

# EVITABLE ITERATES OF THE CONSISTENCY OPERATOR

JAMES WALSH

ABSTRACT. Why are natural theories pre-well-ordered by consistency strength? In previous work, an approach to this question was proposed. This approach was inspired by Martin’s Conjecture, one of the most prominent conjectures in recursion theory. Fixing a reasonable subsystem  $T$  of arithmetic, the goal was to classify the recursive functions that are monotone with respect to the Lindenbaum algebra of  $T$ . According to an optimistic conjecture, roughly, every such function must be equivalent to an iterate  $\text{Con}_T^\alpha$  of the consistency operator “in the limit” within the ultrafilter of sentences that are true in the standard model.

In previous work the author established the first case of this optimistic conjecture; roughly, every recursive monotone function is either as weak as the identity operator in the limit or as strong as  $\text{Con}_T$  in the limit. Yet in this note we prove that this optimistic conjecture fails already at the next step; there are recursive monotone functions that are neither as weak as  $\text{Con}_T$  in the limit nor as strong as  $\text{Con}_T^2$  in the limit. In fact, for every  $\alpha$ , we produce a function that is cofinally equivalent to  $\text{Con}_T^\alpha$  yet cofinally equivalent to  $\text{Con}_T$ .

## 1. INTRODUCTION

Why are natural axiomatic theories pre-well-ordered by consistency strength? It is not clear how to answer this question, nor even how to ask it mathematically, since there is no clear mathematical definition of the “natural” theories. Yet the informal question remains a well-known conceptual open problem.

In [5, 9] an approach to this problem was proposed. The approach in question was inspired by Martin’s approach to an analogous question in recursion theory: Why are the natural Turing degrees well-ordered by Turing reducibility? To state Martin’s Conjecture, let’s introduce some important notions. First, a function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is *degree-invariant* if for all reals  $x, y$ :

$$x \equiv_T y \implies f(x) \equiv_T f(y).$$

It seems that by relativizing the definition of a natural Turing degree one always produces a degree-invariant function on the reals. Second, a *Turing cone* is any set of the form  $\{x \mid x \geq_T y\}$ . Assuming AD, Martin proved that every degree-invariant set of reals (i.e., every set closed under Turing-equivalence) either contains a Turing cone or is disjoint from a Turing cone. This yields a  $\{0, 1\}$ -valued measure on degree-invariant sets; a degree-invariant set has measure 0 if it is disjoint from a Turing cone and measure 1 if it contains a cone. When we say *almost everywhere*, we mean almost everywhere with respect to this measure. Finally, we say that  $f \leq_m g$  if  $f(x) \leq_T g(x)$  for almost all  $x$ . Martin’s Conjecture is stated as follows:

---

Thanks to the referee for extensive and helpful suggestions and corrections.

**Conjecture 1.1** (Martin). *Assume ZF + AD + DC.*

- (1) *If  $f : 2^\omega \rightarrow 2^\omega$  is degree-invariant and it's not the case that  $f$  is increasing almost everywhere, then  $f$  is constant almost everywhere.*
- (2)  *$\leq_m$  pre-well-orders the set of degree-invariant functions that are increasing almost everywhere. If  $f$  has  $\leq_m$ -rank  $\alpha$ , then  $f'$  has  $\leq_m$ -rank  $\alpha + 1$ , where  $f'(x) = f(x)'$  for all  $x$ .*

*Roughly speaking*, Martin's Conjecture says that the only definable degree-invariant functions, up to  $\equiv_m$ , are the constant functions, the identity function, and the iterates of the Turing jump.<sup>1</sup>

The optimistic hope in [5, 9] was that an analogue of Martin's Conjecture holds for axiomatic theories. Let's fix a sound (i.e., true in the standard model) recursively extension  $T$  of elementary arithmetic. Say that a recursive function  $\mathbf{g}$  is *extensional* if for all  $\varphi$  and  $\psi$ :

$$T \vdash \varphi \leftrightarrow \psi \implies T \vdash \mathbf{g}(\varphi) \leftrightarrow \mathbf{g}(\psi).$$

Shavrukov and Visser [7] proved that there is a recursive extensional density function on the Lindenbaum algebra of  $T$ , which precludes any direct analogue Martin's Conjecture in this context. Nevertheless, positive results are available if we replace *extensionality* with the stronger condition of *monotonicity*. Say that a recursive function  $\mathbf{g}$  is *monotone* if for all  $\varphi$  and  $\psi$ :

$$T \vdash \varphi \rightarrow \psi \implies T \vdash \mathbf{g}(\varphi) \rightarrow \mathbf{g}(\psi).$$

We will be exclusively concerned with recursive monotone  $\mathbf{g}$  in this paper. We will be primarily interested in truth-preserving operators with this property, such as the consistency operator.

Monotonicity is an analogue of the recursion-theoretic notion *order-preserving*;  $f$  is order-preserving if for all reals  $x$  and  $y$ :

$$x \leq_T y \implies f(x) \leq_T f(y).$$

Note that Lutz and Siskind [4] proved Part 1 of Martin's Conjecture for order-preserving functions and Slaman and Steel [8] proved Part 2 of Martin's Conjecture for order-preserving Borel functions.

