

COUSIN’S LEMMA IN SECOND-ORDER ARITHMETIC

JORDAN MITCHELL BARRETT, RODNEY G. DOWNEY, AND NOAM GREENBERG

ABSTRACT. Cousin’s lemma is a compactness principle that naturally arises when studying the gauge integral, a generalisation of the Lebesgue integral. We study the axiomatic strength of Cousin’s lemma for various classes of functions, using Friedman and Simpson’s reverse mathematics in second-order arithmetic. We prove that, over RCA_0 :

- (i) Cousin’s lemma for continuous functions is equivalent to WKL_0 ;
- (ii) Cousin’s lemma for Baire class 1 functions is equivalent to ACA_0 ;
- (iii) Cousin’s lemma for Baire class 2 functions, or for Borel functions, are both equivalent to ATR_0 (modulo some induction).

1. INTRODUCTION

For real-valued functions, the most general version of the fundamental theorem of calculus relies on a generalization of both the Lebesgue and (improper) Riemann integrals, called, among other names, the Denjoy integral [Den12, Gor94]. This integral can integrate all derivatives, which the Riemann and Lebesgue integrals fail to do [Gor94].

Denjoy [Den12] first defined this integral in 1912, and shortly after, Luzin [Luz12] and Perron [Per14] gave equivalent characterisations. However, all these definitions were complex and highly nonconstructive, making Denjoy’s integral impractical for applications [Gor97]. In 1957, Kurzweil [Kur57] discovered an equivalent, “elementary” definition, in the style of the Riemann integral. This re-formulation is widely known as the *gauge integral*, or Henstock–Kurzweil integral.

For bounded functions with compact support, the gauge integral coincides with the Lebesgue integral, and hence the gauge integral gives a Riemann-like definition of the Lebesgue integral for a wide class of functions. We also remark that the gauge integral is suitable for integrating classes of highly discontinuous functions, and can be viewed as a mathematically rigorous formalisation of Feynman’s path integral [Mul87].

A *gauge* is simply a function $\delta: [0, 1] \rightarrow \mathbb{R}^+$ from the unit interval to the positive real numbers. A *δ -fine partition* is a finite tagged partition $0 = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq \dots \leq x_n = 1$ for which $\delta(\xi_i) \geq x_{i+1} - x_i$. This generalises the mesh size of a partition: a partition has mesh size $\leq \delta$ if it is δ -fine for the constant map with value δ . The gauge integral is defined exactly like the Riemann integral, except that it allows arbitrary gauges rather than only constant ones: the integral $\int_0^1 f dx$ is K if for every $\varepsilon > 0$ there is a gauge δ such that for every δ -fine partition, the

Date: February 26, 2022.

1991 *Mathematics Subject Classification.* 03B30, 03F35, 03D78, 26A39.

Downey and Greenberg were partially supported by the Marsden Fund of New Zealand. Many of the results in this paper are also contained in Barrett’s honours thesis [Bar20].

associated Riemann sum is within ε of K . So every Riemann integrable function is gauge integrable, by choosing constant gauges. But also the function $x^{-1/2}$ is gauge integrable on $[0, 1]$, by taking, given $\varepsilon > 0$, a gauge that approaches 0 sufficiently rapidly so that the contribution of a leftmost interval to a Riemann sum cannot be unbounded. Similarly, Dirichlet’s function (the indicator function of the rationals) is integrable by taking $\delta(q_n) = \varepsilon 2^{-n}$, where $\langle q_n \rangle$ lists the rationals in $[0, 1]$.

The key fact that enables the theory of the gauge integral is *Cousin’s lemma* [Cou95], which states that for any gauge δ , a δ -fine partition exists. Without Cousin’s lemma, Kurzweil’s definition of the gauge integral would be vacuously satisfied by all functions and all values, trivialising the notion.

The gauge integral and Cousin’s lemma have had metamathematical exploration before; notably in the setting of descriptive set theory by Dougherty and Kechris [DK91], Becker [Bec92], and Walsh [Wal17], for example. Recently, it was also explored in higher-order reverse mathematics by Normann and Sanders [NS19, NS20]. These studies found that the gauge integral requires powerful axioms to prove. For example, working in third-order arithmetic, Normann and Sanders showed that Cousin’s lemma and the existence of the gauge integral were each equivalent to full second-order arithmetic.

The point is that Cousin’s lemma is a statement about *all* functions from $[0, 1]$ to \mathbb{R}^+ ; no definability assumptions are involved. It is thus a statement which is not expressible in second-order arithmetic, which only allows quantification over real numbers (but not sets of real numbers). Second-order arithmetic is the study of countable mathematics, including objects which can be coded by countable objects, for instance separable metric spaces, continuous functions, and Borel sets.

Reverse mathematics of second-order arithmetic, as developed by H. Friedman and S. Simpson, is the project of understanding the proof-theoretic strength of the theorems of mathematics in terms of comprehension and induction strength required to prove them. What axioms of mathematics are actually required to prove a particular statement? The aim is to find optimal proofs, ones which require the least “amount” of axioms; and then show that these proofs are indeed the best that can be found, by finding a “reversal”: a proof of the axioms used, starting with the investigated theorem.

The main insight of the project of reverse mathematics is that almost all theorems of mainstream mathematics are equivalent to one of five axiomatic systems, and that these systems are linearly ordered by logical implication; see, for example [Sim09]. One advantage of working within second-order arithmetic is that the proof-theoretic strength is often aligned with complexity, as defined using the tools of computability theory. Very informally speaking, a theorem which asserts the existence of complicated objects, requires strong axioms to prove, and vice-versa. This connection between computability and proof theory has resulted in a rich body of research.

It is thus natural to investigate the strength of Cousin’s lemma within second-order arithmetic, but in order to do so, we must restrict ourselves to classes of countably-coded functions. In this paper we consider Borel functions, and subclasses of these.

