PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

ARTICLE

Totality, Regularity, and Cardinality in Probability Theory

Paolo Mancosu[1] and Guillaume Massas[2]

[1]University of California, Berkeley, Berkly, CA, USA and [2]University of California, Berkeley, Berkeley, CA, USA
Corresponding author: Guillaume Massas; Email: gmassas@berkeley.edu

Abstract

Recent developments in generalized probability theory have renewed a debate about whether regularity (i.e., the constraint that only logical contradictions get assigned probability 0) should be a necessary feature of both chances and credences. Crucial to this debate, however, are some mathematical facts regarding the interplay between the existence of regular generalized probability measures and various cardinality assumptions. We improve on several known results in the literature regarding the existence of regular generalized probability measures. In particular, we give necessary and sufficient conditions for the existence of regular generalized probability measures defined on the whole powerset of any sample space.

1. Regularity and classical probability theory

The standard Kolmogorovian approach to probability theory on infinite sample spaces is neither regular nor total. In Kolmogorov probability theory, a probability space is given by a triple $(\Omega, \mathfrak{A}, \mu)$, where Ω is a set called the *sample space*, $\mathfrak{A}$ is a σ-algebra of subsets of Ω, and μ is a probability measure assigning a real number in the interval $[0, 1]$ to every set in $\mathfrak{A}$.

Regularity corresponds to the constraint that only logical contradictions get assigned probability 0.[1] It is often understood as a "bridge principle" between modality and probability (Hájek, 2011; Easwaran, 2014), stating that no possible event can receive a probability 0 of occurring. Totality, on the other hand, requires every subset of the sample space Ω to be measurable, that is, to have a probability value.

---

[1] There are many versions of regularity depending on whether one speaks of objective probabilities (chances) or subjective probabilities (credences) and on how one interprets the modalities (doxastic, epistemic, metaphysical, etc.). For our purposes, this can be ignored because we are not concerned with arguing about the plausibility, or lack thereof, of any such principles but only with the mathematical relations between regularity and totality.

A measure μ in a probability space $(\Omega, \mathfrak{A}, \mu)$ satisfies Kolmogorov's axioms if:

- $\mu(U) \geq 0$ for all $U \in \mathfrak{A}$ (nonnegativity);
- $\mu(\Omega) = 1$ (normality); and
- whenever $\{A_i\}_{i \in \omega}$ is a family of pairwise disjoint sets in $\mathfrak{A}, \mu(\cup_{i \in \omega} A_i) = \sum_{i \in \omega} \mu(A_i)$ (countable additivity).

Moreover:

- μ is *regular* if $\mu(A) = 0$ implies $A = \emptyset$ for any $A \in \mathfrak{A}$; and
- μ is *total* if $\mathfrak{A} = \mathscr{P}(\Omega)$.

It is well known that (and this is one example among many) by letting $\Omega = [0, 1]$, we witness a failure of both regularity and totality for many standard Kolmogorovian probability measures. Indeed, suppose μ is a uniform probability measure defined on all singletons in $[0, 1]$; that is, $\mu(\{x\})$ is defined for any $x \in [0, 1]$, and $\mu(\{x\}) = \mu(\{y\})$ for any $x, y \in [0, 1]$. Then a standard argument that uses countable additivity[2] establishes that $\mu(\{x\}) = 0$ for any $x \in [0, 1]$. One of the unpalatable consequences of this fact is that we have to accept that in a Kolmogorovian setup, whenever we throw a perfectly pointed dart at the interval $[0, 1]$ and we assume uniformity, every single point $x \in [0, 1]$ has probability 0 to be hit, even though one of them will be hit (because $\mu([0, 1]) = 1$). But, and we will come back to this point later on, failure of regularity can also occur when the probability distribution is not uniform.

The failure of totality in the Kolmogorovian context often follows from the axiom of choice. For example, Vitali's theorem establishes that there are subsets of the interval $[0, 1]$ that are not Lebesgue measurable.[3] When one does not require the probability measure to be translation invariant, the existence of total probability measures on the interval $[0, 1]$ becomes either trivial (for any $x \in [0, 1]$, the function $\mu_x: \mathscr{P}([0, 1]) \to [0, 1]$ given by $\mu(U) = 1$ if $x \in U, \mu(U) = 0$ otherwise is a two-valued, countably additive, total probability measure) or independent of *ZFC* (the existence of a countably additive total measure μ such that $\mu(\{x\}) = 0$ for all $x \in [0, 1]$ implies that 2^ω is real-valued measurable; see Ciesielski [1989]).

The failure of regularity and totality[4] in the Kolmogorovian case raises the issue of whether there exist theorems relating regularity and totality that can be stated in more general terms and that apply beyond the context of Kolmogorovian probability. For instance, the failure of totality for the Lebesgue measure uses countable additivity. But what if we weaken our probability measures to finitely additive functions?[5]

---

[2] In fact, finite additivity is enough; see Section 3.

[3] The axiom of choice is necessary here: Solovay (1970) showed that if there is an inaccessible cardinal, then there is a model of *ZF* in which all sets of reals are Lebesgue measurable.

[4] We are keeping these two problems separate because in certain accounts, regularity could fail because no value has been assigned to a possible event (see Hájek, 2011, 7). In our account, regularity can fail only when a possible event (a nonempty subset of the sample space) is assigned probability 0.

[5] For example, Banach (1923) constructs a total, finitely additive, and translation-invariant probability measure on [0,1], but that measure is, of course, not regular.

 Philosophers of probability have devoted quite a bit of attention to regularity.[6] The discussion has recently been revitalized by the emergence of powerful infinitesimalist approaches to probability theory, which seem to open new possibilities for a defense of regularity by using infinitesimals in the range of the probability measure. But some philosophers have found this attempt to save regularity wanting. In particular, Williamson (2007), Easwaran (2014), and Hájek (2011) have claimed that the use of infinitesimals is not a satisfactory way of implementing regularity as a constraint on rational credences. For instance, Williamson (2007) provides a thought experiment with a countably infinite sequence of coin tosses and claims that even with infinitesimals at our disposal, the probability of such a sequence is 0 (and yet such a sequence is possible). Meanwhile, Easwaran (2014) and Hájek (2011) argue that some of the issues for orthodox Bayesianism created by the failure of regularity are best addressed by a revision of orthodox Bayesianism altogether.

 Regarding the introduction of infinitesimals as a solution to the regularity problem in standard probability theory, the dialectic that has emerged has been well characterized by Hájek:

> I envisage a kind of arms race: we scotched regularity for real-valued probability measures by canvassing sufficiently large domains: making them uncountable. The friends of regularity fought back, enriching their ranges: making them hyper-realvalued. I counter with a still larger domain: making its values hyperreal-valued. Perhaps regularity can be preserved over that domain by enriching the range again, as it might be, making it hyper-hyperreal-valued. I counter again with a yet larger domain: making its values hyper-hyperreal-valued. And so it goes. Some latter-day Bernstein and Wattenberg would need to keep coming up with constructions that would uphold regularity, however big Ω gets—presumably with ever-richer fields of numbers to provide the values of the probability measures (2011, 21–22).

The thought contained here has received a tentative characterization by Hofweber (2014b) and a more formal one by Pruss (2013). Hofweber advocates "flexibility" in the choice of range and conjectures that by choosing a large enough range, one can preserve properties akin to regularity (he is especially concerned about preserving what he calls the "minimal constraint"):

> The lesson from this for believers in infinitesimal chances is simply this: we don't pick our numbers that measure chance once and for all. We don't use a fixed number system once and for all, suitable for whatever events need their chances measured. Rather, we pick a suitable extension of the real numbers, one that is suitable for the task at hand and the events under consideration that need their chances measured. If we try to measure the chance of the dart to hit a certain point on the real line, or the chance that a countable sequence of coin tosses comes up heads, then any hyperreal extension of the real numbers will do. We can just pick one. If we hope to measure a larger set of events, then we should play it safe and pick a larger non-Archimedean extension of the real numbers.

  [6] Hájek (2011) lists Kemeny (1955), Shimony (1955, 1970), Jeffreys (1922), Edwards et al. (1963), Carnap (1963), Stalnaker (1970), Lewis (1980), Skyrms (1980), Appiah (1985), Jackson (1987), and Jeffrey (1992).

