

Probability Axioms and Set Theory Paradoxes

Ari Herman * and John Caughman 

Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR 97201, USA; caughman@pdx.edu

* Correspondence: ajherman@pdx.edu

Abstract: In this paper, we show that Zermelo–Fraenkel set theory with Choice (ZFC) conflicts with basic intuitions about randomness. Our background assumptions are the Zermelo–Fraenkel axioms without Choice (ZF) together with a fragment of Kolmogorov’s probability theory. Using these minimal assumptions, we prove that a weak form of Choice contradicts two common sense assumptions about probability—both based on simple notions of symmetry and independence.

Keywords: set theory; probability; axiom of choice

1. A Puzzle

We begin with a paradox involving the Axiom of Choice (AC) and an infinite set of fair coins. An early version of this result first appeared as a problem in the American Mathematical Monthly [1]; a version closer to ours can be found in [2].

Let $I = 2^\omega$ denote the set of all binary-valued functions on $\omega = \{0, 1, 2, \dots\}$. Let $\psi : I \rightarrow I$ be a randomly constructed function. By this, we mean that for each $r \in I$, $n \in \omega$, $\psi(r)(n)$ is determined by a fair coin toss. (If you prefer, you may also interpret the elements of I as the binary expansion of a real number $0 \leq r \leq 1$. We note that each dyadic rational, $0 < r < 1$, can occur in two ways; e.g., $\frac{1}{2} = 0.1\bar{0} = 0.0\bar{1}$.) We now ask: “If $\hat{r} \in I$ is chosen at random (i.e., $\hat{r}(n)$ is determined by a fair coin toss for each $n \in \omega$), is it possible to guess the value of $\psi(\hat{r})$ given the values $\psi(r)$ for all $r \neq \hat{r}$?”. The intuitively obvious answer is “no”. Since each value of ψ was chosen independently, the restriction of ψ to $I \setminus \{\hat{r}\}$ should carry no information about $\psi(\hat{r})$. Hence, if we are limited to information about $\psi \upharpoonright (I \setminus \{\hat{r}\})$, the odds of guessing $\psi(\hat{r})$ correctly should be 0, no matter what strategy we employ.

But are they? Consider the following argument, which can be fully formalized within Zermelo–Fraenkel set theory with Choice (ZFC). Define an equivalence relation on I^I by setting $f \sim g$ if and only if $f(r) = g(r)$ for all but finitely many $r \in I$. By AC, there exists an $S \subseteq I^I$ that intersects each \sim -class in one point. For each $g \in I^I$, let g^* be the unique function in $S \cap [g]$, and let Δ_g denote the finite set $\{r \in I : g(r) \neq g^*(r)\}$. Note that g^* is uniquely determined by the restriction of g to any cofinite subset of I . Hence, ψ^* is determined by $\psi \upharpoonright (I \setminus \{\hat{r}\})$, so we can employ the strategy of guessing that $\psi(\hat{r}) = \psi^*(\hat{r})$. This strategy fails if and only if $\hat{r} \in \Delta_\psi$. Since Δ_ψ is a finite set depending only on ψ (not on \hat{r}), any randomly chosen $\hat{r} \in I$ is almost certain not to lie in Δ_ψ . Thus, using only information about $\psi \upharpoonright (I \setminus \{\hat{r}\})$, our strategy is almost certain to guess the correct value of $\psi(\hat{r})$. This paradox invites us to reexamine the assumptions underlying ZFC.

2. Introduction

2.1. Axioms and Mathematical Intuition

Over the last 100 years, set theory (ZFC) has become widely accepted as a foundation for mathematics. If mathematics is to be a search for objective truth, then the correctness of these foundational axioms is essential. Axiomatic systems, such as ZFC, turn our mathematical intuitions into precise statements. Hence, their validity ultimately depends on whether these intuitions are correct. Further, mathematics will always have meaningful questions



Citation: Herman, A.; Caughman, J. Probability Axioms and Set Theory Paradoxes. *Symmetry* **2021**, *13*, 179. <https://doi.org/10.3390/sym13020179>

Received: 15 December 2020

Accepted: 20 January 2021

Published: 22 January 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

that cannot be resolved on the basis of currently accepted axioms (as shown by the MRDP theorem [3], which proves Hilbert's 10th problem to be unsolvable, stating that no consistent formal system can prove all true statements of the form " $\forall \bar{x} \in \omega^m (F(\bar{x}) \neq 0)$ " for $F \in \mathbb{Z}[\bar{X}]$). The desire to move beyond these limitations naturally motivates the invention of new axioms, which can only be justified using informal arguments. In this way, formal mathematics rests atop informal mathematics.

Of course, peoples' mathematical intuitions do not always agree. A prime example of this is the Axiom of Choice, first articulated by Zermelo in 1904. Most mathematicians find this axiom to be self-evident. Indeed, prior to its codification as part of ZFC, mathematicians had already been using it implicitly in their proofs (including even strong opponents of the axiom like Lebesgue [4]). Yet, this axiom has been the subject of rich controversy due to its non-constructive nature and counter-intuitive consequences. Further, accepting (or not accepting) the Axiom of Choice has significant implications for many areas of mathematics. Hence, it shows that the question of whether to accept a foundational principle can be both subtle and important. For a standard reference on the Axiom of Choice, as well as the other ZFC axioms, see [5].

