

# DRAFT: Reverse mathematics of a pigeonhole basis theorem

Jeffrey Hirst, Silva Keohulian, Brody Miller, Jessica Ross

June 6, 2024

## Abstract

The infinite pigeonhole theorem asserts that if  $f : \mathbb{N} \rightarrow m$  is a function with a finite range, then there is a  $j < m$  such that the set  $\{n \in \mathbb{N} \mid f(n) = j\}$  is infinite. This article uses the techniques of reverse mathematics and Weihrauch analysis to examine the strength of a theorem that finds all the values that occur infinitely often in the range of a function.

For a function  $f : \mathbb{N} \rightarrow m$  with a finite range, the pigeonhole basis for  $f$  is the set  $B \subseteq [0, m)$  such that  $c \in B$  if and only if  $c$  appears infinite often in the range. More formally,  $B = \{c < m \mid \forall b \exists n (n > b \wedge f(n) = c)\}$ . (Monin and Patey [8] consider computational basis theorems for finite partitions, a different sort of pigeonhole basis theorem.) The next section examines the strength of the existence of pigeonhole bases in reverse mathematics. The following sections extend the examination via Weihrauch analysis and higher order reverse mathematics.

## 1 Reverse mathematics: Induction and comprehension

The study of reverse mathematics is founded on a hierarchy of subsystems of second order arithmetic, described in detail in the texts of Dzhafarov and Mummert [3] and Simpson [10]. The base system  $\text{RCA}_0$  includes induction restricted to  $\Sigma_1^0$  formulas and a set existence axiom for computable sets (formalized by  $\Delta_1^0$  definability). As a consequence of the restriction on induction,  $\text{RCA}_0$  cannot prove the  $\Pi_1^0$  bounding scheme, defined by

$$\text{B}\Pi_1^0 : (\forall x < a)(\exists y)(\forall z)\theta(x, y, z) \rightarrow (\exists b)(\forall x < a)(\exists y < b)(\forall z)\theta(x, y, z)$$

where  $\theta$  is a  $\Sigma_0^0$  formula. Indeed, over  $\text{RCA}_0$  there is a strict hierarchy of bounding and induction schemes, with  $\text{I}\Sigma_n^0$  weaker than  $\text{B}\Pi_n^0$  weaker than  $\text{I}\Sigma_{n+1}^0$  for all  $n$ . (See Chapter 6 of Dzhafarov and Mummert [3] for details.) The following theorem relates  $\text{B}\Pi_1^0$  to the infinite pigeonhole principle (often called RT1 or Ramsey's theorem for singletons).

**Theorem 1.** ( $\text{RCA}_0$ ) *The following are equivalent:*

- (1)  $\text{B}\Pi_1^0$ .
- (2) RT1: *If  $f : \mathbb{N} \rightarrow m$  then for some  $j < m$ , the set  $\{n \mid f(n) = j\}$  is infinite.*

The proof of Theorem 1 appeared initially in Hirst's thesis [4], but is more readily accessible in the texts of Dzhafarov and Mummert [3] (Theorem 6.5.1) and Weber [11] (Theorem 9.5.1). While RT1 ensures that the pigeonhole basis for a function is not empty, over  $\text{RCA}_0$  the existence of the pigeonhole basis is strictly stronger, as shown by the following theorem.

**Theorem 2.** ( $\text{RCA}_0$ ) *The following are equivalent:*

- (1) PHB: *Every  $f : \mathbb{N} \rightarrow m$  has a pigeonhole basis.*
- (2)  $\text{I}\Sigma_2^0$ : *Induction restricted to  $\Sigma_2^0$  formulas.*

*Proof.* Working in  $\text{RCA}_0$ , by Exercise II.3.13 of Simpson [10], the induction scheme  $\text{I}\Sigma_2^0$  is equivalent to bounded  $\Pi_2^0$  comprehension. Recall that the pigeonhole basis of  $f$  is defined by  $B = \{c < m \mid \forall b \exists n (n > b \wedge f(n) = c)\}$ , which is a bounded  $\Pi_2^0$  set. Thus item (1) follows from item (2).