Gödel's second incompleteness theorem tells us that the monotone function

$$\varphi \mapsto (\varphi \wedge \mathbf{Con}_T(\varphi))$$

produces a strictly stronger sentence whenever  $\varphi$  is true in the standard model. That is, the consistency operator acts like a *jump* on sound finite extensions of  $T$ . Just as with the Turing jump, there are iterates of the consistency operator into the effective transfinite.<sup>2</sup> Indeed, *fixing* a suitable elementary presentation  $<$  of a

<sup>1</sup>Of course, the notion of "definable" is left vague in what I have written. More precisely, note that Martin's Conjecture is stated under the hypothesis AD. So Martin's Conjecture will apply to Borel functions and to ever more capacious notions of "definable function" depending on the large cardinal axioms one assumes.

<sup>2</sup>Unlike with Turing degrees, however, transfinite iterates of the consistency operator depend on the presentation of the base theory and the well-ordering.

well-ordering, we may define the iterates of the consistency operator using the fixed point lemma:

$$T \vdash \text{Con}_T^\alpha(\varphi) \leftrightarrow \forall \beta < \alpha \text{Con}_T(\varphi \wedge \text{Con}_T^\beta(\varphi)).^3$$

In [5, 9] a number of results relating recursive monotone functions to iterates of the consistency operator are established. To state one such theorem, we introduce a definition. Note that whenever we mention sentences being “true,” we mean “true in the standard structure  $\mathbb{N}$ .”

**Definition 1.2.** A *cone* is any set  $\mathfrak{C}$  such that, for some  $\varphi$ ,  $\mathfrak{C} = \{\psi \mid T \vdash \psi \rightarrow \varphi\}$ . A *true cone* is a cone that contains a true sentence.

Now here is a precise statement of a theorem from [9]:

**Theorem 1.3.** Let  $\mathfrak{g}$  be recursive and monotone such that, for some  $k \in \mathbb{N}$ , for all  $\varphi$ ,  $\mathfrak{g}(\varphi)$  is  $\Pi_k$ . Then one of the following holds:

- (1) For all  $\varphi$  in a true cone,  $T + \varphi \vdash \mathfrak{g}(\varphi)$ .
- (2) For all  $\varphi$  in a true cone,  $T + \varphi + \mathfrak{g}(\varphi) \vdash \text{Con}_T(\varphi)$ .

Roughly, this says that any sufficiently nice operator must either be as weak as the identity operator in the limit or as strong as the consistency operator in the limit.

We had conjectured that this would be the first step of a classification of recursive monotone operators. Our hope was to prove—along the lines of Martin’s conjecture—that any function meeting the hypotheses of the theorem would either be as strong as  $\text{Con}_T^\alpha$  in the limit or as weak as  $\text{Con}^\beta$  for some  $\beta < \alpha$  in the limit. That is, we had hoped to prove a result of the following sort:<sup>4</sup>

**Hoped For Result 1.4.** Let  $\mathfrak{g}$  be recursive and monotone such that, for some  $k \in \mathbb{N}$ , for all  $\varphi$ ,  $\mathfrak{g}(\varphi)$  is  $\Pi_k$ . Then, for every  $\alpha > 0$ , one of the following holds:

- (1) For all  $\varphi$  in a true cone, there is a  $\beta < \alpha$  such that  $T + \varphi + \text{Con}_T^\beta(\varphi) \vdash \mathfrak{g}(\varphi)$ .
- (2) For all  $\varphi$  in a true cone,  $T + \varphi + \mathfrak{g}(\varphi) \vdash \text{Con}_T^\alpha(\varphi)$ .

Of course, Theorem 1.3 already covers the  $\alpha = 1$  case of the hoped for result. Some results from [5] provided optimism for the  $\alpha > 1$  cases. Before stating these results, let’s introduce some definitions. We say that *cofinally many* true sentences belong to a set  $\mathfrak{A}$  if  $\mathfrak{A}$  is cofinal in the ultrafilter of true sentences, i.e., for every true  $\varphi$  there is a true  $\psi$  such that  $T \vdash \psi \rightarrow \varphi$  and  $\psi \in \mathfrak{A}$ .  $[\varphi]_T$  is the equivalence class of  $\varphi$  modulo  $T$ -provable equivalence. That is:

$$[\varphi]_T = \{\psi \mid T \vdash \varphi \leftrightarrow \psi\}.$$

If  $T \vdash \varphi \rightarrow \psi$ , let’s say that  $\varphi$   *$T$ -provably implies*  $\psi$ . If  $\varphi$   $T$ -provably implies  $\psi$  and  $\psi$  does not  $T$ -provably imply  $\varphi$  then we say that  $\varphi$  *strictly  $T$ -provably implies*  $\psi$ . Here are the positive results from [5]:

<sup>3</sup>A definition of suitable ordinal notations is given in §2.2.

<sup>4</sup>Note that Hoped For Result 1.4 follows from Conjecture 1.8 in [9].

**Theorem 1.5.** *Let  $\mathbf{g}$  be recursive and monotone. Suppose that for all  $\varphi$  both of the following hold:*

- (1)  $\varphi \wedge \text{Con}_T^\alpha(\varphi)$  *T-provably implies*  $\mathbf{g}(\varphi)$ .
- (2) If  $[\mathbf{g}(\varphi)]_T \neq [\perp]_T$ , then for all  $\beta < \alpha$ ,  $\mathbf{g}(\varphi)$  *strictly T-provably implies*  $\text{Con}_T^\beta(\varphi)$ .