Below we will observe that Cousin’s lemma is a form of a compactness principle. We would thus naturally start by looking at the system WKL_0 (weak König’s lemma), which informally is understood to be equivalent to the compactness of

the unit interval. But more specifically, it is equivalent to *countable compactness*, for example, to the statement that any countable open cover of $[0, 1]$ has a finite sub-cover [Sim09, Thm.IV.1.2]. For continuous functions, Cousin's lemma should be equivalent to its version for constant functions. Despite the fact that this equivalence seems to use a form of compactness, namely that continuous functions on $[0, 1]$ obtain a minimum, we are able to show:

- Cousin's lemma for continuous gauges is equivalent to WKL_0 .

However, once we go up the hierarchy of Borel functions (as measured say by the Baire class of a function), it turns out that this form of uncountable compactness is significantly stronger:

- Cousin's lemma for Baire class 1 gauges is equivalent to ACA_0 .
- Cousin's lemma for all Borel gauges is provable in $ATR_0 + \Delta_2^1$ -induction.
- Cousin's lemma for Baire class 2 gauges implies ATR_0 .

We leave open the question of whether the extra induction is required. For background on reverse mathematics, and the definitions of these three subsystems of second-order arithmetic, see [Sim09].

2. COUSIN'S LEMMA AND COMPACTNESS

Let $\delta: [0, 1] \rightarrow \mathbb{R}^+$ be a gauge. We can view δ as assigning, to each $x \in [0, 1]$, the open interval $(x - \delta(x), x + \delta(x))$. Cousin's lemma is very close to saying that this open cover of the unit interval has a finite sub-cover. It will be easier to work with this simplified version, rather than with tagged partitions, partly because this notion can be extended to other compact metric spaces; we will be working with Cantor space as well as the unit interval.

The metamathematical twist is that for our reversals, we will be working over RCA_0 . Some of the gauges that we will work with may be definable in some models of RCA_0 , but not have values in these models. For example, we will work with Baire class 1 functions, coded in a model by a sequence of continuous functions, where the values of the functions are Cauchy sequences, which need not converge. Nonetheless, relations such as $f(x) > r$ will be still definable in the model, and this will be enough to formalise Cousin's lemma.

In light of this, it is important to note that a δ -fine tagged partition $0 = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq \dots \leq x_{n-1} \leq \xi_{n-1} \leq x_n = 1$, with $x_{i+1} - x_i \leq \delta(\xi_i)$ gives us not only the finite set of points ξ_i giving us a cover, say $[0, 1] \subseteq \bigcup_i (\xi_i - 2\delta(\xi_i), \xi_i + 2\delta(\xi_i))$, but also the distances $x_{i+1} - x_i$, bounded by $\delta(\xi_i)$; and while $\delta(\xi_i)$ may not be in the model, the value $x_{i+1} - x_i$ is. Thus we define:

Definition 2.1. Let X be a compact metric space, and let $\delta: X \rightarrow \mathbb{R}^+$ be a gauge on X . A δ -fine cover of X is a finite set $P \subseteq X$ and a function $p \mapsto r_p$ from P to \mathbb{R}^+ satisfying:

- (i) $X = \bigcup_{p \in P} B(p, r_p)$; and
- (ii) For all $p \in P$, $r_p \leq \delta(p)$.

Assuming that δ is a function coded in the model for which the relations $\delta(x) \in B$ (for an open or closed ball B) are defined, we have:

Lemma 2.2 (RCA_0). Let $\delta: [0, 1] \rightarrow \mathbb{R}^+$ be a gauge.

- (a) If there is a δ -fine tagged partition then there is a 2δ -fine cover.
- (b) If there is a δ -fine cover then there is a 2δ -fine tagged partition.

Proof. For (a), suppose that $0 = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq x_2 \leq \dots \leq \xi_{n-1} \leq x_n = 1$ is a δ -fine partition: for all $i < n$, $\delta(\xi_i) \geq x_{i+1} - x_i$. Let $P = \{\xi_i : i < n\}$, and for $i < n$ let $r_{\xi_i} = 2(x_{i+1} - x_i)$. Then for all $i < n$, $r_{\xi_i} \leq 2\delta(\xi_i)$ and trivially $[x_i, x_{i+1}] \subset (\xi_i - r_{\xi_i}, \xi_i + r_{\xi_i})$, so P, \bar{r}_p is a 2δ -fine cover.

For (b), suppose that $P = \{p_0 < p_1 < \dots < p_{n-1}\}$ equipped with $i \mapsto r_{p_i}$ is a δ -fine cover. For brevity, for $p \in P$ let $U(p) = (p - r_p, p + r_p)$. By applying a reverse greedy algorithm, we may assume that P is minimal in that for no distinct p, q in P do we have $U(p) \subseteq U(q)$. This implies that for all $i < n - 1$, $p_{i+1} - p_i < r_{p_i} + r_{p_{i+1}}$; otherwise, a point between p_i and p_{i+1} would not be covered by $U(p_i)$ or by $U(p_{i+1})$; if $x \in U(p_j)$ then either $U(p_i) \subseteq U(p_j)$ (if $j < i$) or $U(p_{i+1}) \subseteq U(p_j)$ (if $j > i + 1$). So we choose

$$x_{i+1} \in (p_i, p_{i+1}) \cap U(p_i) \cap U(p_{i+1}),$$

as well as $x_0 = 0$ and $x_n = 1$; then $x_0 \leq p_0 \leq x_1 \leq p_1 \leq \dots$ is a 2δ -fine partition, as $[x_i, x_{i+1}] \subseteq U(p_i)$. \square

In light of Lemma 2.2, we state:

Definition 2.3 (Cousin's lemma). Let X be a compact metric space, and let \mathcal{K} be a class of functions. Cousin's lemma for \mathcal{K} on X states that every $\delta: X \rightarrow \mathbb{R}^+$ in \mathcal{K} has a δ -fine cover.

If not stated otherwise, we work with $X = [0, 1]$. For many of our arguments, though, it will be more convenient to work in Cantor space, under the metric $d(x, y) = 2^{-n}$ for the largest n satisfying $x \upharpoonright n = y \upharpoonright n$. We state one connection between Cousin's lemma on Cantor space and on the unit interval informally. Once we work with particular classes of functions, we will see that the following argument, in each particular case, can be made to hold in RCA_0 .