All that is needed, I conjecture, is one that has larger cardinality than the set of events to be measured. As long as the range of the probability measure is larger than the domain, we will be fine. Less might be fine in most cases, but more is a safe bet. Thus when we measure the chance of events in a certain set of events we pick some extension of the real numbers with infinitesimals that has a larger size than the set of events. Any one of them will do. But you can't pick one first to be suitable for any possible sets of events that need their chances measured.

In keeping with the dialectic outlined so far, the formal results discussed in the literature on the topic can be roughly divided into two categories. On the one hand, *existence theorems* establish that under certain conditions on the sample space Ω and the range of probabilities V, regular probability measures from $\mathscr{P}(\Omega)$ to V can be constructed. On the other hand, *impossibility theorems* identify conditions on Ω and V under which there can be no regular probability measure from $\mathscr{P}(\Omega)$ to V. For results of the first kind, Hofweber (2014a) provides, for instance, the following result:

Theorem 1.1. *Let Ω be any infinite sample space. There is a hyperreal field $\Re^*$ of at most cardinality $2^{|\Omega|}$ and a regular probability measure from $\mathscr{P}(\Omega)$ into $\Re^*$.*

In other words, given any sample space Ω of cardinality κ, one can always obtain a total and regular probability measure on $\mathscr{P}(\Omega)$ by finding an appropriate hyperreal field of cardinality 2^{κ}.

By contrast, the best impossibility results existing in the literature are given by Pruss (2013) (see also sec. 2). Unlike Hofweber (2014a), who focuses on the size of the algebra of subsets of the sample space, Pruss focuses on the size of the sample space. Pruss (2013) shows that whenever Ω is strictly larger in cardinality than the range of a probability measure μ, μ cannot be regular.

Our main contribution to the existing literature is an improvement of Pruss's impossibility theorem. Using a notion of generalized probability measure inspired by Pruss's setting, we establish the following:

Theorem 3.1. *For any generalized probability range $(V, \leq, +, 0)$ and any set Ω such that $|\Omega| \geq |V|$, there is no total regular V-probability measure for Ω.*

Theorem 3.1 is a clear strengthening of Pruss's result because it establishes that the assumption that $|\Omega| > |V|$ can be weakened to $|\Omega| \geq |V|$. We also combine our novel impossibility theorem with existence theorems such as the ones noted by Hofweber (2014a) and Benci et al. (2018) to address the following question: Given a sample space Ω and a cardinal κ, is there a generalized probability range V (see Definition 2.1) of size κ and a regular generalized V-probability measure defined on $\mathscr{P}(\Omega)$? We show that under the assumption of the generalized continuum hypothesis (GCH), the conjunction of the known existence theorems and our impossibility result yields a complete answer to the question:

Theorem 5.7. (GCH) *Let κ be a cardinal and Ω a set. Then there is a hyperreal field $\Re$ of size κ and a total and regular finitely additive probability measure $\mu: \mathscr{P}(\Omega) \to \Re$ if and only if (iff) $|\Omega| < \kappa$.*

Moreover, a related problem is to give necessary and sufficient conditions for when, given a sample space Ω and a field V, there exists a regular probability measure defined on $\mathscr{P}(\Omega)$. Because we believe that the first question is the more relevant one for the current philosophical debate on regularity, we do not pursue this question here. However, we consider a similar question in the case where the algebra of events is the set of finite and cofinite subsets of Ω rather than the full powerset. Regarding this issue, we establish the following existence result:

Theorem 4.7. *Let Ω be an infinite set, let $Fin(\Omega)$ be the algebra of finite and cofinite subsets of Ω, and let V be a countable non-Archimedean field. Then there is a regular V-probability measure $\mu: Fin(\Omega) \to V$.*

We then use this lemma together with well-known impossibility results to establish the following:

Theorem 4.8. *Let Ω be an uncountable set and F an infinite field. Then there is a regular generalized F-probability measure defined on the algebra of finite and cofinite subsets of Ω iff F is non-Archimedean.*

The rest of the article is organized as follows. In Section 2, we introduce a framework for generalized probability measures inspired by the one in Pruss (2013) and connect the issue of regularity in probability theory to an old result of Zermelo (1904) regarding functions from the powerset of a set X into X itself. In Section 3, we show how Zermelo's theorem can be used to establish several impossibility results about total and regular probability measures. In particular, we show how to improve the cardinality assumption in Pruss's impossibility theorem. In Section 4, we investigate what happens when the requirement that a generalized probability measure be defined on the full powerset of an infinite sample space is relaxed, and we show that even in this case, Zermelo's theorem allows us to prove some impossibility theorems. Finally, in Section 5, we discuss the connection between our results and some of the hyperreal fields that have been constructed as ranges of regular generalized probability measures.

2. Part-whole and regularity

In this section, we first introduce the general probabilistic setting in which we will work. Throughout this article, we will not assume that the range of a probability measure μ is contained in the real unit interval $[0, 1]$. The motivation for this is twofold. First, Hájek's (2011) "arms race" description of the dialectic between proponents and opponents of regularity involves considering domains and codomains of ever-increasing cardinality. Second, as we shall see later on, the relationship between regularity, totality, and cardinality in the classical setting is muddled by the fact that the standard unit interval $[0, 1]$ is Archimedean. We may therefore hope to gain a clearer understanding of the possible obstacles to the existence of a total and regular probability measure in a more general context in which Archimedeanity does not play a role anymore. However, given the importance of probability measures with a range contained in the real interval $[0, 1]$ in the literature, we will often point out how the results we discuss apply in this specific case. We start with the following definitions adapted from Pruss (2013).

Definition 2.1. *A generalized probability range is a tuple* $(V, \leq, +, 0)$ *satisfying the following conditions:*

- $(V, \leq)$ *is a partial order, with* 0 *being its least element;*
- $(V, +, 0)$ *is a commutative monoid—that is,* $+$ *is a commutative and associative operation on* V, *and* $x + 0 = x$ *for any* $x \in V$; *and*
- *for any* $x, y, z \in V, x < y$ *implies* $x + z < y + z$.

Definition 2.2. *Let* $(V, \leq, +, 0)$ *be a generalized probability range. A generalized V-probability space is a triple* $(\Omega, \mathfrak{A}, \mu)$ *such that:*

- Ω *is a set, and* $\mathfrak{A}$ *is a collection of subsets of* Ω *closed under finite unions and complements; and*
- $\mu : \mathscr{P}(\Omega) \to V$ *is a function satisfying the following two conditions:*
  - $0 \leq \mu(A)$ *for any* $A \in \mathfrak{A}$ *(nonnegativity); and*
  - *whenever* $A, B \in \mathfrak{A}$ *are such that* $A \cap B = \emptyset, \mu(A \cup B) = \mu(A) + \mu(B)$ *(finite additivity).*

Given a generalized probability space, $(\Omega, \mathfrak{A}, \mu)$, the generalized V-probability measure μ is *total* if $\mathfrak{A} = \mathscr{P}(\Omega)$, and it is *regular* if $\mu(A) = 0$ implies $A = \emptyset$ for any $A \in \mathfrak{A}$.

For any ordered field F, the nonnegative elements of F form a generalized probability range. For this reason, we will sometimes consider fields as generalized probability ranges, despite the slight abuse of language. Note that in the classical setting, a finitely additive probability measure is a generalized $\mathbb{R}$-probability measure $(\Omega, \mathfrak{A}, \mu)$ satisfying the additional condition that $\mu(\Omega) = 1$.

Pruss (2013) shows a version of the following:

Theorem 2.3. *Let* $(V, \leq, +, 0)$ *be a generalized probability range and* Ω *a set such that* $|\Omega| > |V|$. *Then there is no total and regular generalized V-probability measure for* Ω.