*Then for cofinally many true  $\varphi$ ,  $[\mathbf{g}(\varphi)]_T = [\varphi \wedge \text{Con}_T^\alpha(\varphi)]_T$ .*

This result says that if the range of a recursive monotone  $\mathbf{g}$  is sufficiently constrained, then  $\mathbf{g}$  must coincide *cofinally* with an iterate of the consistency operator. In particular, if  $\mathbf{g}$  is everywhere as weak as  $\text{Con}_T^\alpha$  but everywhere strictly stronger than  $\text{Con}_T^\beta$  for all  $\beta < \alpha$ , then  $\mathbf{g}$  must coincide *cofinally* with  $\text{Con}_T^\alpha$ . As a corollary we infer that:

**Corollary 1.6.** *There is no recursive monotone  $\mathbf{g}$  such that for all  $\varphi$  such that  $[\varphi \wedge \text{Con}_T^\alpha(\varphi)]_T \neq [\perp]_T$ , both of the following hold:*

- (1)  $\varphi \wedge \text{Con}_T^\alpha(\varphi)$  *strictly T-provably implies*  $\mathbf{g}(\varphi)$ .
- (2) For all  $\beta < \alpha$ ,  $\mathbf{g}(\varphi)$  *strictly T-provably implies*  $\varphi \wedge \text{Con}_T^\beta(\varphi)$ .

This is just to say that there is no recursive monotone function of *strictly intermediate* strength, i.e., a function that is everywhere strictly stronger than  $\text{Con}_T^\beta$  for all  $\beta < \alpha$  yet strictly weaker than  $\text{Con}_T^\alpha$ . This rules out the most obvious sort of counter-example one might expect to Hoped For Result 1.4.

Nevertheless, in this paper we will see that this hoped for result fails. In fact, it fails already at the very next step, i.e.,  $\alpha = 2$ . For any  $\alpha > 0$ , we can construct a function that is cofinally as strong as  $\text{Con}_T^\alpha$  yet cofinally as weak as  $\text{Con}_T$ :

**Theorem 1.7.** *For every  $\alpha > 0$ , there is a recursive monotone  $\mathbf{g}$  such that, for all  $\varphi$ ,  $\mathbf{g}(\varphi)$  is  $\Pi_1$ , and both of the following hold:*

- (1) *For cofinally many true  $\varphi$ :*

$$[\varphi \wedge \mathbf{g}(\varphi)]_T = [\varphi \wedge \text{Con}_T^\alpha(\varphi)]_T.$$

- (2) *For cofinally many true  $\varphi$ :*

$$[\varphi \wedge \text{Con}_T(\varphi)]_T = [\varphi \wedge \mathbf{g}(\varphi)]_T.$$

This the most dramatic possible failure of the Hoped For Result 1.4 that is consistent with Theorem 1.3.

Of course, this result is somewhat *negative* considering the context in which it is proved. That is, it shows that the hoped for analogue of Martin's Conjecture fails. Yet this result also highlights something rather surprising about the consistency operator. Indeed, the consistency operator stands apart from its iterates with respect to its inevitability.

Here is our plan for the rest of the paper. In §2, we will cover some preliminaries. In §3 we will cover the main technical aspect of our result, which is the construction of pathological recursively enumerable sets. In particular, these sets contain cofinally

many true sentences but do not contain any true cones. In §4 we will present the proof of the main theorem. In §5 we will conclude with some observations; in particular, we will discuss the relationship between the negative Theorem 1.7 and the positive Theorem 1.3.

## 2. PRELIMINARIES

In this section we will fix some notation and cover some preliminaries. The main goal of this section is to introduce the iterations of the consistency operator (relative to a fixed ordinal notation system) and prove that they are monotone.

**2.1. Base Theory.** The theories we will be interested in are extensions of elementary arithmetic or EA. The signature of EA is the usual signature of PA with a function symbol for  $2^x$ . EA has as axioms the  $\text{PA}^-$  axioms plus induction for all formulas bounded in an exponential term. EA is just strong enough to carry out the standard arithmetization of syntax in the usual manner. For details about EA see [3].

One of the crucial features of EA is  $\Sigma_1$ -completeness, which we will use. This is just to say that for any  $T$  extending EA and any  $\Sigma_1$  sentence  $\varphi$ :

$$\mathbb{N} \models \varphi \Rightarrow T \vdash \varphi.$$

From here on out we will fix a sound, elementarily axiomatized extension  $T$  of EA. By saying that  $T$  is sound we mean that for all  $\varphi$ :

$$T \vdash \varphi \Rightarrow \mathbb{N} \models \varphi.$$

Of course, this means that for every  $\Sigma_1$  sentence  $\varphi$  we have:

$$\mathbb{N} \models \varphi \Leftrightarrow T \vdash \varphi.$$

**2.2. Ordinal Notations.** There are many ways of reasoning with ordinal notations in elementary arithmetic. We will not need to work with ordinal notation systems in much detail; much of what we do can be done given any “reasonable” choice. For present purposes, it is enough to briefly mention a few properties our ordinal notations will have. We will call our presentations *suitable* presentations.

Every suitable ordinal notation system is a pair  $(D, <)$  of elementary formulas, such that:

- (1) the relation  $<$  well-orders  $D$  in the standard model of arithmetic;
- (2)  $D$  is provably closed under successor;
- (3) EA proves that  $<$  linearly orders the elements satisfying  $D$ ;<sup>5</sup>
- (4) the elementary formulas defining the initial ordinal 0 (which need not be the natural number 0), the set of limit ordinals, and the successor relation provably in EA satisfy their corresponding first order definitions in terms of  $<$ .

---

<sup>5</sup>The referee has pointed out that provable linearity is not required; rather, we only need that the relation is  $T$ -provably transitive and irreflexive.