Compared with other scientists, it may be especially important for mathematicians to take an active interest in the foundations of their subject. Since many mathematical statements cannot be empirically tested, incorrect mathematics will not necessarily "self-correct" as an incorrect theory of physics might. Therefore, the validity of mathematics rests on our willingness to critique our own basic assumptions. In this paper, we study two axioms that, while highly intuitive, are incompatible with ZFC. These axioms stem from simple, intuitive arguments concerning probability, and we believe they merit the consideration of anyone interested the foundations of mathematics.

2.2. Mathematical Background

The probability axioms explored in this article bear a close connection to Freiling's famous Axiom of Symmetry [6], concerning functions from $[0, 1]$ to countable subsets of $[0, 1]$:

Axiom 1 (cf. Axiom A_{\aleph_0} from [6]). *If $f : [0, 1] \rightarrow \mathcal{P}_{\leq \aleph_0}([0, 1])$, then there exist $x, y \in [0, 1]$ such that $x \notin f(y)$ and $y \notin f(x)$.*

Freiling motivates this axiom as follows: Suppose we choose $x, y \in [0, 1]$ uniformly at random. Since $f(y)$ is a countable set, it is almost certain that $x \notin f(y)$. By symmetry, it is almost certain that $y \notin f(x)$. Hence, for a randomly chosen pair, x, y , it is almost certain that both $x \notin f(y)$ and $y \notin f(x)$; and we conclude that some pair must satisfy this condition. Incredibly, under ZFC, this axiom is equivalent to the negation of the continuum hypothesis.

We are not the first to observe that this type of reasoning also has the potential to contradict ZFC itself. If we make the (plausible) assumption that any subset of $[0, 1]$ with cardinality $< 2^{\aleph_0}$ must have probability 0, we might suggest the following stronger principle:

Axiom 2 (cf. Axiom $A_{<2^{\aleph_0}}$ from [6]). *If $f : [0, 1] \rightarrow \mathcal{P}_{<2^{\aleph_0}}([0, 1])$, then there exist $x, y \in [0, 1]$ such that $x \notin f(y)$ and $y \notin f(x)$.*

However, as Freiling notes, this axiom implies that $[0, 1]$ cannot be well-ordered (assuming the other ZF axioms). Ultimately, Freiling rejects Axiom 2, saying that he knows of no compelling reason to support the belief that small sets have probability 0.

A second way in which Freiling's intuition may be used to argue against AC is explored in [7]. Van Lambalgen's approach is to add a new "almost all" quantifier, Q , to the language of set theory. Among van Lambalgen's axioms for this quantifier is the following, which directly formalizes Freiling's symmetry argument:

Axiom 3 (cf. Axiom Q6 from [7]). For any formula ϕ , $QxQy\phi \leftrightarrow QyQx\phi$.

In [7], van Lambalgen shows that his axioms for Q , together with ZF, imply that 2^ω is not well-orderable.

The results of this paper show further ways in which Freiling's reasoning can conflict with ZFC. We note two important ways in which our arguments differ from those just mentioned. First, both Freiling and van Lambalgen derive a contradiction with ZFC by arguing that a set equinumerous with the continuum cannot be well-ordered. By contrast, our contradictions require only the following (even weaker) special case of AC.

Axiom 4 (Weak Axiom of Choice). Every partition of 2^ω into countable sets has a set of representatives.

Second, we will reconsider the usual axioms of probability. We will argue that some of Kolmogorov's axioms do not have a clear justification, and that this calls into question any argument against ZFC based on them. We will then present a more parsimonious framework for probability and derive our contradictions within that framework.

2.3. Freiling's Argument for $\neg CH$

We are greatly indebted to Freiling for his innovative work on the continuum hypothesis. In [6], he proposed a negative solution to CH based on intuitive probability axioms, such as Axiom 1. We view our axioms as natural extensions of his.

The eventual aim of this paper is to show that Freiling's reasoning (taken to its logical conclusion) is incompatible with ZFC, which he assumes in his argument. Therefore, we believe that Freiling's approach to CH does not ultimately work. Nonetheless, his argument is delightful and we include a version of it here for context. The remainder of this section is independent of the rest of the paper and may be skipped without any loss of understanding.