Lemma 2.4. For reasonable classes \mathcal{K} of functions, Cousin's lemma for \mathcal{K} on Cantor space implies Cousin's lemma for \mathcal{K} on the unit interval.

Proof. We use the standard map $\varphi: 2^\omega \rightarrow [0, 1]$ defined by

$$\varphi(x) = \sum x(n)2^{-n-1}.$$

Then φ is continuous and for all $x, y \in 2^\omega$,

$$|\varphi(x) - \varphi(y)| \leq d(x, y).$$

Hence, given $\delta: [0, 1] \rightarrow \mathbb{R}^+$ in \mathcal{K} , we let $\hat{\delta} = \varphi \circ \delta$, which is a gauge on Cantor space. Suppose that (\hat{P}, \bar{r}_p) is a $\hat{\delta}$ -fine cover. Let $P = \{\varphi(p) : p \in \hat{P}\}$, and let $r_{\varphi(p)} = r_p$; this is a δ -fine cover. \square

3. THE BOREL CASE

The "classical" proof of Cousin's lemma can be carried out in the system of $\text{ATR}_0 + \Sigma_1^1$ -induction. Note that while we can formalise Borel codes and functions in RCA_0 , in the system ATR_0 we can prove that for every Borel function f and every x , there is a point $f(x)$ (this follows from [Sim09, Lem.V.3.7]).

We need the following:

Lemma 3.1 ($\text{ATR}_0 + \Sigma_1^1$ -induction). If $T \subseteq 2^{<\omega}$ is an infinite Π_1^1 -definable tree, then T has a path.

Proof. Let

$$S = \{\sigma \in T : (\exists^\infty \tau \succeq \sigma) \tau \in T\}.$$

Then S is Π_1^1 (this uses Σ_1^1 -choice, which is provable in ATR_0 ; see [Sim09, VIII.3.21]). By assumption, the empty sequence is in S ; and every $\sigma \in S$ has a child $\sigma \hat{\ } i$ in S . So we can apply Π_1^1 -dependent choice on numbers, which is provable in $\text{ATR}_0 + \Sigma_1^1$ -induction, [Sim09, VIII.4.10]. \square

Proposition 3.2 ($\text{ATR}_0 + \Delta_2^1$ -induction). Cousin's lemma holds for all Borel functions.

Proof. The proof of Lemma 2.4 holds in RCA_0 for the class \mathcal{K} of Borel functions. Hence, we work with Cantor space. Let $\delta: 2^\omega \rightarrow \mathbb{R}^+$ be a Borel function. Let

$$G = \{\sigma \in 2^{<\omega} : (\exists x \in [\sigma]) [\sigma] \subseteq B(x, \delta(x))\},$$

where $[\sigma] = \{x \in 2^\omega : \sigma \prec x\}$ is the clopen set determined by σ . Note that $[\sigma] \subseteq B(x, \delta(x))$ if and only if $\delta(x) > 2^{-|\sigma|+1}$, so this is a Borel condition. However, the quantification over all $x \in [\sigma]$ means that G is Σ_1^1 (relative to a code for δ).

Now let

$$T = \{\sigma \in 2^{<\omega} : (\forall \tau \preceq \sigma) \tau \notin G\}.$$

Then T is a Π_1^1 tree. We claim that it is finite. If not, then by Lemma 3.1, T has a path y . Let n be sufficiently large so that $2^{-n} < \delta(y)$; this contradicts $y \upharpoonright n \in T$.

So we know that there is some n which bounds the length of all the strings in T . By Δ_2^1 induction, T exists, as a finite set: by induction on the length-lexicographic order of strings of length $\leq n$, we show that for all $\tau \in 2^{\leq n}$, $T \cap \{\sigma : \sigma \text{ precedes } \tau\}$ is a finite set in the model. The formula over which we induct is of the form $(\exists m \leq 2^n)(\Pi_1^1 \ \& \ \Sigma_1^1)$, so Δ_2^1 -induction certainly suffices.

It follows that the set of leaves of T exists (again as a finite set). By taking all immediate extensions of the leaves of T , we obtain a finite antichain $R \subseteq G$ of strings with $2^\omega = [R]^\preceq = \bigcup \{[\sigma] : \sigma \in R\}$.

Finally, there is a sequence $\langle x_\sigma : \sigma \in R \rangle$ witnessing that each $\sigma \in R$ is in G , i.e., $x_\sigma \in [\sigma]$ and $[\sigma] \subseteq B(x_\sigma, \delta(x_\sigma))$. This follows from Σ_1^1 -choice (or by Σ_1^1 -induction, as R is finite). Letting $P = \{x_\sigma : \sigma \in R\}$ and $r_{x_\sigma} = \delta(x_\sigma)$ gives a δ -fine cover. \square

4. THE CONTINUOUS CASE

In this section we show that over the base system RCA_0 of recursive comprehension, Cousin's lemma for continuous functions is equivalent to weak König's lemma WKL_0 . If δ is continuous then for all x , $\delta(x)$ exists, and so for a δ -fine cover (P, \bar{r}_p) , we can always choose $r_p = \delta(p)$.

In one direction, the argument is quick.

Proposition 4.1 (WKL_0). Cousin's lemma holds for continuous functions.

Proof. Let $\delta: [0, 1] \rightarrow \mathbb{R}^+$ be continuous. Weak König's lemma implies that δ obtains a minimum r ([Sim09, Th.IV.2.2]), which must be positive. A partition of $[0, 1]$ into intervals of length $2^{-n} < r$, and choosing any points as tags, gives a δ -fine partition. \square

Note that the argument works for Cantor space as well.

In the other direction, we could argue directly by coding a Π_1^0 class by an effectively closed subset of $[0, 1]$; but the argument is combinatorially simpler by passing via the Heine-Borel theorem.

Theorem 4.2 (RCA₀). Cousin's lemma for continuous functions implies WKL_0 .