**2.3. Iterated Consistency.** Fixing a suitable elementary presentation  $<$  of a well-ordering, we may define the iterates of the consistency operator using the fixed point lemma. Technically, we find a formula  $\mathbf{Con}^*$  in two variables:

$$T \vdash \mathbf{Con}^*(\alpha, \varphi) \leftrightarrow \forall \beta < \alpha \text{Con}_T(\varphi \wedge \mathbf{Con}^*(\beta, \varphi)).$$

We write  $\text{Con}_T^\alpha(\varphi)$  as an abbreviation for  $\mathbf{Con}^*(\alpha, \varphi)$ .

*Remark 2.1.* We warn the reader that there is some discrepancy between our notation and the notation used by other authors. Our iteration scheme

$$\text{Con}_T^{\alpha+1}(\varphi) \equiv \text{Con}_T(\varphi \wedge \text{Con}_T^\alpha(\varphi))$$

is sometimes denoted  $\text{Con}(T + \varphi)_\alpha$ ; see, e.g., [1]. Moreover, the notation  $\text{Con}_T^{\alpha+1}(\varphi)$  is sometimes used to denote  $\text{Con}_T(\text{Con}_T^\alpha(\varphi))$ ; see, e.g., [2].

*Remark 2.2.* Note that all of the iterates of the consistency operator we have defined are *sentences*. This is appropriate for the current investigation since we are interested in functions on the Lindenbaum algebra of  $T$ . For instance, given our definition, for a limit  $\lambda$ , the sentence  $\text{Con}_T^\lambda(\varphi)$  says  $\forall \alpha < \lambda \text{Con}_T(\varphi \wedge \text{Con}_T^\alpha(\varphi))$ . This clashes with the conventions adopted in some other papers, wherein  $\text{Con}_T^\lambda(\varphi)$  is used as a name for the infinite set  $\{\text{Con}_T^\alpha(\varphi) \mid \alpha < \lambda\}$ .

Note that:<sup>6</sup>

$$\begin{aligned} T \vdash \text{Con}_T^0(\varphi) &\leftrightarrow \top \\ T \vdash \text{Con}_T^1(\varphi) &\leftrightarrow \text{Con}_T(\varphi) \end{aligned}$$

Since suitable ordinal notations define well-orderings over the standard structure  $\mathbb{N}$ , it follows by induction that for true  $\varphi$ ,  $\text{Con}_T^\alpha(\varphi)$  is also true. Thus, for true  $\varphi$  the hierarchy  $\text{Con}_T^\alpha(\varphi)$  is proper by Gödel's second incompleteness theorem.

It is immediate from the definition that, for any  $\varphi$ , the sentence  $\text{Con}_T^\alpha(\varphi)$  is  $\Pi_1$ .

To verify that each function  $\varphi \mapsto \text{Con}_T^\alpha(\varphi)$  is monotone, we rely on Schmerl's [6] technique of reflexive induction. Reflexive induction is a way of simulating large amounts of transfinite induction in weak theories. It is particularly useful for proving claims about iterated reflection principles. The technique is facilitated by the following theorem; we include the proof, since it is so short.

**Theorem 2.3** (Schmerl). *Let  $T$  be a recursively axiomatized theory containing EA. Suppose  $T \vdash \forall \alpha (\text{Pr}_T(\forall \beta < \alpha A(\beta)) \rightarrow A(\alpha))$ . Then  $T \vdash \forall \alpha A(\alpha)$ .*<sup>7</sup>

*Proof.* Suppose that  $T \vdash \forall \alpha (\text{Pr}_T(\forall \beta < \alpha A(\beta)) \rightarrow A(\alpha))$ . We infer that:

$$\begin{aligned} T \vdash \text{Pr}_T(\forall \alpha A(\alpha)) &\rightarrow \forall \alpha \text{Pr}_T(\forall \beta < \alpha A(\beta)) \\ &\rightarrow \forall \alpha A(\alpha). \end{aligned}$$

Löb's Theorem then yields  $T \vdash \forall \alpha A(\alpha)$ . □

<sup>6</sup>Here we use the fourth clause of our discussion of suitable ordinal notation systems.

<sup>7</sup>Schmerl proves his result using the base theory PRA. Beklemishev has shown that EA suffices.

We now will verify that the function  $\varphi \mapsto \text{Con}_T^\alpha(\varphi)$  is monotone.

**Lemma 2.4.** *For any  $\varphi$  and  $\psi$ , if  $T \vdash \varphi \rightarrow \psi$  then  $T \vdash \forall \alpha (\text{Con}_T^\alpha(\varphi) \rightarrow \text{Con}_T^\alpha(\psi))$ .*

*Proof.* Suppose that:

$$(1) \quad T \vdash \varphi \rightarrow \psi.$$

We prove the claim by reflexive induction.