Theorem 1 (Adapted from [6]). For all $n < \omega$, there exists $A_n \subseteq \omega_n^{n+2}$ such that:

- (i) For all $\bar{x} \in \omega_n^{n+1}$, $|\{y \in \omega_n : \bar{x} \cap y \in A_n\}| < \infty$
- (ii) $\bigcup_{\sigma \in \text{Sym}(n+2)} \sigma \cdot A_n = \omega_n^{n+2}$

Proof. We proceed by induction. For $n = 0$, we may take $A_0 = \{(m, n) \in \omega^2 : n \leq m\}$. Now, assume the result holds for $n - 1$. For each $\alpha < \omega_n$, there exists an injection $f_\alpha : \omega_{n-1} \rightarrow \omega_n$ such that $\alpha + 1 \subseteq \text{im}(f_\alpha)$. Let $A_n = \{(\alpha, f_\alpha(\beta_0), \dots, f_\alpha(\beta_n)) : \alpha < \omega_n \wedge (\beta_0, \dots, \beta_n) \in A_{n-1}\}$. First, let $G \leq \text{Sym}(n+2)$ be the stabilizer of 0. Then it follows from the inductive hypothesis that $\bigcup_{\sigma \in G} \sigma \cdot A_n \supseteq \{(\alpha, \gamma_0, \dots, \gamma_n) \in \omega_n^{n+2} : \forall i (\alpha \geq \gamma_i)\}$. Therefore, $\bigcup_{\sigma \in \text{Sym}(n+2)} \sigma \cdot A_n = \omega_n^{n+2}$ as desired. \square

The following is adapted from [6]:

Axiom 5. Let $I = [0, 1]$ (or 2^ω). For every positive integer, n , and for any $A \subseteq I^n$ the following cannot simultaneously be true:

- (i) For all $\bar{x} \in I^{n-1}$, $|\{y \in 2^\omega : \bar{x} \cap y \in A\}| < \infty$
- (ii) $\bigcup_{\sigma \in \text{Sym}(n)} \sigma \cdot A = I^n$

The justification for Axiom 5 is similar to that given for Axiom 1. Suppose we choose $(x_0, \dots, x_{n-1}) \in I^n$ by choosing each x_i uniformly at random and independently. If we choose x_0, \dots, x_{n-2} first, then by (i), it is almost certain that x_{n-1} will be such that $(x_0, \dots, x_{n-1}) \notin A$. Yet, the order in which we choose the x_i should be irrelevant, so we conclude that (x_0, \dots, x_{n-1}) is almost certainly not in A , period. By symmetry, for any $\sigma \in \text{Sym}(n)$, $(x_{\sigma(0)}, \dots, x_{\sigma(n-1)})$ is almost certain not to lie in A . It follows that (x_0, \dots, x_{n-1}) is almost certain not to lie in $\bigcup_{\sigma \in \text{Sym}(n)} \sigma \cdot A$, which implies this union does not equal I^n .

Corollary 1. *ZFC + Axiom 5 implies $2^{\aleph_0} \geq \aleph_{\omega+1}$*

Proof. Suppose there is a bijection between 2^ω and \aleph_n , for some $n < \omega$. This, together with Theorem 1, implies that there is a counter-example to Axiom 5 (for the $n + 2$ case). Hence, $2^{\aleph_0} \geq \aleph_\omega$. By König's theorem (a standard result in set theory [5]), 2^{\aleph_0} has uncountable cofinality. Hence, $2^{\aleph_0} \neq \aleph_\omega$. \square

2.4. Revisiting Kolmogorov's Axioms

If we wish to use probability as a source of foundational principles, as Freiling does, we are obliged to examine the assumptions underlying probability theory itself. Modern probability was formalized in the 1930's by Kolmogorov, using the concepts of measure theory. Despite its great success, the measure-theoretic framework is not beyond scrutiny. In particular, we believe that Kolmogorov's axioms cannot be reasonably deduced from first principles. According to Kolmogorov, a **probability space** is a triple (X, Σ, \mathbb{P}) , where X is a set, $\Sigma \subseteq \mathcal{P}(X)$ is a σ -algebra and $\mathbb{P} : \Sigma \rightarrow \mathbb{R}$ is a (σ -additive) measure satisfying $\mathbb{P}(X) = 1$ (for more on probability spaces, see [8]).

Kolmogorov makes two significant assumptions that we believe are not well justified. The first is that probabilities are elements of \mathbb{R} , rather than some other ordered field extending \mathbb{Q} , such as a hyperreal field (this is explored in [9]). The modern conception of \mathbb{R} was motivated by the desire to give a rigorous foundation to analysis (which was primarily developed for physics). Thus, the applicability of real numbers to probability is by no means self-evident. In fact, there are examples of events which *prima facie* appear to require nonzero, infinitesimal probabilities; an impossibility in Archimedean fields such as \mathbb{R} . Hence, if real numbers are well-suited to probability, this ought to be argued for, rather than taken as a hypothesis.

Secondly, we can examine Kolmogorov's assumptions regarding the structure of measurable sets. It is a point in his favor that he does not assume that all subsets of X have defined probabilities (a reasonable precaution given results such as the Banach–Tarski theorem [10]). However, it would then be logical to call a set “measurable” precisely when its probability can be calculated using some set of basic assumptions (e.g., additivity for disjoint sets).

