0pt 998pt 998pt

RANDOM REALS, THE RAINBOW RAMSEY THEOREM, AND ARITHMETIC CONSERVATION

CHRIS J. CONIDIS AND THEODORE A. SLAMAN

Abstract. We investigate the question "To what extent can random reals be used as a tool to establish number theoretic facts?" Let *2-RAN* be the principle that for every real X there is a real R which is 2-random relative to X. In Section 2, we observe that the arguments of Csima and Mileti [3] can be implemented in the base theory RCA_0 and so $RCA_0 + 2\text{-}RAN$ implies the Rainbow Ramsey Theorem. In Section 3, we show that the Rainbow Ramsey Theorem is not conservative over RCA_0 for arithmetic sentences. Thus, from the Csima-Mileti fact that the existence of random reals has infinitary-combinatorial consequences we can conclude that *2-RAN* has non-trivial arithmetic consequences. In Section 4, we show that *2-RAN* is conservative over $RCA_0 + B\Sigma_2$ for Π_1^1-sentences. Thus, the set of first-order consequences of *2-RAN* is strictly stronger than $P^- + I\Sigma_1$ and no stronger than $P^- + B\Sigma_2$.

§1. Introduction. One of the benefits to having a precise formulation of the concept "random infinite binary sequence" is that one can ask and precisely answer questions about what types of objects can be computed from random input and about what sorts of theorems can be proven from the existence of a random sequence.

In [3], Csima and Mileti gave an intriguing example of the first type. Their example concerns the Rainbow Ramsey Theorem, which states that if C is a coloring of size-k subsets of $\mathbb{N}$ such that there is a uniform finite bound on the number of sets assigned to any particular color, then there is an infinite set X such that C is injective on the size-k subsets of X, i.e. X is a C-rainbow. As is described below, they give a proof of the Rainbow Ramsey Theorem for pairs ($k = 2$) by showing that if R is a sufficiently random sequence relative to C, then R can be used to compute a C-rainbow. In a sentence, Csima and Mileti show how a random source can be used to produce a solution to a infinitary-combinatorial problem. Further, since they also show that there is a recursive such C with no recursive rainbow, any general method to produce rainbows for colorings must be driven by some such non-recursive data.

In this article, we analyze the second question with respect to theorems about finite sets. For this question, we shift the setting from the examination of the computation of infinite sets from a random source to the examination of proofs from the assertion that there is a random source.

As is documented in Simpson's text [15], there is a well-established methodology to frame the question. Namely, fix a base theory T, intended to capture a collection of tools with which to work with real numbers. Obtain a new theory T^* by adding to T the statement that for every real number X there is a real R which is random relative to X. Then ask the question, "Are there statements of first order arithmetic which are provable from T^* but not from T?" If not, then we say that T^* is conservative over T for arithmetic sentences.

We take as given the minimal base theory RCA_0, which formalizes the theory of first order arithmetic through induction for Σ_1^0 sets and the formal assertion that the real numbers are closed under relative computability.

Let *2-RAN* be the principle that for every real X there is a real R which is 2-random relative to X. We will review the hierarchy of arithmetical randomness below, which specifies 2-randomness exactly. We will show that *2-RAN* has number-theoretic consequences that go beyond those of RCA_0. In fact, *2-RAN* proves the principle $C\Sigma_2$ defined by Seetapun and Slaman in [14], a version of the pigeon-hole principle.

It is possible to prove $C\Sigma_2$ from *2-RAN* by directly manipulating the conditions on being 2-random. However, we will proceed in two steps. First, we will implement the Csima-Mileti argument in the theory $RCA_0 + $ *2-RAN* to derive the Rainbow Ramsey Theorem for pairs, RRT_2^2. Then, we will show that $RCA_0 + RRT_2^2$ proves $C\Sigma_2$, an apparently stronger non-conservation theorem.

Finally, we give an upper bound on the first order part of $RCA_0 + $ *2-RAN*. Since any model of first order Peano Arithmetic (*PA*) can be extended to a model of arithmetic comprehension, which is then a model of $RCA_0 + $ *2-RAN*, $RCA_0 + $ *2-RAN* is arithmetically conservative over *PA*. A more involved but still familiar argument, using the existence of 2-random reals that are low$_2$, would show that $RCA_0 + $ *2-RAN* is arithmetically conservative over Peano Arithmetic with induction limited to Σ_2-formulas [2].

We will argue for a sharper upper bound and show that if a ϕ is a Π_1^1-sentence which is provable from $RCA_0 + $ *2-RAN*, then ϕ is provable from $RCA_0 + B\Sigma_2$, the principle of bounding for relativized Σ_2^0 formulas.

1.1. Effective Randomness. We begin by reviewing the prerequisites on effective randomness. Downey and Hirschfeldt [4] and Nies [13] give thorough treatments of this material. We assume that the reader is familiar with the basics of recursion theory, such as the first few chapters of [4, 13, 16].

Let $\omega = \{0, 1, 2, \dots\}$ denote the standard set of natural numbers. Unless otherwise indicated, lower case letters $a, b, c, \dots, y, z$ (possibly with sub/superscripts) will always denote natural numbers. Let $2^{<\omega}$ denote the set of all finite binary strings, and $2^n \subset 2^{<\omega}$ denote the set of finite binary strings of length n. Let 2^ω be the set of all infinite binary strings. Elements of 2^ω are called *reals*. We will identify every real $f \in 2^\omega$ with the unique set $X_f \subseteq \omega$ such that f is the characteristic function of X_f. Unless otherwise indicated, lower case Greek letters $\alpha, \beta, \dots, \psi$ (possibly with sub/superscripts) will always denote finite binary sequences. We write $f \upharpoonright n$ to denote the first n bits of f.