To show the converse, suppose  $m \in \mathbb{N}$  and  $\theta(c, b, n)$  is a  $\Sigma_0^0$  formula. Our goal is to use PHB to prove that the set  $\{c < m \mid \forall b \exists n \theta(c, b, n)\}$  exists. Using a bijection identifying triples  $(c, b, n)$  in  $m \times \mathbb{N} \times \mathbb{N}$  with integer codes, define  $f : \mathbb{N} \rightarrow m + 1$  by

$$f(c, b, n) = \begin{cases} c & \text{if } n \text{ is the least } t \leq n \text{ such that } (\forall j \leq b)(\exists k \leq t)\theta(c, j, k) \\ m & \text{otherwise.} \end{cases}$$

Recursive comprehension proves the existence of  $f$ . Note that for a fixed  $c_0$ , if  $\forall b \exists n \theta(c_0, b, n)$ , then  $\text{RCA}_0$  proves that for each  $b$  there is a unique least  $t$  such that  $(\forall j \leq b)(\exists k \leq t)\theta(c_0, j, k)$ . In this situation,  $c_0$  appears in the range of  $f$  once for each value of  $b$ , and so  $c_0$  is in the pigeonhole basis for  $f$ .

On the other hand, for any fixed  $c_1$  satisfying  $\neg\forall b\exists n\theta(c_1, b, n)$ , if  $b_1$  witnesses  $\forall n\neg\theta(c_1, b_1, n)$ , then  $c_1$  appears in the range of  $f$  no more than  $b_1$  times. In this situation,  $c_1$  is not in the pigeonhole basis for  $f$ . Summarizing, the values less than  $m$  that are in the pigeonhole basis for  $f$  are exactly the set  $\{c < m \mid \forall b\exists n\theta(c, b, n)\}$  as desired.  $\square$

In light of the literature on reverse mathematics of matroids, the connection of the pigeonhole basis theorem and  $\text{I}\Sigma_2^0$  is not so surprising. Matroids capture the fundamental notions of basis and dimension in a combinatorial setting. Theorem 5 of Hirst and Mummert's [5] shows the equivalence of a matroid basis theorem and  $\text{I}\Sigma_2^0$ . Informally, a matroid resembles the vectors in a vector space, and an e-matroid as defined below is an enumeration of dependent sets.

**Definition.** An e-matroid is a pair  $(M, e)$  consisting of a non-empty set  $M$  and a function  $e : \mathbb{N} \rightarrow [M]^{<\mathbb{N}}$  enumerating the finite dependent subsets of  $M$ . The enumeration  $e$  satisfies the following conditions:

- (1) The empty set is independent. Formally,  $\forall n(e(n) \neq \emptyset)$ .
- (2) Finite supersets of dependent sets are dependent. Formally,

$$(\forall n)(\forall Y \in M^{<\mathbb{N}})(e(n) \subseteq Y \rightarrow \exists m(e(m) = Y)).$$

- (3) (Exchange principle) If  $X$  and  $Y$  are independent with  $|X| < |Y|$ , then  $Y$  contains an element that is independent of  $X$ . That is, if  $X$  and  $Y$  are independent and  $|X| < |Y|$ , then  $(\exists y \in Y)(\forall n)(e(n) \neq X \cup \{y\})$ .

The set  $M$  is often used as a shorthand for the matroid  $(M, e)$ . A finite set  $B$  spans  $M$  if every proper extension is dependent. Formally,  $B$  spans  $M$  means

$$(\forall x \in M)(x \notin B \rightarrow (\exists n)(e(n) = B \cup \{x\})).$$

A finite subset  $B$  is a *basis* for  $M$  if  $B$  spans  $M$  and  $B$  is independent.

The e-matroid terminology can be used to add another equivalence to Theorem 2.