*Reason in  $T$ :* Let  $\alpha$  be arbitrary. Assume the reflexive induction hypothesis:

$$(2) \quad \text{Pr}_T(\forall \beta < \alpha (\text{Con}_T^\beta(\varphi) \rightarrow \text{Con}_T^\beta(\psi))).$$

We reason as follows:

$$\begin{aligned} \text{Con}_T^\alpha(\varphi) &\rightarrow \forall \beta < \alpha \text{Con}_T(\varphi \wedge \text{Con}_T^\beta(\varphi)) \text{ by definition of } \text{Con}_T^\alpha; \\ &\rightarrow \forall \beta < \alpha \text{Con}_T(\psi \wedge \text{Con}_T^\beta(\varphi)) \text{ by 1 and monotonicity of } \text{Con}_T; \\ &\rightarrow \forall \beta < \alpha \text{Con}_T(\psi \wedge \text{Con}_T^\beta(\psi)) \text{ by 2;} \\ &\rightarrow \text{Con}_T^\alpha(\psi) \text{ by definition of } \text{Con}_T^\alpha. \end{aligned}$$

*Reasoning externally now:* Let  $A(\gamma)$  denote the claim:

$$\text{Con}_T^\gamma(\varphi) \rightarrow \text{Con}_T^\gamma(\psi)$$

We have shown that:

$$T \vdash \forall \alpha (\text{Pr}_T(\forall \beta < \alpha A(\beta)) \rightarrow A(\alpha)).$$

By applying Lemma 2.3, we infer that  $T \vdash \forall \alpha A(\alpha)$ . □

We want to see not only that  $\alpha$ -iterated consistency is monotone but also that it is provably monotone in  $T$ . We carry out this argument in two steps.

**Corollary 2.5.**  $T \vdash \forall \varphi \forall \psi \left( \text{Pr}_T(\varphi \rightarrow \psi) \rightarrow \text{Pr}_T(\forall \alpha (\text{Con}_T^\alpha(\varphi) \rightarrow \text{Con}_T^\alpha(\psi))) \right)$ .

*Proof.* Since the proof of Lemma 2.4 can be carried out in  $T$ . □

**Corollary 2.6.**  $T \vdash \forall \varphi \forall \psi \left( \text{Pr}_T(\varphi \rightarrow \psi) \rightarrow \forall \alpha (\text{Con}_T^\alpha(\varphi) \rightarrow \text{Con}_T^\alpha(\psi)) \right)$ .

*Proof. Reason in  $T$ :* Let  $\varphi$  and  $\psi$  be arbitrary. Suppose that:

$$(3) \quad T \vdash \varphi \rightarrow \psi.$$

Let  $\alpha$  be arbitrary. We reason as follows:

$$\begin{aligned} \text{Con}_T^\alpha(\varphi) &\rightarrow \forall \beta < \alpha \text{Con}_T(\varphi \wedge \text{Con}_T^\beta(\varphi)) \text{ by definition of } \text{Con}_T^\alpha; \\ &\rightarrow \forall \beta < \alpha \text{Con}_T(\psi \wedge \text{Con}_T^\beta(\varphi)) \text{ by 3 and monotonicity of } \text{Con}_T; \\ &\rightarrow \forall \beta < \alpha \text{Con}_T(\psi \wedge \text{Con}_T^\beta(\psi)) \text{ by Corollary 2.5;} \\ &\rightarrow \text{Con}_T^\alpha(\psi) \text{ by definition of } \text{Con}_T^\alpha. \end{aligned}$$

This completes the proof of the corollary. □

### 3. CONSTRUCTING PATHOLOGICAL SETS

The main technical aspect of our result is the construction of recursively enumerable sets that contain arbitrarily strong true sentences but that have wide gaps. In this section (and the next) we will work with a fixed suitable ordinal notation system  $<$ ; see §2.2. We will define a set  $\mathfrak{A}_\alpha$  for each ordinal notation  $\alpha \geq 0$ ; the size of the gaps that we leave in the set will depend on  $\alpha$ .

Similar constructions of recursively enumerable sets appear in [5, 9]. The goal in the construction of these sets was merely to include arbitrarily strong true sentences and to omit arbitrarily strong true sentences. These sets did not leave large enough gaps for present purposes.

Here is how we define the set  $\mathfrak{A}_\alpha$ :

Let  $\varphi_0, \varphi_1, \dots$ , be an effective Gödel numbering of arithmetical sentences. We describe the construction of  $\mathfrak{A}_\alpha$  in stages. During a stage  $n$  we may activate finitely many sentences; if  $\psi$  is some such sentence we say that  $\psi$  is *active* until  $\psi$  is deactivated at the later stage  $n + 1$ .

**Stage 0:** Numerate  $\top$  into  $\mathfrak{A}_\alpha$ ; activate  $\top \wedge \text{Con}_T^{\alpha+1}(\top)$ .

**Stage  $n+1$ :** There are finitely many active sentences. For each active sentence  $\psi$ , numerate  $\theta_0 := \psi \wedge \varphi_n$  and  $\theta_1 := \psi \wedge \neg\varphi_n$  into  $\mathfrak{A}_\alpha$ . Deactivate the sentence  $\psi$  and activate the sentences  $\theta_0 \wedge \text{Con}_T^{\alpha+1}(\theta_0)$  and  $\theta_1 \wedge \text{Con}_T^{\alpha+1}(\theta_1)$ .

It can be useful to visualize, along with the construction of  $\mathfrak{A}_\alpha$ , the construction of a tree that is binary branching. The tree has  $\top$  as its root. The nodes in the tree are the sentences that are numerated into  $\mathfrak{A}_\alpha$ . Informally, the immediate descendants of any sentence  $\psi$  are the sentences that are numerated into  $\mathfrak{A}_\alpha$  immediately after  $\psi$  on account of  $\psi$ . More formally, for any sentence  $\psi$  numerated into  $\mathfrak{A}_\alpha$  at stage  $n$ , say that  $\psi < \psi \wedge \text{Con}_T^{\alpha+1}(\psi) \wedge \pm\varphi_n$ ; the tree ordering is the transitive closure of the ordering  $<$  (note that this ordering is defined only on sentences numerated into  $\mathfrak{A}_\alpha$ ).