Proof. WKL_0 is equivalent to the Heine-Borel theorem, which states the countable compactness of the unit interval: if $\langle U_n \rangle_{n \in \mathbb{N}}$ is a sequence of open intervals and $[0, 1] \subseteq \bigcup_n U_n$, then there is finite sub-cover: there is some $N \in \mathbb{N}$ such that $[0, 1] \subseteq \bigcup_{n < N} U_n$ (see [Sim09, Lem IV.1.1]).

Let $\langle U_n \rangle$ be such an open cover of $[0, 1]$.

For $n \in \mathbb{N}$, let $d_n(x) = d(x, U_n^c)$ be the distance from a point x to the complement of U_n . A code for d_n as a continuous function can be obtained uniformly from the (end-points of) the interval U_n . Let

$$\delta(x) = \frac{1}{4} \sum_{n \in \mathbb{N}} 2^{-n} d_n(x);$$

Then δ is a continuous function on $[0, 1]$ (a code can be obtained effectively from the codes for the d_n 's). Since $\langle U_n \rangle$ covers the unit interval, for all x there is some n such that $d_n(x) > 0$, so δ is positive on $[0, 1]$.

We claim: for all $p \in [0, 1]$ and $k \in \mathbb{N}$,

$$(*) \text{ if } \delta(p) > 2^{-k} \text{ then } (p - \delta(p), p + \delta(p)) \subseteq \bigcup_{n \leq k} U_n.$$

Suppose not: then for all $n \leq k$, $d_n(p) < \delta(p)$. As a result,

$$\delta(p) < \frac{1}{4} \sum_{n \leq k} 2^{-n} \delta(p) + \frac{1}{4} \sum_{n > k} 2^{-n} d_n(p) \leq \frac{1}{2} \delta(p) + \frac{1}{4} 2^{-k} < \delta(p),$$

which is impossible; of course we use $d_n(p) \leq 1$ for all p .

Now let $P \subset [0, 1]$ be a δ -fine cover (as mentioned we can choose $r_p = \delta(p)$). For each $p \in P$ find some k with $\delta(p) > 2^{-k}$; by Σ_1^0 bounding (see for example the proof of [Sim09, Lem.IX.1.8]), there is some k such that for all $p \in P$, $\delta(p) > 2^{-k}$. Hence $\langle U_n \rangle_{n \leq k}$ covers $[0, 1]$, as required. \square

Remark 4.3. The proof above may obscure the main intuition behind it. Consider a model in which WKL_0 fails, witnessed by a ‘‘cover’’ $\langle U_n \rangle$ which is not really a cover: outside the model, there are points not covered by any U_n , but there are none such inside it. Let $C = [0, 1] \setminus \bigcup_n U_n$. This is an effectively closed set (relative to an oracle in the model), and the rough idea is to let $\delta(x) = d(x, C)$. Then no $x \in C$ can be covered by a δ -fine partition. The fact that for any *finite* P , $[0, 1]$ is not covered by $\bigcup_{p \in P} (p - \delta(p), p + \delta(p))$ reflects down to the original model. The function δ is not really a gauge, but the model doesn't see this, as it thinks that C is empty.

The only difficulty is that $x \mapsto d(x, C)$, while continuous, is not computable relative to an oracle in the model; it is only lower semi-computable. In other words, the model does not contain a code for this function. The definition of δ in the proof above gives a continuous lower bound to $d(x, C)$ with a code in the model.

5. BAIRE CLASS 1

A function is *Baire class 1* if it is the pointwise limit of continuous functions. The Baire classes were introduced by Baire in his PhD thesis [Bai99], as a natural generalisation of the continuous functions. One motivation for Baire functions is that many functions arising in analysis are not continuous, such as step functions [Hea93], Walsh functions [Wal23], or Dirichlet's function [Dir29]. However, all such "natural" functions generally have low Baire class; for example, the derivative of any differentiable function is Baire class 1, as are functions arising from Fourier series [KL90].

The Baire classes have previously been studied with respect to computability [KT14, PDD17]. In particular, Kuyper and Terwijn showed that a real number x is 1-generic if and only if every *effective* Baire 1 functions is continuous at x [KT14]. Indeed, the Turing jump function is in some sense a universal Baire class 1 function on Baire space.

In second-order arithmetic, we can formalise this notion as follows:

Definition 5.1. A code for a *Baire class 1* function is a sequence $\langle f_n \rangle$ of (codes of) continuous functions such that for all x , $\langle f_n(x) \rangle$ is a Cauchy sequence.

Note that we do not require the limit of this Cauchy sequence to exist; this would follow from arithmetic comprehension, but the definition still makes sense in RCA_0 . If $\langle f_n \rangle$ is a code in the model for a Baire class 1 function f , the relation " $f(x) \in B$ " for an open or closed ball B is definable, even if $f(x)$ is not an object in the model; for example, we say that $f(x) \in B(y, r)$ if for some $s < r$, for all but finitely many n , $f_n(x) \in B(y, s)$. Thus, it is meaningful to say that $\langle f_n \rangle$ is a code for a gauge.

In this section we show that Cousin's lemma for Baire class 1 functions is equivalent to ACA_0 . We start with:

Lemma 5.2 (RCA_0). If $f: X \rightarrow Y$ is Baire class 1, then for open sets $U \subseteq Y$, $f^{-1}[U]$ is F_σ , uniformly.

Proof. What this means: if $f = \langle f_n \rangle$ is a code for a Baire class 1 function, and $\langle U_n \rangle$ is a sequence of (codes of) open subsets of Y , then there is a sequence $\langle F_n \rangle$ of (codes of) F_σ subsets of X with

$$F_n = \{x \in X : f(x) \in U_n\},$$

where again $f(x) \in U_n$ is a definable relation on x and n .

The lemma is well-known. The f -preimage of a closed ball by a continuous function is closed, and an open ball is the union of closed balls. Hence

$$f(x) \in B(y, r) \Leftrightarrow (\exists s < r)(\exists n)(\forall m \geq n) d(f_n(x), q) \leq s. \quad \square.$$

Remark 5.3. In fact, classically, a function is Baire class 1 if and only if the pull back of open sets is F_σ . The other direction requires a little bit of work; it is like Shoenfield's limit lemma, but for arbitrary spaces requires some topological considerations.