**Theorem 3.** ( $\text{RCA}_0$ ) *The following are equivalent:*

- (1) **EMB:** *If there is a bound  $b$  for the dimension of an e-matroid  $(M, e)$ , that is, if every set of size greater than  $b$  is dependent, then  $M$  has a finite basis.*

(2) PHB: *Every  $f : \mathbb{N} \rightarrow m$  has a pigeonhole basis.*

(3)  $\text{I}\Sigma_2^0$ : *Induction restricted to  $\Sigma_2^0$  formulas.*

*Proof.* The shortest proof is to note that Theorem 2 shows the equivalence of PHB and  $\text{I}\Sigma_2^0$ , and Theorem 5 of Hirst and Mummert [5] shows the equivalence of EMB and  $\text{I}\Sigma_2^0$ .  $\square$

Of course, direct proofs of the equivalence of the first two items of Theorem 3 are possible. In particular, see the comment following the proof of Theorem 5 below.

The subsystem  $\text{ACA}_0$  includes a set comprehension axioms that asserts the existence of arithmetically definable sets. Many results in reverse mathematics prove equivalences between familiar mathematical theorems and  $\text{ACA}_0$ . Finding pigeonhole bases for sequences of functions yields such a result.

**Theorem 4.** ( $\text{RCA}_0$ ) *The following are equivalent:*

(1)  $\text{ACA}_0$ .

(2) *If  $\langle f_i \rangle_{i \in \mathbb{N}}$  is a sequence of functions with finite ranges, then there is a function  $g : \mathbb{N} \rightarrow \mathbb{N}$  such that for every  $n$ ,  $g(n)$  is (the code for) the pigeonhole basis for  $f_n$ .*

(3) *If  $\langle f_i \rangle_{i \in \mathbb{N}}$  is a sequence of functions from  $\mathbb{N}$  to  $\{0, 1\}$ , then there is a function  $g : \mathbb{N} \rightarrow \mathbb{N}$  such that for every  $n$ ,  $g(n)$  is (the code for) the pigeonhole basis for  $f_n$ .*

*Proof.* We work in  $\text{RCA}_0$  throughout. To prove that item (1) implies item (2), assume  $\text{ACA}_0$  and let  $\langle f_i \rangle_{i \in \mathbb{N}}$  satisfy the hypotheses of item (2). Then for each  $i$ , there is a unique (code for a) finite set  $B_i$  which is a pigeonhole basis for  $f_i$ . The set  $B_i$  satisfies the arithmetical formula

$$j \in B_i \leftrightarrow \forall m \forall n (m < n \wedge f_i(n) = j).$$

Thus arithmetical comprehension suffices to prove the existence of the function  $g$  which maps each  $i$  to (the code for)  $B_i$ .

Item (3) is a special case of item (2), so the proof can be completed with a proof of item (1) from item (3). By Lemma III.1.3 of Simpson [10], it suffices

to use item (3) to find the range of an injection  $h : \mathbb{N} \rightarrow \mathbb{N}$ . For each  $i$ , define  $f_i$  by:

$$f_i(n) = \begin{cases} 0 & \text{if } (\forall t \leq n)(h(t) \neq i) \\ 0 & \text{if } (\exists t \leq n)(h(t) = i) \wedge (\exists m \leq n)(2m = n) \\ 1 & \text{if } (\exists t \leq n)(h(t) = i) \wedge (\forall m \leq n)(2m \neq n) \end{cases}$$

The existence of the sequence  $\langle f_i \rangle_{i \in \mathbb{N}}$  is provable in  $\text{RCA}_0$ . Intuitively, if  $i$  has not appeared in the range of  $h$  by  $n$ , then  $f_i(n) = 0$ . If  $i$  has appeared in the range of  $h$ , then  $f_i(n)$  is the parity of  $n$ . Thus the pigeonhole basis for  $f_i$  is  $\{0\}$  if  $i$  is not in the range of  $h$  and the basis is  $\{0, 1\}$  if  $i$  is in the range. Apply item (3) to find a function  $g$  such that  $g(i)$  is the pigeonhole basis for  $f_i$  for all  $i$ . Then the range of  $h$  is  $\{i \in \mathbb{N} \mid g(i) = \{0\}\}$ , and exists by recursive comprehension.  $\square$

## 2 Weihrauch Analysis

This section uses Weihrauch analysis to examine the pigeonhole basis theorem. Introductions to the Weihrauch analysis can be found in the texts of Dzhfarov and Mummert [3] and Weihrauch [12], and the works of Brattka and Gherardi [1]. The article by Dorais et al [2] includes Weihrauch analysis of many problems related to RT1.