We will call the branches through this tree  $\mathfrak{A}_\alpha$ -*branches*. If  $\varphi$  and  $\psi$  share an  $\mathfrak{A}_\alpha$ -branch and  $\varphi$  was numerated into  $\mathfrak{A}_\alpha$  at stage  $n$  and  $\psi$  was numerated into  $\mathfrak{A}_\alpha$  at stage  $k$  where  $k > n$ , we say that  $\psi$  is a *descendant* of  $\varphi$ . If  $k = n + 1$  we say that  $\psi$  is an immediate descendant of  $\varphi$ .

*Remark 3.1.* Note that for each ordinal notation  $\alpha$  we get a set  $\mathfrak{A}_\alpha$ . The gaps that we leave in  $\mathfrak{A}_\alpha$  depend on  $\alpha$  in the following sense: If  $\varphi$  is numerated into  $\mathfrak{A}_\alpha$ , then  $\varphi$ 's immediate descendants imply  $\text{Con}_T^{\alpha+1}(\varphi)$ .

We can easily check some basic properties of this set  $\mathfrak{A}_\alpha$ .

**Lemma 3.2.**  $\mathfrak{A}_\alpha$  is recursively enumerable.

*Proof.* By construction. □

*Remark 3.3.* An important consequence of the recursive enumerability of  $\mathfrak{A}_\alpha$  is that for any  $\varphi$ , if  $\varphi \in \mathfrak{A}_\alpha$  then  $T \vdash \varphi \in \mathfrak{A}_\alpha$ . This follows since  $T$  is  $\Sigma_1$ -complete.



**Lemma 3.4.**  $\mathfrak{A}_\alpha$  contains arbitrarily strong true sentences.

*Proof.* Let  $\varphi_n$  be a true sentence. By induction it is easy to see that there is one true active sentence at each stage. Let's say that going into stage  $n + 1$  the true active sentence is  $\psi$ . Then at stage  $n + 1$  we numerate  $\psi \wedge \varphi_n$  into  $\mathfrak{A}_\alpha$ .  $\square$

$\mathfrak{A}_\alpha$ -branches are sets of formulas, so reasoning about  $\mathfrak{A}_\alpha$ -branches might seem to require second-order expressive resources. Yet we have assumed only that  $T$  contains elementary arithmetic. Nevertheless, elementary arithmetic suffices for reasoning about the descendant relation. The claims we make about  $\mathfrak{A}_\alpha$ -branches in this paper could be translated into  $T$ -intelligible claims about the descendant relation. In the following lemma, for instance, we will prove a claim *within*  $T$  about  $\mathfrak{A}_\alpha$ -branches. All such claims can be translated into claims about the descendant relation, though we will not give an explicit translation here.

**Lemma 3.5.** *Provably in  $T$ , if  $\psi$  and  $\theta$  both belong to  $\mathfrak{A}_\alpha$  but do not share an  $\mathfrak{A}_\alpha$ -branch, then  $\psi$  and  $\theta$  are jointly  $T$ -inconsistent.*

*Proof. Reason in  $T$ :* First observe that for any two sentences  $\varphi$  and  $\psi$  in the tree, if  $\varphi$  is a descendant of  $\psi$  then  $T + \varphi \vdash \psi$ .

Now let  $\psi$  and  $\theta$  be arbitrary sentences in  $\mathfrak{A}_\alpha$  that do not share an  $\mathfrak{A}_\alpha$ -branch. Then there is some node  $\zeta_0$  that has immediate descendants

$$\zeta_1 := \zeta_0 \wedge \text{Con}_T^{\alpha+1}(\zeta_0) \wedge \varphi_n$$

and

$$\zeta_2 := \zeta_0 \wedge \text{Con}_T^{\alpha+1}(\zeta_0) \wedge \neg\varphi_n$$

such that  $T + \psi \vdash \zeta_1$  and  $T + \theta \vdash \zeta_2$ . But  $\zeta_1$  and  $\zeta_2$  are jointly inconsistent whence  $\psi$  and  $\theta$  are too.  $\square$

**Lemma 3.6.** *Provably in  $T$ , some  $T$ -refutable sentence  $\theta$  belongs to  $\mathfrak{A}_\alpha$ .*

*Proof.* Let  $T \vdash \neg\psi$ . Note that  $\psi$  is  $\varphi_n$  for some  $n$ . At stage  $n + 1$ , we numerate a sentence  $\theta$  that  $T$ -provably implies  $\psi$  into  $\mathfrak{A}_\alpha$ . Note that  $T$  proves both that  $T \vdash \neg\theta$  and that  $\theta \in \mathfrak{A}_\alpha$ , since  $T$  is  $\Sigma_1$ -complete.  $\square$

#### 4. THE PROOF

Now we are ready to prove the main theorem. In this section we provide an example of a recursive monotone function that oscillates between behaving like  $\text{Con}_T^\alpha$  and behaving like  $\text{Con}_T$ . Note that this refutes the optimistic Hoped For Result 1.4. Indeed, no function that oscillates cofinally between behaving like  $\text{Con}_T^\alpha$  and behaving like  $\text{Con}_T$  can converge on either in the limit (assuming that  $\alpha > 1$ ).

For convenience, we restate Theorem 1.7 here.