For every finite binary string $\sigma \in 2^{<\omega}$, let $[\sigma] \subseteq 2^\omega$ denote the set of all infinite binary strings extending σ. More generally, for any set $A \subseteq 2^{<\omega}$, let $[A] \subseteq 2^\omega$ denote set of all infinite binary strings extending some element of A –i.e. $[A] = \cup_{\sigma \in A}[\sigma]$. The sets $[\sigma]$, $\sigma \in 2^{<\omega}$, form the basis of a topology on 2^ω. The resulting topological space is known

as *Cantor space*. In this context, an open set U is recursively enumerable if there is a recursively enumerable set of finite strings A such that $U = [A]$. Furthermore, for each $\sigma \in 2^{<\omega}$, let $|\sigma| \in \omega$ denote the number of bits of σ, and define the measure of $[\sigma]$ to be $2^{-|\sigma|}$. This measure extends uniquely to Cantor space, and is known as the *Lebesgue measure*. If $X \subseteq 2^\omega$ is measurable, then we denote its (Lebesgue) measure by $\mu(X)$.

A *Martin-Löf test* is a uniformly recursively enumerable sequence of open sets, $\{U_n\}_{n \in \omega}$, $U_n \subseteq 2^\omega$, such that $\mu(U_n) \leq 2^{-n}$ for all n. We say that $f \in 2^\omega$ is *1-random* whenever $f \notin \cap_{n \in \omega} U_n$ for any Martin-Löf test $\{U_n\}_{n \in \omega}$. More generally, we can relativize the definition of Martin-Löf test to any oracle $A \subseteq \omega$, by considering uniformly recursively enumerable open sets relative to A, and we say that $f \in 2^\omega$ is *A−random* whenever $f \notin \cap_{n \in \omega} U_n$ for any Martin-Löf test relative to A. We say that f is $(n + 1)$-*random* whenever f is random relative to $\emptyset^{(n)}$ (i.e. the nth Turing jump). One may also speak of n-randomness relative to an oracle.

We will mostly be concerned with the class of 2-random reals in the context of second-order arithmetic and RCA_0 (see Section 1.3 for more details). Since RCA_0 does not necessarily imply the existence of incomputable sets [15], in order to define the notion of a 2-random real in RCA_0 we need a characterization of 2-randomness that mentions only recursively definable sets/functions. Using the limit lemma, as well as the fact that $\Sigma_1^{\emptyset'}$ subsets of ω are the same as Σ_2 subsets of ω and the standard recursive approximation theorem for Σ_2 sets, it follows that $f \in 2^\omega$ is a 2-random real iff for all $\{0, 1\}$-valued total recursive approximation functions $g(n, \sigma, s)$ such that

$$U_n^g = \{\sigma : \liminf_{s \to \infty} g(n, \sigma, s) = 1\}$$

satisfies $\mu([U_n^g]) \leq 2^{-n}$, we have that $f \notin \cap_{n \in \omega} [U_n^g]$. In other words, for all such g there exists n such that for all $m \in \omega$ there exist infinitely many s such that $g(n, f \restriction m, s) = 0$.

A *tree* is a set $T \subseteq 2^{<\omega}$ that is closed under initial segments. If T is a tree then $[T] \subseteq 2^\omega$ denotes the set of infinite paths through T. For any given oracle $A \subseteq \omega$, we say that $X \subseteq 2^\omega$ is a Π_1^0-class relative to A, and write X is a $\Pi_1^{0,A}$-class, whenever there exists a tree $T_X \leq_T A$ such that $X = [T_X]$. $\Pi_1^{0,A}$-classes are exactly the effectively closed sets relative to the oracle A. We say that σ is *on* T whenever $\sigma \in T$. By taking complements while thinking of a $\Pi_1^{0,\emptyset'}$-class as the complement of a uniformly recursively enumerable sequence of open sets relative to $\emptyset'$ (i.e. a $\Sigma_1^{0,\emptyset'}$-class) and replacing g (above) with $1 - g$, the characterization of 2-randomness of the previous paragraph can be reworded to the following equivalent definition.

DEFINITION 1.1. A real $f \in 2^\omega$ is *2-random* if for all $\{0, 1\}$-valued recursive approximation functions $h(n, \sigma, s)$ such that for all n, σ, $\lim_{s \to \infty} h(n, \sigma, s)$ exists and

$$T_n^h = \{\sigma : \lim_{s \to \infty} h(n, \sigma, s) = 1\}$$

is a tree with $\mu([T_n^h]) > 1 - 2^{-n}$, we have that $f \in [T_n^h]$, for some n.

From now on we will take this to be the definition of 2-randomness. When we say $\mu([T_n^h]) > 1 - 2^{-n}$ we mean that for all r the fraction of finite binary strings of length r on T_n^h is strictly greater than $1 - 2^{-n}$. When we say that $f \in [T_n^h]$ we mean that for all r the string $f \restriction r$ is on T_n^h.

In the above, it would be sufficient to restrict to sequences of trees which are uniformly computable from $X^{(k)}$ in polynomial time, or within any other reasonable

time bound. For example, given a recursive tree T, we could define the new tree T^* by letting σ belong to T^* if and only if for every initial segment τ of σ, there is no computation that $\tau \notin T$ which converges in time $|\sigma|$. Then, $[T] = [T^*]$ and whether $\sigma \in T^*$ can be computed in a polynomial of $|\sigma|$ many steps.

Finally, we recall the following fact.

PROPOSITION 1.2 (Martin-Löf [12]). *There is a recursive tree T such that every infinite path in T is 1-random and such that $[T]$ has positive measure. Furthermore, this fact holds relative to any set.*

1.2. Effective Infinitary Combinatorics. Recently, effective infinitary combinatorics has played a prominent role in recursion theory and reverse mathematics. One of the earliest works in effective infinitary combinatorics is [7], in which Jockusch examined the effectivity of *Ramsey's Theorem* (for $n-$tuples and k colors). Fix natural numbers n, k, and let $[\omega]^n$ denote the set of unordered $n-$tuples of natural numbers. Ramsey's Theorem for $n-$tuples and k colors says that for any given function/coloring $c : [\omega]^n \to \{0, 1, \dots, k-1\}$ that assigns a color to each $n-$tuple there exists an infinite set $H \subseteq \omega$ such that c is constant on $[H]^n \subseteq [\omega]^n$ –i.e. the set of $n-$tuples all of whose members are contained in H. We say that H is *homogeneous* for (or with respect to) the coloring c. It is not difficult to see that if $n = 1$ and c is recursive coloring then there is a recursive homogeneous set for c. In [7], Jockusch constructs a recursive coloring of pairs $c_2 : [\omega]^2 \to \{0, 1\}$ with no recursive homogeneous set. Jocksuch also constructs a recursive coloring of triples $c_3 : [\omega]^3 \to \{0, 1\}$ such that every infinite homogeneous set for c_3 computes Turing's Halting Set. For more details consult [7].