On the other hand, there seems to be no clear reason why measurable sets should be closed under unions and intersections. There is no general method for assigning a probability to the union or intersection of two events in terms of their individual probabilities. Hence, the assumption that measurable sets form an algebra (let alone a σ -algebra) needlessly constrains the universe of allowed probability spaces, and may exclude some desirable spaces

There are clear practical advantages to having measurable sets be closed under unions and intersections. However, we have good reason not to accept axioms simply because they are plausible and convenient. In the case of \mathbb{R}^3 , it is quite reasonable to think that ordinary volume ought to be invariant under the operation of breaking a set into two pieces and applying a Euclidean transformation to one of those pieces such that it remains disjoint from the other. Hence, one may be tempted to assert that the family of measurable sets be closed under this type of operation. Yet, it is well-known that (assuming AC) this cannot hold for any measure that assigns the usual volumes to all boxes. Such examples urge us to use extreme caution when asserting closure properties for measurable sets.

3. Minimalist Probability and New Axioms

3.1. Uniform Probability on 2^ω

For the reasons given above, we choose for our basic framework a formulation of probability that differs somewhat from Kolmogorov's. In particular, we will not insist that measurable sets form an algebra, nor will we insist that probabilities be real numbers. We will then show that the conflicts between ZFC and our probability axioms arise, even within this simplified framework. (Another approach to probability, which is more philosophically

cautious than Kolmogorov's, is *qualitative probability*. In this formulation, a “not more likely” relation on events, \leq , is axiomatized, rather than a probability function. For one possible axiomatization, see [11]. We note that our axioms and results can be adapted to this setting.)

Our definitions and axioms are motivated by a prototypical example: selecting a random element from 2^ω by flipping a fair coin for each $n \in \omega$; and associating 0 and 1 to heads and tails, respectively. We refer to this *intuitive* idea as the **fair coin space**.

For at least some $A \subseteq 2^\omega$, the probability that a random element of the fair coin space lies in A seems to have an exact answer. For example, if $A = \{x \in 2^\omega : x_0 = x_3 = 0\}$, then membership in A depends only on the outcomes of two of the coins and thus appears to be $\frac{1}{4}$.

In the remainder of this section, we present several new definitions and axioms, which are motivated by further considering the fair coin space and exploiting its symmetry (the coins are fair and identical) and independence (the coins do not influence each other).

Definition 1. Let (R, \leq) be any total order extending (\mathbb{Q}, \leq) . A **minimalist probability space (MPS)** is a pair, $\Sigma \subseteq \mathcal{P}(2^\omega)$, $\mathbb{P} : \Sigma \rightarrow R$, satisfying the conditions below.

(i) If $\Delta \subseteq \omega$ is finite, and $\delta \in 2^\Delta$, then

$$\mathbb{P}(\{z \in 2^\omega : z \upharpoonright \Delta = \delta\}) = 1/2^{|\Delta|}.$$

(ii) If $A, B \in \Sigma$ and $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$.

We refer to 2^ω as the **sample space**, Σ as the **event space** and \mathbb{P} as the **probability function**. Elements of Σ and $\mathcal{P}(2^\omega) \setminus \Sigma$ are called **measurable** and **non-measurable** sets, respectively.

For convenience, we will avoid explicit mention of the set Σ when the meaning is still clear. Thus, expressions such as, “ $\mathbb{P}(A) = \mathbb{P}(B)$ ”, should be understood to mean that either both sides of the equation are defined and equal or that they are both undefined.

Let $I \subseteq \omega$ and let $A = \{x \in 2^\omega : x \upharpoonright I \in A_0\}$ for some $A_0 \subseteq 2^I$. Given $x \in 2^\omega$, whether or not $x \in A$ depends *only* on the restriction of x to I . Therefore, we may reasonably identify “the probability that a random element of 2^I is in A_0 ” with the “probability that a random element of 2^ω is in A ”. In fact, with a slight abuse of notation, we will write $\mathbb{P}(A_0)$ to mean $\mathbb{P}(A)$.

Notation 1. Let $I \subseteq \omega$, and let $A \subseteq 2^I$. We will write $\mathbb{P}(A)$ as a shorthand for $\mathbb{P}(\{x \in 2^\omega : x \upharpoonright I \in A\})$.

Notation 2. Let $I, J \subseteq \omega$ be disjoint, and let $A \subseteq 2^{I \cup J}$. For $x \in 2^I$, we will use A_x to denote the set $\{y \in 2^J : x \cup y \in A\}$.

Remark 1. It will frequently be convenient to replace the set 2^ω with 2^N , where N is some countably infinite set. We will call the resulting space an **MPS on 2^N** . Similarly, if I and J are countable sets, we can speak of an MPS on $2^I \times 2^J$, using the natural identification between this space and $2^{I \sqcup J}$.

3.2. Two Axioms

We now introduce two new intuitively appealing axioms. In Sections 4.1 and 4.2, we show that each of these is incompatible with Axiom 4.

Definition 2. An MPS has the **Freiling property** if the following holds: Let I and J be disjoint subsets of ω . Let $A \subseteq 2^{I \cup J}$. If for every $x \in 2^I$, $\mathbb{P}(A_x) = q$, then $\mathbb{P}(A) = q$.