We denote the Weihrauch problem related to the pigeonhole principle by PHB. An instance of the problem PHB is a pair  $(f, m)$  where  $m$  is a natural number and  $f : \mathbb{N} \rightarrow m$ . The solution for the problem is (the integer code for) the pigeonhole basis for  $f$ . Similarly, an instance of the Weihrauch problem EMB is a triple  $(M, e, b)$  where  $(M, e)$  is an e-matroid in which every set of size  $b$  is dependent, and the solution is (an integer code for) a basis of  $(M, e)$ . This form of the problem is denoted by  $\text{EMB}_{<\omega}$  by Hirst and Mummert [5].

A realizer for a Weihrauch problem is a function that inputs instances of the problem and outputs solutions. Because instances can have many solutions, realizers are not unique. If  $P$  and  $Q$  are Weihrauch problems, we say  $P$  is (weakly) Weihrauch reducible to  $Q$  and write  $P \leq_W Q$  if there is a computable preprocessing procedure  $\Phi$  and a computable postprocessing procedure  $\Psi$  such that for any realizer  $R_Q$  for problem  $Q$ , the composition  $\Psi(R_Q(\Phi(f)), f)$  is a realizer for  $P$ . Informally,  $\Phi$  converts any instance  $f$  of the problem  $P$  into an instance of  $Q$ , and  $\Psi$  converts any solution of  $\Phi(f)$  into a solution for  $f$ , referring to  $f$  in the conversion, if necessary. Using

this terminology, the next theorem relates the Weihrauch problems PHB and EMB.

**Theorem 5.**  $\text{PHB} \leq_W \text{EMB}$ .

*Proof.* The preprocessing procedure for an instance  $(f, M)$  of PHB consists of two steps. First, define  $f' : \mathbb{N} \rightarrow m$  by  $f'(j) = j$  for  $j < m$  and  $f'(j) = f(j)$  for  $j \geq m$ . Note that the range of  $f'$  includes all of  $[0, m)$  and the pigeonhole basis of  $f'$  matches that of  $f$ . Second, compute an instance of EMB for  $f'$ . Let  $h : \mathbb{N} \rightarrow \mathbb{N}^{<\mathbb{N}}$  computably enumerate the finite subsets of  $\mathbb{N}$ , repeating each subset infinitely often. Define the matroid  $(e, \mathbb{N})$  as follows. For each  $n$ , suppose  $h(n) = \{x_0, \dots, x_k\}$ . If  $f'$  assigns the same value to two elements of  $h(n)$ , or if for some  $x_j \in h(n)$  there is a  $t \leq n$  such that  $t > x_j$  and  $f(t) = f(x_j)$ , then set  $e(n) = h(n)$ , otherwise, set  $e(n) = \{m\}$ .

Now we will describe the postprocessing procedure. If  $S$  is any independent set for  $(e, \mathbb{N})$  and  $s \in S$ , then  $s$  is the largest number for which  $f'$  takes the value  $f'(s)$ . Let  $B$  be a basis for  $(e, \mathbb{N})$ . The set  $\{f'(x) \mid x \in B\}$  is exactly those values in the range of  $f'$  which appear finitely often in the range of  $f'$ . Because  $f'$  is onto  $[0, m)$ ,  $B' = \{j < m \mid (\forall x \in B) f'(x) \neq j\}$  is the pigeonhole basis for  $f'$  and thus for  $f$ .  $\square$