In second-order arithmetic, to make sense of the other direction, we would need to define codes for functions which pull back open sets to F_σ sets, similar to how continuous functions are coded by collection of pairs indicating that the pull back of

an open set contains some open set. For such a code, we can computably construct a Baire class 1 code for the function, but the argument seems to require Σ_2^0 induction.

However, even stating the consistency of such a ‘‘Borel class 2’’ code seems complicated, as the containment relation between F_σ sets is Π_1^1 complete.

Proposition 5.4 (ACA_0). Cousin’s lemma holds for Baire class 1 gauges.

Proof. The argument of Lemma 2.4 holds for Baire class 1 functions, and so it suffices to work in Cantor space. Let $\delta: 2^\omega \rightarrow \mathbb{R}^+$ be a Baire class 1 gauge. By ACA_0 , for all x , $\delta(x)$ exists.

Now the main two points are:

- (i) In ACA_0 , we can tell which F_σ subsets of 2^ω are empty;
- (ii) In ACA_0 , given a sequence $\langle H_n \rangle$ of nonempty F_σ subsets of 2^ω , there is a choice sequence $\langle x_n \rangle$ so that $x_n \in H_n$ for all n .

The point is that whether a tree has a path is an arithmetic question, and so whether the union of a countable sequence of closed subsets of Cantor space is empty or not is also arithmetic. For (ii), using ACA_0 we can obtain a sequence $\langle S_n \rangle$ of binary trees, each of which has a path, such that $[S_n] \subseteq H_n$ (where $[S_n]$ is the set of paths of S_n), and then apply WKL_0 (or in ACA_0 , just take the leftmost path of each S_n).

We can therefore repeat the proof of Proposition 3.2. Let δ be a Baire class 1 gauge on Cantor space. We define the set G and the tree T in exactly the same way; G and T exists since they are arithmetically definable: $\delta(x) < 2^{-|\sigma|+1}$ is Σ_2^0 by Lemma 5.2. The same argument shows that T cannot have a path, and so by WKL_0 , T is finite. As above, from T we obtain a finite subset R of G which covers all of Cantor space. By (ii), we can find a sequence $\langle x_\sigma : \sigma \in R \rangle$ as required. \square

In the other direction:

Theorem 5.5 (RCA_0). Cousin’s lemma for Baire class 1 functions implies ACA_0 .

The idea of the proof is as follows. Suppose we have a Cauchy sequence $\langle z_n \rangle$ with no limit. The sequence of functions $x \mapsto |x - z_n|$ determine a Baire class 1 function δ , which is a gauge since $\langle z_n \rangle$ has no limit. But there cannot be any δ -fine partition, since no such partition P can cover the gap where $\lim z_n$ should be. Here are the details.

Proof of Theorem 5.5. If ACA_0 fails then there is an increasing Cauchy sequence $\langle z_n \rangle_{n \in \mathbb{N}}$ of real numbers in the unit interval that has no limit ([Sim09, Thm.III.2.2]). For each $n \in \mathbb{N}$, define

$$\delta_n(x) = |x - z_n|.$$

This is a sequence of continuous functions. This sequence is pointwise Cauchy: for all x , n and m ,

$$|\delta_m(x) - \delta_n(x)| = \left| |x - z_m| - |x - z_n| \right| \leq |z_n - z_m|;$$

so we use the fact that $\langle z_n \rangle$ is Cauchy. Note that in the ‘‘real world’’, $\langle \delta_n \rangle$ converges uniformly to the continuous function $x \mapsto |x - z^*|$ where $z^* = \lim_n z_n$, but since z^* does not exist in the model, this function does not have a code in the model. However in RCA_0 we have just shown that $\langle \delta_n \rangle$ is a code for a Baire class 1 function, which we call δ .

We claim that δ is a gauge. For all $x \in [0, 1]$, since $x \neq \lim_n z_n$, and since $\langle z_n \rangle$ is Cauchy, there is some $\varepsilon > 0$ such that for all but finitely many n , $|x - z_n| \geq \varepsilon$. Then $\delta(x) \geq \varepsilon$.

Now by Lemma 2.2 and Cousin's lemma applied to the gauge $\frac{1}{2}\delta$, there is a δ -fine tagged partition $0 = x_0 \leq \xi_0 \leq x_1 \leq \xi_1 \leq x_2 \leq \dots \leq \xi_{n-1} \leq x_n = 1$; recall that this means that $x_{i+1} - x_i \leq \delta(x_i)$ for all $i < n$.

We derive a contradiction. By Σ_1^0 -induction, there is a least i such that for all n , $z_n \leq x_{i+1}$. Since $\langle z_n \rangle$ is increasing, for all but finitely many n , $z_n \in [x_i, x_{i+1}]$. Since none of x_i , ξ_i or x_{i+1} are the limit of $\langle z_n \rangle$, there is some $\varepsilon > 0$ such that one of things happens:

- for all but finitely many n , $z_n \in [x_i + \varepsilon, \xi_i - \varepsilon]$; or
- for all but finitely many n , $z_n \in [\xi_i + \varepsilon, x_{i+1} - \varepsilon]$.

Without loss of generality, assume the former. Then for all but finitely many n ,

$$\delta_n(\xi_i) = \xi_i - z_n \leq (\xi_i - x_i) - \varepsilon \leq (x_{i+1} - x_i) - \varepsilon,$$

whence $\delta(\xi_i) < x_{i+1} - x_i$, contradicting our assumption. The argument in the other case is the same. \square

6. HIGHER BAIRE CLASSES

Baire class 1 functions aren't closed under taking pointwise limits. Indeed, by iterating the operation of taking pointwise limits, we obtain a transfinite hierarchy of functions, which exhausts all Borel functions.

