Under consideration for publication in Math. Struct. in Comp. Science

A Constructive Interpretation of Ramsey's Theorem via the Product of Selection Functions

Paulo Oliva and Thomas Powell

Queen Mary University of London

School of Electronic Engineering and Computer Science

London E1 4NS

Received June 15, 2012

We use Gödel's dialectica interpretation to produce a computational version of the well known proof of Ramsey's theorem by Erdős and Rado. Our proof makes use of the product of selection functions, which forms an intuitive alternative to Spector's bar recursion when interpreting proofs in analysis. This case study is another instance of the application of proof theoretic techniques in mathematics.

1. Introduction

In a fundamental paper of the 1950s (Kreisel, 1951; Kreisel, 1952), Kreisel first suggested utilising proof interpretations to systematically 'unwind' non-constructive proofs and discover their constructive content. Kreisel's pioneering work forms the foundation of modern applied proof theory (in the sense of (Kohlenbach, 2008)) which has seen variants of Gödel's functional interpretation used to produce improved results in areas such as numerical analysis and ergodic theory through the extraction of computational content from classical proofs.

This *proof mining* program, as it is known today, has generally focused on developing general metatheorems that guarantee the extractability of effective uniform bounds from proofs of theorems of a specific logical form - usually relatively simple Π_2 theorems for which direct computational data can be found. In other words, on the whole proof interpretations have been applied to extract purely quantitative information from a fairly restricted class of theorems. However, the last decade has seen proof interpretations employed much more widely, with an increasing emphasis on understanding the qualitative aspects of interpreted proofs.

There are two main reasons for this. The first is a greater appreciation of the mathematical significance of proof interpretations. It was recently observed (e.g. (Kohlenbach, 2008)) that the monotone variant of Gödel's dialectica interpretation is closely related to the so-called 'correspondence principle' between finite and infinite dynamical systems

as discussed by T. Tao in (Tao, 2008, Chapter 1.3). This observation lies behind current applications of the dialectica interpretation in ergodic theory (see e.g. Avigad (Avigad, 2009)), which in particular explore the dialectica interpretation of Cauchy convergence, known to mathematicians as *metastability*.

The second reason is an improved understanding of the *semantics* of proof interpretations. Formal translations on proofs are highly syntactic and in particular the functional interpretation of proofs that make use of full arithmetic comprehension traditionally involves Spector's abstruse bar recursion schema. Consequently, realisers for interpreted proofs are often stated as almost unreadable higher type terms. This issue is addressed in recent work by the authors and M. Escardó (Escardó and Oliva, 2010a; Escardó and Oliva, 2011), who show that the product of selection functions provides us with an intuitive alternative to bar recursion that can be understood in terms of the computation of optimal strategies in a certain class of sequential games. This makes it easier to appreciate the operational behaviour of realisers of interpreted theorems in analysis.

Therefore the authors believe that it is both practical and meaningful to apply proof interpretations to classical proofs with the object of producing a *mathematical* proof of a new, finitary theorem, as opposed to just extracting a new piece of quantitative information. In this article we apply Gödel's dialectica interpretation to Erdős and Rado's proof of Ramsey's theorem for pairs, similarly to what has been done in (Kohlenbach and Kreuzer, 2009; Kreuzer, 2009). Our main aim here, however, is to produce an intuitive combinatorial proof of the finitary form of the Ramsey's theorem given in Section 1.2. For that purpose, we endeavour to strip our proof of heavy logical syntax in order to understand it in mathematical terms. In a broader sense our aim is to portray the dialectica interpretation as an intelligent translation on proofs as opposed to just a syntactic translation on logical formulas.

The paper is organised as follows. We begin by formulating Ramsey's theorem and its proof in the language of formal arithmetic. We then briefly discuss the main building block of our extracted proof, the product of selection functions, and in Section 4 we prove our finitary version of the theorem. Finally, we discuss a game theoretic reading of our proof.

1.1. *Preliminaries*

In this article we assume that the reader is familiar with Gödel's *dialectica interpretation* of classical proofs (cf. (Avigad and Feferman, 1998; Kohlenbach, 2008) and the original paper (Gödel, 1958)), by which we implicitly mean Gödel's dialectica interpretation combined as usual with the negative translation[†]. We do not assume familiarity with the authors' recent work on the product of selection functions - although the reader is encouraged to consult (Escardó and Oliva, 2011) for a more detailed treatment of the results mentioned in Section 3.

The theory PA^ω is Peano arithmetic in all finite types, and T is Gödel's quantifier-free

[†] As in (Kohlenbach, 2008) we adopt Kuroda's variant of the negative translation.

theory of higher-type primitive recursive functionals (see (Avigad and Feferman, 1998) for full definitions). We make informal use of types like the Booleans $\mathbb{B} = \{0, 1\}$ and finite sequence types X^*.

Notation. We use the following abbreviations:

$s * t$ is the concatenation of sequences s and t.
$\widehat{s} \equiv s * \mathbf{0}^{X^{\mathbb{N}}}$ a canonical infinite extension of a finite sequence s.
$[\alpha^{X^{\mathbb{N}}}](n) \equiv \langle \alpha 0, \ldots, \alpha(n-1) \rangle$ is the initial segment of α of length n.

Finally, we make use of the following key logical principles. Π^0_1 *countable choice* is given by the schema

$$\Pi^0_1\text{-}\mathsf{AC}^0 \ : \ \forall n \exists x^X \forall y^Y A_n(x, y) \to \exists \alpha^{\mathbb{N} \to X} \forall n, y A_n(\alpha n, y),$$

where the A_n are quantifier-free, *weak König's lemma* is the statement that any infinite decidable binary tree T has an infinite branch:

$$\mathsf{WKL} \ : \ \forall n \exists s^{\mathbb{B}^*} (|s| \wedge T(s)) \to \exists \alpha^{\mathbb{N} \to \mathbb{B}} \forall n T([\alpha](n)),$$

and the *infinite pigeonhole principle* states that for any n-colouring c of the natural numbers, at least one colour x is used infinitely often:

$$\mathsf{IPHP} \ : \ \forall c^{\mathbb{N} \to [n]} \exists x, p^{\mathbb{N} \to \mathbb{N}} \forall k (pk \geq k \wedge c(pk) = x).$$

Note that of these only IPHP is provable in PA^ω

1.2. *Ramsey's theorem for pairs*

In this article we only consider Ramsey's theorem for pairs and 2-colourings on the basis that our results can be extended to the more general theorem, although in the course of our program extraction we hint at how key steps can be generalised for the n-colour case.

Let $[\mathbb{N}]^2$ denote the set of subsets of $\mathbb{N}$ of size two, and suppose we are given a colouring $c \colon [\mathbb{N}]^2 \to \mathbb{B}$ of $[\mathbb{N}]^2$ with two colours. Ramsey's theorem says that for any such colouring there exists an infinite pairwise monochromatic subset of $\mathbb{N}$ i.e. an infinite set $S \subseteq \mathbb{N}$ such that all elements of $[S]^2$ have colour x for some $x \in \mathbb{B}$. Formally, we write Ramsey's theorem as

$$\mathsf{RT}^2_2(c) \ : \ \exists x^{\mathbb{B}} \exists F^{\mathbb{N} \to \mathbb{N}} \forall k (Fk \geq k \wedge \forall i, j \leq k (Fi < Fj \to c(\{Fi, Fj\}) = x)).$$

Here the infinite monochromatic set is encoded by the function F and is given by $S_F = \{Fk \ : \ k \in \mathbb{N}\}$. Our main result is a constructive proof of the dialetica interpretation of $\mathsf{RT}^2_2(c)$:

Main Theorem 1. Suppose the colouring c is fixed. For any functional $\eta \colon \mathbb{B} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ there exists a colour $x \colon \mathbb{B}$ and a function $F \colon \mathbb{N} \to \mathbb{N}$ satisfying

$$\forall k \leq \eta_x F (Fk \geq k \wedge \forall i, j \leq k (Fi < Fj \to c(\{Fi, Fj\}) = x)). \tag{1}$$

As with all Σ_2 theorems, the functional interpretation of Ramsey's theorem coincides

with Kreisel's *no-counterexample* interpretation (n.c.i.). The intuition is that the counterexample functions η_0 and η_1 attempt to show that for any F the set S_F cannot be pairwise monochromatic, and we are challenged to effectively refute any such counterexample functions. Alternatively, we can view η as a function that specifies in advance how we want to "use" Ramsey's theorem in a specific computation: while in general there is no effective way of realising Ramsey's theorem, given η we can produce an *approximation* to a monochromatic set that is sufficient for the computation we have in mind.

1.3. *Comparison to existing work*

Ramsey's theorem has been extensively studied in logic, so it is important to outline how our work contrasts to related papers on the constructive content of the theorem.

In (Bellin, 1990) Bellin uses proof theoretic techniques to produce a proof of a finitary version of Ramseys' theorem similar to (1). However, his proof differs from ours in two important respects. Firstly, his is based on Ramsey's original proof as opposed to the one analysed here by Erdős and Rado, and secondly he uses cut-elimination and Herbrand's theorem as opposed to the dialectica interpretation.

A formalisation of Erdős and Rado's proof was recently given by Kreuzer and Kohlenbach in (Kohlenbach and Kreuzer, 2009), and a bar-recursive realizer for its functional interpretation was stated in (Kreuzer, 2009). The main achievement of these works is to calibrate the proof theoretic strength of $\mathsf{RT}^2_2(c)$ and establish its contribution to the complexity of extracted programs in certain cases, whereas our goal is to produce an intuitive constructive version of the Erdős-Rado proof using the product of selection functions that can be understood in mathematical terms.