The proof of Theorem 5 can easily be formalized in  $\text{RCA}_0$ , providing a direct proof of one direction of Theorem 3. Our original reverse mathematics proof (not presented here) applied the preprocessing procedure to the function  $f$ , using bounded comprehension in the postprocessing stage to delete the values not in the range of  $f$  from the complement of the image of the matroid basis. The application of bounded comprehension barred a direct conversion to a Weihrauch reduction, so the use of  $f'$  was added to the preceding proof to address this issue. Our direct proof of the converse in Theorem 5 (not presented here) is more convoluted. The next theorem shows that no Weihrauch reduction can be extracted from that proof.

**Theorem 6.**  $\text{EMB} \not\leq_W \text{PHB}$ .

*Proof.* Suppose by way of contradiction that  $\text{EMB} \leq_W \text{PHB}$ . Let  $\Phi$  and  $\Psi$  be the witnessing computable preprocessing and postprocessing procedures. For each e-matroid  $(\mathbb{N}, e)$  of dimension 1,  $\Psi((\mathbb{N}, e), 1)$  yields a PHB problem of the form  $(f, m)$  where  $f : \mathbb{N} \rightarrow m$ . Consider the e-matroid with  $e_0 : \mathbb{N} \rightarrow \mathbb{N}$  defined by  $e_0(n) = \{n + 1\}$ . Suppose  $\Phi((\mathbb{N}, e_0), 1) = (f, m_0)$ . The procedure  $\Phi$  is computable, so the value of  $m_0$  is determined by some finite stage using

a finite initial segment of  $e_0$ . Call the length of this segment  $u_0$ . For every e-matroid  $(\mathbb{N}, e)$  of dimension 1 that agrees with  $e_0$  up to  $u_0$ ,  $\Phi((\mathbb{N}, e), 1)$  will be a pair  $(f, m_0)$  where  $f : \mathbb{N} \rightarrow m_0$ . The pigeonhole basis for any such  $f$  is one of the finitely many subsets of  $[0, m_0)$ .

We claim that there is an e-matroid  $(\mathbb{N}, e)$  of dimension 1 such that for every  $j$  there is an e-matroid  $(\mathbb{N}, e_1)$  of dimension 1 with  $e(n) = e_1(n)$  for all  $n \leq j$ , the basis of  $(e_1, \mathbb{N})$  is  $\{k\}$  for some  $k > j$ , and  $\Phi((\mathbb{N}, e), 1)$  and  $\Phi((\mathbb{N}, e_1), 1)$  have the same pigeonhole basis. To see this, suppose it is not the case and consider  $(\mathbb{N}, e_0)$  from the preceding paragraph. Then there is a  $j_0$  such that for every  $(\mathbb{N}, e_1)$  of dimension 1, if  $e_0(n) = e_1(n)$  for all  $n \leq j_0$  and the basis of  $(\mathbb{N}, e_1)$  is  $\{k\}$  for some  $k > j_0$  then  $\Phi((\mathbb{N}, e_0), 1)$  and  $\Phi((\mathbb{N}, e_1), 1)$  have different pigeonhole bases. For example, consider the matroid  $(\mathbb{N}, e_1)$  where  $e_1 : \mathbb{N} \rightarrow \mathbb{N}$  is defined by  $e_1(n) = e_0(n)$  for  $n \leq j_0$ ,  $e_1(j_0 + 1) = \{0\}$ , and  $e_1(n) = \{n + 1\}$  for  $n \geq j_0 + 1$ . The basis of  $(\mathbb{N}, e_1)$  is  $\{j_0 + 1\}$ , so  $\Phi((\mathbb{N}, e_0), 1)$  and  $\Phi((\mathbb{N}, e_1), 1)$  must have distinct pigeonhole bases. Indeed, for any e-matroid  $(\mathbb{N}, e)$  of dimension 1 matching  $e_1$  up to  $j_0 + 1$ ,  $\Phi((\mathbb{N}, e), 1)$  and  $\Phi((\mathbb{N}, e_0), 1)$  will have distinct pigeonhole bases. Iterating the construction, we can find  $e_0, e_1, \dots, e_{2^m}$  defining e-matroids so that the pigeonhole bases for  $\Phi((\mathbb{N}, e_0), 1), \dots, \Phi((\mathbb{N}, e_{2^m}), 1)$  are  $2^m + 1$  distinct subsets of  $[0, m)$ , yielding a contradiction.

