le \varepsilon$$

Taking the supremum over all $x \in X$ with $\|x\|_X \le 1$, we arrive at $\|f - f_n\|_{X'} \le \varepsilon$. This shows that $f_n \to f$ in X', which finishes the proof. ☐

In order to prove the Hahn-Banach theorem (theorem 1.9), we first prove the existence of a 'one-dimensional' extension of f_W to a slightly bigger domain. We use the following notation. If X is a vector space and $V, W \subseteq X$ are subspaces such that $V \cap W = \{0\}$, then every element $x \in V + W$ can be written uniquely as $x = v + w$ for $v \in V, w \in W$, and we define $V \oplus W := V + W = \{v + w : v \in V, w \in W\}$.

Lemma A.4 (One-dimensional extension)**.** *Let X be a real vector space and let $W \subseteq X$ be a proper subspace. Let $p : X \to \mathbb{R}$ be a sublinear functional and suppose $f_W : W \to \mathbb{R}$ is a linear functional on W that satisfies*

$$f_W(w) \le p(w) \quad \forall w \in W.$$

Suppose $z^* \in X \setminus W$ and define

$$W^* = \text{span}\{z^*\} \oplus W = \{cz^* + w : c \in \mathbb{R}, w \in W\}.$$

Then there exists a linear functional $f_{W^*} : W^* \to \mathbb{R}$ that extends f_W and moreover satisfies

$$f_{W^*}(y) \le p(y) \quad \forall y \in W^*.$$

Proof. First note that for all $v, w \in W$, we have

$$f_W(w) + f_W(v) = f_W(w + v) \le p(w + v) \le p(w - z^*) + p(z^* + v),$$

and therefore we obtain

$$f_W(w) - p(w - z^*) \le -f_W(v) + p(z^* + v). \tag{24}$$

Now define

$$\xi := \inf_{v\in W} (-f_W(v) + p(z^* + v))$$

which is a real number because (24). Define $f_{W^*} : W^* \to \mathbb{R}$ by $f_{W^*}(cz^* + w) = c\xi + f_W(w)$. Then clearly f_{W^*} is linear and extends f_W. Now let $y = cz^* + w \in W^*$. We show $f_{W^*}(y) \le p(y)$. The case $c = 0$ follows immediately from the assumptions. For $c > 0$, we have that for $v := \frac{1}{c}w$,

$$c\xi \le -cf_W(v) + cp(z^* + v) = -f_W(cv) + p(cz^* + cv) = -f_W(w) + p(cz^* + w),$$

and the desired inequality follows. For $c < 0$, we find from (24) and the definition of ξ that

$$c\xi \leq cf_W\left(\frac{w}{|c|}\right) - cp\left(\frac{w}{|c|} - z^*\right) = -f_W(w) + |c|p\left(\frac{w}{|c|} - z^*\right) = -f_W(w) + p(cz^* + w),$$

which yields the desired inequality. □

Lemma A.4 might give us a hint how to prove the Hahn-Banach theorem (1.9). The idea is roughly as follows. If X is a *separable* real normed vector space, then we can find a dense sequence $(z_n)_{n=1}^{\infty}$ in X (with some additional properties) and recursively define suitable extensions f_n on $\text{span}\{z_1, ..., z_n\} \oplus W$ in 'one-dimensional steps' using lemma A.4, and then by density we can use these f_n to construct an extension defined on X.

However, we do not have a norm at our disposal and even if that is the case, this method does not work for nonseparable spaces. The solution is to use Zorn's lemma.

Theorem A.5 (Hahn-Banach theorem)**.** *Let X be a real vector space. Let $p : X \to \mathbb{R}$ be a sublinear functional on X. Suppose $W \subseteq X$ is a linear subspace of X and $f_W : W \to \mathbb{R}$ is a linear functional on W satisfying*

$$f_W(w) \leq p(w) \quad \forall w \in W.$$

Then f_W has a linear extension $f_X : X \to \mathbb{R}$ on X such that

$$f_X(x) \leq p(x) \quad \forall x \in X$$

Proof. Let us denote the domain of a function f by D_f. Let $\mathcal{M}$ be the collection of all linear functionals $f : D_f \to \mathbb{R}$ defined on a linear subspace D_f of X satisfying $W \subseteq D_f \subseteq X$ with the following properties:

- $f(w) = f_W(w)$ for all $w \in W$;
- $f(y) \leq p(y)$ for all $x \in D_f$.

We will apply Zorn's lemma to show that $\mathcal{M}$ has a maximal element (under an appropriate partial order), which will be the desired extension f_X with domain X.

Observe that $\mathcal{M} \neq \varnothing$ as $f_W \in \mathcal{M}$. Now define a partial order on $\mathcal{M}$ by setting $f \preceq g$ if and only if $D_f \subseteq D_g$ and $f(x) = g(x)$ for all $x \in D_f$. We now check the conditions in Zorn's lemma. Let $\mathcal{N} \subseteq \mathcal{M}$ be a nonempty chain.