More recently, Csima and Mileti [3] studied the effectivity of the *Rainbow Ramsey Theorem* (for n tuples and $k-$bounded colorings). Fix numbers n, k. We say that a coloring of $n-$tuples $c : [\omega]^n \to \omega$ is $k-$*bounded*, $k \geq 1$, whenever $|c^{-1}(x)| \leq k$ for all $x \in \omega$. The Rainbow Ramsey Theorem for $n-$tuples and $k-$bounded colorings, $k \geq 1$, says that if $c : [\omega]^n \to \omega$ is a $k-$bounded coloring of $n-$tuples then there exists an infinite set $R \subseteq \omega$ such that c is injective on $[R]^n$. We call R a rainbow for c. Among other things, Csima and Mileti construct a recursive $2-$bounded coloring of pairs $c_2 : [\omega]^2 \to \omega$ for which there is no recursive rainbow for c_2. They also show that if $c : [\omega]^2 \to \omega$ is a recursive k-bounded coloring of pairs and X is a 2-random real, then X computes a rainbow for c. For more details, consult [3] or the discussion in Section 2.

1.3. Subsystems of Second Order Arithmetic. We now give the necessary background on second-order arithmetic. Simpson [15] gives a thorough treatment of this material, and much more.

The basic symbols of first-order logic are $\forall, \exists, \neg, \vee, \wedge, \to, \leftrightarrow, =, (,)$, and variables $x_1, x_2, \dots$ along with the standard symbols from arithmetic $0, 1, +, \cdot, <$. Formulas and sentences are constructed in the usual way.

Let P^- denote the axioms of Peano arithmetic, without the induction scheme. In other words, P^- consists of the axioms for the nonnegative part of a discretely ordered ring. All of the theories we consider will be extensions of P^-, and all of our models will satisfy P^-.

The language of second-order arithmetic is an extension of the language of first-order arithmetic via *set variables $X_1, X_2, \dots$*. More specifically, the language of second order arithmetic, denoted L_2, is a *two-sorted* language consisting of number variables

$x_1, x_2, \ldots$, and set variables $X_1, X_2, \ldots$. The number variables $x_1, x_2, \ldots$ range over natural numbers, while the set variables $X_1, X_2, \ldots$ range over sets of natural numbers. Formulas and sentences are constructed in the usual way. Boldface characters such as $\mathbf{x}, \mathbf{Y}$ represent sequences of type x, Y, respectively.

The axioms of second-order arithmetic include P^- (above), along with the induction axiom

$$0 \in X \wedge (\forall x)[x \in X \to x + 1 \in X] \to (\forall x)[x \in X]$$

and the comprehension scheme

$$(\exists X)(\forall x)[x \in X \leftrightarrow \varphi(x)],$$

where φ is an L_2-formula in which X does not occur freely.

Second-order arithmetic consists of the formal system consisting of the axioms of second-order arithmetic, together with all formulas deducible from these axioms via the usual logical rules of inference. A *subsystem of second-order arithmetic* is a formal system consisting of some subset of the theorems of second order arithmetic.

RCA_0 is the subsystem of second-order arithmetic consisting of P^-, along with induction for Σ_1 formulas, and comprehension for Δ_1-definable sets of natural numbers (see Section 1.3.1 below). RCA_0 will be our base theory throughout the rest of this article. *2-RAN* denotes the theorem of second-order arithmetic that says "for every set X there exists Y such that Y is 2-random relative to X" (recall our definition of 2-randomness given in the final paragraph of Section 1.1 above). *2-RAN* will play a prominent role in Sections 2 and 4 below. We will use $\mathbb{N}$ to denote the first-order part of a model of RCA_0, and when working in a subsystem of second-order arithmetic we will sometimes use the term "bounded" to mean Σ_0-definable and "recursive" to mean Δ_1-definable.

1.3.1. *Arithmetic Formulas.* We review some basic well-known facts concerning arithmetic formulas in second-order arithmetic. More information on the material presented here and more can be found in [5, 8, 15].

Recall that a formula $\varphi(x, \mathbf{p}, \mathbf{X})$ is Σ_n, $n \geq 1$, if it is of the form

$$\varphi(x, \mathbf{p}, \mathbf{X}) = (\exists x_1)(\forall x_2) \cdots [\phi(x, \mathbf{p}, \mathbf{X}, x_1, x_2, \ldots)],$$

where there are $n - 1$ alternations of quantifiers and ϕ has only bounded quantifiers and parameters $\mathbf{p}, \mathbf{X}$. Similarly, a formula $\varphi(x, \mathbf{p}, \mathbf{X})$ is Π_n, $n \geq 1$, if it is of the form

$$\varphi(x, \mathbf{p}, \mathbf{X}) = (\forall x_1)(\exists x_2) \cdots [\phi(x, \mathbf{p}, \mathbf{X}, x_1, x_2, \ldots)],$$

where there are $n - 1$ alternations of quantifiers and ϕ has only bounded quantifiers and parameters $\mathbf{p}, \mathbf{X}$. A set $X \subseteq \omega$ is Δ_n-definable whenever it is defined by both a Σ_n formula and a Π_n formula.

$I\Sigma_n$ represents the induction scheme for all Σ_n formulas. More specifically, $I\Sigma_n$ consists of all the sentences

$$(\forall \mathbf{X})(\forall \mathbf{p})[(\varphi(0, \mathbf{p}, \mathbf{X}) \wedge (\forall x)(\varphi(x, \mathbf{p}, \mathbf{X}) \to \varphi(x + 1, \mathbf{p}, \mathbf{X}))) \to (\forall x)\varphi(x, \mathbf{p}, \mathbf{X})]$$

where φ is Σ_n. More intuitively, $I\Sigma_n$ says that if a set A is defined by a Σ_n formula, $0 \in A$, and A is closed under the successor function, then every number is in A. It is known that $I\Sigma_n$ is equivalent to $I\Pi_n$, the induction scheme for Π_n formulas.