We believe that this definition, which is based on Freiling's Axiom of Symmetry [6], ought to hold in the fair coin space. To argue this, we will apply a variant of Freiling's own argument. Suppose I, J, A and q are as in Definition 2. Since the fair coin space is symmetric, it should

not matter in what order we flip the coins when choosing a random element. Therefore, let us start with coins corresponding to elements of I , followed by those corresponding to elements of J . Once $x \in 2^I$ has been determined, our hypothesis tells us that the probability that we will choose $y \in 2^J$ such that $x \cup y \in A$ is exactly q . This is true regardless of the value of x , and hence we may say that it is true *before x is chosen*. Thus, $\mathbb{P}(A) = q$ as claimed.

Definition 3. Let N be a countable set. We let $\mathcal{G}(2^N)$ denote the group of transformations on 2^N generated by functions, $\sigma : 2^N \rightarrow 2^N$, such that either

- (i) there exists an $r \in 2^N$ such that $\sigma(x) = x \oplus r$ for all $x \in 2^N$, where \oplus denotes the bitwise XOR operation, or
- (ii) there exists a permutation $\pi : N \rightarrow N$ such that $\sigma(x) = x \circ \pi$ for all $x \in 2^N$.

A **probability symmetry** of 2^N is an element of $\mathcal{G}(2^N)$. We let \cdot denote the natural action of $\mathcal{G}(2^N)$ on 2^N .

Definition 4. An MPS on 2^ω has the **dependent symmetry property** if the following holds: Let $I \cup J$ be any bipartition of ω ; we will identify 2^ω with $2^I \times 2^J$ in the natural way. Then $\mathbb{P}([A \times B] \cup C) = \mathbb{P}([A \times \sigma \cdot B] \cup C)$, for all $A \subseteq 2^I$, $B \subseteq 2^J$, $C \subseteq \overline{A} \times 2^J$ and $\sigma \in \mathcal{G}(2^J)$.

Again, we will argue that this definition ought to be satisfied by the fair coin space. Fix $\sigma \in \mathcal{G}(2^J)$. Consider randomly choosing an element $(x, y) \in 2^I \times 2^J$ using two different methods. The first is just that of the fair coin space. For the second, suppose we choose x and y just as before, but if $x \in A$, we replace y with $\sigma^{-1} \cdot y$ (so the choice of whether to apply σ^{-1} does not depend on y). From the point of view of probability, there ought to be no difference between these methods of choosing (x, y) . Hence, the probabilities that points chosen by each method lie in $(A \times B) \cup C$ should be equal if they are defined. This motivates the equality in Definition 4.

Since we believe that the fair coin space is an MPS having the properties given in Definitions 2 and 4, we propose the following axioms.

Axiom 6. There exists an MPS with the Freiling property.

Axiom 7. There exists an MPS with the dependent symmetry property.

Both of these axioms assert the existence of sets with specific properties. Moreover, we have given informal arguments as to why these properties ought to be satisfiable. Hence, our axioms may be viewed as examples of the *maximize* principle explored by Maddy in [4]. This principle essentially states that the set theoretic universe ought to be very “full”, containing as many sets as possible without generating contradictions (e.g., as one gets by declaring $\{x : x \notin x\}$ a set).

Remark 2. The usual product measure on 2^ω satisfies Definition 1 (though not Axioms 6 or 7). Hence, our arguments may be adapted to that setting.

4. Results

4.1. A Vitali-Type Paradox

In this and the following sections, we will assume the axioms of Zermelo–Fraenkel set theory without Choice or Replacement. When axioms beyond this are used, we will say so explicitly.

Lemma 1. Let (Σ, \mathbb{P}) be an MPS with the Freiling property. Let $\Delta \subseteq \omega$ be finite, and let $A \subseteq 2^\omega$ have the following property:

For every $f : \omega \setminus \Delta \rightarrow 2$, there exists a (unique) $\delta_f : \Delta \rightarrow 2$, such that for all $g : \Delta \rightarrow 2$, $f \cup g \in A \Leftrightarrow g = \delta_f$.

Then $\mathbb{P}(A) = 1/2^{|\Delta|}$.

Proof. This is immediate from Definitions 1 (i) and 2. \square

Theorem 2. Axiom 4 and Axiom 6 are incompatible.

Proof. Let (Σ, \mathbb{P}) be an MPS on $2^\omega \times 2^\omega$ with the Freiling property. Let \sim be the equivalence relation on 2^ω given by

$$f \sim g \Leftrightarrow |\{n \in \omega : f(n) \neq g(n)\}| < \infty. \quad (1)$$

By Axiom 4, there exists a set of representatives, $\Gamma \subseteq 2^\omega$, for the equivalence classes of 2^ω . We will view ω as a subset of 2^ω by identifying elements of ω with their binary representations (padding with 0's). Then for every $r \in 2^\omega$, there are unique $r^* \in \Gamma$, $r' \in \omega$ such that $r = r' \oplus r^*$. Given $r \in 2^\omega$ and $m \in \omega$, let r_m denote the m^{th} bit of r . Let $A = \{(m, n) \in \omega \times \omega : n_{m+1} = 1\}$ and $B = \{(m, n) \in \omega \times \omega : m_n = m_{n+1} = 0\}$. Let $\Omega_A = \{(r, s) \in 2^\omega \times 2^\omega : (r', s') \in A\}$ and $\Omega_B = \{(r, s) \in 2^\omega \times 2^\omega : (r', s') \in B\}$.