Our first main result (Theorem 3.1) is an improvement of Pruss's result. In this section, we lay the groundwork for this theorem by establishing a connection between generalized regular probability measures and part-whole–preserving functions. We start by recalling a result that is well known in the literature (e.g., it is stated en passant in Benci et al. [2018], 516) but is worth proving expressly. We first need the following definition:

Definition 2.4. *Let* V *be a generalized probability range and* $(\Omega, \mathfrak{A}, \mu)$ *be a generalized V-probability space. The generalized V-probability measure* μ *preserves part-whole if for any* $A, B \in \mathfrak{A}$ *such that* $A \subsetneq B, \mu(A) < \mu(B)$.

Lemma 2.5. *Let* $\mu : \mathfrak{A} \to V$ *be a generalized V-probability measure for some generalized probability range* V. *Then* μ *satisfies regularity iff* μ *preserves part-whole.*

Proof. First, it is a simple exercise to verify that any generalized V-probability space $(\Omega, \mathfrak{A}, \mu)$ has the following two properties: $\mu(\emptyset) = 0$, and $\mu(A) \leq \mu(B)$ whenever $A, B \in \mathfrak{A}$ and $A \subseteq B$.

Fix $V, \Omega, \mathfrak{A},$ and $\mu \colon \mathfrak{A} \to V$. If $A \subsetneq B$ for some $A, B \in \mathfrak{A}$, then finite additivity implies that $\mu(B) = \mu(A) + \mu(B \backslash A)$. If μ is regular, then we have $0 < \mu(B \backslash A)$, from which it follows that

$$\mu(A) = \mu(A) + 0 < \mu(A) + \mu(B \backslash A) = \mu(B).$$

Hence, μ preserves part-whole.

Now suppose μ preserves part-whole, and let A be any nonempty set in $\mathfrak{A}$. Because $\emptyset \subsetneq A$, by part-whole we have $\mu(\emptyset) < \mu(A)$. Thus, $\mu(A) \neq \mu(\emptyset) = 0$ for every nonempty subset A, which means that μ is regular. $\square$

We now introduce a theorem that will be of great use for our proofs. Its significance for the study of abstraction operators that are used in the context of neologicism has been highlighted by Mancosu and Siskind (2019) and for issues related to the axiom of choice in second-order logic by Siskind et al. (2023). The theorem, in its set-theoretic version, is implicit in Zermelo (1904) and was brought out explicitly by Kanamori (1997).

Theorem 2.6 (Zermelo, 1904). *Given a set X and an arbitrary function $f \colon \mathscr{P}(X) \to X$, there are sets $A, B \in \mathscr{P}(X)$ such that $A \subsetneq B$ and $f(A) = f(B)$.*

The theorem can be proved without the axiom of choice, and it can be formalized in a second-order theory with only the additional symbol for f (see Mancosu and Siskind, 2019). Here is a set-theoretic proof that uses transfinite induction:

Proof. Fix a set X and a function $f \colon \mathscr{P}(X) \to X$. In order to violate part-whole, we must find two sets $A, B \in \mathscr{P}(X)$ such that A is a proper subset of B and $f(A) = f(B)$. Using transfinite recursion, define a function $G \colon Ord \to \mathscr{P}(X)$, where Ord is the class of ordinals, as follows:

$$G(0) = f(\emptyset);$$

$$G(\alpha) = f(ran(G|\alpha)), \text{where } ran(G|\alpha) = \{G(\beta) | \beta < \alpha\}.$$

One easily verifies that for any ordinal $\alpha, G(\alpha) \subseteq X$. Because $\mathscr{P}(X)$ is a set and Ord is a proper class, there must be $\alpha < \beta \in Ord$ such that $G(\alpha) = G(\beta)$. Letting $A = G|\alpha$ and $B = G|\beta$, this means that $f(A) = G(\alpha) = G(\beta) = f(B)$. At the same time, we clearly have that $A \subsetneq B$, which shows that A and B are the required counterexamples to part-whole. $\square$

The connection between the previous applications and the probabilistic setting is the following. Just as in neologicism we are concerned with functions defined by abstraction principles mapping concepts (conceived extensionally as subsets of the domain of individuals) on a certain domain into the domain itself, in many applications of probability theory, we are also mapping a collection of subsets (or the entire power set) of the sample space Ω into Ω itself.

The following result is the first immediate application of Zermelo's (1904) result in the context of probability theory:

Theorem 2.7. *Let X be a subset of $[0, 1]$, and let μ from $\mathscr{P}([0, 1])$ into X be a finitely additive probability measure. Then μ is not regular.*

Proof. Fix a finitely additive probability measure $\mu : \mathscr{P}([0, 1]) \rightarrow X$, and notice that we may view μ as a finitely additive measure with codomain $[0, 1]$. By Lemma 2.5, μ is regular only if it preserves part-whole. But by Theorem 2.6, μ must violate part-whole and therefore cannot be regular. $\qquad\square$

Theorem 2.7 only uses a special case of Zermelo's theorem—namely, when the codomain of the function f is a subset of the unit interval. As we shall now see, the generality of Zermelo's theorem allows for a much more general impossibility result regarding the existence of regular probability measures.

3. Zermelo's theorem and generalized probability measures

Our first generalization of Theorem 2.7 allows us to consider a much wider class of sample spaces and probability ranges.

Theorem 3.1. *For any generalized probability range $(V, \leq, +, 0)$ and any set Ω such that $|\Omega| \geq |V|$, there is no total regular V-probability measure for Ω.*

Proof. Suppose toward a contradiction that μ is a regular generalized V-probability measure for Ω. Because $|\Omega| \geq |V|$, there is a surjection $f : \Omega \rightarrow V$. Consider the inverse image lift $f_* : \mathscr{P}(V) \rightarrow \mathscr{P}(\Omega)$ of the surjection f, given by $f_*(A) = \{y \in \Omega | f(y) \in A\}$ for any $A \subseteq V$. One easily verifies the following:

- $f_*(\emptyset) = \emptyset, f_*(A) \subseteq f_*(B)$ whenever $A, B \subseteq V$ and $A \subseteq B$; and
- for any $A, B \subseteq V, f_*(A \cup B) = f_*(A) \cup f_*(B), f_*(A \cap B) = f_*(A) \cap f_*(B)$;
- f_* is an injection.

Now consider the composition $\nu = \mu \circ f_* : \mathscr{P}(V) \rightarrow V$. We then verify that ν is a regular generalized V-probability measure. Indeed:

- For any $A \subseteq V, \nu(A) = \mu(f_*(A)) \geq 0$ by nonnegativity of μ.
- For any $A, B \subseteq V$ such that $A \cap B = \emptyset$:

$$\nu(A \cup B) = \mu(f_*(A \cup B)) = \mu(f_*(A) \cup f_*(B)) = \mu(f_*(A)) + \mu(f_*(B))$$

by finite additivity of μ, and because $f_*(A) \cap f_*(B) = f_*(A \cap B) = \emptyset$. Hence,

$$\nu(A \cup B) = \nu(A) + \nu(B),$$

which shows that ν is finitely additive.

- For any $A \subseteq V, (A) = 0$ implies that $\mu(f_*(A)) = 0$. Because μ is regular, this means that $f_*(A) = \emptyset$. Because f_* is injective and $f_*(A) = f_*(\emptyset) = \emptyset$, it follows that $A = \emptyset$, which establishes that ν is regular.

By Lemma 2.5, this means that there is a part-whole–preserving function $\mu \colon \mathscr{P}(V) \to V$, which contradicts Zermelo's theorem. Hence, there is no generalized V-probability measure $\mu \colon \mathscr{P}(\Omega) \to V$.   □

Theorem 3.1 is a clear generalization of Pruss's (2013) result. By appealing to Zermelo's theorem, we weakened the assumption that Ω has size strictly greater than V to merely assuming the existence of a surjection from Ω to V.

Moreover, Pruss's (2013) result only shows that there must be at least one failure of regularity for any V-probability measure. Using Theorem 3.1, however, we can actually show that there must be many failures of regularity, even in the generalized setting.

