A Paradox Arising from the Elimination of a Paradox

Alan D. Taylor and Stan Wagon

Abstract. We present a result of Mycielski and Sierpiński—remarkable and under-appreciated in our view—showing that the natural way of eliminating the Banach–Tarski Paradox by assuming all sets of reals to be Lebesgue measurable leads to another paradox about division of sets that is just as unsettling as the paradox being eliminated. The *Division Paradox* asserts that the reals can be divided into nonempty classes so that there are more classes than there are reals.

1. INTRODUCTION. The Zermelo–Fraenkel axioms (ZF) were introduced a century ago to avoid logical paradoxes, notably Bertrand Russell's: Does the set consisting of every set that is not a member of itself contain itself? The Axiom of Choice (AC)—first stated by Zermelo in 1904—asserts that for every set M of disjoint nonempty sets, there exists a set consisting of exactly one element from each set in M. When the AC is added to ZF, the resulting system is called ZFC. Excellent books about the historical and technical aspects of AC are [**6**, **7**, **11**].

Working mathematicians don't construct formal proofs in ZFC or any other system. They prove theorems by finding arguments that their peers find mathematically convincing. Nevertheless, these proofs always correspond to formal derivations in ZFC (or, rarely, other logical systems). And while we now know hundreds of reasonable assertions (e.g., the Continuum Hypothesis) that ZFC can neither prove nor disprove, there is no assertion that the set-theoretic community feels is so compelling that it should be added to ZFC as a new axiom. In short, ZFC appears to be strong enough to prove everything that is "obvious" (an attempt to define this term is in Section 7).

Now, some mathematicians harbor a nagging fear that AC might be too powerful. The issue is not that AC might yield an inconsistency—Gödel proved that it does not—but rather that it leads to inconveniences such as nonmeasurable sets and more to the point, strikingly counterintuitive results. A recent example is the work of Hardin and Taylor [**4**, **5**], but really nothing underscores the point more than the Banach–Tarski Paradox, a theorem of ZFC proved in 1924; see [**25**]. This is hardly surprising: the idea that a solid ball can be divided into five pieces that are moved by rigid motions to form two balls congruent to the first is unsettling; it defies all reason, at least for those who are uncomfortable with the concept of a nonmeasurable set. And the culprit is indeed AC. For example, noted physicist and mathematician Roger Penrose wrote [**16**, pp. 14–15]:

> "Most mathematicians would probably regard the Axiom of Choice as 'obviously true', while others may regard it as a somewhat questionable assertion which might even be false (and I am myself inclined, to some extent, towards this second viewpoint)."

He then (p. 366) goes on to say:

> "It's not altogether uncontroversial that the Axiom of Choice should be accepted as something that is universally valid. My own position is to be cautious about it. The trouble with this axiom is that it is a pure 'existence' assertion, without any hint of a rule whereby the set might be specified. In fact, it has a number of alarming consequences. One of these is the Banach–Tarski theorem."

And this by the philosopher and mathematician Solomon Feferman [**3**]:

> "I think there is still something very disturbing about the Banach–Tarski paradox…The conflict between common-sense geometrical intuition and the Banach–Tarski paradox seems so egregious that it may force one to question the very basic intuitions about arbitrary sets which lead one to accept the principles lying behind the paradox, namely the principles of Zermelo–Fraenkel Set Theory together with the Axiom of Choice—or, if not that, then at least the relevance of those principles to applicable mathematics."

To eliminate the paradox, one might take one of the following approaches: (1) Eliminate AC as an axiom and just use ZF (or a modest extension of ZF), or (2) Add to ZF the assertion that all sets are Lebesgue measurable. With either approach the Banach–Tarski Paradox evaporates (see note 1 in Section 8). Our purpose here is to present and prove a result of Mycielski and Sierpiński that is not generally known and shows that options (1) and (2) each have a serious drawback involving a statement that is itself a paradox. Since the use of ZFC has not proven to be a problem for mathematics—we have learned to live with nonmeasurable sets—this underscores the accepted view that ZFC is the proper foundation for mathematics.

The modest extension alluded to in (1) refers to the need of a weakened version of AC so that, for example, Lebesgue measure (denoted by λ) works as expected: ZF alone does not yield a proof that λ is countably additive. Typically one adds the Axiom of Dependent Choice (DC): If $*$ is a binary relation on a nonempty set X and for every $x \in X$ there is $y \in X$ with $x * y$, then there is a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n * x_{n+1}$ for every $n \in \mathbb{N}$ (see note 2). Countable additivity of λ follows from the Axiom of Choice for countable families of arbitrary nonempty sets, a consequence of DC.

To set the stage for what we will present, consider the National Football League, which has 32 teams and $53 \cdot 32 = 1696$ players. If the players were assigned to teams in some other way, subject only to the conditions that a team cannot have zero players and each player can be on only one team, then there can certainly be more than 32 teams. But could there be more than 1696 teams? Of course not. The idea of grouping players into nonempty teams so that there are more teams than players is ludicrous. It's like finding a country that has more populated provinces than it has people. Yet this is the essence of the phenomenon that we will present: this sort of thing can arise in a mathematical world without AC.

The main assertion we study here is analogous to the sports league example. The players are the real numbers with two reals placed on the same team if and only if they differ by a rational. That is, we will look at $\mathbb{R}/\mathbb{Q}$, the quotient group of the additive group of reals using the subgroup of rationals. Consider the statement that $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$ (where $|\cdot|$ is cardinality; see Section 2). This says that there are more equivalence classes of reals than there are reals; we call this assertion the *Division Paradox*.

2. PRELIMINARIES. The $\mathbb{R}/\mathbb{Q}$ equivalence relation has $x \sim y$ if and only if $x - y \in \mathbb{Q}$; we use $[x]$ for the equivalence class $x + \mathbb{Q}$ of a real x. This relation was used by Vitali to construct the first nonmeasurable set. He used AC to get a set containing exactly one real from each class in $\mathbb{R}/\mathbb{Q}$; such a set is not Lebesgue measurable. Our interest is in sets that are somewhat the opposite of what Vitali considered. A set A of reals is *Q-invariant* if $A = A + q$ for every rational q; such a set is a union of some classes in $\mathbb{R}/\mathbb{Q}$.

Cardinality is denoted by $|\cdot|$; in a choice-challenged world its definition involves pairs of sets: $|X| \le |Y|$ means that there is a one-one function $f : X \to Y$; $|X| = |Y|$ means that $|X| \le |Y|$ and $|Y| \le |X|$ (the classic Schröder–Bernstein Theorem, which is quite constructive and a theorem of ZF, implies that there is then a bijection from X to Y); $|X| < |Y|$ means that $|X| \le |Y|$ and $|X| \ne |Y|$. Under the Axiom of Choice every set has a cardinality from the well-ordered collection $0, 1, 2, \ldots, \aleph_0, \aleph_1, \aleph_2, \ldots$; but without AC, there can be incomparable sets: X and Y such that both $|X| \le |Y|$ and $|Y| \le |X|$ are false.

A subset of $\mathbb{R}$ is *open* if it is the union of open intervals; it is *nowhere dense* if every nonempty open interval contains a nonempty subinterval disjoint from it. A *meager set* is a countable union of nowhere dense sets; a *comeager* set is one whose complement is meager. A set A of reals has the *property of Baire* if A differs from some open set G by a meager set M (meaning $A = G \triangle M$, where $\triangle$ is symmetric difference). If one has DC, then a countable union of meager sets is meager.

Every nested sequence of bounded, closed intervals has a nonempty intersection; this is Cantor's Intersection Theorem. It is a theorem of ZF, because the least upper bound of the left endpoints is in the intersection. A consequence of this is that the real line is not meager; indeed, given countably many nowhere dense sets, we can construct a nested sequence of closed intervals with rational endpoints so that the nth interval is disjoint from the nth nowhere dense set and then intersect them all. The use of rational endpoints means that this can be done without any form of AC.