Now suppose  $(\mathbb{N}, e)$  is an e-matroid satisfying the claim of the first sentence of the preceding paragraph. Let  $\{b\}$  be the basis of  $(\mathbb{N}, e)$ , and suppose the pigeonhole basis of  $\Phi((\mathbb{N}, e), 1)$  is  $S$ . We know that  $\Psi(S, (\mathbb{N}, e)) = \{b\}$ . This computation uses only a finite initial segment of  $e$ , call the length of this segment  $u$ . Applying the claim, Let  $j = \max\{u, b\}$  and choose  $e_1$  such that  $e(n) = e_1(n)$  for all  $n \leq j$ , the basis of  $(\mathbb{N}, e_1)$  is  $\{k\}$  where  $k > j$ , and the pigeonhole basis of  $\Phi((\mathbb{N}, e_1), 1)$  is  $S$ . By the choice of  $j$ ,  $\Psi(S, (\mathbb{N}, e_1)) = \Psi(S, (\mathbb{N}, e))$ . But if  $\Psi$  is correct,  $\Psi(S, (\mathbb{N}, e)) = \{b_0\}$  and  $\Psi(S, (\mathbb{N}, e_1)) = \{k\}$ , where  $k > j \geq b_0$ . Thus no computable preprocessing and postprocessing procedures can exist.  $\square$

The next two results show that PHB is Weihrauch stronger than the limited principle of omniscience. The Weihrauch problem LPO accepts inputs of the form  $f : \mathbb{N} \rightarrow 2$ , outputs 1 if the range of  $f$  contains no zeros, and outputs 0 if 0 is in the range of  $f$ .

**Theorem 7.**  $\text{LPO} \leq_W \text{PHB}$ .

*Proof.* Given an instance of LPO of the form  $f : \mathbb{N} \rightarrow 2$ , define  $\Phi(f)$  by:

$$\Phi(f)(n) = \begin{cases} 1 & \text{if } (\forall t \leq n)(f(t) = 1) \\ 1 & \text{if } (\exists t \leq n)(f(t) = 0) \text{ and } n \text{ is odd} \\ 0 & \text{if } (\exists t \leq n)(f(t) = 0) \text{ and } n \text{ is even} \end{cases}$$

For any  $f : \mathbb{N} \rightarrow 2$ , the pigeonhole basis of  $\Phi(f)$  is  $\{1\}$  if  $f$  is never 0, and  $\{0, 1\}$  if 0 is in the range of  $f$ . For  $S$  a set of size at most 2, define  $\Psi(S) = 2 - |S|$ . Then if  $S$  is the pigeonhole basis of  $\Phi(f)$ ,  $\Psi(S)$  calculates the LPO output for  $f$ .  $\square$

**Theorem 8.**  $\text{PHB} \not\leq_W \text{LPO}$ .

*Proof.* Suppose by way of contradiction that  $\Phi$  and  $\Psi$  are procedures that witness  $\text{PHB} \leq_W \text{LPO}$ . Suppose first that for every pigeonhole instance  $f : \mathbb{N} \rightarrow 2$ ,  $\Phi(f)$  is the LPO instance that is constantly 1. Suppose that  $f$  is the constant 0 function. Because 1 is the LPO solution of  $\Phi(f)$ , we must have  $\Psi(f, 1) = \{0\}$ , the pigeonhole basis of  $f$ . The computation of  $\Psi$  uses only a finite initial segment of  $f$ , say of length  $u$ . Define  $g(t) = 0$  if  $t \leq u$  and  $g(t) = 1$  otherwise. Then  $\Psi(g, 1) = \Psi(f, 1) = \{0\}$  although  $\{1\}$  is the pigeonhole basis for  $g$ . Thus, there must be some pigeonhole instance whose corresponding LPO instance is not constantly 1.