Similarly, $B\Sigma_n$ represents the bounding scheme for all Σ_n formulas. More specifically, $B\Sigma_n$ consists of all the sentences

$$(\forall \mathbf{X})(\forall \mathbf{p})(\forall a)[(\forall x < a)(\exists y)\varphi(x, y, \mathbf{p}, \mathbf{X}) \rightarrow (\exists b)(\forall x < a)(\exists y < b)\varphi(x, y, \mathbf{p}, \mathbf{X})]$$

where φ is Σ_n (with parameters $\mathbf{p}, \mathbf{X}$). More intuitively, $B\Sigma_n$ says that if for every number x less than a, there are numbers y that satisfy a Σ_n-property relative to x, then there is a bound b such that for all x less than a there is such a y that is less than b. It is known that $B\Sigma_n$ is equivalent to $B\Pi_{n-1}$, the bounding scheme for Π_{n-1} formulas.

Working in the formal system P^-, Kirby and Paris [11] proved that the schemes $I\Sigma_n$ and $B\Sigma_m$, $n, m \geq 1$, form a strict hierarchy as follows (no arrow is reversible):

$$\cdots \longrightarrow B\Sigma_{n+1} \longrightarrow I\Sigma_n \longrightarrow B\Sigma_n \longrightarrow I\Sigma_{n-1} \longrightarrow \cdots \longrightarrow I\Sigma_1 \longrightarrow B\Sigma_1.$$

Note that, in the context of Reverse Mathematics, Definition 1.1 above only makes sense in the context of $B\Sigma_2$, since it mentions $f \in [T_n^h]$ for some limit computable tree T_n^h.

1.3.2. *Formalizing effective randomness.*

PROPOSITION 1.3. *Suppose that $(\mathcal{M}, R) \models RCA_0 + B\Sigma_2$. Then, for every $X \in R$, there is an unbounded X'-recursive tree T_u such that if Y is an infinite path in T_u, then Y is 2-random relative to X.*

PROOF. Here, one need only check that the proof of Proposition 1.2 can be carried out relative to X' using $B\Sigma_2(X)$, which we omit.   ⊣

PROPOSITION 1.4. *Suppose that $(\mathcal{M}, R) \models RCA_0 + B\Sigma_2$ and $X \in R$. Then, if Y is 2-random relative to X, then $(\mathcal{M}, X, Y) \models X' + Y \geq_T Y'$. Further, for any Σ_2-formula $\theta_2(X, Y)$ there is a Σ_1-formula $\theta_1(X', Y)$ such that*

$$(\mathcal{M}, X' + Y) \models \theta_2(X, Y) \leftrightarrow \theta_1(X', Y).$$

PROOF. First, we recall the argument that if Y is 2-random, then $Y + 0' \geq_T Y'$.

Consider an $\epsilon > 0$ and a $\Sigma_1^0(Z)$-formula $(\exists w)\theta(w, Z \restriction w)$, where θ has only bounded quantifiers. There is a recursive enumeration of the set A consisting of those finite sequences τ at which this formula first appears true, i.e. of the set of τ satisfying the following formula.

$$\theta(|\tau|, \tau) \, \& \, (\forall w < |\tau|)\neg\theta(w, \tau \restriction w)$$

Further, $0'$ can compute a cut-off point s such that all but measure less than ϵ of $[A]$ is contained in $[A \cap 2^{\leq s}]$. Namely, $0'$ can compute the least s in a non-empty Π_1^0-subset of $\mathbb{N}$. Now, uniformly in e, let U_e be the measure less than 2^{-e} subset of those strings at which the conjunction of the first e-many Σ_1^0-formulas

$$(\exists w_1)\theta_1(w_1, Z \restriction w_1) \wedge (\exists w_2)\theta_2(w_2, Z \restriction w_2) \wedge \cdots \wedge (\exists w_e)\theta_e(w_e, Z \restriction w_e)$$

first appears true and let s_e be the associated cut-off stage after which we begin to enumerate U_e. The sequence $(U_e : e \in \mathbb{N})$ yields a Martin-Löf test relative to $0'$ and the function $e \mapsto s_e$ is recursive in $0'$. So, if Y is 2-random, then there is an n such that for all $e > n$, $Y \notin [U_e]$. Further, for $e > n$, $(\exists w)\theta_e(w, Y \restriction w)$ holds if and only if $(\exists w < s_e)\theta_e(w, Y \restriction w)$. Thus, the complete $\Sigma_1^0(Y)$ subset of $\mathbb{N}$ is recursive in $Y + 0'$, as required.

The above argument goes through unchanged in the formal theory $RCA_0 + B\Sigma_2$ relativized to a real parameter. Formally, one has to note that the definition of A is valid, a use of bounded induction, and to demonstrate that the measure of a Σ_1^0 set $[A]$ can be approximated within arbitrary accuracy, a use of Σ_1^0-induction.

Now, consider the second claim, that Σ_2^0-statements about Y are equivalent to Σ_1^0-statements about $Y + 0'$. The above analysis gives a translation from $\Sigma_1^0(Y)$-statements to Δ_1^0-statements about $Y + 0'$, which also translates $\Sigma_2^0(Y)$ to $\Sigma_1^0(Y + 0')$. The other direction follows from the observation that Σ_1^0-induction is sufficient to ensure that the enumeration of any initial segment of $0'$ is bounded. $\dashv$

1.3.3. *Formalizing effective combinatorics.* We now define the formal subsystems of second order arithmetic RT_2^2 (Ramsey's Theorem for pairs and 2 colors) and RRT_2^2 (Rainbow Ramsey for pairs and $2-$bounded colorings), and state a well-known conservation theorem for RT_2^2 due to Cholak, Jockusch, and Slaman [1].