**Theorem 4.1.** *For every  $\alpha > 0$ , there is a recursive monotone  $\mathfrak{g}$  such that, for all  $\varphi$ ,  $\mathfrak{g}(\varphi)$  is  $\Pi_1$ , and both of the following hold:*

(1) For cofinally many true  $\varphi$ :

$$[\varphi \wedge \mathbf{g}(\varphi)]_T = [\varphi \wedge \text{Con}_T^\alpha(\varphi)]_T.$$

(2) For cofinally many true  $\varphi$ :

$$[\varphi \wedge \text{Con}_T(\varphi)]_T = [\varphi \wedge \mathbf{g}(\varphi)]_T.$$

*Proof.* Given  $\alpha > 0$ , let:

$$\mathbf{g}(\varphi) := \forall \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T^\alpha(\theta)).$$

Note that  $\mathbf{g}$  is clearly recursive. It is routine to check that, for all  $\varphi$ ,  $\mathbf{g}(\varphi) \in \Pi_1$  and (using Corollary 2.6) that  $\mathbf{g}$  is monotone.

Let  $\varphi \in \mathfrak{A}_\alpha$ . By  $\Sigma_1$ -completeness of  $T$ :

$$(\Delta) \quad T \vdash \varphi \in \mathfrak{A}_\alpha.$$

We reason as follows:

$$\begin{aligned} T + \mathbf{g}(\varphi) &\vdash \forall \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T^\alpha(\theta)) \text{ by choice of } \mathbf{g}; \\ T + \mathbf{g}(\varphi) &\vdash \text{Pr}_T(\varphi \rightarrow \varphi) \rightarrow \text{Con}_T^\alpha(\varphi) \text{ by } (\Delta); \\ T + \mathbf{g}(\varphi) &\vdash \text{Con}_T^\alpha(\varphi) \text{ by } \Sigma_1\text{-completeness.} \end{aligned}$$

On the other hand, the monotonicity of  $\text{Con}_T^\alpha$  is provable in  $T$  (see Corollary 2.6). Whence:

$$\begin{aligned} T + \text{Con}_T^\alpha(\varphi) &\vdash \forall \theta (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T^\alpha(\theta)); \\ T + \text{Con}_T^\alpha(\varphi) &\vdash \mathbf{g}(\varphi). \end{aligned}$$

Since cofinally many true sentences belong to  $\mathfrak{A}_\alpha$ , this already takes care of (1).

Now we pick some true  $\psi \in \mathfrak{A}_\alpha$ . We consider the sentence  $\varphi := \psi \wedge \text{Con}_T^\alpha(\psi)$ .

**Claim.**  *$T$  proves that if  $\varphi$  is consistent, then  $\psi$  is the strongest sentence in  $\mathfrak{A}_\alpha$  that  $\varphi$   $T$ -provably implies.*

To see that the claim is true, *we reason in  $T$* : Suppose that  $\varphi$  is consistent. Note that  $\varphi$   $T$ -provably implies  $\psi$ . Note, moreover, that  $\psi$  is inconsistent with every sentence in  $\mathfrak{A}_\alpha$  with which  $\psi$  does not share an  $\mathfrak{A}_\alpha$ -branch by Lemma 3.5. So, since  $\varphi$  is consistent, the only  $\mathfrak{A}_\alpha$  sentences that  $\varphi$   $T$ -provably implies must share an  $\mathfrak{A}_\alpha$ -branch with  $\psi$ . By construction of  $\mathfrak{A}_\alpha$ , every descendant of  $\psi$   $T$ -provably implies  $\text{Con}_T^{\alpha+1}(\psi)$ . But, since  $\varphi$  is consistent,  $\varphi$  does not  $T$ -provably imply  $\text{Con}_T^{\alpha+1}(\psi)$  by Gödel's second incompleteness theorem. This delivers the claim.

We then reason as follows:

$$\begin{aligned} T + \varphi + \text{Con}_T(\varphi) &\vdash \forall \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Pr}_T(\psi \rightarrow \theta)) \text{ by the claim}; \\ T + \varphi + \text{Con}_T(\varphi) &\vdash \forall \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow (\text{Con}_T^\alpha(\psi) \rightarrow \text{Con}_T^\alpha(\theta))) \text{ by Corollary 2.6}; \\ T + \varphi + \text{Con}_T(\varphi) &\vdash \forall \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T^\alpha(\theta)) \text{ since } T + \varphi \vdash \text{Con}_T^\alpha(\psi); \\ T + \varphi + \text{Con}_T(\varphi) &\vdash \mathbf{g}(\varphi) \text{ by the definition of } \mathbf{g}. \end{aligned}$$

For the converse:

$$\begin{aligned} T &\vdash \exists \theta \in \mathfrak{A}_\alpha \neg \text{Con}_T(\theta) \text{ by Lemma 3.6;} \\ T + \neg \text{Con}_T(\varphi) &\vdash \exists \theta \in \mathfrak{A}_\alpha (\text{Pr}_T(\varphi \rightarrow \theta) \wedge \neg \text{Con}_T^\alpha(\theta)); \\ T + \neg \text{Con}_T(\varphi) &\vdash \neg \mathfrak{g}(\varphi) \text{ by the definition of } \mathfrak{g}. \end{aligned}$$

This takes care of (2).  $\square$

## 5. OBSERVATIONS

Theorem 1.7 refutes Hoped For Result 1.4. Yet what happens in the case  $\alpha = 0$ ? That is, why can't the proof of Theorem 1.7 be adapted to the  $\alpha = 0$  case, thereby contradicting Theorem 1.3? The answer exhibits an important feature that the notion of consistency does not share with its iterates.