Now suppose that there is a pigeonhole instance  $f$  such that  $\Phi(f)$  is an LPO instance with a 0 in its range. The first zero of  $\Phi(f)$  is calculated using only a finite initial segment of  $f$ , say of length  $u_0$ . Suppose  $\Psi(f, 0)$  calculates the basis of  $f$  using only an initial segment of  $f$  of length  $u_1$ . Let  $u = \max\{u_0, u_1\}$ . Let  $g$  be a function that matches  $f$  up to  $u$ , but has a different pigeonhole basis for  $f$ . Then  $\Phi(g)$  must contain a 0, and  $\Psi(g, 0)$  matches the pigeonhole basis of  $f$ , yielding an incorrect value for  $g$ .  $\square$

The proofs of Theorem 7 and 8 actually only use the restriction of the pigeonhole basis problem to functions from  $\mathbb{N}$  into 2. If we write  $\text{PHB}_2$  for the restricted problem, we have shown that  $\text{LPO} \leq_W \text{PHB}_2$  and  $\text{PHB} \not\leq_W \text{LPO}_2$ .

### 3 Higher order reverse mathematics

Reverse mathematics can be extended from numbers and sets of numbers to higher types, such as functions from sets to numbers or from sets to sets. A base theory  $\text{RCA}_0^\omega$  and early results are presented in Kohlenbach's article [7].



This framework has been used in many articles by Normann and Sanders and by Hirst and Mummert (e.g. [9] and [6]). With the more expressive language, principles can be formulated asserting the existence of realizers for Weihrauch problems. For example, in the next theorem, the principle (LPO) asserts the existence of a realizer for the Weihrauch problem LPO. Over  $\text{RCA}_0^\omega$ , (LPO) is identical to Kohlenbach's principle  $(\exists^2)$ , which is related to Kleene's functional E2.

**Theorem 9.** ( $\text{RCA}_0^\omega$ ) *The following are equivalent:*

- (1) (LPO) *there is a functional LPO such that for all  $f : \mathbb{N} \rightarrow 2$ ,  $\text{LPO}(f) = 0$  if and only if  $\exists t(f(t) = 0)$ . This principle is sometimes denoted  $\text{ACA}_0^\omega$ .*
- (2) ( $\text{PHB}_2$ ) *There is a function  $\text{PHB}_2$  such that for all  $f : \mathbb{N} \rightarrow 2$ ,  $\text{PHB}_2(f)$  is the pigeonhole basis of  $f$ .*

*Proof.* To prove that item (2) implies item (1), note that  $\text{RCA}_0^\omega$  proves that there is a function PRE such that for all  $f : \mathbb{N} \rightarrow 2$ , PRE( $f$ ) is a function that is constantly 1 until a zero appears in the range of  $f$  and constantly 0 afterwards. The function LPO( $f$ ) is the element appearing in  $\text{PHB}_2(\text{PRE}(f))$ .

The underlying idea of the proof that item (1) implies (2) is that given the LPO function,  $\text{RCA}_0^\omega$  can iterate it. Suppose (LPO) holds. Let  $f : \mathbb{N} \rightarrow 2$  be an input for  $\text{PHB}_2$ . Define the function  $Z(f, n)(k)$  by setting  $Z(f, n)(k) = 1$  unless  $k$  is the  $n^{\text{th}}$  number where  $f$  equals 0, in which case  $Z(f, n)(k) = 0$ . Note that  $f$  has at least  $n$  zeros if and only if  $\text{LPO}(Z(f, n)) = 0$ . If  $f$  has finitely many zeros, then for all values  $n$  larger than some bound  $m$ ,  $\text{LPO}(Z(f, n)) = 1$ . The function  $g(f, n) = 1 - \text{LPO}(Z(f, n))$  has zeros in its range if and only if  $f$  has only finitely many zeros. Thus the function  $Z'(f) = \text{LPO}(g(f, n))$  takes the value 0 if  $f$  has finitely many zeros in its range and 1 if  $f$  has infinitely many zeros. Define a similar function  $U'(f)$  that counts ones, so that  $U'(f) = 0$  if  $f$  has finitely many ones in its range and 1 if  $f$  has infinitely many ones. The function  $B(f)$  defined by