We note that while our formalisation of the Erdős-Rado proof is influenced by theirs in that we also encode the Erdős-Rado min-monochromatic tree as as a binary Σ^0_1 tree, our treatment differs substantially from (Kohlenbach and Kreuzer, 2009; Kreuzer, 2009). In particular we encode min-monochromatic branches using a different Σ^0_1 tree, and in our program extraction we use new interpretations of WKL and Π^0_1-AC using the product of selection functions, as opposed to the standard bar-recursive interpretations of Howard and Spector used in (Kreuzer, 2009).

Veldman and Bezem (Veldman and Bezem, 1993) discovered an interesting constructive variant of Ramsey's theorem. That formulation and proof have been simplified by Coquand in (Coquand, 1994b) (see also (Coquand, 1994a)). Coquand's proof makes use a recursion on well-founded trees similar to Spector's bar recursion. The main difference being that in our algorithm the well-founded tree is not given explicitly as part of the problem, as it is in Coquand's formulation of Ramsey's theorem.

To summarise, then, in comparison to existing work our analysis of Ramsey's theorem combines the following key benefits:

1 Our **constructive interpretation** of the theorem is based on Gödel's dialectica interpretation. The advantage of this is that our theorem is more 'computational' than e.g. that of Veldman and Bezem (Veldman and Bezem, 1993) in that we explicitly prove the existence of arbitrarily large approximations to a monochromatic set.

Moreover, as indicated previously, our finitary Ramsey's theorem can be related to the finitisations of infinitary theorems in the sense of Tao (Tao, 2008).

2 Our **constructive proof** of Ramsey's theorem is based on the product of selection functions as opposed to Spector's bar recursion and can be given a clear game theoretic interpretation (Section 5).

2. A Formal Proof of Ramsey's Theorem

Notation. For simplicity we encode a colouring $c\colon [\mathbb{N}]^2 \to \mathbb{B}$ as a map $c\colon \mathbb{N}^2 \to \mathbb{B}$ with the property that $c(i,j) = c(j,i)$ for all i, j.

We now present a formal proof of Ramsey's theorem based on that of Erdős and Rado ((Erdős et al., 1984), Section 10.2). In doing so we show that $\mathsf{RT}^2_2(c)$ can be formalised in $\mathrm{PA}^\omega + \mathsf{WKL} + \Pi^0_1\text{-}\mathsf{AC}$, and therefore its functional interpretation can theoretically be witnessed using Spector's bar recursion (Kohlenbach, 2008; Spector, 1962), or alternatively (as we demonstrate in Section 3) the product of selection functions. Of course, actually constructing this witness is non-trivial – the soundness theorem for the dialectica interpretation gives a syntactic algorithm which would be impractical to carry out by hand. Therefore, we make use of the soundness theorem as a very rough guide on how to proceed but use shortcuts whenever possible.

The main idea behind the classical proof is, given a colouring c, to organise the natural numbers into a tree (described as an ordering $\prec$ on $\mathbb{N}$) whose branches are *min-monochromatic*, in the sense that $c(i,j) = c(i,k)$ for $i \prec j \prec k$, where $i \prec j$ says that node i precedes j in the tree. This is the so-called Erdős-Rado (E/R) tree. By König's lemma the E/R tree has an infinite min-monochromatic branch $a\colon \mathbb{N}^{\mathbb{N}}$, so by the infinite pigeonhole principle applied to the colouring $c^a(i) = c(a(i), a(i+1))$ there exists an infinite subset of the branch that is pairwise monochromatic.

Our formal proof proceeds, in a similar fashion to (Kohlenbach and Kreuzer, 2009), as follows. We encode branches of the E/R tree by an infinite Σ^0_1 binary tree T (Definition 2.0.2). We then reduce T to an infinite *decidable* binary tree using $\Pi^0_1\text{-}\mathsf{AC}$ (Lemma 2.0.4), which by WKL has an infinite branch. We then show that an infinite branch of T does indeed encode an infinite branch of the E/R tree (Lemma 2.0.7). Hence, we are finally able to complete the proof using IPHP (Theorem 2.0.11). Because we are only considering here the case of two colours, we do not need the full IPHP but only a very simple instance of it (case $n = 2$). Nevertheless, we discuss the whole construction in terms of the full IPHP so that a generalisation to the case of finitely many colours is more straightforward. We sketch our formal proof in Figure 1. Here $\mathsf{E/R}(c)$ abbreviates the statement that the E/R tree defined by c has an infinite branch.

$$\dfrac{\mathsf{IPHP} \qquad \dfrac{\mathsf{WKL} \qquad \Pi^0_1\text{-}\mathsf{AC}}{\mathsf{E/R}(c)}\,{\scriptstyle 2.0.4\text{-}2.0.10}}{\mathsf{RT}^2_2(c)}\,{\scriptstyle 2.0.11}$$

Figure 1. Formal proof of Ramsey's theorem.

It is important to remark why we have chosen this proof over Ramsey's seemingly simpler proof in (Ramsey, 1930). Ramsey constructs an infinite min-monochromatic branch directly using dependent choice: First we define $a0 = 0$, then we use IPHP to produce an infinite set $A_1 \subseteq \mathbb{N}\backslash 0$ that is monochromatic under $c_0(i) = c(0, i)$ and define $a(1) = \min A_0$. Next use IPHP to produce an infinite set $A_2 \subseteq A_1 \backslash a(1)$ that is monochromatic under $c_{a(1)}(i) = c(a(1), i)$ and define $a(2) = \min A_1$ and so on. It is easy to see that the resulting a is min-monochromatic. However, Ramsey's construction uses dependent choice of type 1 (Simpson shows in (Simpson, 1999) that it cannot be formalised in the subsystem ACA_0), therefore its computational interpretation would seemingly involve bar recursion/product of selection functions of level 1. Our interpretation of the Erdős-Rado proof, on the other hand, makes use of the product of selection functions of lowest type only, meaning that our construction is computationally simpler.

Definition 2.0.1 (Erdős/Rado tree). Given a colouring $c\colon \mathbb{N}^2 \to \mathbb{B}$, define a partial order $\prec$ on $\mathbb{N}$ recursively as follows:

1 $0 \prec 1$

2 Given that $\prec$ is already defined on the initial segment of the natural numbers $[j]$, for $j < i$ define

$$j \prec i \quad \text{iff} \quad c(k, i) = c(k, j), \text{ for all } k \prec j$$

It is easy to show that $\prec$ defines a tree on $\mathbb{N}$, the so-called Erdős/Rado tree, and that its branches are min-monochromatic i.e. $c(k, i) = c(k, j)$ for $k \prec i \prec j$. Moreover, the tree is binary branching because i and j are successors of k if and only if $c(k, i) \neq c(k, j)$. For proofs of these facts see (Kohlenbach and Kreuzer, 2009, Section 4). We consider the following Σ^0_1 tree.

Definition 2.0.2 (Binary Erdős/Rado tree). Define the Σ^0_1-predicate T on $\mathbb{B}^*$ by

$$T(s) := \exists k(\underbrace{\exists k' \in [|s|, k]\, \forall i < |s|(s_i = 0 \;\leftrightarrow\; i \prec k')}_{T'(s,k)}).$$

A 0-1 sequence s belongs to T if it is the characteristic function of a finite branch of the Erdős/Rado tree. We use a k and a k' in order to make $T(s)$ a Σ^0_1-predicate *monotone* on unbounded quantifier k. This will simplify the construction.

Lemma 2.0.3. The following are simple properties of T

(i) T as defined above is an infinite tree.
(ii) The branches of T are characteristic functions of branches of the E/R tree.
(iii) T satisfies the following monotonicity conditions[‡]:

$$\text{(M1)}\;\; T'(s * t, k) \to T'(s, k) \qquad \text{and} \qquad \text{(M2)}\;\; T'(s, k) \to T'(s, k + l).$$

Proof. (i) Clearly T is prefix closed. Moreover, for all n, $T(s)$ has a branch s of length n given by $s_i = 0$ iff $i \prec n$, for $i < n$. (ii) $T(s)$ implies that the set defined by s is an initial segment of the branch of the predecessors of k', denoted $pd(k')$, of the E/R

[‡] It will become clear in Section 4 why we require our tree to have these properties.

tree. Therefore is also a branch of the E/R tree. (*iii*) (M1) is obvious, and (M2) follows because we only ask for a *bound* on k'. □

The first step in our proof is to prove the existence of a function β which will allow us to turn the Σ^0_1-tree $T(s)$ into a decidable tree.

Lemma 2.0.4 (Monotone Skolem function). There exists a function β such that

$$\forall n \forall s(|s| = n \wedge \exists k T'(s,k) \to T'(s, \beta n)). \tag{2}$$

Proof. Classically we have that

$$\forall n \forall s(|s| = n \to \exists k'(\exists k T'(s,k) \to T'(s,k'))).$$

By *bounded collection* and monotonicity of T' we have

$$\forall n \exists k' \forall s(|s| = n \wedge (\exists k T'(s,k) \to T'(s,k'))).$$

Finally, by *countable choice* for Π^0_1-formulas we obtain the function β. □

The Skolem function β allows us to turn the Σ^0_1-predicate $T(s)$ into a decidable predicate:

Corollary 2.0.5. Given a function β satisfying (2) we have that $T(s)$ is equivalent to

$$\underbrace{\exists k \in [|s|, \beta(|s|)]\, \forall i < |s| (s_i = 0 \;\leftrightarrow\; i \prec k)}_{T^\beta(s)}.$$

Once we have a decidable infinite finitely branching tree $T^\beta(s)$ we can apply weak König's lemma to obtain an infinite path in the tree.