Let RT_2^2 denote the L_2-sentence that says

> "For every coloring of (unordered) pairs $c : [\mathbb{N}]^2 \to \{0, 1\}$ there exists an infinite set $H \subseteq \mathbb{N}$ such that c is constant on $[H]^2 \subseteq [\mathbb{N}]^2$."

In [1], Cholak, Jockusch, and Slaman show that $RCA_0 + I\Sigma_2 + RT_2^2$ is conservative over $RCA_0 + I\Sigma_2$ for Π_1^1 statements.

We define RRT_2^2 to be the L_2-sentence that says

> "For every $k-$bounded coloring of unordered pairs $c : [\mathbb{N}]^2 \to \mathbb{N}$, $k \geq 1$, there exists an infinite set $R \subseteq \mathbb{N}$ such that c is injective on $[R]^2 \subseteq [\omega]^2$."

In the next section, we show that $RCA_0 + 2\text{-}RAN$ implies RRT_2^2. In Section 3, we prove that RRT_2^2 is not conservative over RCA_0 for arithmetic statements.

§2. The Csima-Mileti Argument in $RCA_0 + 2\text{-}RAN$. We assume that the reader is familiar with [3]. Our main goal in this section is to sketch a proof, *in the formal system $RCA_0 + 2\text{-}RAN$*, of the theorem of Csima and Mileti [3, Section 3] that says for any recursive 2-bounded coloring of pairs $c : [\mathbb{N}]^2 \to \mathbb{N}$ and any 2-random real X, there is a rainbow $R \subseteq \mathbb{N}$ for c that is recursive in X. We reason in $RCA_0 + 2\text{-}RAN$. Fix a recursive 2-bounded coloring of pairs $c : [\mathbb{N}]^2 \to \mathbb{N}$.

Since the proof of [3, Proposition 3.3] gives an effective algorithm that constructs an infinite set of numbers $N_0 \subseteq \mathbb{N}$ such that c is a *normal* coloring when restricted to pairs in N_0, and uses at most Σ_0-induction to verify this fact, it follows that [3, Proposition 3.3] holds in RCA_0 and so without any loss of generality we can assume that c is normal. Similarly, since the proof of [3, Proposition 3.5] can be carried out effectively (its essence lies in an application of $B\Sigma_1$ and the Δ_1 finitary pigeonhole principle), it follows that [3, Proposition 3.5] holds in RCA_0. Also, for any given rational number $p < 1$, the proof of [3, Proposition 3.9] can be carried out uniformly in p to produce a binary tree T_p such that $\mu([T_p]) \geq p$. Finally, recall that the trees constructed in [3, Section 3] are all recursively enumerable (i.e. Σ_1-definable).

For each $n \in \mathbb{N}$ let $\epsilon_0^n, \epsilon_1^n, \ldots \in \mathbb{Q}$ be a recursive sequence of rational numbers less than one such that for each $m \in \mathbb{N}$, $\prod_{i=0}^m \epsilon_i^n > 1 - 2^{-n}$. Let $\beta_0, \beta_1, \ldots \in \mathbb{N}$ be defined as $\beta(0), \beta(1), \ldots$ are in the proof [3, Proposition 3.9]. By recursively and uniformly approximating the method of Csima and Mileti described in [3, pages 1316-1318] we can

construct, in the formal system $RCA_0 + 2\text{-}RAN$, a $\{0,1\}$-valued recursive approximation function $t(n, \sigma, s)$, $\sigma \in 2^{<\mathbb{N}}$, $n, s \in \mathbb{N}$, such that:

1. for all n, σ, $t(n, \sigma, s)$ nondecreasing in s; (hence $\lim_{s\to\infty} t(n, \sigma, s) = 1$ iff $(\exists s)[t(n, \sigma, s) = 1]$)
2. for all n, s, $t(n, \sigma, s) = 1$ whenever $t(n, \sigma z, s) = 1$ for some $z \in \{0, 1\}$;
3. for all $i, n \in \mathbb{N}$ and $\sigma \in 2^{\beta_i}$ such that $t(n, \sigma, s) = 1$ for some $s \in \mathbb{N}$, the fraction of strings $\tau \in 2^{\beta_{i+1}}$ extending σ such that $t(n, \sigma, t) = 1$ for some $t \in \mathbb{N}$ is at least ϵ_i^n.

Intuitively speaking, we think of $t(n, \sigma, s)$ as a uniform recursive approximation to a sequence of trees $\{T_n\}_{n\in\mathbb{N}}$ such that for all $n \in \mathbb{N}, \sigma \in 2^{<\mathbb{N}}$,

$$\sigma \in T_n \text{ iff } \lim_{s\to\infty} t(n, \sigma, s) = 1 \text{ iff } (\exists s)[t(n, \sigma, s) = 1].$$

Property 2 says that T_n is a tree for all $n \in \mathbb{N}$. Property 1 follows from the fact that the Csima-Mileti trees are recursively enumerable, and property 3 follows from [3, Proposition 3.5]. In the next paragraph we will use properties 1, 2, and 3 to sketch an argument that says for all $n \in \mathbb{N}$ the measure of T_n is at least $\prod_{i\in\mathbb{N}} \epsilon_i \geq 1 - 2^{-n}$.

We now sketch a proof that says for all $m, n \in \mathbb{N}$ the fraction of strings $\sigma \in 2^{\beta_m}$ for which there exists a stage $s \in \mathbb{N}$ such that $t(n, \sigma, s) = 1$ is at least $\prod_{i=0}^m \epsilon_i^n > 1 - 2^{-n}$. By property 2 it will then follow that for all $m, n \in \mathbb{N}$ the fraction of $\sigma \in 2^m$ for which there exists a stage $s \in \mathbb{N}$ such that $t(n, \sigma, s) = 1$ is strictly greater than $1 - 2^{-n}$. Suppose that we are given $m, n \in \mathbb{N}$. Then, by property 1 and RCA_0, there exists a large stage s_0 such that our recursive approximation $t(n, \sigma, s)$ has stabilized on all nodes of length at most β_m by stage s_0.[1] Now, using property 3 above and Σ_1-induction on the length of nodes up to length β_m, one can show that the fraction of nodes $\sigma \in 2^{\beta_m}$ such that $t(n, \sigma, s_0) = 1$ is at least $\prod_{i=0}^m \epsilon_i^n > 1 - 2^{-n}$ (this is essentially the argument that Csima and Mileti give in [3, Proposition 3.9]).