Lemma 3.2. *Let Ω be a set and V be a generalized probability range such that $|\Omega| = |V| = \kappa$. Then for any V-probability measure $\mu \colon \mathscr{P}(\Omega) \to V, \mu(\{x\}) = 0$ for κ many elements $x \in \Omega$.*

Proof. Let $C = \{x \in \Omega | \mu(\{x\}) = 0\}$, and let $\Omega' = \Omega \backslash C$. Consider the function $\nu \colon \mathscr{P}(\Omega') \to V$ determined by $\nu(A) = \mu(A)$ for any $A \subseteq \Omega'$. One easily verifies that ν is a V-probability measure for Ω'. Now suppose that $A \subsetneq B \in \mathscr{P}(\Omega')$. This implies that there is $x \in B \backslash A$, and note that because B is disjoint from $C, \nu(\{x\}) = \mu(\{x\}) > 0$, which implies that $\nu(B \backslash A) > 0$. Because $\nu(B) = \nu(A) + \nu(B \backslash A)$, it therefore follows that $\nu(A) < \nu(B)$. Thus ν is a regular V-probability measure for Ω'. By Zermelo's theorem (Theorem 3.1), it follows that $|\Omega'| < |V| = \kappa$. Now, because $\kappa = |\Omega| = |\Omega'| + |C|$, from $|\Omega'| < \kappa$ and basic cardinal arithmetic, it follows that $|C| = \kappa$. Thus, $\mu(\{x\}) = 0$ for κ many elements x of Ω.

Finally, we conclude this section by discussing in some detail the special case of real-valued finitely additive probability measures. As an immediate corollary of Theorem 3.1, we obtain the following:

Corollary 3.3. *Let Ω be a set such that $|\Omega| = 2^{\aleph_0}$. Then no finitely additive probability measure $\mu \colon \mathscr{P}(\Omega) \to [0, 1]$ is regular.*

Corollary 3.3 is a weaker form of a well-known result establishing that there can be no finitely additive regular probability measure from a set Ω into $[0, 1]$ when Ω is uncountable. The argument has been presented neatly by both Hájek (2003) and Williamson (2007). Let Ω be a set, and suppose that μ is a finitely additive probability measure into $[0, 1]$ and defined on all singletons. For any natural number n, let A_n be the set of all $x \in \Omega$ such that $\mu(\{x\}) \geq \frac{1}{n}$. By finite additivity, each A_n must be finite, which means that there can only be countably many singletons mapped to a nonzero value in $[0, 1]$. If Ω is uncountable, then this means that uncountably many singletons must be mapped to 0.

The Hájek–Williamson argument has three advantages. The first one is that it shows that the failure of regularity must occur in many places because uncountably

many singletons must be mapped to 0. By contrast, Corollary 3.3 only guarantees that regularity must fail in some place, which is enough to conclude that any standard probability measure must assign probability 0 to $\{x\}$ for *some x*, but nothing more. Here, however, it is easy to see that one can adapt the proof of Lemma 3.2 in a straightforward way so as to obtain continuum many failures of regularity in the case of a probability measure defined on the full powerset of a sample space of size continuum. Second, although our application of Zermelo's theorem requires the sample space Ω to have size continuum, one only needs Ω to be uncountable for the Hájek–Williamson argument to go through. In the absence of the continuum hypothesis, the Hájek–Williamson result is therefore stronger than ours. Finally, the Hájek–Williamson argument only requires the probability measure to be defined on all singletons,[7] whereas our results assume that the probability measure is total. However, the last two differences between the two results are connected to a third significant one: the Hájek–Williamson argument relies in an essential way on the fact that the real interval $[0, 1]$ is Archimedean. By finite additivity, only finitely many singletons can have probability greater than $\frac{1}{n}$ for a fixed natural number n, which means that uncountably many singletons must have a value smaller than $\frac{1}{n}$ for any n. But this implies that uncountably many singletons must have probability 0 only because 0 is the unique element smaller than $\frac{1}{n}$ for any n (i.e., $[0, 1]$ is Archimedean). The Hájek–Williamson argument does not say anything about the failure of regularity in a non-Archimedean field, which contrasts with the generality of Theorem 3.1. We will return to the issue of Archimedeanity in the next section, in which we investigate to what extent totality is necessary to establish the impossibility results mentioned so far.

4. Relaxing totality

In this section, we investigate the consequences of relaxing the requirement that the algebra of events must be the entire powerset of Ω for the existence of regular measures. Let us start with an easy result.

Lemma 4.1. *Let Ω be a set, and assume that there exists a total and regular generalized V-probability measure μ on $\mathscr{P}(\Omega)$ for some generalized probability range V. Then for any infinite algebra $\mathbb{A}$ of subsets of Ω, there is a generalized probability range V' of size at most $|\mathbb{A}|$ and a regular probability measure $\nu \colon \mathbb{A} \to V'$. Moreover, if V is a field, then so is V'.*

Proof. Let $\Omega, \mathbb{A}, V$, and $\mu \colon \mathbb{A} \to V$ be as in the statement of the lemma, and let S be the range of the restriction of μ to $\mathbb{A}$. Let V' be the substructure of V generated by S, and notice that because $|\mathbb{A}| \geq \max\{\aleph_0, |S|\}$, we have that $|V'| \leq \mathbb{A}$. But clearly, the restriction $\nu \colon \mathbb{A} \to V'$ of μ is the required regular probability measure. Moreover, if V is a field, then we can take V' to be the substructure of V generated by S in the language of fields, and thus we may assume that V' is a field.

---

[7] In fact, it requires an even weaker assumption on the algebra $\mathfrak{A}$ of events—namely, that it contains an uncountable antichain (recall that an antichain in a field of sets is a collection of pairwise disjoint sets).

In the next section (Theorem 5.5), we will recall existing results in the literature establishing that for any set Ω there is a hyperreal field V and a regular generalized V-probability measure defined on $\mathscr{P}(\Omega)$. This entails that the assumption in the statement of Lemma 4.1 can in fact always be satisfied.

For now, let us investigate some consequences of this lemma for countable sample spaces and real-valued probability measures. In order to do this, we start by recalling a classical result of measure theory (e.g., it is a special case of theorem 2.5 in Horn and Tarski [1948]):

Lemma 4.2. *Let Ω be a countable set. Then there is a regular finitely additive probability measure $\mu \colon \mathscr{P}(\Omega) \to [0, 1]$.*

Proof. Let us first show this in the case $\Omega = \mathbb{N}$. Recall that any real number in the interval $[0, 1]$ has a binary expansion; that is, it can be written as a (possibly infinite) sequence of negative powers of 2. For example, the rational number 0.625 can be written as $0.5 + 0.125 = 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3}$. Conversely, any countable sequence s of 0s and 1s determines the binary expansion of some real number $r(s) \in [0, 1]$. For any subset U of $\mathbb{N}$, let $\mu_0(U) = r(\chi_U)$; that is, let $\mu_0(U)$ be the real number whose binary expansion is determined by the characteristic function of U. Formally, this means that for any $U \subseteq \mathbb{N}$:

$$\mu_0(U) = \sum_{i \in \mathbb{N}} \chi_U(i) \cdot 2^{-i}.$$

For example, $\mu_0(\{1, 3\}) = 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-3} = 0.625$. Then observe the following:

- $\mu_0(U) = 0$ iff χ_U is constantly 0 iff $U = \emptyset$.
- Whenever $U \cap V = \emptyset$, we have that:

$$\begin{aligned}
\mu_0(U) + \mu_0(V) &= \sum_{i \in \mathbb{N}} \chi_U(i) \cdot 2^{-i} + \sum_{i \in \mathbb{N}} \chi_V(i) \cdot 2^{-i} \\
&= \sum_{i \in N}(\chi_U(i) + \chi_V(i)) \cdot 2^{-i} \\
&= \sum_{i \in N} \max(\chi_U(i), \chi_V(i)) \cdot 2^{-i} \\
&= \sum_{i \in N} \chi_{U \cup V} \cdot 2^{-i} \\
&= \mu_0(U \cup V),
\end{aligned}$$

which shows that μ_0 is finitely additive (in fact, the same argument shows that it is σ-additive).