Lemma 2.0.6. There exists an infinite sequence α such that

$$\forall n \underbrace{\exists k \in [n, \beta n]\, \forall i < n (\alpha(i) = 0 \;\leftrightarrow\; i \prec k)}_{T^\beta([\alpha](n))}. \tag{3}$$

Proof. By *weak König's lemma.* □

However, it remains to show that an infinite branch of T encodes an infinite branch of the Erdős/Rado tree.

Lemma 2.0.7. The sequence α has infinitely many zeros, i.e. it is the characteristic function of an infinite set. More specifically, we can construct a function $a\colon \mathbb{N} \to \mathbb{N}$ that returns the first $k \geq n$ with $\alpha(k) = 0$.

Proof. Define $a(n)$ as

$$a(n) = \begin{cases} 0 & \text{if } n = 0 \\ k & \text{for least } k \in [n, \beta(\beta n + 1)] \text{ such that } \alpha(k) = 0. \end{cases}$$

We show that a is well-defined, so that in fact $\alpha(a(n)) = 0$ for all n. Because the image of a is unbounded the result follows. We must have that $\alpha(0) = 0$ by definition of $\prec$. Now given $n > 0$, let $i < n$ be the largest such that $\alpha(i) = 0$. Consider $k \in [n, \beta n]$ which

by (3) satisfies $\forall i < n(\alpha(i) = 0 \;\leftrightarrow\; i \prec k)$; and hence $i \prec k$. Now, let n be $\beta n + 1$ in (3) so that we have a $k' \in [\beta n + 1, \beta(\beta n + 1)]$ satisfying $\forall i < n(\alpha(i) = 0 \;\leftrightarrow\; i \prec k')$; and hence $i \prec k'$ as well. Finally, let n be $\beta(\beta n + 1) + 1$ in (3) so that we have a $k'' \in [\beta(\beta n+1)+1, \beta(\beta(\beta n+1)+1)]$ satisfying $\forall i < n(\alpha(i) = 0 \;\leftrightarrow\; i \prec k'')$; so that also $i \prec k''$. Since we have $i \prec k$ and $i \prec k'$ and $i \prec k''$, it follows that either $k \prec k'$ or $k \prec k''$ or $k' \prec k''$, since the Erdős/Rado tree is binary branching. Hence, either $\alpha(k) = 0$ or $\alpha(k') = 0$, and either way there is some $l \in [n, \beta(\beta n + 1)]$ with $\alpha(l) = 0$. □

Remark 2.0.8. Note that in verifying that $\alpha(a(n)) = 0$ we have only used (3) up to the point $\max\{n, \beta n + 1, \beta(\beta n + 1) + 1\}$. We use this fact later to show that a sufficiently large approximation of α is sufficient for an approximation of a.

Remark 2.0.9. For the n colour case the Erdös/Rado tree is still *finitely* branching but not *binary* branching as it is for case of two colours $n = 2$. This in particular means that generalising the proof of Lemma 2.0.7 for arbitrarily many colours is non-trivial (although still routine), and the construction of the function a and the bound in Remark 2.0.8 are more complex (involving further iterations of β). Note, however, that the tree $T(s)$ would still be *binary* branching, even in the case of n colours, as $T(s)$ means that s is the "characteristic function" of a branch in the Erdös/Rado tree. In particular, only the weak form of König's lemma is required in the general case as well.

Corollary 2.0.10. There exists an infinite set that is min-monochromatic under the colouring $c\colon \mathbb{N}^2 \to \mathbb{B}$.

Proof. Clearly the set $\{an \;:\; n \in \mathbb{N}\}$ is infinite. Moreover for $ak < ai < aj$ it follows from (3) for $n = aj + 1$ that $ak \prec ai \prec aj$, and therefore $c(ak, ai) = c(ak, aj)$. □

All that remains is to apply the infinite pigeonhole principle to the min-monochromatic branch given by a.

Theorem 2.0.11 (Ramsey's theorem). For every colouring $c\colon \mathbb{N}^2 \to \mathbb{B}$

$$\exists x^{\mathbb{B}} \exists F^{\mathbb{N}\to\mathbb{N}} \forall k (Fk \geq k \wedge \forall i, j \leq k (Fi < Fj \to c(Fi, Fj) = x)).$$

Proof. Let a be as in the previous lemma. Clearly $an \geq n$, so the image of a is an infinite set. Moreover,

$$c(a(k), a(i)) = c(a(k), a(j)), \tag{4}$$

whenever $a(k) < a(i)$ and $a(k) < a(j)$, by (3) and definition of a. Finally, define a couloring $c'\colon \mathbb{N} \to \mathbb{B}$ as $c'(n) = c(a(n), a(a(n) + 1))$. By the *infinite pigeon-hole principle* we have a p and an x such that $p(n) \geq n$ and

$$x = c'(pi) = c(a(pi), a(a(pi) + 1)) \overset{(4)}{=} c(a(pi), a(pj)),$$

for $a(pi) < a(pj)$. Hence, $F(i) = a(pi)$ does the job. □

3. The Product of Selection Functions

It is well known that just as Peano arithmetic has a dialectica interpretation in the primitive recursive functionals of finite type T, classical analysis (i.e. $\mathrm{PA}^\omega + \mathsf{AC}^0$) has a dialectica interpretation in the bar recursive functionals $\mathrm{T} + \mathsf{BR}$, where BR is the bar recursor introduced by Spector in his fundamental paper (Spector, 1962).

Spector's bar recursion is rather abstruse, and the operational behaviour of programs that make use of this kind of recursion tends to be quite difficult to understand. This was not originally an issue, as Spector's aim was simply to obtain a relative consistency proof for analysis. However, when using the dialectica interpretation to extract programs from proofs in analysis, it is sensible to ask whether there is a more intuitive alternative to bar recursion that facilitates a better understanding of these programs.

In (Escardó and Oliva, 2010a), the first author and Escardó propose the *product of selection functions* as a (computationally equivalent) alternative to bar recursion. In contrast to bar recursion, the product of selection functions is a versatile construction that seems to appear naturally in a variety of different contexts in mathematics and computer science, such as fixed point theory, algorithms and game theory. As such, extracted programs that make use of the product tend to be more illuminating.

In this section we briefly outline the main results that will be used in Section 4, and provide some motivation as to why we prefer the product over bar recursion. The reader is encouraged to consult the survey paper (Escardó and Oliva, 2011) and a recent paper on the extraction of programs from proofs using selection functions (Oliva and Powell, 2012) for further details and discussion.

We call *selection function* any element of type $J_R X := (X \to R) \to X$. Given a selection function $\varepsilon\colon (X \to R) \to X$ we denote by $\overline{\varepsilon}\colon (X \to R) \to R$ the functional $\overline{\varepsilon}(p) \overset{R}{=} p(\varepsilon p)$.

Definition 3.0.12 (Binary product of selection functions (Escardó and Oliva, 2010b)). Given a selection function $\varepsilon\colon J_R X$ and family of selection functions $\delta_x\colon J_R Y$ and a predicate $q\colon X \times Y \to R$, let

$$\begin{aligned} B[x^X] &\overset{Y}{:=} \delta(x, \lambda y.q(x,y)) \\ a &\overset{X}{:=} \varepsilon(\lambda x.q(x, B[x])). \end{aligned}$$

The binary product $\varepsilon \otimes \delta$ of the selection functions ε and δ is another selection function, of type $J_R(X \times Y)$, defined by

$$(\varepsilon \otimes \delta)(q) \overset{X \times Y}{:=} \langle a, B[a] \rangle.$$

As described in (Escardó and Oliva, 2010b), we can iterate the binary product of selection functions an unbounded number of times, where the length of the iteration is dependent on the output of the product in the following sense.

Definition 3.0.13 (Iterated product of selection functions (Escardó and Oliva, 2010b)). Suppose we are given a family of selection functions $(\varepsilon_s\colon J_R X)$, where $s\colon X^*$. The *explicitly controlled unbounded product* of the selection functions ε_s is defined by the

recursion schema

$$\mathsf{EPS}^{\omega}_{s}(\varepsilon)(q) \overset{X^{\mathbb{N}}}{=} \begin{cases} \mathbf{0} & \text{if } \omega(\widehat{s}) < |s| \\ (\varepsilon_s \otimes \lambda x.\mathsf{EPS}^{\omega}_{s*x}(\varepsilon))(q) & \text{otherwise} \end{cases} \tag{5}$$

where $s\colon X^*$, $q\colon X^{\mathbb{N}} \to R$ and $\omega\colon X^{\mathbb{N}} \to \mathbb{N}$.

When ω is a constant function, say $\omega\alpha = n$, this corresponds to a finite iteration of the binary product. The functional ω acts as a control, terminating the recursion once it has produced a sequence s satisfying $\omega(\widehat{s}) < |s|$. By simply unwinding the definition of the binary product in (5) we obtain an equivalent equation

$$\mathsf{EPS}^{\omega}_{s}(\varepsilon)(q) \overset{X^{\mathbb{N}}}{=} \begin{cases} \mathbf{0} & \text{if } \omega(\widehat{s}) < |s| \\ a_s * \mathsf{EPS}^{\omega}_{s*a_s}(\varepsilon)(q_{a_s}) & \text{otherwise} \end{cases} \tag{6}$$

where $a_s = \varepsilon_s(\lambda x.\overline{\mathsf{EPS}^{\omega}_{s*x}(\varepsilon)}(q_x))$, with $q_x(\alpha) = q(x * \alpha)$ and $\overline{\delta}(p) \overset{R}{=} p(\delta p)$.