Now, the axiom $2\text{-}RAN$ says that there exists a real $X \in 2^{\mathbb{N}}$ such that $X \in [T_n]$, for some $n \in \mathbb{N}$, and via the same decoding procedure described in [3, Proposition 3.9] there is an X-recursive rainbow R for c. Thus, we have essentially used the Csima-Mileti method [3, Section 3] to prove the following theorem.

THEOREM 2.1 ($RCA_0 + 2\text{-}RAN$). *Let* $c : [\mathbb{N}]^2 \to \mathbb{N}$ *be a 2-bounded coloring of pairs and X be a 2-random real. Then there is an X-recursive rainbow R for c.*

As in [3, Section 3], one can easily modify our method to prove the following more general theorem.

THEOREM 2.2 ($RCA_0 + 2\text{-}RAN$). *Let* $c : [\mathbb{N}]^2 \to \mathbb{N}$ *be a k-bounded coloring of pairs, $k \in \mathbb{N}$, $k \geq 2$, and let X be a 2-random real. Then there is an X-recursive rainbow R for c.*

§3. Non-conservativity of the Rainbow Ramsey Theorem over RCA_0. In this section, we show that the Rainbow Ramsey Theorem for Pairs, in the form RRT_2^2, is not arithmetically conservative over $P^- + I\Sigma_1$. Our argument is based on the Seetapun

---

[1] Here we are using the fact that recursively enumerable (i.e. Σ_1-definable) sets of natural numbers are amenable in RCA_0. A set defined by a formula φ is *amenable* in a theory T whenever every initial segment of that set exists in every model of T (though the set itself might not exist). The amenability of recursively enumerable sets in RCA_0 is well-known and follows from the fact that any recursive (i.e. Δ_1) approximation to an r.e. (i.e. Σ_1) set changes at most 2^k times on the first k bits.

and Slaman [14] proof of the same fact for Ramsey's Theorem for Pairs, RT^2_2. It is open whether Hirst's [6] theorem, that RT^2_2 implies $B\Sigma_2$ over RCA_0, applies to RRT^2_2.

3.1. The Cardinality Scheme. Now, we recall a restricted form of the bounding principle.

DEFINITION 3.1 (Seetapun and Slaman, [14]). We let Γ be a set of formulas and define the *cardinality scheme* $C\Gamma$ for Γ. If $\phi(x, y) \in \Gamma$ then the universal closure of the following formula is in $C\Gamma$: If $\phi(x, y)$ *defines an injective function, then its range is unbounded.*

Instances of the cardinality scheme follow from instances of the bounding scheme by a pigeon hole argument. A Skolem hull argument, see Slaman and Seetapun [14], shows that $P^- + I\Sigma_1 + C\Sigma_2$ is strictly stronger than $P^- + I\Sigma_1$. Kaye [9] and [10] study versions of $C\Gamma$ and compare them with the hierarchy of induction and collection. In particular, [10] exhibited structures to establish that $P^- + B\Sigma_n + \bigcup_{k\in\omega} C\Sigma_k$ does not prove $B\Sigma_{n+1}$. More recently Slaman has shown that $P^- + I\Sigma_n + \bigcup_{k\in\omega} C\Sigma_k$ does not prove $B\Sigma_{n+1}$. The case of interest in the context of *2-RAN* and RRT^2_2 is that $P^- + I\Sigma_1 + C\Sigma_2$ does not prove $B\Sigma_2$.

THEOREM 3.2 (Slaman, unpublished). $P^- + I\Sigma_n + \bigcup_{k\in\omega} C\Sigma_k \nvdash B\Sigma_{n+1}$

3.2. RRT^2_2 implies $C\Sigma_2$.

THEOREM 3.3. *Working over the base theory RCA_0, RRT^2_2 proves $C\Sigma_2$.*

PROOF. We assume that $\mathcal{M}$ is a model of $P^- + I\Sigma_1$ and that $C\Sigma_2$ fails in $\mathcal{M}$. We then adapt an argument from Seetapun and Slaman [14] to exhibit a recursive 2-bounded coloring of pairs c for which there can be no rainbow X such that $I\Sigma_1$ holds relative to X.

By the failure of $C\Sigma_2$ in $\mathcal{M}$, choose an element a of $\mathcal{M}$ and a Σ_2 formula

$$\phi(x, y) = (\exists w)(\forall z)\phi_0(x, y, w, z),$$

in which ϕ_0 is bounded and possibly mentions parameters from $\mathcal{M}$, such that $\phi(x, y)$ defines an injection G from the entirety of $\mathcal{M}$ into the set of elements of $\mathcal{M}$ which are less than a.

First, we define the stage-s recursive approximation $g(-, s)$ to G^{-1} as follows. Let $\prec$ be the ordering on pairs defined by $(x_1, w_1) \prec (x_2, w_2)$ if the maximum of $\{x_1, w_1\}$ is less than the maximum of $\{x_2, w_2\}$ or the maxima are equal and (x_1, w_1) is smaller than (x_2, w_2) lexicographically. Then, $\prec$ is a total order of $\mathcal{M} \times \mathcal{M}$ of order-type $\mathcal{M}$. Now, for $y < a$, let $g(y, s)$ be x, if there is a $w < s$ such that $(\forall z < s)\phi_0(x, y, w, z)$ and also (w, x) is the $\prec$-least pair such that $(\forall z < s)\phi_0(x, y, w, z)$. If there is no x as above, then we let $g(y, s) = 0$.

If $G(x) = y$, then by $I\Sigma_1$ there is a least w such that $(\forall z)\phi_0(x, y, w, z)$. Further, since G is injective, if $x_1 \neq x$ or if $w_1 < w$, then $\neg(\forall z)\phi_0(x_1, y, w_1, z)$. Consequently, if $(x_1, w_1) \prec (x, w)$ then $(\exists z)\neg\phi_0(x, y, w, z)$. By $B\Sigma_1$, there is a b such that for all $(x_1, w_1) \prec (x, w)$, $(\exists z < b)\neg\phi_0(x, y, w, z)$. But then for all $s > b$, $g(y, s) = x$. In other words, if y is in the range of G then $\lim_{s\to\infty} g(y, s)$ exists and is equal to $G^{-1}(y)$.

