TERM EXTRACTION AND RAMSEY’S THEOREM FOR PAIRS

ALEXANDER P. KREUZER AND ULRICH KOLLENBACH

Abstract. In this paper we study with proof-theoretic methods the functionals provably recursive relative to Ramsey’s theorem for pairs and the cohesive principle (COH).

Our main result on COH is that the type $2$ functionals provably recursive from $\text{RCA}_0 + \text{COH} + \Pi^1_0$-$\text{CP}$ are primitive recursive. This also provides a uniform method to extract bounds from proofs that use these principles. As a consequence we obtain a new proof of the fact that $\text{WKL}_0 + \Pi^1_0$-$\text{CP} + \text{COH}$ is $\Pi^0_2$-conservative over $\text{PRA}$.

Recent work of the first author showed that $\Pi^0_1$-$\text{CP} + \text{COH}$ is equivalent to a weak variant of the Bolzano-Weierstraß principle. This makes it possible to use our results to analyze not only combinatorial but also analytical proofs.

For Ramsey’s theorem for pairs and two colors ($\text{RT}^2_2$) we obtain the upper bounded that the type $2$ functionals provable recursive relative to $\text{RCA}_0 + \Sigma^0_2$-$\text{IA} + \text{RT}^2_2$ are in $T_1$. This is the fragment of Gödel’s system $T$ containing only type $1$ recursion — roughly speaking it consists of functions of Ackermann type. With this we also obtain a uniform method for the extraction of $T_1$-bounds from proofs that use $\text{RT}^2_2$. Moreover, this yields a new proof of the fact that $\text{WKL}_0 + \Sigma^0_2$-$\text{IA} + \text{RT}^2_2$ is $\Pi^0_1$-conservative over $\text{RCA}_0 + \Sigma^0_1$-$\text{IA}$.

The results are obtained in two steps: in the first step a term including Skolem functions for the above principles is extracted from a given proof. This is done using Gödel’s functional interpretation. After this the term is normalized, such that only specific instances of the Skolem functions are used. In the second step this term is interpreted using $\Pi^0_1$-comprehension. The comprehension is then eliminated in favor of induction using either elimination of monotone Skolem functions (for COH) or Howard’s ordinal analysis of bar recursion (for $\text{RT}^2_2$).

1. Introduction

The aim of this paper is to develop a technique of program extraction for proofs that use Ramsey’s theorem for pairs, the cohesive principle and other principle weaker than Ramsey’s theorem for pairs. As a consequence it also gives a proof theoretic account of conservation results for those principles. This paper extends our previous treatment of Ramsey’s theorem for pairs in [34], where only single instances of Ramsey’s theorem are discussed, to the full second order closure of those principles.

*Ramsey’s theorem for pairs* ($\text{RT}^2_2$) is the statement that every coloring of pairs of natural numbers ($[\mathbb{N}]^2$) with $n$ colors has an infinite homogeneous set. A simple colorblindness argument shows that

$$\text{RT}^2_2 \leftrightarrow \text{RT}^2_n$$

for every fixed $n$. 


2010 Mathematics Subject Classification. Primary: 03F10, 03F35, 03B30; Secondary: 05D10.

Key words and phrases. Ramsey’s theorem, cohesive principle, proof mining, functional interpretation, conservation.

The authors gratefully acknowledge the support by the German Science Foundation (DFG Project KO 1737/5-1).
Ramsey’s theorem for pairs and arbitrary large colorings ($\RT_n^2 \leq \infty$) is defined as $\forall n \RT_n^2$. This principle is proof-theoretically stronger than $\RT_2^2$.

A coloring $c$ of pairs is called stable if $c(\{x, y\})$ eventually becomes constant for every $x$. The restriction of $\RT_n^2$ to stable colorings is denoted by $\SRT_n^{\geq \infty}$. Here a similar colorblindness argument can be applied.

A set $G$ is called cohesive for a sequence $(R_i)_{i \in \mathbb{N}}$ of subsets of $\mathbb{N}$ if

$$\forall i \, (G \subseteq^* R_i \lor G \subseteq^* \overline{R_i}),$$

where $X \subseteq^* Y :\equiv (X \setminus Y$ is finite). The cohesive principle (COH) states that for every $(R_i)_{i \in \mathbb{N}}$ an infinite cohesive set exists. It is in some way the counterpart to $\SRT_n^{\geq \infty}$ since

$$\text{RCA}_0 \vdash \RT_n^2 \iff \SRT_n^{\geq \infty} \land \text{COH}$$

for $2 \leq n \leq \infty$, see [7, 8].