For fixed ω, ε and q we should think of $\mathsf{EPS}^{\omega}_{s}(\varepsilon)(q)$ as computing an infinite extension to any given finite sequence s. Hence, we are interested in the sequence $s * \mathsf{EPS}^{\omega}_{s}(\varepsilon)(q)$. The fundamental property of EPS is that the infinite extension of an initial segment $[\alpha](n)$ of a previous infinite extension α is identical to the original infinite extension. Formally:

Lemma 3.0.14 (Main lemma on EPS). If $\alpha = \mathsf{EPS}^{\omega}_{s}(\varepsilon)(q)$ then, for all n,

$$\alpha = [\alpha](n) * \mathsf{EPS}^{\omega}_{s*[\alpha](n)}(\varepsilon)(q_{[\alpha](n)}). \tag{7}$$

Proof. Induction on n. See (Escardó and Oliva, 2011) for details. □

This lemma is the main building block behind the proof of the following fundamental theorem about EPS.

Theorem 3.0.15 (Main theorem on EPS). Let $q\colon X^{\mathbb{N}} \to R$ and $\omega\colon X^{\mathbb{N}} \to \mathbb{N}$ and $\varepsilon_s\colon J_R X$ be given. Define

$$\begin{aligned} \alpha &\overset{X^{\mathbb{N}}}{=} \mathsf{EPS}^{\omega}_{\langle\,\rangle}(\varepsilon)(q) \\ p_s(x) &\overset{R}{=} \overline{\mathsf{EPS}^{\omega}_{s*x}(\varepsilon)}(q_{s*x}). \end{aligned}$$

For $n \leq \omega(\alpha)$ we have

$$\begin{aligned} \alpha(n) &\overset{X}{=} \varepsilon_{[\alpha](n)}(p_{[\alpha](n)}) \\ q\alpha &\overset{R}{=} \overline{\varepsilon_{[\alpha](n)}}(p_{[\alpha](n)}). \end{aligned}$$

Proof. Assume $n \leq \omega(\alpha)$. We argue that $(*)$ $n \leq \omega([\alpha](n) * \mathbf{0})$. Otherwise, assuming $n > \omega([\alpha](n) * \mathbf{0})$ we would have, by Lemma 3.0.14, that $\alpha = [\alpha](n) * \mathbf{0}$. And hence, $n > \omega([\alpha](n) * \mathbf{0}) = \omega(\alpha) \geq n$, which is a contradiction. Hence, we have that

$$\begin{aligned} \alpha(n) &\overset{\text{L3.0.14}}{=} \mathsf{EPS}^{\omega}_{[\alpha](n)}(\varepsilon)(q_{[\alpha](n)})(0) \\ &\overset{(*)}{=} \varepsilon_{[\alpha](n)}(\lambda x.\overline{\mathsf{EPS}^{\omega}_{[\alpha](n)*x}(\varepsilon)}(q_{[\alpha](n)*x})) \\ &= \varepsilon_{[\alpha](n)}(p_{[\alpha](n)}). \end{aligned}$$

For the second identity, we have

$$\begin{aligned} q\alpha &\overset{\text{L3.0.14}}{=} q_{[\alpha](n+1)}(\mathsf{EPS}^{\omega}_{[\alpha\langle\,\rangle](n+1)}(\varepsilon)(q_{[\alpha](n+1)})) \\ &= p_{[\alpha](n)}(\alpha(n)) \\ &= \overline{\varepsilon_{[\alpha](n)}}(p_{[\alpha](n)}), \end{aligned}$$

where the last equality uses that $\alpha(n) = \varepsilon_{[\alpha](n)}(p_{[\alpha](n)})$ is already shown. □

The significance of Theorem 3.0.15 is that it shows how the product of selection functions computes a sequence α that represents, in some sense, a sequential equilibrium between the selection functions up to the point $\omega\alpha$. This kind of equilibrium appears in a variety of different contexts, most notably the following.

3.1. *Optimal strategies in sequential games*

As discussed in (Escardó and Oliva, 2011), the parameters ε, q and ω of EPS naturally define a sequential game

$$\mathcal{G}^{X,R} = (\varepsilon, q, \omega)$$

of type (X, R). We imagine X as a set of possible moves at each round, and R as a set of possible outcomes. A finite sequence $s\colon X^*$ can be thought of as a position in the game determined by the first $|s|$ moves, while an infinite sequence $\alpha\colon X^{\mathbb{N}}$ can be thought of as a play of the game. We then make the following associations:

— $\varepsilon_s\colon J_R X$ determines an optimal move at position s given that the outcome of each possible move $X \to R$ is known.
— $q\colon X^{\mathbb{N}} \to R$ determines the outcome of each play α.
— $\omega\colon X^{\mathbb{N}} \to \mathbb{N}$ determines the relevant part of a play. A position s is relevant if $\omega\hat{s} \geq |s|$.

We refer to q and ω as the outcome function and control function, respectively. In general these games can be thought of as *unbounded* games in which we only care about a finite initial segment of any play, as determined by ω. In the context of game theory Theorem 3.0.15 can be rephrased as the following.

Theorem 3.1.1. The sequence $\alpha = \mathsf{EPS}^{\omega}_{\langle\rangle}(\varepsilon)(q)$ is an *optimal play* in the game $\mathcal{G}^{X,R} = (\varepsilon, q, \omega)$.

We do not go into details on exactly what constitutes an optimal play, or how Theorem 3.1.1 is proved (for this see (Escardó and Oliva, 2010a)) but the main idea is not difficult to see. We imagine the function p_s defined in Theorem 3.0.15 as giving outcome of playing x at position s, under the assumption that all subsequent moves are played optimally, and thus $\varepsilon_s(p_s)$ is the best move at position s.

The product of selection functions carries out a backtracking algorithm and eventually computes a sequence α such that $\alpha(n) = \varepsilon_{[\alpha](n)}(p_{[\alpha](n)})$ for all $n \leq \omega\alpha$. In other words $\alpha(0)$ is the best move at position $\langle\rangle$, $\alpha(1)$ the best move at position $\langle\alpha(0)\rangle$ and more generally $\alpha(n)$ the best move at position $[\alpha](n)$ for as long as $[\alpha](n)$ is relevant. In this sense α forms an optimal play of $\mathcal{G}$. We remark that the strategy profile arising from this notion of optimal play coincides with the *Nash equilibrium* of a sequential game (see (Escardó and Oliva, 2012)).

3.2. *The dialectica interpretation of the axiom of choice*

Sequential games provide us with perhaps the most illuminating instance of the equilibrium computed by the product of selection functions. Remarkably, another instance is the dialectica interpretation of the axiom of choice.

The functional interpretation of Π^0_1-AC is equivalent to

$$\forall \varepsilon, q, \omega(\forall n, p A_n(\varepsilon_n p, p(\varepsilon_n p)) \to \exists \alpha \forall n \leq \omega \alpha A_n(\alpha n, q\alpha)).$$

It challenges us, given a collection of 'pointwise' strategies ε_n that witness the no-counterexample interpretation of the A_n, to combine them into a global strategy α that witnesses the n.c.i. of $\forall n A_n$. It is clear by Theorem 3.0.15 that the product of selection functions does the job.

Theorem 3.2.1. The functional $\lambda \varepsilon, q, \omega.\mathsf{EPS}^{\omega}_{\langle\rangle}(\varepsilon)(q)$ realises the dialectica interpretation of Π^0_1-AC.

Again, we do not go into detail, this time we refer the reader to (Escardó and Oliva, 2010a; Oliva and Powell, 2012). It can be shown more generally that EPS directly witnesses the dialectica interpretation of dependent choice for arbitrary formulas, and that a finite form of EPS with ω constant directly interprets finite choice or *bounded collection*. Moreover, EPS is primitive recursively equivalent to Spector's bar recursion (Escardó and Oliva, 2010b), and its finite form is equivalent to primitive recursion over a weak base theory (Escardó et al., 2011).

The key point we emphasise is that, as a computational analogue of choice, the product of selection functions is an extremely useful recursion schema to have at our disposal when it comes to extracting programs from proofs in both arithmetic and analysis. The fact that it also computes optimal strategies in sequential games means that extracted programs can be given an intuitive game-theoretic semantics, in the sense that we can often informally identify the "classical" dialectica interpretation A^{ND} of a theorem A with a partially defined sequential game:

$$A^{ND} \quad \sim \quad \mathcal{G}_A,$$

where a realizer for A^{ND} can given in terms of optimal strategies in $\mathcal{G}_A$. This gives the product of selection functions a clear advantage over bar recursion when interpreting theorems in analysis.

We now extract a program from the formal proof of Ramsey's theorem described in Section 2 using the product of selection functions EPS. We apply the product directly, appealing only to the main Theorem 3.0.15. The other results in this section were mentioned simply to provide some motivation as to why the product appears naturally in proof theory and why it is preferred to the more traditional modes of recursion.