Suppose now that A is a $\mathbb{Q}$-invariant set of reals that is measurable or has the Baire property. Remarkably, in the measure case (assuming λ is countably additive) either A or $\mathbb{R} \setminus A$ has measure zero and in the Baire case either A or $\mathbb{R} \setminus A$ is meager (we prove both in a moment). We need the fact that any Lebesgue measurable A is contained in a union of open intervals with rational endpoints for which the sum of all the interval lengths is arbitrarily close to $\lambda(A)$. This is true because $\lambda(A)$ equals the outer measure of A, which is the greatest lower bound of the aforementioned sums for countable sets of intervals that cover A. For more on zero-one laws, see Section 5, and also [**15**, chap. 21].

Zero-One Law For $\mathbb{R}/\mathbb{Q}$ (ZF). *Let $A \subseteq \mathbb{R}$ be $\mathbb{Q}$-invariant. If A has the Baire property, then A is either meager or comeager. If A is measurable, then either (a) A intersects all bounded intervals in measure zero; or (b) A intersects all bounded intervals J in measure $\lambda(J)$; if λ is countably additive, then either A or $\mathbb{R} \setminus A$ has measure zero. If λ is restricted to* $[0,1]$, *then any set that is $\mathbb{Q}$-invariant (modulo* 1*) has measure* 0 *or* 1.

Proof. For the first, assume that A is nonmeager and M and G witness A having the Baire property; then $G \ne \emptyset$. We claim that if $y \notin A$, then $y \in \bigcup_{q\in\mathbb{Q}} M + q$, which proves that A is comeager (this countable union is constructive, so DC is not used). To prove the claim, choose $q \in \mathbb{Q}$ so that $y \in G + q$. This is possible because the rational translates of a fixed open interval cover $\mathbb{R}$. Thus $y - q \in G$. Since $y \notin A$ and A is $\mathbb{Q}$-invariant, we have $y - q \notin A$; so $y - q \in G \setminus A \subseteq M$. Hence $y \in M + q$.

Now suppose A is Lebesgue measurable and $\mathbb{Q}$-invariant; let $B = A \cap [0,1]$ and $\alpha = \lambda(B)$. We will show that $\lambda(A \cap J) = \alpha\,\lambda(J)$ for any interval J with rational ends. Letting $m, n \in \mathbb{N}$, the invariance property means that this holds for $[m, m+1]$; it then extends to $[0,m]$ by subdividing into unit intervals. Division of $[0, m]$ into n equal subintervals then gives the property for $[0, m/n]$, from which one gets it for all intervals with rational ends having length m/n. Now suppose $0 < \alpha < 1$. There is a family of intervals with rational endpoints $\{K_i\}_{i=0}^{\infty}$ covering B and having $\sum_{i=0}^{\infty} \lambda(K_i) = \beta$, with $\alpha \le \beta < 1$; let $\epsilon = \alpha\,(1-\beta)$. Because the tail of a series approaches 0, there is a finite union $\bigcup_{i=0}^{n} K_i$ that covers B except for a set of measure less than ϵ. Finite subadditivity of λ then gives the following contradiction

$$\alpha = \lambda(B) < \epsilon + \lambda\left(\bigcup_{i \le n} B \cap K_i\right) \le \epsilon + \sum_{i \le n} \lambda(B \cap K_i) = \epsilon + \alpha \sum_{i \le n} \lambda(K_i) \le \alpha(1-\beta) + \alpha\beta = \alpha. \ \blacksquare$$

We use LM for the assertion that all sets of reals are Lebesgue measurable. The theory ZF + DC + LM is consistent, provided one assumes the consistency of the existence of an inaccessible cardinal (note 3). This is a remarkable connection, especially because the inaccessible is both necessary and sufficient for this [**17**, **23**]. Because it appears that inaccessibles do not introduce a contradiction, we will treat ZF + DC + LM as we do ZF: they are assumed to be consistent. Similarly, we use PB for the assertion that all sets of reals have the property of Baire. The theory ZF + DC + PB is equiconsistent with ZF [**24**]; the contrast to the connection of LM to large cardinals is surprising.

The *harmonic expansion* [**26**] of a real will be useful: any $x \in \mathbb{R}$ can be expressed as $x = [d_1; d_2, d_3, \ldots] = \sum_{i=1}^{\infty} \frac{d_i}{i!}$, where $d_i \in \mathbb{Z}$ and $0 \le d_i \le i-1$ for $i \ge 2$. Uniqueness is desirable, so define the harmonic digits as follows: For integers $i \ge 1$, let $c_i = \lfloor i!\,x \rfloor$, $d_1 = \lfloor x \rfloor$, and, for $i \ge 2$, $d_i = c_i - i\,c_{i-1}$. An easy identity for integers $i \ge 1$ is $\lfloor i\,\alpha \rfloor + 1 \le i\,\lfloor \alpha \rfloor + i$. This and the fact that $i\,c_{i-1}$ is an integer no greater than $i!\,x$ yield $i\,c_{i-1} \le c_i \le i\,c_{i-1} + i - 1$, which gives $0 \le d_i \le i-1$. Because $\frac{d_i}{i!} = \frac{c_i}{i!} - \frac{c_{i-1}}{(i-1)!}$, telescoping gives $\sum_{i=1}^{m} \frac{d_i}{i!} = \frac{c_m}{m!}$. But $c_m \le m!\,x < c_m + 1$, so $0 \le x - \frac{c_m}{m!} < \frac{1}{m!}$ and $x = \sum_{i=1}^{\infty} \frac{d_i}{i!}$. A key fact is that $x - y \in \mathbb{Q}$ if and only if the harmonic digits d_i of x and y are eventually equal. The reverse direction is immediate; the other is easily proved by taking i greater than the rational's denominator.

3. THE DIVISION PARADOX.

We give here a self-contained and short proof of the Division Paradox in the context of the familiar additive group $\mathbb{R}$ and its rational subgroup. More precisely, we show that $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$ is a theorem of ZF (with no assumption of any form of AC) when either LM or PB is assumed. Theorem 1 is due to Mycielski [**12**, **14**], Theorem 2 to Sierpiński [**19**; **20**, §8; **21**, **22**].

Theorem 1 (ZF). $|\mathbb{R}| \le |\mathbb{R}/\mathbb{Q}|$.

Theorem 2 (ZF). *If* $\mathbb{R}/\mathbb{Q}$ *has a linear ordering* $\preccurlyeq$, *then* $A = \{x \in \mathbb{R} : [x] \preccurlyeq [-x]\}$ *is* $\mathbb{Q}$*-invariant, is not Lebesgue measurable, and does not have the property of Baire.*

An injection from $\mathbb{R}/\mathbb{Q}$ to $\mathbb{R}$ induces a linear ordering on $\mathbb{R}/\mathbb{Q}$ and so these theorems immediately yield the following.

Corollary 3. *The Division Paradox* $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$ *is a theorem in either* ZF + LM *or* ZF + PB.

Under AC one can find a choice set V for the equivalence classes in $\mathbb{R}/\mathbb{Q}$; this yields $|\mathbb{R}/\mathbb{Q}| \le |\mathbb{R}|$. Being nonmeasurable, V cannot exist under LM. Theorem 2 is stronger: in ZF + (LM or PB), not only is there no choice set, but there is no injection of any sort from $\mathbb{R}/\mathbb{Q}$ into $\mathbb{R}$, nor any linear ordering of $\mathbb{R}/\mathbb{Q}$.

The proof of Theorem 2 given here, with its use of a linear order, is based on the ideas used by Sierpiński (for alternate views see Section 5). Under AC there is a well-ordering of the reals; that yields a choice set for the classes by choosing the least element in each class. So the well-ordering gives a nonmeasurable set. Theorem 2 shows that a linear ordering of the classes is enough to get a nonmeasurable set.

The proof of Theorem 1 that follows will construct a linear ordering of a large subset of $\mathbb{R}/\mathbb{Q}$. The set of reals $\{y_x : 0 < x < 1\}$ in the proof is a constructively defined choice set for continuum many classes in $\mathbb{R}/\mathbb{Q}$ and it induces a linear ordering for the classes $[y_x]$.