In second-order arithmetic, recall that in RCA_0 , for a Baire class 1 function f , the relation $f(x) \in B$ for open or closed balls B is definable even if $f(x)$ does not exist. We can therefore iterate. By (external) induction on standard n we define:

Definition 6.1. For each $n \geq 1$, a code for a Baire class $(n + 1)$ function f is a sequence $\langle f_n \rangle$ of codes for Baire class n functions such that for all x , $\langle f_n(x) \rangle$ is Cauchy, in the sense that for all $\varepsilon > 0$ there is an open ball B of radius $< \varepsilon$ such that for all but finitely many n , $f_n(x) \in B$.

If $\langle f_n \rangle$ is such a code, then for any open ball $B = B(y, r)$, we say that $f(x) \in B$ if for some $s < r$, for all but finitely many n , $f_n(x) \in B(y, s)$.

Of course by taking constant sequences, we see that every Baire class n function is also Baire class $n + 1$, so the classes are increasing.

ACA_0 implies, for each n , that if f is Baire class n then $f(x)$ exists for all x . Similarly to Lemma 5.2, the inverse images of open sets by Baire class n functions are Σ_{1+n}^0 . We can further extend the definition to Baire class α functions for ordinals α ; a code in this case is an α -ranked well-founded tree, where each non-leaf has full splitting, the leaves are labeled by continuous functions, and each non-leaf node represents the pointwise limit of the functions represented by its children. The relation $f(x) \in U$ is then defined by transfinite recursion on the rank of a node, and so it makes most sense to use ATR_0 as a base system for this kind of development; we do not pursue it here. We remark that ATR_0 implies that every Borel function is Baire class α for some α .¹

Now the proof of Proposition 5.4 cannot be replicated for Baire class n functions for any for $n > 1$. While ACA_0 suffices to determine if a given Σ_2^0 subset of Cantor

¹This follows from the fact that every well-founded tree is ranked; see [Hir00].

space is empty or not, it is Π_1^1 -complete to determine whether a given Π_2^0 set is empty.² And indeed we show that Cousin's lemma for Baire class 2 functions is much stronger than ACA_0 : it implies ATR_0 .

It would again be more convenient to work in Cantor space, and so we need a converse of Lemma 2.4.

Lemma 6.2 (RCA_0). Cousin's lemma for Baire class 2 functions on the unit interval implies Cousin's lemma for Baire class 2 functions on Cantor space.

Of course there is nothing special for the case $n = 2$, the lemma holds for all Baire classes.

Proof. We use the standard embedding of Cantor space into the unit interval: for $x \in 2^\omega$ let

$$\psi(x) = \sum_n 2x(n)3^{-n-1},$$

and let $C = \psi[2^\omega]$ be the ternary Cantor set. We use the following facts: ψ is continuous; the function $z \mapsto d(z, C)$ is computable on $[0, 1]$; and: for $x, y \in 2^\omega$, if $d(x, y) = 2^{-n}$ then $|\psi(x) - \psi(y)| \geq 3^{-n}$, i.e.,

$$|\psi(x) - \psi(y)| \geq 3^{\log_2 d(x, y)}$$

For all $x, y \in 2^\omega$.

Let δ be a Baire class 2 gauge on Cantor space. Define $\hat{\delta}: [0, 1] \rightarrow \mathbb{R}$ by letting, for $z \in [0, 1]$,

$$\hat{\delta}(z) = \begin{cases} 3^{\log_2 \psi^{-1}(z)}, & \text{if } z \in C; \text{ and} \\ d(x, C), & \text{otherwise.} \end{cases}$$

Then $\hat{\delta}$ is positive on $[0, 1]$, and it is Baire class 2; we use the fact that 1_C (the characteristic function of C) is Baire class 1, indeed with a very nice, uniformly computable approximation $\langle g_n \rangle$ of piecewise linear functions with $g_n(z) = 1$ for all n when $z \in C$, and $g_n(z) = 0$ for all but finitely many n when $z \notin C$.

Suppose that (P, \bar{r}_p) is a $\hat{\delta}$ -fine cover. Define $Q = \{x \in 2^\omega : \psi(x) \in P\}$ and for $x \in Q$ let $s_x = 2^{\log_3 r_{\psi(x)}}$. Then (Q, \bar{s}_x) is a δ -fine cover. The main point is that $z \in C$ cannot be covered by any $p \in P \setminus Q$, as for such p we have $\delta(p) = d(p, C) \leq |p - z|$. \square

Before we give the reversal, we motivate our technique by showing something weaker: that ACA_0 does not imply Cousin's lemma for Baire class 2 functions. We will show that every ω -model of $\text{RCA}_0 + \text{Cousin's lemma for Baire class 2 functions}$ must contain $0^{(\omega)}$ (in fact $0^{(\alpha)}$ for any computable ordinal α), in particular Cousin's lemma for Baire class 2 functions fails in the model consisting of the arithmetic sets.

We review some definitions. The Turing jump x' of $x \in 2^\omega$ is the complete $\Sigma_1^0(x)$ set, identified with an element of 2^ω . As mentioned above, the function $x \mapsto x'$ is Baire class 1, but more importantly, its graph (the relation $y = x'$) is Π_2^0 . We will use the following well-known fact, which is provable in ACA_0 :

Lemma 6.3 (ACA_0). Let $f: 2^\omega \rightarrow \mathbb{N}$ be a function. The following are equivalent:

²Every Σ_1^1 set is the projection of a Π_2^0 set in Cantor space (see for example [Sac90, Thm.I.1.5]), and the question of whether a Σ_1^1 set is empty or not is equivalent to a question about well-foundedness of a tree, which is Π_1^1 -complete.

- (1) f is Δ_3^0 -definable;
- (2) There is a computable function $\varphi: 2^\omega \rightarrow \mathbb{N}$ such that for all x , $f(x) = \varphi(x'')$;
- (3) f is effective Baire class 2.