4. A Constructive Proof of Ramsey's Theorem

Before launching into the full interpretation of the classical proof, it is instructive to look at the overall structure of our extracted program. Let us first look at the computational

interpretation of the final part of the classical proof – Theorem 2.0.11. Here we show that $\mathsf{RT}^2_2(c)$ follows directly from IPHP given that we have already constructed a min-monochromatic set. Suppose we have interpreted the lemma IPHP, in other words: for any $\varepsilon\colon \mathbb{B} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ and c we can construct x and p satisfying

$$\forall n \leq \varepsilon_x p(pn \geq n \wedge c(pn) = x). \tag{8}$$

Assuming that we have already (ineffectively) produced the min-monochromatic set given by a, if c^a is defined as in Theorem 2.0.11 and we set $\varepsilon^a_x p = \eta_x(a \circ p)$ (where we recall that η is a counterexample function for the finitary Ramsey's theorem as in (1)), then by (8) there exist x^a and p^a satisfying

$$\forall n \leq \eta_{x^a}(a \circ p^a)(p^a n \geq n \wedge c^a(p^a n) = x^a).$$

It is easy to see that our main theorem follows since setting $F = a \circ p^a$, for $k \leq \eta_x F$ we have

$$Fk = a(p^a k) \geq p^a k \geq k$$

and, given $Fi < Fj$

$$c(Fi, Fj) = c(a(p^a i), a(p^a j)) = c^a(p^a i) = x^a.$$

So what about a? The key observation is that we do not really need to have constructed the whole of a for the above argument to work, only a finite approximation of a is necessary. By inspection, provided that a is min-monochromatic up to

$\varphi a = \max_{i \leq \eta_{x^a}(a \circ p^a)} p^a(i)$

the claim above still holds. Therefore, if in addition we have interpreted the lemma $\mathsf{E/R}(c)$, running it on the counterexample function φ gives us a sufficiently large approximation of the min-monochromatic branch needed for an approximation of Ramsey's theorem on η. Denoting the quantifier-free matrix of the dialectica interpretation of $A \equiv \exists x \forall y A_D(x; y)$ as $|A|^x_y$, we illustrate this construction, very informally, with the inference

$$\frac{\lambda a \,.\, |\mathsf{IPHP}[a]|^{p^a, x^a}_{\varepsilon^a} \qquad |\mathsf{E/R}(c)|^a_\varphi}{|\mathsf{RT}^2_2(c)|^{a \circ p^a, x^a}_\eta}$$

making clear that the realiser for IPHP is computed relative to the parameter a. In practise this means that we run our program for $\mathsf{E/R}(c)$ once, calling on the interpretation of $\mathsf{IPHP}[a]$ each time we wish to check that a candidate a is suitable.

An entirely analogous procedure is involved, in turn, for interpreting $\mathsf{E/R}(c)$ itself. $\mathsf{E/R}(c)$ follows from WKL assuming the existence of a monotone Skolem function β making the tree T decidable. Therefore we need to calibrate exactly how much of β is required in order to successfully run the computational interpretation of WKL. As we will see, this part is rather more involved. A rough map of our whole construction is given in Figure 2.

By comparison with our proof tree in Section 2 it is clear – as expected – that the structure of the interpreted proof reflects that of the classical proof.

As mentioned in Remark 2.0.9, generalising our construction to the n-colour case

$$\dfrac{\overset{\text{L. 4.3.1}}{\lambda a\,.\,|\mathsf{IPHP}[a]|^{p^a,x^a}_{\varepsilon^a}} \qquad \dfrac{\overset{\text{Th. 4.1.4}}{\lambda\beta\,.\,|\mathsf{WKL}[\beta]|^{\alpha^\beta}_{\omega_\beta}} \qquad \overset{\text{L. 4.2.2}}{|\Pi^0_1\text{-}\mathsf{AC}|^{\beta}_{\tilde{q},\tilde{\omega}}}}{|\mathsf{E/R}(c)|^{a^{\alpha,\beta}}_{\varphi}}\ \text{Th. 4.2.3, L. 4.2.4}}{|\mathsf{RT}^2_2(c)|^{a\circ p^a,x^a}_{\eta}}\ \text{Th. 4.3.2}$$

Figure 2. Interpreted proof of Ramsey's theorem

becomes non-trivial in the construction of the min-monochromatic branch, as the E/R tree is no longer binary branching for $n > 2$ and therefore calibrating how much of β we require is a little more intricate. Also, in the n-colour case full use of IPHP would be made. That is explained in Lemma 4.3.1 below.

We now proceed with our formal program extraction. We interpret each of the main ineffective lemmas IPHP, WKL and Π^0_1-AC in turn using the product of selection functions, and combine these realisers as described above in order to produce an approximation of Ramsey's theorem. In Section 5 we discuss the aforementioned link with sequential games, and give our program a game-theoretic reading.

4.1. *Interpreting* WKL

The first ineffective step in the proof we examine is the use of weak König's lemma to produce the infinite sequence α given a Skolem function β, as in Lemma 2.0.6. We will show how to witness the no-counterexample interpretation of this lemma. As before, let T be the Σ^0_1-predicate on $\mathbb{B}^*$ defined as

$$T(s) := \exists k(\underbrace{\exists k' \in [|s|, k]\, \forall i < |s|(s_i = 0 \ \leftrightarrow\ i \prec k')}_{T'(s,k)}).$$

Let us assume we have an ideal Skolem function β satisfying

$$\forall n, k \forall s(|s| = n \wedge T'(s,k) \to T'(s, \beta n)). \tag{9}$$

Because the existence of β is ineffective, we will keep track of exactly when we call on β by highlighting it with a $\boxed{\text{box}}$. This means that we know how much of β is needed to construct an approximation of α, so that later we can in turn produce an approximation to β sufficient for the construction of α.

Recall that we use the abbreviation $T^\beta(s) = T'(s, \beta(|s|))$. The n.c.i. of Lemma 2.0.6 is as follows

$$\forall \omega^{\mathbb{B}^{\mathbb{N}} \to \mathbb{N}} \exists \alpha T^\beta([\alpha](\omega\alpha)). \tag{10}$$

Therefore, let us show how to witness α as a function of β and ω.

Lemma 4.1.1. Let β be a function satisfying (9). The tree T^β has branches of arbitrary length, i.e. for all n there exists s such that $|s| = n$ and

$$\underbrace{\exists k' \in [n, \beta n]\, \forall i < |s|(s_i = 0 \ \leftrightarrow\ i \prec k')}_{T^\beta(s)},$$

Proof. Given n define s as the sequence of length n such that, for $i < n$, $s_i = 0$ if and only if $i \prec n$. We then have $T'(s, n)$. By (9) with $\boxed{k = n}$, we can conclude $T'(s, \beta n)$. $\square$

Lemma 4.1.2. Let $\mathsf{Depth}_n(T) \equiv \exists s(|s| = n \wedge T(s))$. Let also β be a function satisfying (9), and $\varepsilon\colon J_{\mathbb{B}}\mathbb{B}$ be defined as

$$\varepsilon_s p \overset{\mathbb{B}}{=} \begin{cases} 0 & \text{if } \mathsf{Depth}_{p(0)+1}(T_s^\beta) \to \mathsf{Depth}_{p(0)}(T_{s*0}^\beta) \\ 1 & \text{otherwise.} \end{cases}$$

Then

$$\forall s, p(\underbrace{\mathsf{Depth}_{p(\varepsilon_s p)+1}(T_s^\beta)}_{(i)} \to \mathsf{Depth}_{p(\varepsilon_s p)}(T_{s*\varepsilon_s p}^\beta)). \tag{11}$$

Proof. Fix s and p and assume (i). If

$$\mathsf{Depth}_{p(0)+1}(T_s^\beta) \to \mathsf{Depth}_{p(0)}(T_{s*0}^\beta)$$

holds, then $\varepsilon_s p = 0$ and we are done. If, on the other hand,

$$\underbrace{\mathsf{Depth}_{p(0)+1}(T_s^\beta)}_{(ii)} \wedge \underbrace{\neg\mathsf{Depth}_{p(0)}(T_{s*0}^\beta)}_{(iii)}$$

holds, then $\varepsilon_s p = 1$. Hence, the assumption (i) implies (iv) $\mathsf{Depth}_{p(1)+1}(T_s^\beta)$. Now we consider two cases:

Case 1: $p(0) \geq p(1)$**.** By (ii) and (iii) we have $\mathsf{Depth}_{p(0)}(T_{s*1}^\beta)$. Therefore by (M1) we have $\exists t^{\mathbb{B}^{p(1)}} T'(s*1*t, \beta(|s|+p(0)+1))$, and, with $\boxed{n = |s| + p(1) + 1 \text{ and } k = \beta(|s| + p(0) + 1)}$, applying (9) we obtain

$$\exists t^{\mathbb{B}^{p(1)}} T'(s*1*t, \beta(|s| + p(1) + 1)) \equiv \mathsf{Depth}_{p(1)}(T_{s*1}^\beta).$$

Case 2: $p(0) < p(1)$**.** Applying (9) on $\boxed{n = |s| + p(0) + 1 \text{ and } k = \beta(|s| + p(1) + 1)}$ and (iii) we obtain

$$\forall t^{\mathbb{B}^{p(0)}} \neg T'(s*0*t, \beta(|s| + p(1) + 1)).$$

By (M1) we have

$$\forall r^{\mathbb{B}^{p(1)}} \neg T'(s*0*r, \beta(|s| + p(1) + 1)) \equiv \neg\mathsf{Depth}_{p(1)}(T_{s*0}^\beta).$$

But then by (iv) we obtain $\mathsf{Depth}_{p(1)}(T_{s*1}^\beta)$ and we are done. □

Remark 4.1.3. By inspecting the above proof we see that to verify that the selection functions ε satisfy (11) for given s, p it is sufficient that the Skolem function β satisfies (9) only up to

$$n = |s| + \max\{p(0), p(1)\} + 1 \text{ and } k = \max_{i \leq n} \beta(i).$$

In order to construct a witness for (10) we shall first build a sequence α satisfying

$$\forall k < \omega\alpha(\mathsf{Depth}_{\omega\alpha-k}(T_{[\alpha](k)}^\beta) \to \mathsf{Depth}_{\omega\alpha-k-1}(T_{[\alpha](k+1)}^\beta)). \tag{12}$$

We will then obtain (10) by a simple induction on k.

Theorem 4.1.4. Let β be a function satisfying (9), and $\omega\colon \mathbb{B}^{\mathbb{N}} \to \mathbb{N}$ be given. Define $q^\omega\alpha$ as $\omega\alpha - k - 1$ where $k < \omega\alpha$ is the least refuting (12), and 0 if no such k exists. Also, let ε be as defined in Lemma 4.1.2. The sequence

$\alpha = \mathsf{EPS}^{\omega}_{\langle\rangle}(\varepsilon)(q^{\omega})$
satisfies $T^{\beta}([\alpha](\omega\alpha))$.