Proof of Theorem 1. Suppose the binary form of $x \in (0,1)$ is $0.abcd\ldots$. Use harmonic expansion to define $y_x = [0; a, a, b, a, b, c, a, b, c, d, a, \ldots]$. If $x \ne z$, then $y_x - y_z \notin \mathbb{Q}$ because the harmonic digits of y_x and y_z differ infinitely often. Thus $x \mapsto y_x$ injects $(0,1)$ into $\mathbb{R}/\mathbb{Q}$. Because $|(0,1)| = |\mathbb{R}|$ (via $\tan\left(\frac{\pi}{2}(2x-1)\right)$), this suffices. $\blacksquare$

Proof of Theorem 2. Because $x \in \mathbb{Q}$ if and only if $2x \in \mathbb{Q}$, we have $x \sim -x$ if and only if $x \in \mathbb{Q}$. This means that $\mathbb{Q} \subseteq A$ and also that $\rho(x) = -x$ defines a bijection from the irrationals in A to $\mathbb{R} \setminus A$ that preserves measure and meagerness. Because $x \sim x + q$ and $-x \sim -(x + q)$, A is $\mathbb{Q}$-invariant. Assume A has the Property of Baire. By the Zero-One Law, A is either meager or comeager. But A is meager if and only if $A \setminus \mathbb{Q}$ is meager if and only if $\rho(A \setminus \mathbb{Q})$ is meager if and only if $\mathbb{R} \setminus A$ is meager, contradiction. Similar reasoning, with meager replaced by measure 0 and working with $A \cap [-1, 1]$, works for the measure case. ∎

To see why the Division Paradox is so surprising, recall that cardinality is a partial order and so, in ZF, there are exactly four possibilities for the cardinality relation between $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$:

1. $|\mathbb{R}| = |\mathbb{R}/\mathbb{Q}|$
2. $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$ are incomparable: $|\mathbb{R}/\mathbb{Q}| \nleq |\mathbb{R}|$ and $|\mathbb{R}| \nleq |\mathbb{R}/\mathbb{Q}|$
3. $|\mathbb{R}/\mathbb{Q}| < |\mathbb{R}|$
4. $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}|$

The first is a viable choice because it follows from AC: a choice set for the classes means $|\mathbb{R}/\mathbb{Q}| \le |\mathbb{R}|$; then Theorem 1 and the Schröder–Bernstein Theorem give equality. At first glance, one might expect (2) and (3) to be consistent with ZF. If AC is false, there will be incomparable sets, so perhaps $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$ could be such a pair; and because the Continuum Hypothesis can fail, there could be room beneath the continuum for a set such as $\mathbb{R}/\mathbb{Q}$, which would give (3). But Theorem 1 proves $|\mathbb{R}| \le |\mathbb{R}/\mathbb{Q}|$, showing directly that (2) and (3) are always false. So if (1) is false, as it might be if AC is abandoned, then (4) must hold; and (4) says that the set that one expects to be smaller, $\mathbb{R}/\mathbb{Q}$, is in fact strictly larger. Corollary 3 shows that (4) can be true in some situations.

The conclusion of Theorem 2 holds under the hypothesis of AC_2, which states that if one has a collection of pairs of socks, one can choose one from each pair (Sierpiński, see [**7**, p. 7]; see also Theorem 6). The set $\{[x], [-x]\}$ corresponds to a pair of socks.

One can eliminate the Division Paradox by using a surjective definition of cardinality instead of the standard injective one. Define $|X| \le^* |Y|$ if there is a surjection from Y onto X. But this approach has a fatal flaw: there is no Schröder–Bernstein Theorem in ZF. For consider what happens under ZF + PB. We have a trivial surjection from $\mathbb{R}$ to $\mathbb{R}/\mathbb{Q}$ and the proof of Theorem 1 yields a surjection in the other direction: use $y_x \mapsto x$, sending all other reals to 0. So a Schröder–Bernstein Theorem for $\le^*$ would give a bijection from $\mathbb{R}/\mathbb{Q}$ to $\mathbb{R}$, contradicting Theorem 2. For more on the surjective approach, see [**1**]. A consequence of these ideas is the interesting result that either LM or PB negates the Generalized Continuum Hypothesis [**12**]. The two surjections just given, via inverse images, induce injections between $\mathcal{P}(\mathbb{R})$ and $\mathcal{P}(\mathbb{R}/\mathbb{Q})$ in both directions; these two power sets therefore have the same cardinality. Corollary 3 then gives $|\mathbb{R}| < |\mathbb{R}/\mathbb{Q}| < |\mathcal{P}(\mathbb{R}/\mathbb{Q})| = |\mathcal{P}(\mathbb{R})|$, showing that GCH fails. This is not a surprise since it is known that $\neg$ AC $\Rightarrow \neg$ GCH, but is a concrete example of the failure.

Those who find Banach–Tarski duplications unpalatable but do not want to give up the many useful consequences of the Axiom of Choice can work in ZFC and try, whenever possible, to restrict themselves to measurable sets. A constructive point of view also elucidates the Division Paradox, which can be studied profitably in ZFC provided one reinterprets cardinality. The key idea is *Borel cardinality*: a naive view would define $|X| \le_B |Y|$ to mean that there is a one-one Borel function $F: X \to Y$. But there is no natural topology on $\mathbb{R}/\mathbb{Q}$ and so instead this definition is used: Suppose X and Y are complete separable metric spaces (known as *Polish spaces*), each endowed with an equivalence relation ($\sim$ denotes either one) and with the collections of equivalence classes denoted $\hat{X}$ and $\hat{Y}$; then $|\hat{X}| \le_B |\hat{Y}|$ means that there is a Borel function $F: X \to Y$ such that $a \sim b \Leftrightarrow F(a) \sim F(b)$. The proof of Corollary 3, with no essential change, then yields the nonparadoxical and interesting result that $|\mathbb{R}| <_B |\mathbb{R}/\mathbb{Q}|$ (where the equivalence relation on $\mathbb{R}$ is just equality). In short, there is a nicely definable injection from $\mathbb{R}$ to $\mathbb{R}$ that respects the rational equivalence relation in the codomain, but no such injection that respects the rational relation in the domain. More precisely, $|\mathbb{R}/\mathbb{Q}| \nleq_B |\mathbb{R}|$, so there is no Borel F from $\mathbb{R}$ to $\mathbb{R}$ such that $x - y \in \mathbb{Q} \Leftrightarrow F(x) = F(y)$. For the theory and applications of Borel cardinality and Borel equivalence relations see [**2**, **10**].

4. CURIOUSER AND CURIOUSER. In Section 1, we motivated the Division Paradox by imagining a sports league having more teams than players. We could equally well have phrased this in terms of more conferences than teams: the conferences divide up the teams just as the teams partition the players. But can there be more conferences than teams *and* more teams than players? We'll refer to any example of this as a Double Division Paradox, which we define as follows.

Definition. A *Double Division Paradox* is a triple (X, Y, Z) such that $|X| < |Y| < |Z|$ with surjections $f: X \to Y$ and $g: Y \to Z$.

If (X, Y, Z) is a Double Division Paradox, then we can think of X as the player pool, with x_1 and x_2 on the same team if $f(x_1) = f(x_2)$, and with teams y_1 and y_2 in the same conference if $g(y_1) = g(y_2)$. The argument after the proof of Theorem 2 easily extends to show that a Double Division Paradox yields a double failure of the GCH: $|X| < |Y| < |Z| < |\mathcal{P}(X)|$.

As this section's title suggests, a Double Division Paradox can exist (in the absence of AC). In fact, it can be much, much worse than this (note 4). The following simple example, which like our main example starts with the reals, was provided by Asaf Karagila and is included with his kind permission.

Theorem 4 (ZF). *If no uncountable set of reals can be well-ordered, then there is a Double Division Paradox starting with* $\mathbb{R}$.

Proof. As usual, ω_1 is the smallest uncountable ordinal. We'll show that $(\mathbb{R}, \mathbb{R} \bigcup \omega_1, \mathbb{R} \times \omega_1)$ is a Double Division Paradox. We need Lebesgue's classic surjection from $\mathbb{R}$ to ω_1: identify $\mathbb{R}$ with $\mathcal{P}(\mathbb{N} \times \mathbb{N})$ (both have the cardinality of $2^{\mathbb{N}}$) and map each set of pairs that is a well-ordering to its order type, sending other sets to 0. There is a surjection from $\mathbb{R}$ to $\mathbb{R} \times \mathbb{R}$ (e.g., by the method in note 1) and using Lebesgue's map on the second coordinate turns it into a surjection from $\mathbb{R}$ to $\mathbb{R} \times \omega_1$. This induces the two surjections needed for the paradox.