Proof. This is a double application of Shoenfield's limit lemma and Post's hierarchy theorem: a function $f: \mathbb{N} \rightarrow \mathbb{N}$ is Δ_3^0 -definable if and only if it is \emptyset'' -computable if and only if $f = \lim_s \lim_t g(-, s, t)$ for some computable $g: \mathbb{N}^3 \rightarrow \mathbb{N}$ (see for example [Soa87, Lem.III.3.3;Thm.IV.2.2]). The lemma now follows from a uniform relativisation to oracles $x \in 2^\omega$. \square

As a result we get:

Lemma 6.4 (RCA_0). Suppose that $f: 2^\omega \rightarrow \mathbb{N}$ is Δ_3^0 . Suppose that Cousin's lemma for Baire class 2 functions holds. Then there is a finite $P \subset 2^\omega$ such that

$$2^\omega = \bigcup_{p \in P} [p \upharpoonright f(p)].$$

Proof. Define $\delta(x) = 2^{-f(x)}$. Then δ is Δ_3^0 . By Lemma 6.3 (note that we already know that Cousin's lemma for Baire class 2 functions implies ACA_0), δ is a Baire class 2 gauge. A δ -fine cover $P \subset 2^\omega$ is as required. \square

Recall the definition of a transfinite iteration of the Turing jump. Suppose that α is a linear ordering of a subset of \mathbb{N} , which will usually be well-founded. An iteration of the Turing jump along α is a set $H \subseteq \alpha \times \mathbb{N}$ satisfying, for all $\beta < \alpha$,

$$H^{[\beta]} = \left(H^{[<\beta]} \right)',$$

where $H^{[\beta]} = \{k : (\beta, k) \in H\}$ and $H^{[<\beta]} = \{(\gamma, k) \in H : \gamma < \beta\}$. Again note that we view the domain of α as a subset of \mathbb{N} , so both $H^{[\beta]}$ and $H^{[<\beta]}$ are identified with elements of Cantor space, and so taking the Turing jump makes sense.

If α is indeed an ordinal (is well-founded), then ACA_0 implies there is at most one iteration of the Turing jump along α , and if H is such, we write $H = 0^{(\alpha)}$. When this is relativised to an oracle x , we write $H = x^{(\alpha)}$. The relation “ H is an iteration of the Turing jump along α ” is Π_2^0 . ATR_0 is equivalent to the statement that for all ordinals α and all x , $x^{(\alpha)}$ exists ([Sim09, Thm.VIII.3.15]).

Now let \mathcal{M} be an ω -model of ACA_0 , let α be a computable ordinal, and suppose that $0^{(\omega)} \notin \mathcal{M}$; we show that Cousin's lemma for Baire class 2 functions fails in \mathcal{M} . Let $X \in 2^\omega$. If $X \neq 0^{(\omega)}$, there is some $n < \omega$ such that

$$X^{[n]} \neq (X^{[<n]})';$$

let $n(X)$ be the least such n . The function $X \mapsto n(X)$ (as well as its domain) is Δ_3^0 . Further, for such X , we let $k(X)$ be the least k such that

$$X^{[n(X)]}(k) \neq (X^{[<n(X)]})'(k);$$

the function $X \mapsto k(X)$ is Δ_3^0 as well. By their definition, for all $X \neq 0^{(\omega)}$,

$$X(n(X), k(X)) \neq 0^{(\omega)}(n(X), k(X)),$$

as $X^{[<n(X)]} = (0^{(\omega)})^{[<n(X)]}$. Hence, we let $f(X) = \langle n(X), k(X) \rangle + 1$; for all $X \neq 0^{(\omega)}$,

$$0^{(\omega)} \notin [X \upharpoonright f(X)].$$

Since $0^{(\omega)} \notin \mathcal{M}$, f is total on $2^\omega \cap \mathcal{M}$, and since \mathcal{M} is an ω -model, the same Δ_3^0 definition holds in \mathcal{M} , so \mathcal{M} believes that f is a Baire class 2 gauge on Cantor space. Let $P \subset 2^\omega \cap \mathcal{M}$ be finite. Since $0^{(\omega)} \notin P$, we have

$$0^{(\omega)} \notin \bigcup_{X \in P} [X \upharpoonright f(X)],$$

and so

$$2^\omega \neq \bigcup_{X \in P} [X \upharpoonright f(X)].$$

But since P is finite, the latter fact is absolute for \mathcal{M} (the set $\bigcup_{X \in P} [X \upharpoonright f(X)]$ is a clopen set and so its complement is a non-empty clopen set; \mathcal{M} contains an element of any non-empty clopen set). Hence Lemma 6.4 fails in \mathcal{M} .

Exactly the same argument shows that for any computable ordinal α , $0^{(\alpha)}$ is an element of every ω -model of Cousin's lemma for Baire class 2 functions. The more general argument below builds on this technique. In the meantime, we note that this argument gives a purely computability-theoretic result:

Theorem 6.5. For every computable ordinal α , there is an effective Baire class 2 gauge δ such that every δ -fine partition must contain $0^{(\alpha)}$.

Proof. Extend the function f above to a function on all of 2^ω by defining $f(0^{(\alpha)}) = 0$, and transform to a gauge as in the proof of Lemma 6.4. This is still a Δ_3^0 -definable function since the relation $X = 0^{(\alpha)}$ is Δ_3^0 (indeed it is Π_2^0). \square

Theorem 6.6 (RCA_0). Cousin's lemma for Baire class 2 functions implies ATR_0 .

Proof. By Theorem 5.5, we may work in ACA_0 .

Let α be an ordinal and let $x \in 2^\omega$; and suppose that $x^{(\alpha)}$ does not exist, i.e., there is no iteration of the Turing jump (relativised to x) along α . For simplicity of notation, suppose $x = 0$.

Let

$$\mathcal{I} = \left\{ \beta < \alpha : 0^{(\beta)} \text{ exists} \right\}.$$

The set \mathcal{I} does not exist (as a set in the model), but is definable. Again note that since ACA_0 holds, for all $\beta \in \mathcal{I}$ there is exactly one iteration of the Turing jump along β , which is the set we call $0^{(\beta)}$. We observe that \mathcal{I} cannot have a greatest element; if 0^β exists then by ACA_0 , so does $0^{(\beta+1)}$.