Hence, μ_0 is a regular finitely additive probability measure defined on the whole of $\mathscr{P}(\mathbb{N})$. Now if Ω is any countable set, there is a surjection $f \colon \mathbb{N} \to \Omega$, which can be lifted to an injection $f_* \colon \mathscr{P}(\Omega) \to \mathscr{P}(\mathbb{N})$, as in the proof of Theorem 3.1. But then $\mu_0 \circ f_* \colon \mathscr{P}(\Omega) \to [0, 1]$ is also a regular finitely additive probability measure defined on $\mathscr{P}(\Omega)$. $\qquad\square$

Combined with a similar construction as in Lemma 4.1, we get the following as an immediate corollary of Lemma 4.2:

Corollary 4.3. *Let Ω be a countable set, and let $Fin(\Omega)$ be the algebra of finite and cofinite subsets of Ω. Then there is a regular probability measure $\mu \colon Fin(\Omega) \to [0,1] \cap \mathbb{Q}$.*

Proof. Let f be a bijection between Ω and ω. This induces an isomorphism of fields of sets $\varphi \colon Fin(\Omega) \to Fin(\omega)$. Note, moreover, that for any finite or cofinite subset A of ω, $\mu_0(A)$ is rational, where μ_0 is the regular probability measure defined in the proof of Lemma 4.2. Then letting μ be the restriction of μ_0 to $Fin(\omega)$ and $\nu = \mu \circ \varphi$, we have that ν is a regular $\mathbb{Q}$-probability measure. □

Moreover, Lemmas 4.1 and 4.2 also have some immediate consequences regarding the interaction of probability theory and computability theory. For example, it is well known that there are only countably many recursive subsets of ω and that they form a Boolean subalgebra $\mathbb{A}$ of $\mathscr{P}(\omega)$. By Lemma 4.1, we may consider the subfield V' of $\mathbb{R}$ generated by the range of the restriction of μ_0 to $\mathbb{A}$, where, again, μ_0 is the function defined in the proof of Lemma 4.2. In fact, it is easy to verify that V' is exactly the field of all *recursive* reals in the interval $[0,1]$. This has the following consequence:

Corollary 4.4. *Let $\mathbb{R}'$ be the field of recursive reals in the interval $[0,1]$. Then there is a countable algebra $\mathfrak{A} \subseteq \mathscr{P}(\mathbb{R}')$ and a regular probability measure $\nu \colon \mathfrak{A} \to \mathbb{R}'$.*

Proof. Because $\mathbb{R}'$ is countable, fix a bijection $f \colon \mathbb{R}' \to \omega$. As usual, this yields an isomorphism $f_* \colon \mathscr{P}(\omega) \to \mathscr{P}(\mathbb{R}')$, so let $\mathfrak{A} = f_*[\mathbb{A}]$; that is, $\mathfrak{A}$ is the range of the restriction of f to $\mathbb{A}$. Clearly, $\mathfrak{A}$ is a countable subalgebra of $\mathbb{R}'$. Moreover, let $\nu \colon \mathfrak{A} \to \mathbb{R}'$ be defined as $\nu(f_*(A)) = \mu_0(A)$. Notice that this is well defined because, by construction, any element in $\mathfrak{A}$ is of the form $f_*(A)$ for some $A \in \mathbb{A}$ such that $\mu_0(A) \in \mathbb{R}'$. But clearly, ν is regular, which completes the proof. □

In other words, when restricting our attention to a countable subalgebra of the powerset of the recursive reals in the interval $[0,1]$, we are able to escape the conclusion of Zermelo's theorem. Of course, this is only because $\mathfrak{A}$ is not the full powerset of the set of recursive reals because there are uncountably many subsets of the recursive reals. Let us also note that the regular probability measure thus constructed is not itself recursive, nor is $\mathfrak{A}$ the algebra of recursive subsets of $\mathbb{R}'$.

The previous results establish that relaxing the assumption of totality may sometimes allow for the existence of regular probability measures. One is therefore led to wonder whether totality is in fact a *necessary* condition for the conclusion of Zermelo's theorem to apply. The proof of Zermelo's original theorem certainly seems to make essential use of the fact that the function $f \colon \mathscr{P}(X) \to X$ that ends up violating part-whole is defined on all *arbitrary* subsets of X. However, the obstacle to the existence of regular probability measures is not totality per se because it is sufficient that the algebra of events be merely isomorphic to the full powerset of the sample space, as the following straightforward lemma shows:

Lemma 4.5. *Let Ω be an infinite set and V a generalized probability range such that $|\Omega| \geq |V|$. Then for any algebra of events $\mathfrak{A}$ such that $\mathfrak{A}$ and $\mathscr{P}(\Omega)$ are isomorphic as fields of sets, there is no regular V-probability measure $\mu \colon \mathfrak{A} \to V$.*

Proof. Suppose $\mathfrak{A}$ is an algebra of events isomorphic to $\mathscr{P}(\Omega)$, and let $\varphi \colon \mathscr{P}(\Omega) \to \mathfrak{A}$ be such an isomorphism. Clearly, if $\mu \colon \mathfrak{A} \to V$ is a regular V-probability measure, then so is $\mu \circ \varphi \colon \mathscr{P}(\Omega) \to V$. Therefore, by Theorem 3.1 there can be no regular V-probability measure defined on $\mathfrak{A}$. ◻

As a direct consequence of Lemma 4.5, we can now prove the following impossibility theorem for algebras of events that are far from being the full powerset of the sample space:

Corollary 4.6. *Let Ω be an infinite set and V be a generalized probability range such that $|\Omega| \geq |V|$. Then there is a subalgebra $\mathfrak{A}$ of $\mathscr{P}(\Omega)$ such that $|\mathfrak{A}| = |\mathscr{P}(\Omega) \backslash \mathfrak{A}| = |\mathscr{P}(\Omega)|$ and for which there is no regular V-probability measure defined on $\mathfrak{A}$.*

Proof. Fix a bijection f from Ω into $\Omega \times \{0, 1\}$, and let $\pi \colon \Omega \times \{0, 1\} \to \Omega$ be defined as $\pi(x, i) = x$ for any $x \in \Omega, i \in \{0, 1\}$. Finally, let $g \colon \Omega \to \Omega$ be the composition $\pi \circ f$. Because g is a surjection, it induces an embedding g^* of $\mathscr{P}(\Omega)$ into $\mathscr{P}(\Omega)$. Let $\mathfrak{A} = g^*[\mathscr{P}(\Omega)] = \{g^*(U) | U \subseteq \Omega\}$. It is straightforward to verify that the map $\varphi \colon \mathfrak{A} \to \mathscr{P}(\Omega)$ given by $\varphi(g^*(U)) = U$ for any $U \in \mathscr{P}(\Omega)$ is an isomorphism of fields of sets. By Lemma 4.5, this implies that there can be no regular V-probability measure defined on $\mathfrak{A}$. Now for any $x \in \Omega$, let $x_0 = f^{-1}(x, 0)$ and $x_1 = f^{-1}(x, 1)$. Note that $x_0 \neq x_1$ yet $g(x_0) = g(x_1) = x$ for any $x \in \Omega$. Thus, for any $x \in \Omega$ and $U \subseteq \Omega, x_0 \in g^*(U)$ iff $g(x_0) \in U$ iff $g(x_1) \in U$ iff $x_1 \in g^*(U)$. From this, it follows at once that $\{x_0\} \notin \mathfrak{A}$ for any $x \in \Omega$. Because $x_0 \neq y_0$ whenever $x \neq y$, we can therefore conclude that $\mathfrak{A}$ is a proper subalgebra of $\mathscr{P}(\Omega)$ such that $|\mathscr{P}(\Omega) \backslash \mathfrak{A}| = |\mathscr{P}(\Omega)|$. ◻

This result establishes that totality is not, after all, a necessary condition for the conclusion of Zermelo's theorem to hold. Indeed, to guarantee that there can be no regular generalized V-probability measure defined on a subalgebra $\mathfrak{A}$ of $\mathscr{P}(\Omega)$ for some set Ω of size $|V|$, it is enough to assume that $\mathfrak{A}$ is isomorphic to $\mathscr{P}(\Omega)$, which, as corollary 4.6 shows, does not coincide with $\mathfrak{A}$ being the full powerset of Ω. Moreover, the real obstacle to the existence of a regular function also does not lie in the constraints put on the cardinality of the algebra of events, as the following result shows:

Lemma 4.7. *Let Ω be an infinite set, let $Fin(\Omega)$ be the algebra of finite and cofinite subsets of Ω, and let V be a countable non-Archimedean field. Then there is a regular V-probability measure $\mu \colon Fin(\Omega) \to V$.*