Recall that we construct  $\mathfrak{A}_\alpha$  so that whenever  $\varphi$  is numerated into  $\mathfrak{A}_\alpha$ , then  $\varphi$ 's immediate descendants  $T$ -provably imply  $\text{Con}_T^{\alpha+1}(\varphi)$ . In particular, whenever  $\varphi$  is numerated into  $\mathfrak{A}_0$ , then  $\varphi$ 's immediate descendants  $T$ -provably imply  $\text{Con}_T(\varphi)$ . Let's consider the function:

$$\mathfrak{g}_0(\varphi) := \forall \theta \in \mathfrak{A}_0 (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T(\theta)).$$

Surprisingly,  $\mathfrak{g}_0$  is actually equivalent to the consistency operator. That is:

**Proposition 5.1.** *For every  $\varphi$ ,  $T \vdash \text{Con}_T(\varphi) \leftrightarrow \mathfrak{g}_0(\varphi)$ .*

*Proof.* Left to right:

$$\begin{aligned} T \vdash \text{Con}_T(\varphi) &\rightarrow \forall \theta (\text{Pr}_T(\varphi \rightarrow \theta) \rightarrow \text{Con}_T(\theta)); \\ &\rightarrow \mathfrak{g}_0(\varphi). \end{aligned}$$

Right to left:

$$\begin{aligned} T \vdash \neg \text{Con}_T(\varphi) &\rightarrow \exists \theta \in \mathfrak{A}_0 (\text{Pr}_T(\varphi \rightarrow \theta) \wedge \neg \text{Con}_T(\theta)) \text{ by Lemma 3.6;} \\ &\rightarrow \neg \mathfrak{g}_0(\varphi). \end{aligned}$$

This completes the proof.  $\square$

Note the appeal to Lemma 3.6. Here we use that  $\mathfrak{A}_0$  is guaranteed to contain an inconsistent sentence; the important point is that if  $\varphi$  is  $T$ -inconsistent, then some sentence that  $\varphi$   $T$ -provably implies belongs to  $\mathfrak{A}_0$ . Indeed, the fact that all  $T$ -inconsistent sentences  $T$ -provably imply each other is used in the proof of the positive Theorem 1.3 (see the proof of Theorem 2.4 Case 2 in [9]). By contrast, if we merely knew  $\neg \text{Con}_T^{\alpha+1}(\varphi)$ , we would not be able to conclude that some  $T$ -consequence of  $\varphi$  belongs to  $\mathfrak{A}_\alpha$ . Nor for any of the other iterates of the consistency operator.

There are ways of modifying  $\mathfrak{g}_0$  to avoid Proposition 5.1. Rather than quantifying over the  $T$ -implications of  $\varphi$  in  $\mathfrak{A}_0$  and saying that they are *all* consistent, we can let  $\mathfrak{g}(\varphi)$  merely assert the conjunction according to which *each* is consistent. That is:

$$\mathfrak{g}_0^*(\varphi) := \begin{cases} \bigwedge \{\text{Con}_T(\zeta) \mid \zeta \in \mathfrak{A}_0 \text{ and } T + \varphi \vdash \zeta\} & \text{if } [\varphi]_T \neq [\perp]_T, \\ \perp & \text{otherwise} \end{cases}$$

In previous work we have shown that  $\mathfrak{g}_0^*$  oscillates between behaving like the identity and the consistency operator (see Theorem 3.7 in [9]). However, computing the function  $\mathfrak{g}_0^*$  requires access to the oracle  $0'$ . Indeed, to calculate  $\mathfrak{g}_0^*(\varphi)$  we must know whether  $[\varphi]_T \neq [\perp]_T$ , which requires  $0'$ . So  $\mathfrak{g}_0^*$  demonstrates that recursiveness is also a necessary condition in Theorem 1.3; it cannot be weakened to limit-recursiveness.

#### REFERENCES

- [1] Lev Beklemishev. Iterated local reflection versus iterated consistency. *Annals of Pure and Applied Logic*, 75(1-2):25–48, 1995.
- [2] Lev D Beklemishev. Provability logics for natural Turing progressions of arithmetical theories. *Studia Logica*, pages 107–128, 1991.
- [3] Lev Dmitrievich Beklemishev. Reflection principles and provability algebras in formal arithmetic. *Russian Mathematical Surveys*, 60(2):197–268, 2005.
- [4] Patrick Lutz. Results on Martin’s Conjecture. *Bulletin of Symbolic Logic*, 27(2):219–220, 2021.
- [5] Antonio Montalbán and James Walsh. On the inevitability of the consistency operator. *The Journal of Symbolic Logic*, 84(1):205–225, 2019.
- [6] Ulf R Schmerl. A fine structure generated by reflection formulas over primitive recursive arithmetic. In *Studies in Logic and the Foundations of Mathematics*, volume 97, pages 335–350. Elsevier, 1979.
- [7] V Yu Shavrukov and Albert Visser. Uniform density in Lindenbaum algebras. *Notre Dame Journal of Formal Logic*, 55(4):569–582, 2014.
- [8] Theodore A Slaman and John R Steel. Definable functions on degrees. In *Cabal Seminar 81–85*, pages 37–55. Springer, 1988.
- [9] James Walsh. A note on the consistency operator. *Proceedings of the American Mathematical Society*, 148(6):2645–2654, 2020.

SAGE SCHOOL OF PHILOSOPHY, CORNELL UNIVERSITY

*Email address:* jameswalsh@cornell.edu