Proof. By Lemma 4.1.2 we have

$$\forall s, p(\mathsf{Depth}_{p(\varepsilon_s p)+1}(T^{\beta}_{s}) \to \mathsf{Depth}_{p(\varepsilon_s p)}(T^{\beta}_{s*\varepsilon_s p})). \tag{13}$$

By Theorem 3.0.15 we have that α (as above) and $p_{[\alpha](n)}$ (as defined in Theorem 3.0.15) are such that, for $n \leq \omega\alpha$,

$$\begin{aligned} \alpha n &= \varepsilon_{[\alpha](n)} p_{[\alpha](n)} \\ q^{\omega}\alpha &= p_{[\alpha](n)}(\varepsilon_{[\alpha](n)} p_{[\alpha](n)}). \end{aligned}$$

Hence, taking $\boxed{s = [\alpha](\omega\alpha - q^{\omega}\alpha - 1) \text{ and } p = p_s}$ in (13), we obtain

$$\mathsf{Depth}_{q^{\omega}\alpha+1}(T^{\beta}_{[\alpha](\omega\alpha - q^{\omega}\alpha - 1)}) \to \mathsf{Depth}_{q^{\omega}\alpha}(T^{\beta}_{[\alpha](\omega\alpha - q^{\omega}\alpha)}). \tag{14}$$

Therefore, by the definition of q^{ω} we must have that (12) holds. If not, then there is some least $k < \omega\alpha$ refuting (12), but then (14) is equivalent to

$$\mathsf{Depth}_{\omega\alpha - k}(T^{\beta}_{[\alpha](k)}) \to \mathsf{Depth}_{\omega\alpha - k - 1}(T^{\beta}_{[\alpha](k+1)}).$$

Now, by Lemma 4.1.1 we have $\mathsf{Depth}_{\omega\alpha}(T^{\beta})$ (i.e. by taking $\boxed{n = \omega\alpha \text{ and } k = \omega\alpha}$ in (13). Hence, by induction on k, from $k = 0$ to $k = \omega\alpha - 1$, we obtain $\mathsf{Depth}_0(T^{\beta}_{[\alpha](\omega\alpha)})$, i.e. $T^{\beta}([\alpha](\omega\alpha))$. □

Theorem 4.1.4 defines a construction $\beta, \omega \mapsto \alpha^{\beta,\omega}$ that takes a Skolem function β satisfying (9) and a counterexample function ω and produces an "approximately infinite" branch α of T^{β}. But the proof above only requires the selection functions ε to satisfy (13) for the specific s, p outlined, which in turn (Remark 4.1.3) only require β to satisfy (9) for a finite number of inputs. Thus we obtain:

Corollary 4.1.5. Given β and ω, let α and p_s be constructed as in Theorem 4.1.4 and define

$$\begin{aligned} N^{\beta,\omega} &= \max\{\omega\alpha, |\omega\alpha - q^{\omega}\alpha - 1| + \max\{p_s(0), p_s(1)\} + 1\} \\ K^{\beta,\omega} &= \max\{\omega\alpha, \max_{i \leq N^{\beta,\omega}} \beta(i)\}. \end{aligned}$$

If β is an *approximate* Skolem function up to $\boxed{n = N^{\beta,\omega} \text{ and } k = K^{\beta,\omega}}$ then α (from Theorem 4.1.4) satisfies $T^{\beta}([\alpha](\omega\alpha))$.

4.2. *Interpreting* Π^0_1*-countable choice*

We have described a construction $\beta \mapsto \alpha$ which for each oracle for the Skolem function β computes an approximation to the infinite binary branch α. In Corollary 4.1.5 we argued that one only needs an approximation to β in order for our construction to work. We now show how to compute such an approximation. We first need the following lemma:

Lemma 4.2.1. Let $\delta_n \colon J_{\mathbb{N}}\mathbb{N}$ be defined as

$$\delta_n p = p^i(0) \tag{15}$$

where i is the least $\leq 2^n$ such that, for all $s^{\mathbb{B}^n}$, $T'(s, p^{i+1}(0)) \to T'(s, p^i(0))$. We have

$$\forall s^{\mathbb{B}^n} (T'(s, p(\delta_n p)) \to T'(s, \delta_n p)) \tag{16}$$

for arbitrary n, p.

Proof. Note that (16) holds by definition once we show that such $i \leq 2^n$ must exist. Assume, for the sake of a contradiction, that

(I) for all $i \leq 2^n$ there exists an $s^{\mathbb{B}^n}$ such that $T'(s, p^{i+1}(0))$ and $\neg T'(s, p^i(0))$.

By monotonicity of T' on the second argument, (I) clearly implies that

(II)$p^i(0) < p(p^i(0))$, for all $0 \leq i \leq 2^n$.

Since, in (I), we have $2^n + 1$ possible values for i but only 2^n possible values for s, there must be an s and distinct i and j, say $i < i + 1 \leq j$, such that $T'(s, p^{i+1}(0))$ and $\neg T'(s, p^j(0))$. By (II), however, that is a contradiction. □

We now show how to construct an arbitrary approximation to the Skolem function β. The next result can be viewed as the computational analogue of Lemma 2.0.4.

Lemma 4.2.2. Given arbitrary counterexample functionals $\tilde{\omega}, \tilde{q} \colon \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$, define

$$\beta = \mathsf{EPS}^{\tilde{\omega}}_{\langle\rangle}(\delta)(\tilde{q})$$

where δ is defined as in Lemma 4.2.1. Then β satisfies

$$\forall n \leq \tilde{\omega}\beta \; \forall s^{\mathbb{B}^n} (\exists k \leq \tilde{q}\beta \; T'(s, k) \to T'(s, \beta n)). \tag{17}$$

Proof. By the main theorem on EPS and (16) we obtain

$$\forall n \leq \tilde{\omega}\beta \; \forall s^{\mathbb{B}^n} (T'(s, \tilde{q}\beta) \to T'(s, \beta n)).$$

By (M2) we have $\forall k \leq \tilde{q}\beta (T'(s, k) \to T'(s, \tilde{q}\beta))$, therefore (17) follows. □

Now that we can construct approximations to β we are able to construct an approximation to an infinite branch of the Σ^0_1 tree T.

Theorem 4.2.3. For all $\omega \colon \mathbb{B}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ there exists α and β such that

$$\forall n \leq \omega\alpha\beta \underbrace{\exists k \in [n, \beta n] \, \forall i < n (\alpha(i) = 0 \;\leftrightarrow\; i \prec k)}_{T^\beta([\alpha](n))}. \tag{18}$$

Proof. Let $\beta, \omega \mapsto \alpha^{\beta,\omega}$ denote the construction defined in Theorem 4.1.4. Define

$$\tilde{\omega}\beta = N^{\beta,\omega} \quad \text{and} \quad \tilde{q}\beta = K^{\beta,\omega}$$

where $N^{\beta,\omega}$ and $K^{\beta,\omega}$ are defined as in Corollary 4.1.5. Define $\beta = \mathsf{EPS}^{\tilde{\omega}}_{\langle\rangle}(\delta)(\tilde{q})$ and $\alpha = \alpha^{\beta,\omega_\beta}$. We claim that these satisfy (18). By (17) and Corollary 4.1.5 we have $T^\beta([\alpha](\omega\alpha\beta))$. Now suppose that $n \leq \omega\alpha\beta$. Then by (17)

$$T'([\alpha](n), \beta(\omega\alpha\beta)) \to T'([\alpha](n), \beta n) \equiv T^\beta([\alpha](n)),$$

and since by (M1) $T^\beta([\alpha](\omega\alpha\beta)) \to T'([\alpha](n), \beta(\omega\alpha\beta))$ we are done. □

We are now in a position where we can construct an arbitrarily long min-monochromatic sequence a, even if the length of the sequence is determined only after we have built a, as given by ψa, as long as we are allowed to use ψ in the construction of a.

Lemma 4.2.4. For any ψ there exists a function $a\colon \mathbb{N} \to \mathbb{N}$ such that for all $n \leq \psi a$

$$(n \leq an) \wedge \forall i, j, k{<}n(ak < ai \wedge ak < aj \to c(ak, ai) = c(ak, aj)). \tag{19}$$

Proof. First, define a parametrised $a^{\alpha,\beta}$ as in Lemma 2.0.7:

$$\begin{aligned} a^{\alpha,\beta}0 &:= 0 \\ a^{\alpha,\beta}(n+1) &:= \mu k \in [n, \beta(\beta n + 1)]\,(\alpha k = 0)\,. \end{aligned} \tag{20}$$

Then, take (cf. Remark 2.0.8)

$\omega\alpha\beta = \max_{i \leq \psi(a^{\alpha,\beta})}(\max\{i, \beta i + 1, \beta(\beta i + 1) + 1\})$

and let α and β be as the Theorem 4.2.3. It is easy, following the same proof as in Lemma 2.0.7, to check that $a = a^{\alpha,\beta}$ satisfies (19). □

Remark 4.2.5. For the n-colour case, the construction of a is more complicated (cf. Remark 2.0.9) and ω will need to demand a larger approximation to β.

4.3. *Final arguments and* IPHP

Finally, the last non-constructive step in the proof is the use of the infinite pigeon-hole principle. Note that we in fact only make use of a particular instance of IPHP, namely $n = 2$. Nevertheless, we refer to the general IPHP so it is easier to see how our construction can be generalised for arbitrarily many colours.