The computational strength of Ramsey’s theorem has been investigated since the early 70’s. Specker showed 1971 that there exists a computable coloring of $[\mathbb{N}]^2$ that has no computable homogeneous set, see [47]. Jockusch improved this 1972 by showing that in general there is not even a $\Sigma^0_2$ infinite homogeneous set. He also provided an upper bound on the strength of Ramsey’s theorem for pairs and showed that each computable coloring of pairs admits an infinite homogeneous set $H$ with $H' \leq_T \emptyset'$, see [21]. Seetapun and Slaman showed in [43] that $\RT_2^2$ does not solve the halting problem. Cholak, Jockusch and Slaman improved both results by showing that an infinite homogeneous $low_2$ set exists for every computable coloring of pairs, i.e. a set $H$ satisfying $H'' \leq_T \emptyset''$, see [7].

From Specker’s results it is clear that $\text{RCA}_0 \not\vdash \RT_2^2$. Seetapun’s and Slaman’s results immediately yield an upper bound on the proof-theoretic strength, it implies that $\RT_2^2$ does not prove $\Pi^1_1$-comprehension (or, equivalently, $\text{ACA}_0$). Hirst showed 1987 that $\RT_2^2$ implies the infinite pigeonhole principle ($\RT_{\leq \infty}^1$) which is equivalent to the $\Pi^1_1$-bounded collection principle ($\Pi^1_1$-$\text{CP}$)\footnote{In the first order context this principle is usually denoted by $B\Pi^0_1$ which is equivalent to $B\Sigma^0_2$.}, see [16]. Cholak, Jockusch and Slaman showed along their recursion theoretic proof that $\RT_2^2$ is $\Pi^1_1$-conservative over $\text{RCA}_0 + \Sigma^0_2$-$\text{IA}$. This leaves the question whether $\RT_2^2$ implies $\Sigma^0_2$-$\text{IA}$. Despite of many efforts in the last years this question has not been settled yet.

Ramsey’s theorem for triples and bigger tuples is equivalent to $\text{ACA}_0$ and hence fully classified in the sense of reverse mathematics, see [45].

The cohesive principle has been originally considered in recursion theory, see for instance [46]. Its computational strength has been fully determined in [19]. Cholak, Jockusch and Slaman observed in [7] that Ramsey’s theorem for pairs splits nicely into a stable part and the cohesive principle. They also showed that it is $\Pi^1_1$-conservative over $\text{ACA}_0$ and $\text{RCA}_0 + \Sigma^0_2$-$\text{IA}$. In the course of the classification of Ramsey’s theorem the cohesive principle’s logical strength received attention in the last years, see for instance [10] and [9]. In [9] it was shown that the cohesive principle is $\Pi^1_1$-conservative over $\text{RCA}_0 + \Pi^0_1$-$\text{CP}$. We recently showed that over $\text{RCA}_0 + \Pi^0_1$-$\text{CP}$ the cohesive principle is equivalent to a weak form of the Bolzano-Weierstraß principle, see [37]. Thus the cohesive principle also shows up in analytic proofs.

For an extensive survey on the current status of Ramsey’s theorem for pairs and weaker principles, see [15] and [44].

The purpose of this paper is to give an account to the above mentioned conservation results from the perspective of proof mining and program extraction. We provide new proofs for these conservation results which additionally yield realizing...
terms. Since the types of these terms rise with the complexity of the formula this
is naturally bounded to \( \Pi^0_3 \)-sentences.

**Proofwise low.** Define \( \Pi^1_1 \)-comprehension as
\[
(\Pi^1_1\text{-CA}): \forall X \exists Y \forall u \ (u \in Y \leftrightarrow \forall v \langle u, v \rangle \in X).
\]
This covers the full strength of \( \Pi^1_0 \)-comprehension since \( \forall v \langle u, v \rangle \in X \) is a universal \( \Pi^1_1 \)-statement (relative to the parameter \( u \)). Full arithmetical comprehension (\( \text{ACA}_0 \)) follows by iteration. For a primitive recursive term \( t \) we will write \( \Pi^1_1\text{-CA}(t) \) if \( X \) is instantiated with the set \( \{ n \mid t(n) = 0 \} \).\(^2\) For a closed term \( t \) the principle \( \Pi^1_1\text{-CA}(t) \) is also called an instance of \( \Pi^1_1 \)-comprehension.

The union of \( \Pi^1_1\text{-CA}(t) \) for all terms \( t \) containing only number variables free is the same as light-face \( \Pi^1_0 \)-comprehension. In particular, this does not prove \( \text{ACA}_0 \).

Let \( \mathcal{P} \) be a second order principle stating the existence of a set \( G \) relative to a set parameter \( S \) — that is a principle of the form
\[
(\mathcal{P}): \forall S \exists G \left( P(S, G) \land \Pi^0_1\text{-CA}(\varphi SG) \right).
\]

**Definition 1** (proofwise low). Call a principle of the form \( \mathcal{P} \) proofwise low over a system \( T \) if for every provably continuous\(^3\) term \( \varphi \) a provably continuous term \( \xi \) exists such that
\[
(1) \quad T \vdash \forall S \left( \Pi^1_1\text{-CA}(\xi S) \rightarrow \exists G \left( P(S, G) \land \Pi^0_1\text{-CA}(\varphi SG) \right) \right).
\]
If we additionally can prove this for a sequence of solutions, i.e.
\[
(2) \quad T \vdash \forall (S_i)_{i \in \mathbb{N}} \left( \Pi^1_1\text{-CA}(\xi (S_i)) \rightarrow \exists (G_i)_{i \in \mathbb{N}} \left( \forall i P(S_i, G_i) \land \Pi^0_1\text{-CA}(\varphi (S_i), (G_i), i) \right) \right)
\]
then we call \( \mathcal{P} \) proofwise low in sequence over the system \( T \). Here \( (S_i) \) is (a code of) the sequence of sets \( S_i \). It is given by the set \( \{ (i, x) \mid x \in S_i \} \).