Now, we define the recursive 2-bounded coloring $c : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$ by recursion. At stage s of the recursion, we define $c(x, s)$ for all x less than s. Fix an $\mathcal{M}$-recursive Gödel numbering of subsets of $\mathcal{M}$, extending the usual Gödel numbering of the finite

subsets of $\mathbb{N}$. We will say that a set is $\mathcal{M}$-finite if it has a Gödel number in $\mathcal{M}$. By recursion on $y < a$, if $g(y, s)$ is the Gödel number of a set A_y of size greater than $2a$, then choose b_1 and b_2 from that set such that $c(b_1, s)$ and $c(b_2, s)$ are not yet defined. Since the stage-s recursion has length a and A_y has size greater than $2a$, there must be such b_1 and b_2. Define c so that $c(b_1, s) = c(b_2, s)$ with a value not previously used. Once this stage-s recursion is complete, assign unique and previously unused values to $c(x, s)$, for each x such that $c(x, s)$ was not already defined. By, $I\Sigma_1$, $c(x, s)$ is defined for every pair (x, s) with $x < s$. By construction, c is 2-bounded.

By construction, for each $\mathcal{M}$-finite set A of size greater than $2a$, there are only boundedly many elements s of $\mathcal{M}$ such that $A \cup \{s\}$ is a rainbow. If X is an unbounded subset of $\mathcal{M}$ and $(\mathcal{M}, X)$ satisfies $I\Sigma_1(X)$, then the order type of the elements of X is that of $\mathcal{M}$ and every initial segment of X is $\mathcal{M}$-finite. Consequently, X cannot be a c-rainbow.

In conclusion, if $\mathcal{M}$ is the number-theoretic component of a model of $RCA_0 + RRT_2^2$, then $\mathcal{M} \models P^- + I\Sigma_1$ and for every $\mathcal{M}$-recursive 2-bounded coloring, there is a rainbow X such that $(\mathcal{M}, X)$ satisfies $I\Sigma_1(X)$. By the above, $\mathcal{M}$ must satisfy $C\Sigma_2$. ⊣

§4. Π_1^1-Conservativity of *2-RAN* over $RCA_0 + B\Sigma_2$. We follow the now-standard method for establishing Π_1^1-conservativity, originally due to Harrington and described in Simpson [15].

THEOREM 4.1. *Suppose that $(\mathcal{M}, R)$ is a countable model of $RCA_0 + B\Sigma_2$ and that there is an element $X \in R$ such that for every $Y \in R$, $(\mathcal{M}, R) \models Y$ is recursive in X. There is an R^* such that $R \subseteq R^*$ and $(\mathcal{M}, R^*) \models RCA_0 + B\Sigma_2 + 2\text{-RAN}$.*

Before proving the theorem, we draw the immediate corollary.[2]

COROLLARY 4.2. *$RCA_0 + B\Sigma_2 + 2\text{-RAN}$ is conservative over $RCA_0 + B\Sigma_2$ for Π_1^1-sentences.*

PROOF OF COROLLARY 4.2. We show that if a Σ_1^1-sentence is consistent with the theory $RCA_0 + B\Sigma_2$, then it is consistent with the stronger theory $RCA_0 + B\Sigma_2 + 2\text{-RAN}$, which is sufficient to prove the corollary.

Suppose that $\exists X \psi(X)$ is Σ_1^1, where X is a second order variable and $\psi(X)$ is an arithmetic statement about X, and suppose that $\exists X \psi(X)$ is consistent with $RCA_0 + B\Sigma_2$. Let $(\mathcal{M}, R_{\mathcal{M}})$ be a countable model of $RCA_0 + B\Sigma_2 + \exists X \psi(X)$ and take X so that $X \in R_{\mathcal{M}}$ and $(\mathcal{M}, R_{\mathcal{M}}) \models \psi(X)$. Finally, let R be the subset of $R_{\mathcal{M}}$ consisting of those $Y \in R_{\mathcal{M}}$ such that $(\mathcal{M}, R_{\mathcal{M}}) \models X \geq_T Y$. Then,

$(\mathcal{M}, R)$ satisfies the hypothesis of Theorem 4.1. So, let R^* be a superset of R such that $(\mathcal{M}, R^*) \models RCA_0 + B\Sigma_2 + 2\text{-RAN}$. Since $X \in R$ and $R \subseteq R^*$, $X \in R^*$. The satisfaction of $\psi(X)$ within $(\mathcal{M}, R^*)$ depends only on $\mathcal{M}$ and X, hence $(\mathcal{M}, R_{\mathcal{M}}) \models \psi(X)$ implies that $(\mathcal{M}, R^*) \models \psi(X)$. Thus, $(\mathcal{M}, R^*) \models RCA_0 + B\Sigma_2 + 2\text{-RAN} + (\exists X)\psi(X)$, which establishes the required consistency of that theory. ⊣

Now, we turn to the proof of the theorem itself.

PROOF OF THEOREM 4.1. Let $\mathcal{M}$, R, and X be fixed as in the statement of Theorem 4.1.

We extend R to R^* in ω-many steps, so as to close under instances of *2-RAN*. At each step, we will start with a model, like $(\mathcal{M}, R)$, of $RCA_0 + B\Sigma_2$ in which there is a

---

[2]Corollary 4.2 was proven independently by Wei Wang (unpublished).

set X^* of greatest Turing degree, like X. We will adjoin a new subset of $\mathcal{M}$ which is 2-random relative to X^* and which satisfies $B\Sigma_2$ relative to X^*. We complete the step by closing under relative Δ_1^0 definability. It is then sufficient to show that we can carry out the plan for one step. Without loss of generality, we will show that we can carry out the first step and extend $(\mathcal{M}, R)$ by adding a set Y which is 2-random relative to X and which preserves $B\Sigma_2$.