Fix $r \in 2^\omega$, and let $m = r'$ and $(\Omega_A)_r = \{s \in 2^\omega : (r, s) \in \Omega_A\}$. Then for any $s \in 2^\omega$, $(r, s) \in \Omega_A$ iff $(m, s') \in A$ iff $s'_{m+1} = 1$, so $(\Omega_A)_r = \{s \in 2^\omega : s_{m+1} = s^*_{m+1}\}$. By Lemma 1, $\mathbb{P}((\Omega_A)_r) = \frac{1}{2}$. By Definition 2, $\mathbb{P}(\Omega_A) = \frac{1}{2}$. An analogous argument gives $\mathbb{P}(\Omega_B) = \frac{1}{4}$.

We will show that $A \subseteq B$. Let $(m, n) \in A$. Then $n_{m+1} = 1$, which implies $n \geq 2^{m+1}$. If $m_k = 1$ for any $k \geq n$, then $m \geq 2^k \geq 2^n \geq 2^{2^{m+1}}$, which is impossible. Hence, $m_k = 0$ for all $k \geq n$; thus $m_n = m_{n+1} = 0$, so $(m, n) \in B$. This implies $\Omega_A \subseteq \Omega_B$. Therefore, $\mathbb{P}(\Omega_A) \leq \mathbb{P}(\Omega_B)$, a contradiction. \square

4.2. The Banach–Tarski Paradox in 2^ω

In this section, we let $F_2 = F(a, b)$ denote the free group on two generators. The following is an adaptation of the proof of the famous Banach–Tarski paradoxical decomposition [10].

Lemma 2 (Dependent Symmetries). Let $I \cup J$ be a bipartition of ω . Let (Σ, \mathbb{P}) be an MPS on $2^\omega (\approx 2^I \times 2^J)$ with the dependent symmetry property. Let $A_1, \dots, A_n \subseteq 2^I$ be pairwise disjoint, and let $B_1, \dots, B_n \subseteq 2^J$. If $\sigma_i \in \mathcal{G}(2^J)$ for each $i \leq n$, then $\mathbb{P}(\bigcup_{i=1}^n A_i \times B_i) = \mathbb{P}(\bigcup_{i=1}^n A_i \times (\sigma_i \cdot B_i))$.

Proof. This follows immediately from n applications of Axiom 7. \square

Lemma 3. There exists a bijection $\pi : F_2 \rightarrow F_2$ such that $\pi(e) = e$ and for all $g, h \in F_2 \setminus \{e\}$, there exists an integer $n > 1$ such that $\pi(g^n) = h^n$.

Proof. Let $\{(g_i, h_i) : i < \omega\}$ be an enumeration of $(F_2 \setminus \{e\})^2$. First, note that for every $g \in F_2 \setminus \{e\}$, the set $\{g^n : n \in \mathbb{Z}_+\}$ is infinite. For $i < \omega$, recursively define n_i to be the least integer ≥ 2 such that $g_i^{n_i} \notin \{g_j^{n_j} : j < i\}$ and $h_i^{n_i} \notin \{h_j^{n_j} : j < i\}$. Let $\pi' : X \rightarrow X$ be the partial function $\pi' = \{(e, e)\} \cup \{(g_i^{n_i}, h_i^{n_i}) : i < \omega\}$. Since $F_2 \setminus \{e, g_i^{n_i}, h_i^{n_i} : i < \omega\}$ is infinite (e.g., it contains $a^n b$ for all n), we can extend π' to a bijection $\pi : F_2 \rightarrow F_2$. \square

Lemma 4. Let G be a group acting on a set, X . Let $Y = \{x \in X : |G_x| = 1\}$. Then G fixes Y as a set, and G acts freely on Y .

Proof. Routine. \square

Lemma 5 (Axiom 4). There exist sets $A, B \subseteq 2^{F_2}$ and $\alpha, \beta, \gamma \in \mathcal{G}(2^{F_2})$ such that

$$(i) \quad \gamma \cdot (A^C \cap B^C) \subseteq A \cup B,$$

- (ii) $\beta^m \cdot A \cap \beta^n \cdot A = \emptyset$ for all $m \neq n \in \omega$ and
 (iii) $\alpha^m \cdot B \cap \alpha^n \cdot B = \emptyset$ for all $m \neq n \in \omega$

Proof. By Definition 3, every permutation $\pi \in \text{Sym}(F_2)$, induces a probability symmetry of 2^{F_2} , given by $f \mapsto f \circ \pi^{-1}$. Therefore, there is a natural embedding $\text{Sym}(F_2) \hookrightarrow \mathcal{G}(2^{F_2})$. Moreover, via Cayley's theorem, there is a natural embedding $F_2 \hookrightarrow \text{Sym}(F_2)$. Therefore, we have $F_2 \hookrightarrow \text{Sym}(F_2) \hookrightarrow \mathcal{G}(2^{F_2})$. We will identify F_2 and $\text{Sym}(F_2)$ with their images in $\mathcal{G}(2^{F_2})$. We let F_2 act on 2^{F_2} using this identification.