Proof. Let V be a countable non-Archimedean field.[8] Fix a positive infinitesimal $\varepsilon \in V$. By definition, any element in $Fin(\Omega)$ is either a finite subset of Ω or the

---

[8] Recall that such a field V can be constructed, for example, by building a non-Archimedean elementary extension V' of $\mathbb{Q}$ using the compactness theorem and then taking V to be a countable elementary substructure of V'.

complement of such a finite subset. Define $\mu : Fin(\Omega) \to V$ by letting $\mu(A) = \varepsilon|A|$ if A is finite, and $(A) = 1 - \varepsilon|\Omega \backslash A|$ otherwise. Let us show that μ is the required measure. Suppose $A, B \in Fin(\Omega)$ are such that $A \cap B = \emptyset$. We distinguish two cases:

- Case 1: Both A and B are finite. In this case we have that $A \cup B$ is finite and $|A \cup B| = |A| + |B|$, so

$$\mu(A \cup B) = \varepsilon|A \cup B| = \varepsilon(|A| + |B|) = \varepsilon|A| + \varepsilon|B| = \mu(A) + \mu(B).$$

- Case 2: Either A or B is infinite. Without loss of generality, assume A is infinite. Note that this means that $\Omega \backslash A$ is finite. Because $A \cap B = \emptyset$, we have that $B \subseteq \Omega \backslash A$, so B is also finite. By simple finite cardinality reasoning, we have that $|\Omega \backslash A| = |B| + |(\Omega \backslash A) \cap (\Omega \backslash B)|$. Moreover, we have that $\Omega \backslash (A \cup B) = (\Omega \backslash A) \cap (\Omega \backslash B)$. Because A and $A \cup B$ are infinite but B is finite, it follows that $\mu(A) = 1 - \varepsilon|\Omega \backslash A|, \mu(A \cup B) = 1 - \varepsilon|(\Omega \backslash A) \cap (\Omega \backslash B)|$, and $\mu(B) = \varepsilon|B|$. Now we compute:

$$|\Omega \backslash A| = |B| + |(\Omega \backslash A) \cap (\Omega \backslash B)| \Leftrightarrow$$

$$|B| = |\Omega \backslash A| - |(\Omega \backslash A) \cap (\Omega \backslash B)| \Leftrightarrow$$

$$\varepsilon|B| = \varepsilon|\Omega \backslash A| - \varepsilon|(\Omega \backslash A) \cap (\Omega \backslash B)| \Leftrightarrow$$

$$1 - \varepsilon|\Omega \backslash A| + \varepsilon|\Omega \backslash B| = 1 - \varepsilon|(\Omega \backslash A) \cap (\Omega \backslash B)| \Leftrightarrow$$

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

It follows that μ is a V-probability measure. Moreover, by construction, μ is clearly regular. $\square$

Lemma 4.7 shows that the cardinality of the algebra of events does not determine by itself how big the codomain of a regular generalized probability measure must be. Indeed, a countable codomain will suffice for a sample space Ω, provided that the algebra of events under consideration is generated by the singletons. But such an algebra has the same cardinality as Ω, which shows that a countable codomain is enough to find regular measures for arbitrarily large algebras.

We conclude this section by connecting this last result with our discussion of Archimedeanity in the previous section:

Corollary 4.8. *Let Ω be an uncountable set and F be an infinite field. Then there is a regular generalized F-probability measure defined on the algebra of finite and cofinite subsets of Ω iff F is non-Archimedean.*

Proof. Suppose that F is non-Archimedean. Because F is infinite, it contains a countable non-Archimedean subfield. By Lemma 4.7, it follows that there is a regular generalized F-probability measure defined on the Boolean algebra of finite and cofinite subsets of F. Conversely, if F is Archimedean, then the Hájek–Williamson argument shows that there can be no regular generalized F-probability measure defined on all the singletons in Ω.

5. Regularity, cardinality, and hyperreal fields

In this final section, we connect our investigations on the relationship between totality, regularity, and cardinality to the nonstandard approach to probability and to hyperreal extensions of $\mathbb{R}$. We will limit ourselves to the model-theoretic approach to nonstandard analysis because our interest in the cardinality of domains and codomains of regular probability measures does not align with an axiomatic approach à la Nelson. We start by recalling the framework of enlargements that is central in Robinsonian nonstandard analysis.[9]

Definition 5.1. *The universe $\mathbb{U}(\mathbb{X})$ over a set $\mathbb{X}$ is defined as follows:*

- $\mathbb{U}_0(\mathbb{X}) = \mathbb{X}$;
- $\mathbb{U}_{i+1}(\mathbb{X}) = \mathbb{U}_i(\mathbb{X}) \cup \mathscr{P}(\mathbb{U}_i(\mathbb{X}))$ *for $i < \omega$; and*
- $\mathbb{U}(\mathbb{X}) = \cup_{i\,<\,\omega}\mathbb{U}_i(\mathbb{X})$.

Definition 5.2 *An enlargement is a map $^*: \mathbb{U}(\mathbb{X}) \to \mathbb{U}(\mathfrak{X})$ between universes such that:*

- $\mathbb{X} \subsetneq \mathfrak{X}$, *and $^*x = x$ for any $x \in \mathbb{X}$;*
- $\mathfrak{X} =^* \mathbb{X}$ *and $^*\emptyset = \emptyset$;*
- *(Transfer Principle) for any first-order formula $\varphi(x_1, .., x_n)$ in the language of set theory and any $a_1, \ldots, a_n \in \mathbb{U}(\mathbb{X})$:*

$$\mathbb{U}(\mathbb{X})| = \varphi(a_1, \ldots, a_n) \Leftrightarrow{}^* \mathbb{U}(\mathbb{X})| = \varphi(^*a_1, \ldots, {}^* a_n); \text{and}$$

- *(Saturation Principle) whenever $A \in \mathbb{U}(\mathbb{X})$ is collection of sets such that for any $a_1, \ldots, a_n \in A, \cap_{i \leq n}a_i \neq \emptyset$, there is $b \in \mathbb{U}(\mathbb{X})$ such that $b \in \cap_{i \leq n}^* a_i$.*

Whenever $^*: \mathbb{U}(\mathbb{R}) \to \mathbb{U}(\mathfrak{R})$ is an enlargement, $\mathfrak{R}$ is a hyperreal extension of $\mathbb{R}$, meaning that it satisfies the same first-order theory in the language of fields as $\mathbb{R}$, but it contains infinitely small and infinitely large numbers. In particular, one obtains a nonstandard unit interval $[0, 1]_{\mathfrak{R}} =^* [0, 1]_{\mathbb{R}}$. Using a nonstandard framework similar to enlargements, Bernstein and Wattenberg (1969) showed the following:

Theorem 5.3 (Bernstein and Wattenberg, 1969). *There is a hyperreal extension $\mathfrak{R}$ and a total and regular probability measure $\mu: \mathscr{P}([0, 1]_{\mathbb{R}}) \to [0, 1]_{\mathfrak{R}}$ such that for any Lebesgue measurable set $A \subseteq [0, 1]$, the Lebesgue measure of A is equal to the standard part of $\mu(A)$.*

One must exert some care in distinguishing the hyperreal field $\mathfrak{R}$ used by Bernstein and Wattenberg (1969) from the simplest type of hyperreal extensions obtained by taking an ultrapower of the reals modulo a nonprincipal ultrafilter on ω. Indeed, if R is such an ultrapower, then

---

[9] For a comprehensive introduction to enlargements, see Goldblatt (2012, part IV).

$$|[0,1]_R| = |\{f \mid f : \omega \to [0,1]_{\mathbb{R}}\}| = |[0,1]_{\mathbb{R}}|^{\aleph_0} = (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0},$$

and hence $[0,1]_R$ has the same size as $\mathbb{R}$. But then it follows directly from Theorem 3.1 that there can be no total and regular probability measure from $\mathscr{P}(\mathbb{R})$ into $[0,1]_R$. Thus, the hyperreal field $\mathfrak{R}$ considered by Bernstein and Wattenberg cannot be isomorphic to such a hyperreal extension R because it must have size strictly greater than $2^{\aleph_0}$.