Injections from $\mathbb{R}$ to $\mathbb{R} \bigcup \omega_1$ to $\mathbb{R} \times \omega_1$ are trivial, so it remains to show that there are no injections in the reverse direction. An injection of $\mathbb{R} \bigcup \omega_1$ into $\mathbb{R}$ maps ω_1 to an uncountable set of reals admitting a well-ordering, contrary to the theorem's assumption. To conclude, suppose $f: \mathbb{R} \times \omega_1 \to \mathbb{R} \bigcup \omega_1$ is one-one. For real x, let $A_x = f(\{x\} \times \omega_1) \bigcap \omega_1$. If no A_x is empty, then $\mathbb{R}$ is well-ordered by the least ordinal in A_x, contradiction. If $A_x = \emptyset$, then f embeds $\{x\} \times \omega_1$ into $\mathbb{R}$, again a contradiction. ■

Theorem 5. *Under* ZF + DC + LM *there is a Double Division Paradox starting with* $\mathbb{R}$.

Proof. Shelah [**18**, Theorem 5.1B] showed that the hypothesis of Theorem 4 is true in ZF + DC + LM. ■

5. ALTERNATE SETTINGS FOR THE PARADOX. The Division Paradox was first established not for $\mathbb{R}/\mathbb{Q}$ but for the quotient group arising from the group $(\mathcal{P}(\mathbb{N}), \triangle)$ and its subgroup $\mathcal{F}$ consisting of the finite subsets of $\mathbb{N} = \{0,1,2,\ldots\}$. And there are other alternate settings: two important relations are the Bernoulli shift and the tail relation. For the tail relation the underlying set is $\mathcal{P}(\mathbb{N})$ (viewed as $2^{\mathbb{N}}$) with sequences s and t being *tail-equivalent* if there are $m, n \in \mathbb{N}$ so that $s_{m+k} = t_{n+k}$ for all $k \in \mathbb{N}$. For the shift the underlying set is $2^{\mathbb{Z}}$, the integer-indexed binary sequences, with the equivalence given by the *shift map*: $s \sim_{\text{shift}} t$ if and only if there is an integer k so that, for every n, $s_n = t_{n+k}$. An interesting aspect of $\sim_{\text{shift}}$ is that the equivalence class of an anchored sequence $\ldots xyz\,\boldsymbol{a}\,bcd\ldots$ in $2^{\mathbb{Z}}$ (the anchor—the 0-coordinate—is $\boldsymbol{a}$) is just the same object with the anchor omitted: $\ldots xyzabcd\ldots$; the no-origin version may be viewed as the set of all sequences that are shifts of an anchored form of it.

We can visualize the Division Paradox in $2^{\mathbb{Z}}$ by imagining a column with all unanchored nonperiodic sequences. Beside each one, place the set (typically infinite, possibly finite, and never empty) of all the anchored sequences based on it, as follows.

$\ldots abcd\ldots$ | ⋯ $\ldots\boldsymbol{a}bcd\ldots$ $\ldots a\boldsymbol{b}cd\ldots$ $\ldots ab\boldsymbol{c}d\ldots$ $\ldots abc\boldsymbol{d}\ldots$ ⋯

$\ldots 01010101\ldots$ | $\ldots 101\mathbf{0}101\ldots$ $\ldots 010\mathbf{1}010\ldots$

The Division Paradox for $2^{\mathbb{Z}}$ (proved in Cor. 7) says that there are more items in the first column than in all columns to the right, a conclusion that is just as absurd as the paradox for $\mathbb{R}$ and $\mathbb{Q}$.

There are various approaches to the Division Paradox in these alternative settings. An elegant method is to generalize the proofs of Section 3 to apply to all three contexts at once. An alternative is to fashion proofs specific to the context; this can lead to simplifications. And finally one can ask whether the paradoxes in the new settings are logical consequences of Corollary 3. Each approach has its charms and we will comment briefly on each. Everything that follows is part of the folklore that includes work of Sierpiński, Mycielski, and others, as well as more recent work in the area of Borel equivalence relations.

We take the topology on $2^{\mathbb{N}}$ to be the usual product topology from the discrete set $\{0,1\}$ and the measure (denoted λ) on $2^{\mathbb{N}}$ to be the product measure from $\{0,1\}$, where $\{0\}$ and $\{1\}$ each get measure $1/2$. The natural map $f: 2^{\mathbb{N}} \to [0,1]$ via binary expansions is not one-one (a rational of the form $m/2^n$ arises from a sequence ending in only 0s and another ending in 1s), but it does induce a bijection from $2^{\mathbb{N}} \setminus f^{-1}(D)$ to $[0,1] \setminus D$ where D is the set of rationals of the form $m/2^n$. Because countable sets are meager and have measure 0, this bijection allows one to show that LM and PB are equivalent to the corresponding assertions in $2^{\mathbb{N}}$. Topology and measure in $2^{\mathbb{Z}}$ are similar and the same results hold (e.g., a basic open set is the set of sequences extending a fixed finitely specified sequence and the measure of such an open set is defined to be 2^{-m} where m is the number of components specified; the standard outer measure construction then yields the product measure).

We start with a unified approach that yields Corollary 3 for reals, the finite set and tail relations on $2^{\mathbb{N}}$, and the Bernoulli shift. Recall from Section 3 that a *Polish space* is a complete, separable metric space; any uncountable Polish space has the same cardinality as the reals (or $2^{\mathbb{N}}$ or $2^{\mathbb{Z}}$) [**10**, Cor. 6.5]. Throughout this section we assume that $\sim$ is an equivalence relation on X, an uncountable Polish space, and $\mathcal{I} \subseteq \mathcal{P}(X)$ is a countably complete, nonprincipal ideal on X.

Definition. A set $A \subseteq X$ is *almost Borel* (w.r.t. $\mathcal{I}$) if for some Borel set B, $A \triangle B \in \mathcal{I}$; $A \subseteq X$ is *invariant* if A is closed under $\sim$. The relation and ideal satisfy the *Zero-One Law* if whenever A is an invariant set that is almost Borel, either $A \in \mathcal{I}$ or $X \setminus A \in \mathcal{I}$.

The next proof is nicely general and quite different than the construction used in Theorem 1.

Theorem 5 (ZF). *Suppose $\sim$ is meager in $X \times X$. Then $|X| \leq |X/\sim|$.*

Proof. Cover $\sim$ by a countable union $\bigcup N_i$ of nowhere dense subsets of $X \times X$, where we may assume $N_i \subseteq N_j$ when $i \leq j$: Build a tree by taking X as the root. For level n choose two disjoint open balls of diameter at most 2^{-n} contained in each set of level $n-1$. Then move through all possible pairs of these, continually shrinking the rectangle determined by the sets of a pair so to avoid N_n; then no ball at level n will have a point equivalent by the part of the relation in N_n to a point in another ball at this level. Given $x \in [0, 1)$, let $s \in 2^{\mathbb{N}}$ be its base-2 expansion, avoiding sequences that have a tail of 1s. Viewing 0 as left and 1 as right, s determines a branch in the constructed tree and the Cantor intersection theorem, which holds in X, then yields a unique point y_x in the intersection of the sets in the branch. The set $\{y_x : 0 \leq x < 1\}$ has only inequivalent points, as required. Because $|X| = |\mathbb{R}| = |[0,1)|$, the result follows. ■

The relations we are studying all satisfy the hypothesis of Theorem 5. The relation for $\mathbb{R}/\mathbb{Q}$ is a union of countably many lines in the plane of slope 1. For the tail relation define, for each $j, k \in \mathbb{N}$, $N_{j,k}$ to be $\{(s, t) : s$ beyond j equals t beyond $k\}$. The the tail relation is $\bigcup N_{j,k}$ and each $N_{j,k}$ is nowhere dense in $2^{\mathbb{N}} \times 2^{\mathbb{N}}$, as we show next. Working with binary sequences, an open set in the product contains the product of two basic open sets in $2^{\mathbb{N}}$; if they are determined by finite sequences with specified bits on coordinates at most m and n, respectively, extend each to the coordinates in $[0, \max(m, n, j, k) + 1]$ by filling with all 0s in one and all 1s in the other. This handles $2^{\mathbb{N}}/\mathcal{F}$ too because its relation is a subset of the tail relation. The argument for $2^{\mathbb{Z}} \times 2^{\mathbb{Z}}$ is similar, where one appends 0s to both ends of one finite string and 1s to the other. So Theorem 5 shows that the ambient set embeds into the set of classes for each of these three relations.