Let $S = \{y \in 2^{F_2} : \exists h \in F_2 \setminus \{e\} (h \cdot y = y)\}$ and let $S^C = 2^{F_2} \setminus S$. By Lemma 4, F_2 acts freely on S^C . Let Γ be a set of representatives for the F_2 -orbits of S^C . Let $W_c \subseteq F_2$ be the set of reduced words starting with c (for $c \in \{a, b, a^{-1}, b^{-1}\}$). Let $X_a = W_a \cup W_{a^{-1}} \cup \{e\}$, $X_b = W_b \cup W_{b^{-1}} \cup \{e\}$. Let $A = X_a \cdot \Gamma$ and $B = X_b \cdot \Gamma$. Then $S = A^C \cap B^C$.

Let $\pi \in \text{Sym}(F_2) \leq \mathcal{G}(2^{F_2})$ be as in Lemma 3; let $\phi : 2^{F_2} \rightarrow 2^{F_2}$ be the map that flips the e^{th} bit. Let $\gamma = \phi \circ \pi \in \mathcal{G}(2^{F_2})$. For any $y \in S$, we have $g \cdot y = y$ for some $g \in F_2 \setminus \{e\}$. Fix $h \in F_2 \setminus \{e\}$. By Lemma 3, $\pi(g^n) = h^n$ for some $n > 1$. Hence, $(h^n \gamma \cdot y)(e) = (\gamma \cdot y)(h^n) = (\pi \cdot y)(h^n) = y(g^n) = y(e) = (\pi \cdot y)(e) \neq (\gamma \cdot y)(e)$. Therefore, $h^n \cdot (\gamma \cdot y) \neq \gamma \cdot y$, so h does not fix $\gamma \cdot y$. Since h was arbitrary, $\gamma \cdot y \in S^C$. Since $y \in S$ was arbitrary, $\gamma \cdot S \subseteq S^C = A \cup B$.

Let α, β be a, b , respectively. To complete the proof, it is sufficient (by symmetry) to show that $b^m \cdot A \cap b^n \cdot A = \emptyset$ for all $m \neq n \in \omega$. Suppose $y \in b^m \cdot A \cap b^n \cdot A$. Then $y = (b^m g_1) \cdot z_1 = (b^n g_2) \cdot z_2$, for some $z_1, z_2 \in \Gamma$ and $g_1, g_2 \in X_a$. Then z_1 and z_2 are in the same F_2 -orbit of S^C , and so in fact $z_1 = z_2 = z$. We now have $(g_2^{-1} b^{m-n} g_1) \cdot z = z$, which implies $g_2^{-1} b^{m-n} g_1 = e$, since $z \in S^C$. Thus, $b^{m-n} = g_2 g_1^{-1} \in X_a$, which yields $m = n$. \square

Theorem 3. Axioms 4 and 7 are incompatible.

Proof. Let x, y, z be three elements not in F_2 , and let (Σ, \mathbb{P}) be an MPS on $2^{\{x, y, z\}} \times 2^{F_2}$ with the dependent symmetry property. Let $A, B \subseteq 2^{F_2}$ and $\alpha, \beta, \gamma \in \mathcal{G}(2^{F_2})$ be as in Lemma 5; let $S = (A \cup B)^C$. Furthermore, let $S_A = S \cap \gamma^{-1} \cdot A$ and $S_B = S \cap \gamma^{-1} \cdot B$. Let

- (i) $P_0 = S_A, P_1 = S_B \setminus S_A, P_2 = A, P_3 = B \setminus A$
 (ii) $Q_0 = \beta \gamma \cdot S_A, Q_1 = \alpha \gamma \cdot (S_B \setminus S_A), Q_2 = \beta^2 \cdot A, Q_3 = \alpha^2 \cdot (B \setminus A)$
 (iii) $R_0 = \beta^3 \gamma \cdot S_A, R_1 = \alpha^3 \gamma \cdot (S_B \setminus S_A), R_2 = \beta^4 \cdot A, R_3 = \alpha^4 \cdot (B \setminus A)$

Note that $P_0 \cup P_1 \cup P_2 \cup P_3 = 2^{F_2}$ and $Q_0 \cup Q_1 \cup Q_2 \cup Q_3 \cup R_0 \cup R_1 \cup R_2 \cup R_3 \subseteq F_2$ are both pairwise disjoint unions. Using Lemma 5, we have the following:

$$\frac{1}{4} = \mathbb{P}(2^{\{x\}} \cap 00 \cap 2^{F_2}) = \mathbb{P}([000 \cap 2^{F_2}] \cup [100 \cap 2^{F_2}]),$$

by Definition 1 (i). Partitioning the set, this is equal to

$$\mathbb{P}([000 \cap P_0] \cup [000 \cap P_1] \cup [000 \cap P_2] \cup [000 \cap P_3] \cup [100 \cap P_0] \cup [100 \cap P_1] \cup [100 \cap P_2] \cup [100 \cap P_3]),$$

which, by Lemma 2, equals

$$\mathbb{P}([000 \cap P_0] \cup [001 \cap P_1] \cup [010 \cap P_2] \cup [011 \cap P_3] \cup [100 \cap P_0] \cup [101 \cap P_1] \cup [110 \cap P_2] \cup [111 \cap P_3]).$$