Lemma 4.3.1. We have

$$\forall \varepsilon^{\mathbb{B}\times\mathbb{N}^{\mathbb{N}}\to\mathbb{N}} \exists x^{\mathbb{B}}, p^{\mathbb{N}^{\mathbb{N}}} \forall i \leq \varepsilon_x p(pi \geq i \wedge c(pi) = x).$$

Proof. Given ε_x define

$\tilde{\varepsilon}_x p = \mu i \leq \varepsilon_x p \neg(pi \geq i \wedge c(pi) = x).$

Then let $(a_0, a_1) = (\tilde{\varepsilon}_0 \otimes \tilde{\varepsilon}_1)(\max)$ and $N = \max\{a_0, a_1\}$. By the main theorem on the product of selection functions we have p_0 and p_1 such that

$$a_0 = \tilde{\varepsilon}_0 p_0 \qquad a_1 = \tilde{\varepsilon}_1 p_1 \qquad N = p_0(a_0) = p_1(a_1).$$

Let $x = c(N)$ and $p = p_x$. Clearly, $p(\tilde{\varepsilon}_x p_x) = pa_x = p_x a_x = N \geq a_x$. Moreover, $c(p(\tilde{\varepsilon}_x p_x)) = c(pa_x) = c(N) = x$. Hence, by the definition of $\tilde{\varepsilon}_x$ we must have

$$\forall i \leq \varepsilon_x p(pi \geq i \wedge c(pi) = x).$$

Note that essentially the same proof works for the n-colour case, where we have n selection functions $\tilde{\varepsilon}_0, \ldots, \tilde{\varepsilon}_{n-1}$ accounting for each colour, and we take the finite product $(\tilde{\varepsilon}_0 \otimes \ldots \otimes \tilde{\varepsilon}_{n-1})(\max)$. □

The theorem then follows by combining the construction of the min-monochromatic sequence with an application of IPHP.

Theorem 4.3.2. Let a colouring $c\colon \mathbb{N}^2 \to \mathbb{B}$ be fixed. For any pair of selection functions $\eta_x\colon J_{\mathbb{N}}\mathbb{N}$ there exists $F\colon \mathbb{N} \to \mathbb{N}$ and $x\colon \mathbb{B}$ (explicitly given in Section 4.4) such that

$$\forall k \leq \eta_x F(k \leq Fk \wedge \forall i, j \leq k(Fi < Fj \to c(Fi, Fj) = x)).$$

Proof. Assume $c\colon [\mathbb{N}]^2 \to \mathbb{B}$ and $\eta_0\colon J_{\mathbb{N}}\mathbb{N}$ and $\eta_1\colon J_{\mathbb{N}}\mathbb{N}$ are given. For any function a let $c^a(i) = c(a(i), a(i+1))$. Let $\varepsilon^a_x p = \eta_x(a \circ p)$, with $a\colon \mathbb{N} \to \mathbb{N}$ as a parameter. By Lemma 4.3.1 we have that there exists p^a and x^a such that

$$\forall i < \eta_{x^a}(a \circ p^a)(p^a(i) \geq i \wedge c^a(p^a(i)) = x^a). \tag{21}$$

Let $\psi a = \max_{i \leq p^a(\eta_{x^a}(a \circ p^a))} p^a(i)$. By Lemma 4.2.4 there exists an $a\colon \mathbb{N} \to \mathbb{N}$ such that for all $n \leq p^a(\eta_{x^a}(a \circ p^a))$ we have $an \geq n$ and

$$\forall i, j, k < n(ak < ai \wedge ak < aj \to c(ak, ai) = c(ak, aj)). \tag{22}$$

Take $F = a \circ p^a$ and $x = x^a$. Therefore, for $k \leq \eta_x F = \eta_x(a \circ p^a)$ we have

— $p^a k \geq k$ by (21) which, by the above implies that

$$Fk = a(p^a k) \geq p^a k \geq k.$$

— and, for $i, j \leq k$, given that $Fi < Fj$, we have

$$x \overset{(21)}{=} c^a(p^a(i)) = c(a(p^a(i)), a(p^a(i+1))).$$

Hence

$$c(\underbrace{a(p^a i)}_{Fi}, \underbrace{a(p^a j)}_{Fj}) \overset{(22)}{=} c(a(p^a i), a(p^a(i)+1)) \overset{(21)}{=} x.$$

□

4.4. *Summarising the construction of x and F*

From the proof of Theorem 4.3.2 we can read off the construction of F and x which we summarise here. Recall that the input to our problem is a colouring $c\colon \mathbb{N}^2 \to \mathbb{B}$ and a pair of selection functions $\eta_x\colon J_{\mathbb{N}}\mathbb{N}$. Also, recall the abbreviations

$$\begin{aligned} T'(s,k) &\equiv \exists k' \in [|s|, k]\, \forall i < |s|(s_i = 0 \leftrightarrow i \prec k')) \\ T^\beta(s) &\equiv T'(s, \beta(|s|)) \\ \mathsf{Depth}_n(T^\beta_s) &\equiv \exists t(|t| = n \wedge T^\beta(s * t)). \end{aligned}$$

(A) **Construction of x and F given** $a\colon \mathbb{N}^{\mathbb{N}}$. First, assume a function $a\colon \mathbb{N}^{\mathbb{N}}$ given and let $c^a(i) = c(a(i), a(i+1))$ and $\varepsilon^a_x p = \eta_x(a \circ p)$. Define

$$\tilde{\varepsilon}_x p = \mu i \leq \varepsilon^a_x p \neg(pi \geq i \wedge c^a(pi) = x).$$

Take $(k_0, k_1) = (\tilde{\varepsilon}_0 \otimes \tilde{\varepsilon}_1)(\max)$ and $x^a = c(\max\{k_0, k_1\})$ and

$$p^a(k) = \begin{cases} \tilde{\varepsilon}_1(\lambda k'.\max\{k, k'\}) & \text{if } x^a = 0 \\ \max\{k_0, k\} & \text{if } x^a = 1. \end{cases}$$

and $F^a = a \circ p^a$.

(B) **Construction of α given $\beta\colon \mathbb{N}^{\mathbb{N}}$ and $\omega\colon \mathbb{B}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$.** Then, we construct a sequence $\alpha^{\beta,\omega}\colon \mathbb{B}^{\mathbb{N}}$ parametrised by $\beta\colon \mathbb{N}^{\mathbb{N}}$ and $\omega\colon \mathbb{B}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ as follows. Let

$$q^{\beta,\omega}\alpha = \omega\alpha\beta - k - 1,$$

where $k < \omega\alpha\beta$ is the least refuting

$$\forall k < \omega\alpha\beta(\mathsf{Depth}_{\omega\alpha\beta-k}(T^{\beta}_{[\alpha](k)}) \to \mathsf{Depth}_{\omega\alpha\beta-k-1}(T^{\beta}_{[\alpha](k+1)})),$$

and

$$\varepsilon^{\beta}_{s}p \stackrel{\mathbb{B}}{=} \begin{cases} 0 & \text{if } \mathsf{Depth}_{p(0)+1}(T^{\beta}_{s}) \to \mathsf{Depth}_{p(0)}(T^{\beta}_{s*0}) \\ 1 & \text{otherwise.} \end{cases}$$

Define

$$\alpha^{\beta,\omega} = \mathsf{EPS}^{\lambda\alpha.\omega\alpha\beta}_{\langle\,\rangle}(\varepsilon^{\beta})(q^{\beta,\omega}).$$

(C) **Construction of β given $\omega\colon \mathbb{B}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ using (B).** Using $\alpha^{\beta,\omega}$ we construct a sequence $\beta^{\omega}\colon \mathbb{N}^{\mathbb{N}}$ parametrised by $\omega\colon \mathbb{B}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}$ only. Let $\delta_n\colon J_{\mathbb{N}}\mathbb{N}$ be

$$\delta_n p = p^i(0)$$

where i is the least $\leq 2^n$ such that, for all $s^{\mathbb{B}^n}$, $T'(s, p^{i+1}(0)) \to T'(s, p^i(0))$, and

$$\begin{aligned} \tilde{\omega}\beta &= \max\{\omega\alpha^{\beta,\omega}\beta, |\omega\alpha^{\beta,\omega}\beta - q^{\beta,\omega}\alpha^{\beta,\omega} - 1| + \max\{p(0), p(1)\} + 1\} \\ \tilde{q}\beta &= \max\{\omega\alpha^{\beta,\omega}\beta, \max_{i \leq \tilde{\omega}\beta}\beta(i)\}, \end{aligned}$$

where

$$\begin{aligned} p(x) &= \overline{\mathsf{EPS}^{\lambda\alpha.\omega\alpha\beta}_{s*x}(\varepsilon^{\beta})}(q^{\beta,\omega}_{s*x}) \\ s &= [\alpha^{\beta,\omega}](\omega\alpha^{\beta,\omega}\beta - q^{\beta,\omega}\alpha^{\beta,\omega} - 1). \end{aligned}$$

Define

$$\beta^{\omega} = \mathsf{EPS}^{\tilde{\omega}}_{\langle\rangle}(\delta)(\tilde{q}).$$

(D) **Construction of ω using (A).** We now construct the missing ω as

$$\omega\alpha\beta = \max_{i \leq \psi(a^{\alpha,\beta})}(\max\{i, \beta i + 1, \beta(\beta i + 1) + 1\})$$

where $\psi a = \max_{i \leq p^a(\eta_{x^a}(a \circ p^a))} p^a(i)$, with p^a and x^a as defined in (A), and

$$a^{\alpha,\beta}n = \begin{cases} 0 & \text{if } n = 0 \\ \mu k \in [n-1, \beta(\beta(n-1)+1)]\,(\alpha k = 0) & \text{if } n > 0. \end{cases}$$

(E) **Construction of x and F using (A) − (D).** Finally, take $\beta = \beta^{\omega}$ and $\alpha = \alpha^{\beta,\omega}$ and $a = a^{\alpha,\beta}$, so that x and F are defined as $x = x^a$ and $F = a \circ p^a$.