Another important aspect of this nonstandard setting is the distinction between internal and external objects. Given an enlargement $^* : \mathbb{U}(\mathbb{X}) \to \mathbb{U}(\mathfrak{X})$, an element $a \in \mathbb{U}(\mathfrak{X})$ is *internal* if $a \in \cup * \mathbb{U}(\mathbb{X}) = \cup\{^*b \mid b \in \mathbb{U}(\mathbb{X})\}$, and *external* otherwise. The distinction is crucial for many applications of nonstandard analysis because internal objects generally behave in a much more tractable fashion than external objects. An important example of internal objects is the internal subsets $\mathscr{P}_{\mathrm{Int}}(^*A)$ of the image $^*A$ of some standard object A. In particular, one has that for any $A \in \mathbb{U}(\mathbb{X})$,

$$^*\mathscr{P}(A) = \mathscr{P}_{\mathrm{Int}}(^*A).$$

In the context of an enlargement $^* : \mathbb{U}(\mathbb{R}) \to \mathbb{U}(\mathfrak{R})$, one may therefore be interested in considering both internal and external entities when discussing the existence of total and regular probability measures on the nonstandard unit interval $[0,1]_{\mathfrak{R}}$. The following is straightforward to verify:

Lemma 5.4. *Let* $^* : \mathbb{U}(\mathbb{R}) \to \mathbb{U}(\mathfrak{R})$ *be an enlargement. Then:*

1. *There is no external total and regular probability measure* $\mu : \mathscr{P}([0,1]_{\mathfrak{R}}) \to [0,1]_{\mathfrak{R}}$.
2. *There is no internal total and regular probability measure* $\mu : \mathscr{P}_{\mathrm{Int}}([0,1]_{\mathfrak{R}}) \to [0,1]_{\mathfrak{R}}$.

Proof. Item 1 is a direct application of Theorem 3.1. For item 2, note that $\mathbb{U}(\mathbb{R})$ satisfies the first-order statement that there is no total and regular probability measure $\mu : \mathscr{P}([0,1]_{\mathbb{R}}) \to [0,1]_{\mathbb{R}}$, so by the Transfer Principle $^*\mathbb{U}(\mathbb{R})$ satisfies the corresponding statement that there is no total and regular probability measure $\mu :^* \mathscr{P}([0,1]_{\mathbb{R}}) \to [0,1]_{\mathfrak{R}}$. But because $^*\mathscr{P}([0,1]_{\mathbb{R}}) = \mathscr{P}_{\mathrm{Int}}([0,1]_{\mathbb{R}})$, it follows that there is no regular internal probability measure into $[0,1]_{\mathfrak{R}}$ defined on all the internal subsets of $[0,1]_{\mathfrak{R}}$. $\qquad\square$

It is worth mentioning that Lemma 5.4 does not address the issue of whether there can be an *external* $[0,1]_{\mathfrak{R}}$-valued regular probability measure that is defined on all the *internal* subsets of $[0,1]_{\mathfrak{R}}$. Note in particular that the results of Section 4 do not apply here because $\mathscr{P}_{\mathrm{Int}}([0,1]_{\mathfrak{R}})$ has the same size as the full powerset of $[0,1]_{\mathfrak{R}}$, but it is not a complete algebra, and thus it is not isomorphic to $\mathscr{P}([0,1]_{\mathfrak{R}})$.[10] We leave this technical issue as an interesting open problem for the time being and limit ourselves to pointing out that nonstandard frameworks obtained as limits of enlargements might be a promising approach toward a positive answer to this problem.[11]

---

[10] By contrast, the fourth possibility—namely, that there must be an *internal* probability measure defined on the whole powerset of $[0,1]_{\mathfrak{R}}$—is obviously ruled out by Theorem 3.1.

[11] We thank C. Ward Henson for the suggestion, connected to his work on the isomorphism property in enlargements (Henson, 1974).

We conclude this section by discussing the connection between our results and an alternative approach to probability theory inspired by nonstandard analysis—namely, the non-Archimedean probability (NAP) theory developed by Benci et al. (2013, 2018). The main advantage of NAP over classical probability theory is that one can construct, for any set Ω, a total and regular probability measure on $\mathscr{P}(\Omega)$ whose codomain is a reasonably small hyperreal extension of $\mathbb{R}$. We briefly review their construction here.

Let Ω be an infinite set. We write $\mathscr{P}_{\text{Fin}}(\Omega)$ for the set of all finite subsets of Ω, and we write $^{\Omega}\mathbb{R}$ for the ring of functions from $\mathscr{P}_{\text{Fin}}(\Omega)$ into $\mathbb{R}$, where the operations $+$ and $\cdot$ and the order $<$ are defined pointwise. A *fine ideal* on $^{\Omega}\mathbb{R}$ is a maximal ideal containing the set

$$\{\varphi \colon \mathscr{P}_{\text{Fin}}(\Omega) \to \mathbb{R} | \exists A \subseteq \mathbb{R}^{\smile} \forall B \supseteq A \colon \varphi(B) = 0\}.$$

Given a fine ideal I, the quotient $^{\Omega}\mathbb{R}/I$ is a hyperreal field that extends $\mathbb{R}$ and has size

$$|^{\Omega}\mathbb{R}| = |\mathbb{R}|^{|\mathscr{P}_{\text{Fin}}(\Omega)|} = (2^{\aleph_0})^{|\Omega|} = 2^{|\Omega|}.$$

Benci et al. (2013) prove the following.

Theorem 5.5. *For any infinite set Ω, there is a total, regular, finitely additive probability measure $\mu \colon \mathscr{P}(\Omega) \to ^{\Omega}\mathbb{R}/I$.*

Benci et al. (2013) call such functions *NAP measures* and show that they exhibit a number of attractive properties. In particular, they satisfy a notion of additivity that is arguably a generalization of countable additivity in the non-Archimedean context. Combined with Theorem 3.1, this yields the following result regarding the interplay between cardinality, regularity, and totality for hyperreal fields:

Theorem 5.6. *Let Ω be an infinite set and κ be a cardinal. Then:*

1. *If $|\Omega| \geq \kappa$, then for any hyperreal field $\mathfrak{R}$ of size at most κ, there is no total and regular probability measure $\mu \colon \mathscr{P}(\Omega) \to \mathfrak{R}$.*
2. *If $2^{|\Omega|} \leq \kappa$, then there is a hyperreal field $\mathfrak{R}$ of size at most κ and a total and regular probability measure $\mu \colon \mathscr{P}(\Omega) \to \mathfrak{R}$.*

Proof. Item 1 follows directly from Theorem 3.1 because any hyperreal field is a generalized probability range, and item 2 follows immediately from Theorem 5.5. $\square$

Finally, let us note that under the assumption of GCH, Theorem 5.6 yields a complete answer to the question of the compatibility between regularity and totality as raised in Section 1.

Theorem 5.7 [GCH] *Let κ be a cardinal and Ω a set. Then there is a hyperreal field $\mathfrak{R}$ of size κ and a total and regular probability measure $\mu \colon \mathscr{P}(\Omega) \to \mathfrak{R}$ iff $|\Omega| < \kappa$.*

Proof. Suppose that there is a hyperreal field $\Re$ of size κ and a total and regular probability measure $\mu : \mathscr{P}(\Omega) \to \Re$. Then by the contrapositive of Theorem 5.6.1, it follows that $|\Omega| < \kappa$. Conversely, suppose that $|\Omega| < \kappa$. Because GCH holds, it follows that $2^{|\Omega|} = |\Omega|^+ \leq \kappa$. Hence, by theorem 5.6.2, there is a hyperreal field $\Re$ of size at most κ and a probability measure $\mu : \mathscr{P}(\Omega) \to \Re$. But clearly, without loss of generality, we can assume that $\Re$ has size κ, which completes the proof. □