In fact, embeddings as in Theorem 5 can be defined quite directly for many specific relations. Theorem 1 does this for $\mathbb{R}/\mathbb{Q}$ and we now do the same for the finite set, tail, and Bernoulli shift relations. For $2^{\mathbb{N}}$, let $A = \{a, b, c, \ldots\} \subseteq \mathbb{N}$ in increasing order and, using the primes, let $f(A) = \{2^a, 2^a 3^b, 2^a 3^b 5^c, \ldots\}$ and then consider the equivalence class of $f(A)$ in $2^{\mathbb{N}}/\mathcal{F}$. This works because if A and B are distinct, then $f(A)$ and $f(B)$ have finite intersection. Note that f is a continuous function. This shows $|2^{\mathbb{N}}| \leq |2^{\mathbb{N}}/\mathcal{F}|$, and the identical mapping also works for the tail relation on $2^{\mathbb{N}}$.

For the shift, define $f: 2^{\mathbb{N}} \to 2^{\mathbb{Z}}/\text{shift}$ by $f(abc\ldots) = \ldots ccbbabbcc$, which is an equivalence class. Then a is the center of a unique constant block of odd length (the two constant sequences are easily handled) and this locates the origin and allows the recovery of $abc\ldots$ from its image. Therefore f is one-one, proving $|2^{\mathbb{Z}}| = |2^{\mathbb{N}}| \leq |2^{\mathbb{Z}}/\text{shift}|$. This construction shows that the palindromic sequences with even blocks except for a central odd block form a continuum-sized choice set for the corresponding subset of shift-classes.

For the second half of the Division Paradox, there are two approaches. The first requires no assumption of a map from the space to itself (such as $x \mapsto -x$ in Theorem 2), but it requires DC because it appeals to the countable completeness of the ideals (the argument ends by taking the intersection of countably many sets in the ultrafilter). Invariant sets are simply unions of some equivalence classes. An important point is that any invariant set for $2^{\mathbb{N}}/\mathcal{F}$, $2^{\mathbb{N}}/\text{tail}$, or $2^{\mathbb{Z}}/\text{shift}$ obeys the Zero-One Law: an invariant set having the Baire property is meager or comeager and an invariant measurable set has measure 0 or 1. For the measure case this is the same as saying that λ is *ergodic* with respect to the relation. The results for the tail relation follow from the same for $2^{\mathbb{N}}/\mathcal{F}$; so we have four cases of the Zero-One Law to examine.

Case 1. $2^{\mathbb{N}}/\mathcal{F}$ and meager sets. This is the same as the proof for $\mathbb{R}$ in Section 2, with the operation of adding a rational to a real replaced by the change of an initial segment of a sequence. For example, the key relation $y \in \bigcup_{q \in \mathbb{Q}} M + q$ becomes $y \in \bigcup_s M_s$, where M_s is the result of changing every element in M so that it starts with s.

Case 2. $2^{\mathbb{N}}\big/\mathcal{F}$ and measure-zero sets. As in Case 1, the proof of Section 2 works with minimal change. Suppose $A \subseteq 2^{\mathbb{N}}$ is measurable and invariant under finite changes. Then it is easy to show that $\lambda\big(A \cap \hat{s}\big) = \lambda(A)\,\lambda(\hat{s})$ for any basic open set $\hat{s}$. The proof concludes with an argument identical to the one at the end of the proof of the proof in Section 2.

Case 3. $2^{\mathbb{Z}}\big/$shift and meager sets. Here and in Case 4, σ denotes the one-unit shift to the right. Let H consist of sequences in $2^{\mathbb{Z}}$ in which every finite bit sequence occurs somewhere as a block. Then $s \in H$ if and only if every basic open set contains a shift of s. Moreover, H is a Borel set (and so has the Baire property) and H is comeager. For if $2^{\mathbb{Z}} \setminus H$ is not meager, there is a nonempty basic open set $\hat{s}$ so that $\hat{s} \cap H = \bigcup N_k$, where N_k is nowhere dense. Now build t by starting with s and continually extending to ensure that $t \notin N_k$ and s_k occurs in t, where $\{s_k\}$ enumerates all finite bit sequences. Then $t \in \hat{s} \cap H$ but $t \notin \bigcup N_k$, contradiction. To finish, suppose A is a nonmeager shift-invariant set having the form $G \triangle M$, where G is open and M is meager. Suppose $y \notin A$ and $y \in H$. Then some shift $\sigma^n(y)$ of y is in G. So, by invariance, $\sigma^n(y) \in G \setminus A \subseteq M$. But $\sigma^n(y) \in M$ implies that $y \in \sigma^{-n}(M)$. This shows that the complement of A is contained in the meager set $\big(2^{\mathbb{Z}} \setminus H\big) \cup \big(\bigcup_{n \in \mathbb{Z}} \sigma^n(M)\big)$.

Case 4. $2^{\mathbb{Z}}\big/$shift and measure-zero sets. Assume A is a shift-invariant measurable subset of $2^{\mathbb{Z}}$, with $\lambda(A) = \alpha$ and fix $\epsilon > 0$. Choose finitely many basic open sets $\hat{s}_i$ so that, with $E = \bigcup \hat{s}_i$, we have $A \subseteq E$ and $\lambda(E \setminus A) < \epsilon$ (this uses outer measure, as in the proof in Section 2). Let n be larger than the largest coordinate used in any s_i. Then the basic open sets that occur in $F = \sigma^n(E)$ have support disjoint from the basic sets $\hat{s}_i$; this means that $\lambda(E \cap F) = \lambda(E)\,\lambda(F) = \lambda(E)^2$. Because $A \subseteq E$, we have $A = \sigma^n A \subseteq \sigma^n E = F$, so $A \subseteq E \cap F$ and the measure difference of these two sets is at most ϵ. Now, $\left|\alpha - \alpha^2\right| \le |\alpha - \lambda(E \cap F)| + \left|\lambda(E \cap F) - \alpha^2\right| < \epsilon + \left|\alpha^2 - \lambda(E)^2\right| = \epsilon + (\lambda(E) + \alpha)\cdot|\lambda(E) - \alpha| \le \epsilon + 2\epsilon = 3\epsilon$. This proves $\alpha - \alpha^2 = 0$, so α is 0 or 1. ■

Now we can formulate a general result relating the Zero-One Law to the nonexistence of a certain injection.

Theorem 6 (ZF + DC). *Suppose the relation ~ and ideal $\mathcal{I}$ satisfy the Zero-One Law and every invariant set is almost Borel. Then* $|X/{\sim}| \not\le |X|$.

Proof. Define a countably complete ultrafilter U on $X/{\sim}$ by placing A in U if $X \setminus \bigcup A$ is in $\mathcal{I}$; then U is nonprincipal. An injection of $X/{\sim}$ into X would transfer this to an ultrafilter on X and because $|X| = \left|2^{\mathbb{N}}\right|$, this gives an ultrafilter V on $2^{\mathbb{N}}$. For each k, split $2^{\mathbb{N}}$ into the set of sequences having a 0 in the kth position and the ones having a 1 there. One of these, call it A_k, is in V; then $\bigcap A_k$ is a singleton in V, contradiction. This is essentially the argument that a measurable cardinal is strongly inaccessible. ■

Recall that any Lebesgue measurable set differs from a Borel set by a measure-zero set [**15**, §3] and by definition, a set with the Baire Property differs from a Borel set by a meager set. Therefore one can let $\mathcal{I}$ be the measure-zero sets or the meager sets, depending on whether one is assuming LM or PB, and the hypotheses of Theorem 6 are satisfied. In this way, Theorems 5 and 6 give the Division Paradox for the finite set and tail relations on $2^{\mathbb{N}}$ and the Bernoulli shift on $2^{\mathbb{Z}}$, in the form given in Corollary 3.