For any $X \in 2^\omega$, since X is not an iteration of the Turing jump along α , by arithmetic comprehension, there is a least $\beta < \alpha$ such that $X^{[\beta]} \neq (X^{[<\beta]})'$; call this $\beta(X)$. Then $X^{[<\beta(X)]} = 0^{(\beta(X))}$, so $\beta(X) \in \mathcal{I}$. Similar to the argument above, we also let $k(X)$ be the least k such that

$$X^{[\beta(X)]}(k) \neq (X^{[<\beta(X)]})'(k).$$

From $k(X)$ and $\beta(X)$ we can compute some $f(X) \in \mathbb{N}$ such that for all $\gamma > \beta(X)$ in \mathcal{I} ,

$$0^{(\gamma)} \notin [X \upharpoonright f(X)].$$

Let $P \subset 2^\omega$ be finite. Then

$$\beta^* = \max\{\beta(X) : X \in P\}$$

is an element of \mathcal{I} . Since \mathcal{I} does not have a last element, take any $\gamma > \beta^*$ in \mathcal{I} ; then

$$0^{(\gamma)} \notin \bigcup_{X \in P} [X \upharpoonright f(P)].$$

Therefore, Lemma 6.4 fails, and so Cousin's lemma for Baire class 2 functions does not hold. \square

REFERENCES

- [Bai99] René-Louis Baire. Sur les fonctions de variables réelles. *Annali di Matematica Pura ed Applicata (1898-1922)*, 3(1):1–123, 1899.
- [Bar20] Jordan Mitchell Barrett. The reverse mathematics of Cousin's lemma. Honours thesis, Victoria University of Wellington, 2020.
- [Bec92] Howard Becker. Descriptive set-theoretic phenomena in analysis and topology. In *Set theory of the continuum (Berkeley, CA, 1989)*, volume 26 of *Math. Sci. Res. Inst. Publ.*, pages 1–25. Springer, New York, 1992.
- [Cou95] Pierre Cousin. Sur les fonctions de n variables complexes. *Acta Math.*, 19(1):1–61, 1895.
- [Den12] Arnaud Denjoy. Une extension de l'intégrale de M. Lebesgue. *CR Acad. Sci. Paris*, 154:859–862, 1912.
- [Dir29] P. G. L. Dirichlet. Sur la convergence des séries trigonométriques qui servent à représenter une fonction arbitraire entre des limites données. *Journal für die reine und angewandte Mathematik*, 4:157–169, 1829.
- [DK91] Randall Dougherty and Alexander S. Kechris. The complexity of antidifferentiation. *Adv. Math.*, 88(2):145–169, 1991.
- [Gor94] Russell A. Gordon. *The integrals of Lebesgue, Denjoy, Perron, and Henstock*, volume 4 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1994.
- [Gor97] Russell A. Gordon. Is nonabsolute integration worth doing? volume 22, pages 23–33. 1996/97. Summer Symposium in Real Analysis, XX (Windsor, ON, 1996).
- [Hea93] Oliver Heaviside. *Electromagnetic Theory*, volume 1. The Electrician Printing and Publishing Company, London, 1893.
- [Hir00] Jeffrey L. Hirst. Reverse mathematics and rank functions for directed graphs. *Arch. Math. Logic*, 39(8):569–579, 2000.
- [KL90] Alexander S. Kechris and Alain Louveau. A classification of Baire class 1 functions. *Trans. Amer. Math. Soc.*, 318(1):209–236, 1990.
- [KT14] Rutger Kuyper and Sebastiaan A. Terwijn. Effective genericity and differentiability. *J. Log. Anal.*, 6:Paper 4, 14, 2014.
- [Kur57] Jaroslav Kurzweil. Generalized ordinary differential equations and continuous dependence on a parameter. *Czechoslovak Math. J.*, 7(82):418–449, 1957.
- [Luz12] Nikolai N. Luzin. Sur les propriétés de l'intégrale de M. Denjoy. *CR Acad. Sci. Paris*, 155:1475–1478, 1912.
- [Mul87] P. Muldowney. *A general theory of integration in function spaces, including Wiener and Feynman integration*, volume 153 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1987.
- [NS19] Dag Normann and Sam Sanders. On the mathematical and foundational significance of the uncountable. *J. Math. Log.*, 19(1):1950001, 40, 2019.
- [NS20] Dag Normann and Sam Sanders. Open sets in computability theory and reverse mathematics. *J. Logic Comput.*, 30(8):1639–1679, 2020.
- [PDD17] Michelle Porter, Adam Day, and Rodney Downey. Notes on computable analysis. *Theory Comput. Syst.*, 60(1):53–111, 2017.
- [Per14] Oskar Perron. *Über den Integralbegriff*. Sitzungsberichte. Heidelberger Akademie der Wissenschaften, 1914.
- [Sac90] Gerald E. Sacks. *Higher recursion theory*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1990.
- [Sim09] Stephen G. Simpson. *Subsystems of second order arithmetic*. Perspectives in Logic. Cambridge University Press, Cambridge; Association for Symbolic Logic, Poughkeepsie, NY, second edition, 2009.

- [Soa87] Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987. A study of computable functions and computably generated sets.
- [Wal23] J. L. Walsh. A closed set of normal orthogonal functions. *Amer. J. Math.*, 45(1):5–24, 1923.
- [Wal17] Sean Walsh. Definability aspects of the Denjoy integral. *Fund. Math.*, 237(1):1–29, 2017.

VICTORIA UNIVERSITY OF WELLINGTON, NZ

E-mail address: `math@jmbarrett.nz`

URL: `https://jmbarrett.nz/`

VICTORIA UNIVERSITY OF WELLINGTON, NZ

E-mail address: `rod.downey@vuw.ac.nz`

URL: `https://homepages.ecs.vuw.ac.nz/~downey/`

VICTORIA UNIVERSITY OF WELLINGTON, NZ

E-mail address: `greenberg@msor.vuw.ac.nz`

URL: `https://homepages.ecs.vuw.ac.nz/~greenberg/`