By Lemma 2 again, this is

$$\mathbb{P}([000 \cap Q_0] \cup [001 \cap Q_1] \cup [010 \cap Q_2] \cup [011 \cap Q_3] \cup [100 \cap R_0] \cup [101 \cap R_1] \cup [110 \cap R_2] \cup [111 \cap R_3]),$$

which is further equal to

$$\mathbb{P}([000 \smallfrown Q_0] \cup [000 \smallfrown Q_1] \cup [000 \smallfrown Q_2] \cup [000 \smallfrown Q_3] \cup [000 \smallfrown R_0] \cup [000 \smallfrown R_1] \cup [000 \smallfrown R_2] \cup [000 \smallfrown R_3]).$$

By Definition 1 (iii), this is at most $\mathbb{P}(000 \smallfrown 2^{F_2})$, which equals $\frac{1}{8}$, by Definition 1 (i); a contradiction. \square

5. Conclusions

We have shown that the Axiom of Choice is incompatible with highly intuitive assumptions about probability (Axiom 6 or Axiom 7). Our work improves on other similar results in two important ways. First, our contradictions rely on a weaker version of AC than that used in [6,7], which both assume the well-orderability of \mathbb{R} . Second, we analyze the philosophical “weak points” in Kolmogorov’s axiomatization of probability, and show that our results do not depend on these.

While the standard resolution to these paradoxes is to reject Axioms 6 and 7, we believe that there are three plausible options:

- (i) reject the Infinity or Powerset axioms, so 2^ω cannot be constructed,
- (ii) reject Axiom 4, or
- (iii) reject Axioms 6 and 7.

For those with finitistic leanings, option (i) will be the obvious choice, and our paradoxes may be viewed as consequences of invalid reasoning about “completed infinities”. For mathematicians who accept completed infinities, but believe that valid mathematical objects must be explicitly constructed, option (ii) may be appealing.

Finally, if we reject Axioms 6 and 7, we can keep all of ZFC. Unfortunately, these axioms seem to follow immediately from our most basic intuitions about random events. Therefore, the authors believe that rejecting them amounts to a rejection of randomness as a valid mathematical concept. This leads one to question whether probability can be meaningfully formalized at all.

Each of these attempts at a resolution has drawbacks, and it is not clear to the authors what approach is best. The aim of this paper has been to make these issues more widely understood, especially by mathematicians working outside of set theory. We hope that in so doing, we will stimulate an interesting conversation about the conflicts between probability, infinite sets, and the Axiom of Choice.

Author Contributions: Conceptualization, A.H.; writing—original draft preparation, A.H.; writing—review and editing, A.H. and J.C.; supervision, J.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank R. Koncel-Kedziorski, J. D. Hamkins, and the anonymous referees for their helpful feedback on the early drafts of this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Galvin, F. Problem 5348. *Am. Math. Mon.* **1965**, *72*, 1136.
- Hardin, C.S.; Taylor, A.D. An introduction to infinite hat problems. *Math. Intell.* **2008**, *30*, 20–25. [[CrossRef](#)]
- Matijasevic, J.V. Enumerable sets are diophantine. *Soviet Math. Dokl.* **1970**, *11*, 354–358.
- Maddy, P. Believing the axioms. I. *J. Symb. Log.* **1988**, *53*, 481–511. [[CrossRef](#)]
- Jech, T. *Set Theory: The Third Millennium Edition, Revised and Expanded*; Springer: Berlin, Germany, 2003.

6. Freiling, C. Axioms of symmetry: Throwing darts at the real number line. *J. Symb. Log.* **1986**, *51*, 190–200. [[CrossRef](#)]
7. Van Lambalgen, M. Independence, randomness and the axiom of choice. *J. Symb. Log.* **1992**, *57*, 1274–1304. [[CrossRef](#)]
8. Chung, K.L.; Zhong, K. *A Course in Probability Theory*; Academic Press: San Diego, CA, USA, 2001.
9. Hofweber, T.; Schindler, R. Hyperreal-valued probability measures approximating a real-valued measure. *Notre Dame J. Form. Log.* **2016**, *57*, 369–374. [[CrossRef](#)]
10. Banach, S.; Tarski, A. Sur la décomposition des ensembles de points en parties respectivement congruentes. *Fund. Math.* **1924**, *6*, 244–277. [[CrossRef](#)]
11. Villegas, C. On qualitative probability. *Am. Math. Mon.* **1967**, *74*, 661–669. [[CrossRef](#)]