We can eliminate the use of DC by a slightly different argument, one based on the proof of Theorem 2. We use $[x]$ for the equivalence class of x.

Theorem 7 (ZF). *Suppose the relation ~ and ideal $\mathcal{I}$ satisfy the Zero-One Law and there is a function $\rho : X \to X$ such that*

1. $x \sim y$ if and only if $\rho(x) \sim \rho(y)$.
2. *If $A \in \mathcal{I}$, then $\rho(A) \in \mathcal{I}$.*
3. *$\{x \in X : x \nsim \rho(x)\}$ is almost Borel and is not in $\mathcal{I}$.*
4. *$\rho(\rho(x)) = x$ for all x.*

Suppose C is a choice set for $\{\{\gamma, \rho(\gamma)\} : \gamma \in X/{\sim}$ and $\gamma \ne \rho(\gamma)\}$. Then $\bigcup C$ is not almost Borel w.r.t. $\mathcal{I}$.

Proof. Note first that, by (1) and (4), $\rho(\gamma)$ is an equivalence class when γ is. Suppose $A = \bigcup C$ is almost Borel. Let $D = \{x \in X : x \sim \rho(x)\}$; D is almost Borel by (3) and therefore so is $B = \rho(A)$. Because $B = X \setminus (A \cup D)$, A and B are disjoint, and because they are unions of equivalence classes, A and B are ~-invariant. Therefore, by the Zero-One Law, $A \in \mathcal{I}$ or $B \in \mathcal{I}$ (disjointness means that the complements of both cannot by in $\mathcal{I}$). By (2) and (4) this means that both are in $\mathcal{I}$, so $A \cup B \in \mathcal{I}$, contradicting $A \cup B = X \setminus D \notin \mathcal{I}$, which holds by (3). ■

A function ρ as in the theorem exists for our four main examples, with either the meager or measure-zero ideal. For $\mathbb{R}$ use $x \mapsto -x$, while $s \mapsto 1-s$ works for $2^{\mathbb{N}}/\mathcal{F}$ and the tail relation. For the Bernoulli shift, let $\sim$ be the shift relation and define ρ to be the reflection: $\rho(s)_n = s_{-n}$. Then (1), (2), and (4) are clear. For (3), we have $\{s : s \sim \rho(s)\} = \bigcup_{k\in\mathbb{Z}} N_k$, where $N_k = \{s : s \text{ is a } k\text{-shift of } \rho(s)\}$. Then N_0 is the set of palindromic sequences with a single center element at the origin; N_1 is the set of palindromes with a double center, the rightmost of which is at the origin, and, for $k \geq 0$, N_{2k} (resp., N_{2k+1}) is the k-shift of N_0 (resp., N_1); the negative case is similar. It is then not hard to see that each N_k is nowhere dense and has measure 0, and the same is true of their union (in ZF). This verifies (3).

Corollary 7 (ZF). *Under* LM *or* PB, *there is a Division Paradox in* $2^{\mathbb{N}}/\mathcal{F}$, $2^{\mathbb{N}}/tail$, *and* $2^{\mathbb{Z}}/shift$. *That is, in each case* $|X| < |X/\sim|$.

Proof. Assume PB and let $\mathcal{I}$ be the meager ideal. Theorem 5 gives the injection of $|X|$ into $|X/\sim|$. If there is an injection in the other direction, then there is a choice set C as in Theorem 7. Use ρ as given before the corollary and apply Theorem 7 to conclude that C does not differ from a Borel set by a meager set, in contradiction to all sets having the Baire Property. The proof under LM is the same, with $\mathcal{I}$ being the measure-zero sets. ∎

To conclude we will discuss several connections between various versions of the Division Paradox. First, the paradox for $2^{\mathbb{N}}/\mathcal{F}$ is equivalent to the same for $\mathbb{R}/\mathbb{Q}$ [**14**]. For the proper context define, for two Borel equivalence relations (each denoted here by $\sim$), $X/{\sim} \sqsubseteq Y/{\sim}$ to mean that there is a Borel injection $f: X \to Y$ so that $x \sim x'$ if and only if $f(x) \sim f(x')$; this is called *Borel embeddability* of the first relation into the second [**2**]. Theorem 9 shows that $\mathbb{R}/\mathbb{Q}$ and $2^{\mathbb{N}}/\mathcal{F}$ are bi-embeddable. Because the injections on $\mathbb{R}$ and $2^{\mathbb{N}}$ induce injections on the classes, the Schröder–Bernstein Theorem then gives, in ZF, $|2^{\mathbb{N}}/\mathcal{F}| = |\mathbb{R}/\mathbb{Q}|$.

Theorem 9 (ZF). $\mathbb{R}/\mathbb{Q} \sqsubseteq 2^{\mathbb{N}}/\mathcal{F} \sqsubseteq \mathbb{R}/\mathbb{Q}$ *and* $\mathbb{R}/\mathbb{Q} \sqsubseteq 2^{\mathbb{N}}/tail$.

Proof. Given a real x, define $f(x) = s \in 2^{\mathbb{N}}$ as follows. Let the harmonic expansion of x be $[d_1; d_2, d_3, \ldots]$. Use s_0 to encode the sign of d_0 and use $s_2, s_4, \ldots$ to store d_0 many 1s; these even-indexed digits have an all-0 tail. Use $\{s_{2j+1}\}$ to encode each d_k in turn, using $0^2\, 1^{d_k+1}\, 0\, 1^{k-d_k}$ (exponents denote repeated digits); the block for d_k has length $k+4$. It is clear that f is one-one; as in Theorem 1's proof, if $x - y \in \mathbb{Q}$ then $f(x) \sim_{\text{finite}} f(y)$ (which implies the same for $\sim_{\text{tail}}$); and it is easy to see that if $f(x) \sim_{\text{tail}} f(y)$, then $x - y \in \mathbb{Q}$ (which implies the same for $\sim_{\text{finite}}$). This gives the first and last of the three claimed embeddings.

For the remaining one, define g by viewing the ith bit of a binary sequence in $2^{\mathbb{N}}$ as d_{i+2} in the harmonic expansion of a real. Then $g: 2^{\mathbb{N}} \to [0, e-2]$ and proves $2^{\mathbb{N}}/\mathcal{F} \sqsubseteq \mathbb{R}/\mathbb{Q}$. ∎

We have discussed the relations $\mathbb{R}/\mathbb{Q}$, $2^{\mathbb{N}}/\mathcal{F}$, $2^{\mathbb{N}}/\text{tail}$, and $2^{\mathbb{Z}}/\text{shift}$. In a weak sense (Borel bi-embeddability) these are all the same [**2**, Theorem 7.1]; this means that, in ZF, they all have the same cardinality and the Division Paradox for one implies the same for all. But in a strong sense they are not the same. The strong sense is *Borel isomorphism*: the existence of a Borel bijection $f: X \to Y$ so that $x \sim x'$ if and only if $f(x) \sim f(x')$. A remarkable theorem [**2**, Cor. 9.3] asserts that any "nice" Borel equivalence relation—one that is aperiodic (no finite equivalence classes), hyperfinite (a union of equivalence relations having finite equivalence classes), and having no Borel selector for the classes—is isomorphic to exactly one in the following sequence

$$2^{\mathbb{N}}/\text{tail} \;\sqsubset^i\; 2^{\mathbb{N}}/\mathcal{F} \;\sqsubset^i\; 2^{\mathbb{N}}/\mathcal{F} \times 2_= \;\sqsubset^i\; \cdots \sqsubset^i\; 2^{\mathbb{N}}/\mathcal{F} \times \mathbb{N}_= \;\sqsubset^i\; 2^{\mathbb{Z}}_*/\text{shift}$$

Here $X_=$ is the equality relation on X, $2^{\mathbb{Z}}_*$ denotes the nonperiodic sequences in $2^{\mathbb{Z}}$, and $\sqsubseteq^i$ means that there is a function as in the definition of $\sqsubseteq$, but with the additional property that the image is invariant under the relation in Y (X is *invariantly embeddable* in Y). The invariance condition allows the Schröder–Bernstein theorem to be used to deduce that Borel isomorphism is equivalent to $X \sqsubseteq^i Y \sqsubseteq^i X$. If the embedding exists in one direction only, one uses $\sqsubset^i$.