5. A Game-Theoretic Reading of the Proof

Following the discussion in Section 3, we know that each instance of EPS used in our finitisation of Ramsey's theorem corresponds to the computation of an optimal strategy in a partially defined[§] game. We now discuss the specific games corresponding to the main instances of EPS used in our extracted program, and show how our constructive proof Ramsey's theorem can be understood in game-theoretic terms.

[§] We call a game $\mathcal{G}$ partially defined when not all three parameters ε, q and ω are given, and write the open parameters in square brackets e.g. $\mathcal{G}[\varepsilon]$.

Π^0_1 *countable choice:* $\mathcal{G}^{\mathbb{N},\mathbb{N}}_{\Pi^0_1\text{-}\mathsf{AC}}[\tilde{q},\tilde{\omega}] = (\delta,\tilde{q},\tilde{\omega})$ The game central to our interpretation is that corresponding to our use of countable choice. The selection functions δ_n defined in Lemma 4.2.1 implement a 'no new branches' strategy, picking a number $i = \delta_n p$ satisfying

$$\forall s^{\mathbb{B}^n}(T'(s,p(i)) \to T'(s,i))$$

i.e. there are no branches s of T which have a witness bounded by the outcome $p(i)$ which is not already bounded by the move i.

For any outcome function $\tilde{q}$ and control function $\tilde{\omega}$, an optimal strategy in this case is a sequence β satisfying, for all $n \leq \tilde{\omega}\beta$,

$$\forall s^{\mathbb{B}^n}(T'(s,\tilde{q}\beta) \to T'(s,\beta n)).$$

This means that every move βn in the play β (for $n \leq \tilde{\omega}\beta$) already bounds a witness for any branch s of length n which has a witness bounded by the final outcome $\tilde{q}\beta$. This optimal strategy is precisely the approximation to a monotone Skolem function we require.

Weak König's lemma: $\mathcal{G}^{\mathbb{B},\mathbb{N}}_{\mathsf{WKL}}[\omega] = (\varepsilon, q^\omega, \omega)$ The interpretation of WKL applied to the decidable tree T^β is interpreted by a binary game (where the set of possible moves at each round is $\mathbb{B}$). The strategy ε_s at position s defined by the selection functions given in Lemma 4.1.2 is to pick a boolean b such that if s extends to a branch in T of length $|s| + p(b) + 1$ then $s * b$ also extends to a branch of length $|s| + p(b) + 1$.

Given ω, by choosing q^ω suitably as in Theorem 4.1.4, the optimal strategy of $\mathcal{G}_{\mathsf{WKL}}$ determined by these selection functions is a sequence α such that for all $k \leq \omega\alpha$, whenever $[\alpha](k)$ extends to a branch of length $\omega\alpha$, so does $[\alpha](k+1)$. If T^β is infinite then $\langle\rangle$ extends to a branch of length $\omega\alpha$. Hence, by induction the relevant part $[\alpha](\omega\alpha)$ of this optimal play must be in T^β, and is therefore an approximation to an infinite branch.

The infinite pigeonhole principle: $\mathcal{G}^{\mathbb{N},\mathbb{N}}_{\mathsf{IPHP}}[\varepsilon] = (\tilde{\varepsilon}, \max, 2)$ The game corresponding to IPHP is a *finite* game with two rounds (or n rounds for the n-colour Ramsey's theorem). The strategy $\tilde{\varepsilon}$ at each round $x = 0, 1$ is to play the least move $i \leq \varepsilon_x p$ the outcome $p(i)$ of which satisfies

$$p(i) < i \vee c(p(i)) \neq x.$$

We compute the optimal play $\langle a_0, a_1\rangle$, and its outcome is the maximum $N = \max\{a_0, a_1\}$. But then, at round $x = c(N)$ we have

$$p_x(a_x) \geq a_x \wedge c(p_x(a_x)) = x$$

since $p_x(a_x) = N$, which implies that the selection function $\tilde{\varepsilon}_x$ must fail to find a suitable candidate. But since we know that an optimal strategy must exist, the only explanation is that such a candidate does not exist, or in other words, x, p_x form an approximation to the infinite pigeonhole principle.

Following the discussion at the beginning of the section, it is not too hard to visualise how these games combine to witness the functional interpretation of Ramsey's theorem.

We compute an optimal strategy β in the game

$$\mathcal{G}_{\Pi^0_1\text{-}\mathsf{AC}}[\lambda\beta.K^{\beta,\omega}, \lambda\beta.N^{\beta,\omega}]$$

where the outcome and control functions involve computing an optimal strategy α^β in the auxiliary game

$$\mathcal{G}_{\mathsf{WKL}}[\omega_\beta]$$

on T^β. As a result we obtain two optimal strategies β, α^β that combine to form an approximation $a^{\alpha,\beta}$ to a min-monochromatic branch.

In addition, the control function ω_β is defined in terms of $\varphi a^{\alpha,\beta}$, which in turn involves computing an optimal strategy in a *further* auxiliary game

$$\mathcal{G}_{\mathsf{IPHP}}[\lambda x, p.\eta_x(a^{\alpha,\beta} \circ p)]$$

where η is our counterexample function for $\mathsf{RT}^2_2(c)$, in order to produce x^a, p^a required to compute φa.

Therefore our program can be viewed in terms of the computational of optimal strategies in three symbiotic games: one central game corresponding to Π^0_1-AC and two nested auxiliary games that are run each time we call on the relevant counterexample functions.

The computation as a whole returns an optimal strategy β of $\mathcal{G}_{\Pi^0_1\text{-}\mathsf{AC}}$ and an optimal strategy α^β of $\mathcal{G}_{\mathsf{WKL}}$ that combine to form a sequence $a^{\alpha,\beta}$, along with p^a, x^a arising from optimal strategy in $\mathcal{G}_{\mathsf{IPHP}}$. Our realiser for the functional interpretation of Ramsey's theorem $F = a^{\alpha,\beta} \circ p^{a^{\alpha,\beta}}$ and $x = x^{a^{\alpha,\beta}}$ can therefore be written in terms of optimal strategies in these three games.

Acknowledgements. The authors thank the two anonymous referees for their helpful comments and suggestions on the preliminary version of this paper. The first author gratefully acknowledges support of The Royal Society under grant 516002.K501/RH/kk, and the second author acknowledges the support of an EPSRC doctoral training grant.

References

Avigad, J. (2009). The metamathematics of ergodic theory. *Annals of Pure and Applied Logic*, 157:64–76.

Avigad, J. and Feferman, S. (1998). Gödel's functional ("Dialectica") interpretation. In Buss, S. R., editor, *Handbook of proof theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 337–405. North Holland, Amsterdam.

Bellin, G. (1990). Ramsey interpreted: a parametric version of Ramsey's theorem. In *Logic and computation (Pittsburgh, PA, 1987)*, volume 106, pages 17–37. Amer. Math. Soc., Providence, RI.

Coquand, T. (1994a). An analysis of Ramsey's theorem. *Information and Computation*, 110(2):297–304.

Coquand, T. (1994b). A direct proof of Ramsey's theorem. Author's website.

Erdős, P., Hajnal, A., Máté, A., and Rado, R. (1984). *Combinatorial Set Theory: Partition Relations for Cardinals*, volume 106 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Company.

Escardó, M. H. and Oliva, P. (2010a). Computational interpretations of analysis via products of selection functions. In Ferreira, F., Lowe, B., Mayordomo, E., and Gomes, L. M., editors, *Computability in Europe 2010, LNCS 6158*, pages 141–150. Springer.

Escardó, M. H. and Oliva, P. (2010b). Selection functions, bar recursion, and backward induction. *Mathematical Structures in Computer Science*, 20(2):127–168.

Escardó, M. H. and Oliva, P. (2011). Sequential games and optimal strategies. *Royal Society Proceedings A*, 467:1519–1545.

Escardó, M. H. and Oliva, P. (2012). Computing Nash equilibria of unbounded games. In *Proceedings of the Turing 100 Conference*. Turing 100, Manchester.

Escardó, M. H., Oliva, P., and Powell, T. (2011). System T and the product of selection functions. *Proceedings of CSL'11*, LIPIcs 12:233–247.

Gödel, K. (1958). Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287.

Kohlenbach, U. (2008). *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Monographs in Mathematics. Springer.

Kohlenbach, U. and Kreuzer, A. (2009). Ramsey's theorem for pairs and provably recursive functions. *Notre Dame Journal of Formal Logic*, 50:427–444.

Kreisel, G. (1951). On the interpretation of non-finitist proofs, part I. *The Journal of Symbolic Logic*, 16:241–267.

Kreisel, G. (1952). On the interpretation of non-finitist proofs, part II: Interpretation of number theory. *The Journal of Symbolic Logic*, 17:43–58.

Kreuzer, A. (February 2009). Der Satz von Ramsey für Paare und beweisbar rekursive Funktionen. Diploma thesis.

Oliva, P. and Powell, T. (2012). A game-theoretic computational interpretation of proofs in classical analysis. Preprint, available online at http://arxiv.org/abs/1204.5244.

Ramsey, F. (1930). On a problem of formal logic. *Proceedings of the London Mathematical Society*, s2-30(1):264–286.

Simpson, S. G. (1999). *Subsystems of Second Order Arithmetic*. Perspectives in Mathematical Logic. Springer, Berlin.

Spector, C. (1962). Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. In Dekker, F. D. E., editor, *Recursive Function Theory: Proc. Symposia in Pure Mathematics*, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island.

Tao, T. (2008). *Structure and Randomness: pages from year one of a mathematical blog*. American Mathematical Society.

Veldman, W. and Bezem, M. (1993). Ramsey's theorem and the pigeonhole principle in intuitionistic mathematics. *Journal of the London Mathematical Society*, 2(47):193–211.