Step 1. Define

$$A = \bigcup_{f \in \mathcal{N}} D_f$$

This will be the domain of our upper bound for $\mathcal{N}$. First we check that A is a linear subspace of X. It is nonempty as $\mathcal{N}$ is nonempty. Let $x, y \in A$ and $\alpha, \beta \in \mathbb{R}$, then there exist $f, g \in \mathcal{N}$ such that $x \in D_f$, $y \in D_g$. As $\mathcal{N}$ is totally ordered, we have either $f \preceq g$ or $g \preceq f$, say (without loss of generality) $f \preceq g$. Then $D_f \subseteq D_g$, so that $\alpha x + \beta y \in D_g \subseteq A$.

Step 2. Define $f_{\mathcal{N}} : A \to \mathbb{R}$ as follows. Let $y \in A$, then there exists $f \in \mathcal{N}$ such that $y \in D_f$. Define $f_{\mathcal{N}}(y) := f(y)$. This does not depend on the choice of f: if $f, g \in \mathcal{N}$ satisfy $y \in D_f \cap D_g$, then as $\mathcal{N}$ is totally ordered we have that $f \preceq g$ or $g \preceq f$, so that $f(y) = g(y)$.

Step 3. We now check that $f \in \mathcal{M}$. Note that $f_{\mathcal{N}}$ is linear: this follows from a similar reasoning as in step 1. We have that $f_{\mathcal{N}}(w) = f_W(w)$ for all $w \in W$: let $w \in W \subseteq A$, choose $f \in \mathcal{N}$ such that $w \in D_f$, then we have $f_{\mathcal{N}}(w) = f(w) = f_W(w)$. A similar reasoning shows that $f_{\mathcal{N}}(y) \leq p(y)$ for all $y \in A$.

We conclude that $f_{\mathcal{N}} \in \mathcal{M}$, and by construction it is an upper bound for $\mathcal{N}$. Thus the conditions in Zorn's lemma are satisfied, hence there exists a maximal element $F \in \mathcal{M}$. We claim that $D_F = X$. Suppose namely that $D_F \subsetneq X$, then by lemma A.4 there exists an extension of F, defined on a larger domain, which also lies in $\mathcal{M}$. But this contradicts the maximality of F. Thus $D_F = X$. Finally, the function $f_X := F$ is an extension of f_W with the desired properties. □

The next version of the Hahn-Banach theorem addresses the extension of a bounded linear functional defined on a subspace to a bounded linear functional defined on the whole space with the same norm.

Theorem A.6 (Hahn-Banach theorem)**.** *Let X be a real normed vector space, $W \subseteq X$ a linear subspace. Suppose $f_W \in W'$. Then there exists an extension $f_X \in X'$ of f_W such that $\|f_X\|_{X'} = \|f_W\|_{W'}$.*

Proof. Define $p : X \to \mathbb{R}$ by $p(x) = \|f_W\|_{W'} \cdot \|x\|_X$ Then p is a sublinear functional, and $p(x) = p(-x)$. We have that

$$f_W(w) \leq |f_W(w)| \leq \|f_W\|_{W'}\|w\|_X = p(w) \quad \forall w \in W.$$

By theorem A.5, f_W has a linear extension $f_X : X \to \mathbb{R}$ such that $f_X(x) \leq p(x) \quad \forall x \in X$. We also have that $-f_X(x) = f_X(-x) \leq p(-x) = p(x)$, so that $-p(x) \leq f_X(x) \leq p(x)$ hence $|f_X(x)| \leq p(x) = \|f_W\|_{W'} \cdot \|x\|_X$. In particular, we see that $f_X \in X'$ and $\|f_X\|_{X'} \leq \|f_W\|_{W'}$. We also have that $\|f_W\|_{W'} \leq \|f_X\|_{X'}$. This finishes the proof. ☐

References

[1] A. Amenta. (2017). *Fourier Analysis*. Lecture notes, TU Delft.

[2] K.P.S. Bhaskara Rao and M. Bhaskara Rao. (1983). *Theory of Charges: A Study of Finitely Additive Measures*. Academic Press, London.

[3] L. Bukovský, E. Coplakova and K.P. Hart. (2022). *Verzamelingen en van Alles Eromheen.* Epsilon Uitgaven, Amsterdam.

[4] N.L. Carothers. (2000). *Real Analysis*. Cambridge University Press, Cambridge.

[5] E.K. van Douwen. (1992). Finitely additive measures on $\mathbb{N}$. *Topology and its Applications*, 47, 223-268.

[6] J. van Neerven. (2022). *Functional Analysis*. Cambridge University Press, Cambridge.

[7] B.P. Rynne and M.A. Youngson. (2008). *Linear Functional Analysis*. 2nd ed. Springer-Verlag London, London.

[8] M.C. Veraar. (2021). *Measure and Integration*. Lecture notes, TU Delft. Available at https://fa.ewi.tudelft.nl/~veraar/teaching/lecture_notes_KW2.pdf.