Because $\mathbb{R}/\mathbb{Q}$ arises from a nice relation, it must be isomorphic to one in the displayed chain. Which one? It turns out to be $2^{\mathbb{N}}/\text{tail}$. The reason is that these two have a property that the others do not. A relation $\sim$ on X is *paradoxical* if X contains disjoint A_1 and A_2 and there are two Borel bijections $f_i : X \to A_i$ so that, for all x, $x \sim f_i(x)$. Then the tail relation on $2^{\mathbb{N}}$ is paradoxical via the functions f_i that prepend i to each sequence. And $\mathbb{R}/\mathbb{Q}$ is paradoxical via $\mathbb{R} \supseteq (-\infty, 0) \cup (0, \infty)$; $\mathbb{R}$ maps to each unbounded interval by taking $[n, n+1)$ to $[m, m+1)$, using alternation and parity to fit them all in. But $2^{\mathbb{N}}/\mathcal{F}$ and $2^{\mathbb{Z}}_*/\text{shift}$ are not paradoxical, because the product measure is invariant under rational change; this means neither of these can be Borel isomorphic to $\mathbb{R}/\mathbb{Q}$.

It is proved in [**2**, §2] that for paradoxical relations, noninvariant embeddings yield, with the help of the Schröder–Bernstein theorem, invariant embeddings. Theorem 9's proof presents a relation-preserving injection f of the Vitali relation into the tail relation. It is also possible to give a constructive embedding in the other direction. So the whole thing is a little complicated, requiring three calls to Schröder–Bernstein, but one can find an explicit Borel function from $\mathbb{R}$ to $2^{\mathbb{N}}$ that is an isomorphism between the Vitali relation and the tail relation.

6. OPEN QUESTIONS. The Banach–Tarski Paradox violates the intuition one has from physical reality and LM eliminates the paradox. But the connection with topology is more subtle. Dougherty and Foreman (see [**25**, §11.2]) proved that, in ZFC, one can derive the Banach–Tarski Paradox with the pieces all having the property of Baire. But it is not known whether PB eliminates the paradox; it seems reasonable to conjecture that the answer to Question 1 is NO.

Question 1. Is the negation of the Banach–Tarski Paradox a theorem of ZF + DC + PB?

Let GM (for *general measure*) be: For each n, there is a countably additive, isometry-invariant measure on $\mathcal{P}(\mathbb{R}^n)$ that assigns measure 1 to the unit cube. Then DC + LM implies GM. But GM eliminates the Banach–Tarski Paradox and has the advantage that its consistency does not require any large cardinal assumption (see [**25**, §15.1]). Because the Zero-One Law uses outer measure, it is not clear that the proof of Theorem 2 can be modified to work under GM and so we have the following question.

Question 2. Is the Division Paradox a theorem of ZF + DC + GM?

To understand how the Division Paradox relates to more general statements, we recall two classical principles (see [**1**] for the history of these).

Definition. The *Partition Principle,* PP is: If Y is a family of disjoint nonempty subsets of X, then $|Y| \le |X|$. The *Weak Partition Principle*, WPP is: If Y is a family of disjoint nonempty subsets of X, then $|X| \not< |Y|$.

It is easy to see that the existence of Y as in the definition is equivalent to the existence of a surjection $f: X \to Y$. Therefore PP says: If $|Y| \le^* |X|$ then $|Y| \le |X|$; and WPP says that $|X| < |Y| \le^* |X|$ cannot occur. If WPP holds then one cannot have the Division Paradox for $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$, nor any similar phenomenon for other sets. Easy implications are AC $\Rightarrow$ PP $\Rightarrow$ WPP. More subtle are PP $\Rightarrow$ DC and PP $\Rightarrow$ AC for well-ordered families (A. Pelc; see [**11**, p. 10]). These connections lead to two fascinating open questions.

Question 3. (a) Does PP imply AC? (b) Does WPP imply AC?

An affirmative answer to (b) would be of some importance. For then we could abandon AC as an axiom and work with ZF + WPP, in the knowledge that it is no different than ZFC. But WPP feels more fundamentally obvious than the Axiom of Choice. Of course, absent any proofs, this is just speculation and it would not be terribly surprising if WPP was strictly weaker than AC.

7. CONCLUSION. We believe that the Division Paradox should be taken as an obviously false statement because of how seriously it undermines our intuition about how sets work—even more so than the Banach–Tarski Paradox. We therefore consider its negation to be obvious and conclude from Corollary 3 and the consistency of ZF + PB that (1) ZF is not strong enough to prove all obvious statements; (2) ZF + DC + LM is not a suitable axiom system.

We can say what we think it means for a statement to be mathematically obvious. We do not take a Platonist view, where one would mean the assertion is obviously true in an objective reality (the "real world"). We prefer a more pragmatic, generally formalist view, based on the experience of the mathematical community. Say that P is *necessary* if there is a consensus among mathematicians that, for whatever reason (truth, usefulness, convenience), any axiom system for mathematics should include P or yield it as a theorem. And call *P obvious* if it is necessary and has a compelling formulation that is not overly complicated and would be accepted as true even by people with little mathematical experience. This definition of "obvious" is of course subjective. Our view of "necessary" gives mathematicians veto power over the axioms they use. This is a fundamental difference with other sciences: physicists cannot veto the laws of nature.

Here is how we view some examples:

- All axioms of ZF are necessary. Most are obvious (e.g., the empty set axiom, extensionality, pairing, power set, the axiom of infinity, and the axiom of comprehension), but the regularity axiom (every nonempty set has an element that is disjoint from the set) is not obvious because of its somewhat technical nature.
- Fermat's Last Theorem is now agreed to be proved; hence it is necessary, but certainly not obvious.

- λ is countably additive: necessary, and arguably obvious for those familiar with infinite series.
- The Axiom of Dependent Choice: necessary, but not obvious.
- The negation of the Division Paradox, $|\mathbb{R}| \not< |\mathbb{R}/\mathbb{Q}|$: obvious.
- The negation of the Banach–Tarski Paradox is not necessary, despite its being a compelling and simple statement that is true in our physical world.
- The Axiom of Choice is necessary because of its usefulness and convenience (ZFC is the axiom system in current general use). But its statement is unlike the ZF axioms and is not obvious.
- The Partition Principle PP of Section 6 is necessary, but not obvious.
- The Weak Partition Principle WPP of Section 6 is obvious.

To us (and many before us), WPP is obvious. While the instinctive use of AC and DC before they were even formulated argues for their obviousness, their complexity argues against it. We do not view PP as obvious because in the absence of AC, two sets can have incomparable cardinalities and it does not seem out of the question that some relation $\sim$ on some set might yield an incomparable pair X and $X/\sim$. If we did not have Theorem 1 and knew only that, under LM, $|\mathbb{R}/\mathbb{Q}| \not\leq |\mathbb{R}|$ (i.e., $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$ violated PP, as opposed to WPP), it would not seem all that disturbing as it says only that a certain injection does not exist. So the interesting point is Corollary 3, which concludes that $\mathbb{R}$ and $\mathbb{R}/\mathbb{Q}$ are comparable, but the comparison goes the *wrong* way: the set of Vitali classes of real numbers has more elements than the set of reals! This is why we consider WPP to be obvious. Of course, one can take the opposite view and consider the Division Paradox to be a small price to pay in order to be rid of nonmeasurable sets; but there is definitely a price in terms of how the Division Paradox violates our intuition about size.

Mathematics appears to work well under a system that combines the best aspects of Platonism (mathematics describes a real world) and formalism (examine proofs under diverse axiom systems). Our perceived physical world certainly affects how we think of much of mathematics, but formalism combines the precision of proofs with the possibility of imaginatively and profitably examining a variety of axiom systems in realms beyond physical reality.

There are systems other than ZFC that could plausibly serve as a foundation for mathematics. One such alternative is in [**13**]. And even the underlying rules of logic can be questioned: some approaches to a constructive philosophy of mathematics reject the law of the excluded middle. But the consensus today is that ZFC is indeed the best foundation for our current view of mathematics; the results of Section 3 are one reason why the Axiom of Choice should not be abandoned.