Let T_u be the X'-recursive tree of Theorem 1.3, for which any infinite path is 2-random relative to X. We will build our Y by forcing with unbounded X'-recursive subtrees of T_u.

Now, fix an an unbounded X'-recursive tree T (ultimately, to be contained in T_u) and consider the hypothesis $(\forall x < a)(\exists y)\theta(x, y, (X' + Y) \restriction y)$ of an instance of $B\Sigma_1(X' + Y)$, in which θ is bounded.

If for every $x < a$, there is an ℓ such that for every $\sigma \in T$ of length ℓ, $(\exists y < \ell)\theta(x, y, (X'+\sigma) \restriction y)$, then $B\Sigma_2$ implies that there is a single ℓ such that for every $x < a$ and for every $\sigma \in T$ of length ℓ, $(\exists y < \ell)\theta(x, y, (X' + \sigma) \restriction y)$. Thus, if Y is an infinite path in T, then

$$(\mathcal{M}, \{X, Y\}) \models \left(\begin{array}{l} (\forall x < a)(\exists y)\theta(x, y, (X' + Y) \restriction y) \to \\ (\exists \ell)(\forall x < a)(\exists y < \ell)\theta(x, y, (X' + Y) \restriction y). \end{array} \right)$$

If, otherwise, there is an $x < a$ such that there is no ℓ such that for every $\sigma \in T$ of length ℓ, $(\exists y < \ell)\theta(x, y, (X'+\sigma) \restriction y)$, then the set of σ such that $\neg(\exists y < |\sigma|)\theta(x, y, (X'+\sigma) \restriction y)$ is an unbounded X'-recursive subtree T_x of T. Further, if Y is an infinite path in T_x, then

$$(\mathcal{M}, \{X, Y\}) \models \neg(\exists y)\theta(x, y, (X' + Y) \restriction y).$$

Hence, for any instance of $B\Sigma_1(X' + Y)$ and any unbounded X'-recursive tree T, there is an unbounded X'-recursive subtree of T which ensures the satisfaction of that instance of $B\Sigma_1(X' + Y)$ for all unbounded paths in the subtree. So, for Y generic, $(\mathcal{M}, X' + Y) \models B\Sigma_1$.

Since we fix our initial condition to be the tree T_u, a generic Y is 2-random and Proposition 1.4 applies. Thus, $B\Sigma_1(X' + Y)$ is equivalent to $B\Sigma_2(X, Y)$. Consequently, for Y generic as above, adjoining Y to $(\mathcal{M}, R)$ and closing under join and $\geq_T$ produces a model of the appropriate form, with a sequence Y which is 2-random relative to X, which satisfies $B\Sigma_2$, as was required. $\dashv$

§5. Conclusion. The results of the previous sections bracket the set of first-order consequences of *2-RAN* between $P^- + C\Sigma_2$ and $P^- + B\Sigma_2$. This leaves the following problem as still open.

PROBLEM 5.1. Characterize the set of first-order consequences of *2-RAN*.

REFERENCES

[1] PETER A. CHOLAK, CARL G. JOCKSUCH, and THEODORE A. SLAMAN, *On the strength of Ramsey's theorem for pairs*, **J. Symbolic Logic**, vol. 66 (2001), no. 1, pp. 1–55.

[2] PETER A. CHOLAK, CARL G. JOCKUSCH, and THEODORE A. SLAMAN, *On the strength of Ramsey's theorem for pairs*, **J. Symbolic Logic**, vol. 66 (2001), no. 1, pp. 1–55.

[3] BARBARA F. CSIMA and JOSEPH R. MILETI, *The strength of the rainbow Ramsey theorem*, **J. Symbolic Logic**, vol. 74 (2009), no. 4, pp. 1310–1324.

[4] R. G. DOWNEY and D. HIRSCHFELDT, *Algorithmic randomness and complexity*, Theory and Applications of Computability. New York, NY: Springer. xxvi, 855 p., 2010.

[5] P. HÁJEK and P. PUDLÁK, *Metamathematics of first-order arithmetic*, Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1998, Second Printing.

[6] J. L. HIRST, *Combinatorics in subsystems of second order arithmetic*, **Ph.D. thesis**, The Pennsylvania State University, 1987.

[7] CARL G. JOCKSUCH, *Ramsey's theorem and recursion theory*, **J. Symbolic Logic**, vol. 37 (1972), no. 2, pp. 268–280.

[8] RICHARD KAYE, *Models of Peano arithmetic*, Oxford Logic Guides, vol. 15, The Clarendon Press Oxford University Press, New York, 1991.

[9] ——— , *The theory of κ-like models of arithmetic*, **Notre Dame J. Formal Logic**, vol. 36 (1995), no. 4, pp. 547–559.

[10] ——— , *Constructing κ-like models of arithmetic*, **J. London Math. Soc. (2)**, vol. 55 (1997), no. 1, pp. 1–10.

[11] L.A.S. KIRBY and J.B. PARIS, *Initial segments of models of Peano's axioms*, **Set theory and hierarchy theory, V (Proc. third conf., Bierutowice, 1976)** (Berlin), Lecture Notes in Math, vol. 619, Springer, 1977, pp. 221–226.

[12] PER MARTIN-LÖF, *The definition of random sequences*, **Information and Control**, vol. 9 (1966), pp. 602–619.

[13] ANDRÉ NIES, *Computability and randomness*, Oxford Logic Guides, vol. 51, Oxford University Press, Oxford, 2009.

[14] DAVID SEETAPUN and THEODORE A. SLAMAN, *On the strength of Ramsey's theorem*, **Notre Dame J. Formal Logic**, vol. 36 (1995), no. 4, pp. 570–582, Special Issue: Models of arithmetic.

[15] STEPHEN G. SIMPSON, *Subsystems of second order arithmetic*, second ed., Perspectives in Logic, Cambridge University Press, Cambridge, 2009.

[16] ROBERT I. SOARE, *Recursively enumerable sets and degrees*, Perspectives in Mathematical Logic, Springer-Verlag, Berlin, 1987.