We conclude by briefly discussing the significance of our results to the debate regarding regularity in probability theory. In short, we see our results as confirming that the debate has been essentially well posed by Hofweber (2014b) and Hájek (2011). On the one hand, Theorem 5.7 can be viewed as the mathematical fact underlying Hájek's (2011) "arms race." Indeed, as a consequence of Zermelo's theorem, for any possible range of probability values V, any regular probability measure defined on the full powerset of V must take its values in a range of cardinality strictly greater than V. In that sense, no hyperreal field can be "large enough" to play the role of the reals in an alternative to Kolmogorov probability theory that would satisfy the regularity constraint. At the same time, Hofweber is correct in saying that the cardinality of the codomain is the only substantial obstacle to the existence of a regular probability measure for any sample space. In fact, under GCH, it is necessary and sufficient to take the successor cardinal of the size of the domain as the size of the codomain of such a regular probability measure. To make this point differently, this means that, at least under GCH, the existence theorems noted by Hofweber (2014a) and Benci et al. (2018) are the best possible. Interestingly, this also has a consequence for a slightly different question, regarding whether any standard probability function can be approximated by a regular function defined on a hyperreal field.[12] Hofweber and Schindler (2016) give a positive answer to this question. Given a standard probability space $(\Omega, \mathfrak{A}, \overline{\mu})$, they show that $\overline{\mu}$ can be approximated by a regular probability measure μ (in the sense that $\mu(U)$ is infinitesimally close to $\overline{\mu}(U)$ for any $U \in \mathfrak{A}$) defined on a hyperreal field V of size at most $2^{|\Omega|}$. As an easy consequence of Theorem 5.7, we have that, under GCH, the cardinality bound obtained by Hofweber and Schindler is the best possible. In other words, this means that from the point of view of the cardinality of the range of probabilities, the existence problem (i.e., whether there exists a regular probability measure) and the approximation problem (i.e., whether any standard probability function can be approximated by a regular one) are equivalent.

Finally, let us make one final remark on the issue. As the results in Section 4 show, a lot hinges on the particular structure of the algebra of events under consideration. If one agrees with Hájek (2011) that the regularity constraint implicitly entails a totality constraint, then the "arms race" is indeed the only option. But if one only wants to consider events that are finite or cofinite sets of outcomes, then, as Lemma 4.7 shows, any countable non-Archimedean field would suffice. We take this as evidence that there is still much to explore regarding the interplay between regularity, totality, and cardinality in probability theory.

---

[12] We thank an anonymous referee for mentioning this issue.

References

Appiah, A. (1985) *Assertion and Conditionals.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511895975.

Banach, S. (1923) Sur le problème de la mesure. *Fundamenta Mathematicae* 4(1), 7–33. https://doi.org/10.4064/fm-4-1-7-33.

Benci, V., L. Horsten and S. Wenmackers. (2013) Non-archimedean probability. *Milan Journal of Mathematics* 81(1), 121–151. https://doi.org/10.1007/s00032-012-0191-x.

Benci, V., L. Horsten and S. Wenmackers. (2018) Infinitesimal probabilities. *British Journal for the Philosophy of Science* 69(2), 509–552. https://doi.org/10.1093/bjps/axw013.

Bernstein, A. R. and F. Wattenberg. (1969) Nonstandard Measure Theory. In: *Applications of Model Theory to Algebra, Analysis, and Probability: Proceedings of the International Symposium on the Applications of Model Theory to Algebra, Analysis, and Probability, California Institute of Technology, May 23 to 26*, New York: Holt, Rinehart, and Winston, 171–185, *1967*

Carnap, R. (1963) Intellectual Autobiography., In *The Philosophy of Rudolf Carnap*, vol. XI, pp. 1–84. La Salle, IL: Cambridge University Press.

Ciesielski, K. (1989) How good is lebesgue measure? *Mathematical Intelligencer* 11(2), 54–58. https://doi.org/10.1007/BF03023824.

Easwaran, K. (2014) Regularity and hyperreal credences. *Philosophical Review* 123(1), 1–41. https://doi.org/10.1215/00318108-2366479.

Edwards, W., H. Lindman and L. J. Savage. (1963) Bayesian Statistical Inference for Psychological Research., *Psychological Review* 70(3), 193–242. https://doi.org/10.1037/h0044139.

Goldblatt, R. (2012) *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, vol. 188. Berlin: Springer Science & Business Media. https://doi.org/10.1007/978-1-4612-0615-6.

Hájek, A. (2003) What conditional probability could not be. *Synthese* 137(3), 273–323. https://doi.org/10.1023/B:SYNT.0000004904.91112.16.

Hájek, A. (2011) Staying regular?, unpublished typescript.

Henson, C. W. (1974) The isomorphism property in nonstandard analysis and its use in the theory of banach spaces. *Journal of Symbolic Logic* 39(4), 717–731. https://doi.org/10.2307/2272856.

Hofweber, T. (2014a) Cardinality arguments against regular probability measures., thought. *A Journal of Philosophy* 3(2), 166–175. https://doi.org/10.1002/tht3.125.

Hofweber, T. (2014b) Infinitesimal Chances., philosophers. *Imprint* 14(2), 1–34.

Hofweber, T. and R. Schindler. (2016) Hyperreal-valued probability measures approximating a real-valued measure. *Notre Dame Journal of Formal Logic* 57(3), 369–374. https://doi.org/10.1215/00294527-3542210.

Horn, A. and A. Tarski. (1948) Measures in boolean algebras. *Transactions of the American Mathematical Society* 64(3), 467–497. https://doi.org/10.1090/s0002-9947-1948-0028922-8.

Jackson, F. (1987) Conditionals: Oxford, Wiley-Blackwell.

Jeffrey, R. (1992) *Probability and the Art of Judgment.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139172394.

Jeffreys, H.. (1922) *The Theory of Probability.* London: Nature Publishing Group.

Kanamori, A. (1997) The Mathematical Import of Zermelo's Well-Ordering Theorem. *Bulletin of Symbolic Logic* 3(3), 281–311. https://doi.org/10.2307/421146.

Kemeny, J. G. (1955) Fair bets and inductive probabilities. *Journal of Symbolic Logic* 20(3), 263–273. https://doi.org/10.2307/2268222.

Lewis, D. (1980) A Subjectivist's Guide to Objective Chance. In: Lewis, D. (ed.), *Studies in Inductive Logic and Probability*, vol. II, pp. 263–293. Berkeley: University of California Press. https://doi.org/10.1093/0195036468.003.0004.

Mancosu, P. and B. Siskind. (2019). Neologicist Foundations: Inconsistent Abstraction Principles and Part-Whole. In: *Philosophy of Logic and Mathematics: Proceedings of the 41st International Ludwig Wittgenstein Symposium*, Berlin: De Gruyter, 215–248. https://doi.org/10.1515/9783110657883-014.

Pruss, A. R. (2013) Probability, regularity, and cardinality. *Philosophy of Science* 80(2), 231–240. https://doi.org/10.1086/670299.

Shimony, A. (1955) Coherence and the axioms of confirmation. *Journal of Symbolic Logic* 20(1), 1–28. https://doi.org/10.2307/2268039.

Shimony, A. (1970) Scientific Inference. In: Shimony, A. (ed.), *The Nature and Function of Scientific Theories,* pp. 79–172. Pittsburgh: Pittsburgh University Press. https://doi.org/10.1017/CBO9780511621147.009.

Siskind, B., P. Mancosu and S. Shapiro. (2023) A note on choice principles in second-order logic. *Review of Symbolic Logic* 16(2), 339–350. https://doi.org/10.1017/S1755020320000301.

Skyrms, B. (1980) *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws.* New Haven, CT: Yale University Press.

Solovay, R. M. (1970) A model of set-theory in which every set of reals is lebesgue measurable. *Annals of Mathematics* 92(1), 1–56. https://doi.org/10.2307/1970696.

Stalnaker, R. C. (1970) Probability and conditionals. *Philosophy of Science* 37(1), 64–80. https://doi.org/10.1086/288280.

Williamson, T. (2007) How probable is an infinite sequence of heads? *Analysis* 67(3), 173–180. https://doi.org/10.1093/analys/67.3.173.

Zermelo, E. (1904) Beweis, dass jede menge wohlgeordnet werden kann: aus einem an Herrn Hilbert gerichteten briefe. *Mathematische Annalen* 59(4), 514–516. https://doi.org/10.1007/bf01445300.