8. NOTES.

1. LM negates the Banach–Tarski Paradox. Because $1 + 1 \neq 1$, it suffices to show that all subsets of $\mathbb{R}^3$ are Lebesgue measurable. Cantor's classic digit-mixing bijection from $[0,\infty)$ to $[0,\infty)^3$ uses the positions congruent to i (mod 3) ($i = 0, 1, 2$) to split the decimal digits into three infinite strings. This function preserves the measure of intervals (in dimensions 1 and 3) and hence preserves outer measure. Therefore all subsets of the first octant are measurable and the same then holds for all octants and, by finite additivity, for all of $\mathbb{R}^3$. This is also a consequence of more general characterization theorems; see [**10**, Theorem 17.41] or [**7**, Appendix A].

2. A tree equivalent to the Axiom of Dependent Choice. DC is equivalent to: Any nonempty tree with no leaves and a countable infinity of levels has a branch. Here a *tree* is a partially ordered set so that the set of elements above any vertex is well-ordered; a *branch* is a maximal linearly ordered subset. The proof starting from DC is easy: the tree leads to a relation $*$ and DC applies. For the other direction, choose any point x as the root and build a tree from finite sequences $(x, y, z, \ldots)$ for which $x * y$, $y * z$, and so on, using extension for the ordering. This use of sequences means the ordering is a tree. This tree property differs from the more familiar König Tree Lemma, which is about finitely branching trees. The Axiom of Countable Choice (AC_ω) follows from DC and suffices for the countable additivity of λ. But DC is preferred over AC_ω, one reason being that DC is equivalent to the Baire Category Theorem for complete metric spaces.

3. Measure and large cardinals. While there are many similarities between the concepts of Lebesgue measure and meager sets, there are important differences. The most famous is that LM is stronger in that its consistency requires the consistency of IC, the assertion that an inaccessible cardinal exists. To summarize much seminal and difficult work by Solovay, Shelah, and others, let Con(T) mean that the theory T is consistent. Then

- ZF + IC $\Rightarrow$ Con(ZF); hence by Gödel's Second Incompleteness Theorem we cannot derive Con(ZF + IC) from Con(ZF).
- Con(ZF) $\Leftrightarrow$ Con(ZF + DC + PB).
- Con(ZF + IC) $\Leftrightarrow$ Con(ZF + DC + LM).

In short, the consistency strength of LM is strictly greater than that of PB (but see the discussion of GM in Section 6 for how to avoid inaccessibles when negating the Banach–Tarski Paradox).

4. Extreme Division Paradoxes. Define a *κ-Division Paradox* to be a sequence $\{X_\alpha\}_{\alpha\le\kappa}$ so that $|X_0| < |X_1| < \ldots < |X_\kappa|$ and $|X_0| \ge^* |X_1| \ge^* \ldots \ge^* |X_\kappa|$. A Double Division Paradox is the case $\kappa = 2$. A. Karagila (personal communication) has shown that an $\aleph_0$-Division Paradox is consistent with ZF (he adds Cohen reals to a model of ZFC + CH; the proof works for larger κ too). So, for example, there can be a quadruple paradox: more teams than players, more conferences than teams, more leagues than conferences, and more sports than leagues. Multiple Division Paradoxes are also a consequence of the Axiom of Determinacy (AD). First one shows that AD implies $\eta = \aleph_1 < \aleph_\omega \le \Theta$, where η is the least cardinal that does not inject into $\mathbb{R}$ and Θ is the least cardinal that $\mathbb{R}$ does not map onto. This follows from AD $\Rightarrow$ every uncountable set of reals has a nonempty perfect subset $\Rightarrow \omega_1 \not\le |\mathbb{R}|$ and a result known as the soft version of the Moschovakis Coding Lemma [**7**, Theorem 28.15]: if $|\mathbb{R}| \ge^* \kappa$ then $|\mathbb{R}| \ge^* |2^\kappa|$. The cardinal inequality easily implies that $(\mathbb{R}, \mathbb{R} \cup \omega_1, \mathbb{R} \cup \omega_2, \mathbb{R} \cup \omega_3, \ldots)$ is an $\aleph_0$-Division Paradox.

ACKNOWLEDGMENTS. We are grateful to Andreas Blass for drawing our attention to the Division Paradox, to Matt Foreman, Menachem Magidor, Greg Moore, Andrew Marks, and Dan Velleman for much helpful correspondence, to Randall Dougherty, who provided the details of the injections used at the end of Section 5 and to Asaf Karagila, who provided the extension of the division paradoxes.

REFERENCES

1. B. Banaschewski, G. H. Moore, The dual Cantor–Bernstein Theorem and the Partition Principle, *Notre Dame J. Formal Logic* **31** (1990) 375–381.
2. R. Dougherty, S. Jackson, A. S. Kechris, The structure of hyperfinite Borel equivalence relations, *Trans. Amer. Math. Soc.* **341** (1994) 193–225.
3. S. Feferman, Mathematical intuition vs. mathematical monsters, *Synthèse* **125** (2000) 317–332.
4. C. S. Hardin, A. D. Taylor, A peculiar connection between the Axiom of Choice and predicting the future, *Amer. Math. Monthly* **115** (2008) 91–96.
5. ——, *The Mathematics of Coordinated Inference,* NewYork, Springer, 2013.
6. H. Herrlich, *Axiom of Choice*, Berlin, Springer, 2006.
7. D. Kerr, H. Li, *Ergodic Theory: Independence and Dichotomies,* in preparation.
8. T. J. Jech, *The Axiom of Choice,* New York, Dover, 2013.
9. A. Kanamori, *The Higher Infinite*, 2nd. ed., New York: Springer, 2003.
10. A. Kechris, *Classical Descriptive Set Theory,* Graduate Texts in Mathematics, vol. 156, New York, Springer, 1995.
11. G. H. Moore, *Zermelo's Axiom of Choice*, New York, Springer, 1982.
12. J. Mycielski, On the axiom of determinateness, *Fund. Math.* **53** (1964) 205–224.
13. ——, A system of axioms of set theory for the rationalists, *Notices Amer. Math. Soc.* **53** (2006) 206–213.
14. J. Mycielski, B. Osofsky, Problem 5937 solution, *Amer. Math. Monthly* **82** (1975) 308–309.
15. J. C. Oxtoby, *Measure and Category*, Graduate Texts in Mathematics, vol. 2, New York, Springer, 1971.
16. R. Penrose, *The Road to Reality,* New York, Knopf, 2005.
17. J. Raisonnier, A mathematical proof of S. Shelah's theorem on the measure problem and related results, *Isr. J. Math.* **48** (1984) 48–56.
18. S. Shelah, Can you take Solovay's inaccessible away?, *Isr. J. Math.* **48** (1984) 1–47.
19. W. Sierpiński, Sur quelques problèmes qui impliquent des fonctions non-mesurables, *Comptes Rendus Acad. Sci. Paris* **164** (1917) 882–884.
20. ——, L'axiome de M. Zermelo et son rôle dans la théorie des ensembles et l'analyse, *Bull. Intern. Acad. Sci. Cracovie* A (1918) 97–152.
21. ——, Fonctions additives non complètement additive et fonctions non mesurables, *Fund. Math.* **30** (1938) 96–99.
22. ——, Sur une proposition qui entraîne l'existence des ensembles non mesurables, *Fund. Math.* **34** (1947) 1–5.
23. R. M. Solovay, A model of set theory in which every set of reals is Lebesgue measurable, *Ann. Math.* **92** (1970) 1–56.
24. J. Stern, Regularity properties of definable sets of reals, *Ann. Pure App. Logic* **29** (1985) 289–324.
25. G. Tomkowicz, S. Wagon, *The Banach–Tarski Paradox,* 2nd ed., New York: Cambridge, 2016.
26. E. W. Weisstein, Harmonic expansion, *MathWorld*, A Wolfram Web Resource, http://mathworld.wolfram.com/HarmonicExpansion.html

ALAN D. TAYLOR, Union College, Schenectady, New York, 12308; taylora@union.edu

STAN WAGON, Macalester College, St. Paul, Minnesota 55105; wagon@macalester.edu