$$B(f) = \begin{cases} \{0\} & \text{if } U'(f) = 0 \wedge Z'(f) = 1 \\ \{1\} & \text{if } U'(f) = 1 \wedge Z'(f) = 0 \\ \{0, 1\} & \text{if } U'(f) = 1 \wedge Z'(f) = 1 \end{cases}$$

finds the pigeonhole basis for  $f$ . □

While the comment following Theorem 8 indicates that the Weihrauch problems  $\text{PHB}_2$  and  $\text{LPO}$  are not Weihrauch equivalent, Theorem 9 shows that the related higher order principles  $(\text{PHB}_2)$  and  $(\text{LPO})$  are provably equivalent over  $\text{RCA}_0^\omega$ . In this case, the fact that the higher order functionals can be applied sequentially makes them behave like the parallelized versions of the Weihrauch problems, which are Weihrauch equivalent.

## Bibliography

- [1] Vasco Brattka and Guido Gherardi, *Effective choice and boundedness principles in computable analysis*, Bull. Symbolic Logic **17** (2011), no. 1, 73–117, DOI 10.2178/bsl/1294186663. MR2760117
- [2] François G. Dorais, Damir D. Dzhafarov, Jeffrey L. Hirst, Joseph R. Mileti, and Paul Shafer, *On uniform relationships between combinatorial problems*, Trans. Amer. Math. Soc. **368** (2016), no. 2, 1321–1359, DOI 10.1090/tran/6465. MR3430365
- [3] Damir D. Dzhafarov and Carl Mummert, *Reverse mathematics—problems, reductions, and proofs*, Theory and Applications of Computability, Springer, Cham, 2022. MR4472209
- [4] Jeffrey Lynn Hirst, *Combinatorics in subsystems of second order arithmetic*, 1987. Thesis (Ph.D.)—The Pennsylvania State University. MR2635978
- [5] Jeffrey L. Hirst and Carl Mummert, *Reverse mathematics of matroids*, Computability and complexity, Lecture Notes in Comput. Sci., vol. 10010, Springer, Cham, 2017, pp. 143–159. MR3629720
- [6] ———, *Banach’s theorem in higher-order reverse mathematics*, Computability **12** (2023), no. 3, 203–225, DOI 10.3233/com-230453. MR4668958
- [7] Ulrich Kohlenbach, *Higher order reverse mathematics*, Reverse mathematics 2001, Lect. Notes Log., vol. 21, Assoc. Symbol. Logic, La Jolla, CA, 2005, pp. 281–295. MR2185441
- [8] Benoit Monin and Ludovic Patey, *Partition genericity and pigeonhole basis theorems*, J. Symb. Log. **89** (2024), no. 2, 829–857, DOI 10.1017/jsl.2022.69. MR4747983
- [9] Dag Normann and Sam Sanders, *On the uncountability of  $\mathbb{R}$* , J. Symb. Log. **87** (2022), no. 4, 1474–1521, DOI 10.1017/jsl.2022.27. MR4510829
- [10] Stephen G. Simpson, *Subsystems of second order arithmetic*, 2nd ed., Perspectives in Logic, Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, 2009. MR2517689
- [11] Rebecca Weber, *Computability theory*, Student Mathematical Library, vol. 62, American Mathematical Society, Providence, RI, 2012. MR2920681
- [12] Klaus Weihrauch, *Computable analysis*, Texts in Theoretical Computer Science. An EATCS Series, Springer-Verlag, Berlin, 2000. An introduction. MR1795407