The notion of proofwise low is comparable to \( \text{low}_2 \) in the recursion theoretic setting: take for instance \( T = \text{WKL}_0 \), then a proofwise low statement in \( T \) satisfies
\[
\text{RCA}_0 \vdash \forall S \left( \text{WKL} \land \Pi^0_1\text{-CA}(\xi S) \rightarrow \exists G \left( P(S, G) \land \Pi^0_1\text{-CA}(\varphi SG) \right) \right).
\]
The analogous recursion theoretic statement would be that relative to an oracle of Turing degree \( d \gg 0' \) — this resembles the premise — a set \( G \) satisfying the statement \( P(S, G) \) and its Turing jump \( G' \) can be computed. From this follows that \( G' \equiv_T 0'' \) or in other word that \( G \) is \( \text{low}_2 \).

The main results of this paper are divided into two parts:

(i) We show roughly that
- \( \text{RT}_2 \) is proofwise low over \( \text{WKL}_0 \) (Corollary 46) and that
- \( \text{COH} \) is proofwise low in sequence over \( \text{WKL}_0^* \). The system \( \text{WKL}_0^* \) is defined to be \( \text{WKL}_0 \) where \( \Sigma^1_2 \)-induction is replaced by quantifier-free-induction plus the exponential function. (Corollary 33)

(ii) We show for principles \( \mathcal{P} \) that
- if \( P(S, G) \) is \( \Pi^1_1 \) and \( \mathcal{P} \) is proofwise low over \( \text{WKL}_0 \), the system \( \text{WKL}_0 + \Sigma^1_2\text{-IA} + \mathcal{P} \) is \( \Pi^1_1 \)-conservative over \( \Sigma^1_2 \)-induction. (Section 10.3)
- if \( P(S, G) \) is \( \Pi^1_3 \) and \( \mathcal{P} \) is proofwise low in sequence over \( \text{WKL}_0^* \) the system \( \text{WKL}_0 + \Pi^1_1\text{-CP} + \mathcal{P} \) is \( \Pi^1_2 \)-conservative over \( \text{RCA}_0 \) and \( \Pi^1_2 \)-consistent over \( \text{PRA} \). (This covers \( \text{COH} \). See Theorem 36.)

\(^2\)Strictly speaking \( \text{RCA}_0 \) does not contain terms. Here and in the following we silently assume that we work in the conservative extension of \( \text{RCA}_0 \) by all primitive recursive functions.

\(^3\)Continuous means here continuous in the sense of Baire space, i.e. \( \varphi \) is continuous if
\[
\forall f \exists n (\forall x < n f(x) = g(x) \rightarrow \varphi(f) = \varphi(g)).
\]
Such functionals can be coded into primitive recursive functions. For details see definitions 6 and 7 below.
This simplifies the results slightly. The actual results require a suitable finite type extension of $\text{WKL}_0$ and $\text{WKL}_0^∗$, they also allow a function parameter to the $\Pi^0_0$-formula and provide extraction of type 2 functionals, see below.

The first part of the results is based on the proofs by “first jump control” for $\text{SRT}^2_2$ and $\text{COH}$ of Cholak, Jockusch and Slaman, see [7], showing that these principles have low$_2$ solutions. (See proposition 31 with corollary 33 and proposition 44 with corollary 46.) To our knowledge these proofs have not been used before to obtain conservativity results for $\text{RT}^2_2$. Cholak, Jockusch and Slaman developed in this paper a different, more complicated proof needing $\Pi^0_2$-comprehension that can be used in a forcing construction to show conservativity of $\text{RT}^2_2$ over $\Sigma^0_2$-induction.

For the second part we use Gödel’s functional interpretation (always combined with a negative translation) to extract a term $t$ from a proof of an arbitrary statement of the following form

$$\mathcal{P} \rightarrow \forall x \exists y A(x, y),$$

where $A$ is quantifier-free and $\mathcal{P}$ is a proofwise low principle. (See the proof of proposition 35 and proposition 50.) For an oracle solution $\mathcal{P}$ of the functional interpretation of $\mathcal{P}$ this term will then satisfy

$$\forall x A(x, t(\mathcal{P}, x)).$$

We normalize $t$ so that every application of $\mathcal{P}$ in the proof is of a specific form and one can read off from the term and the proof how much of $\mathcal{P}$ is used (section 8). The functional $\mathcal{P}$ is then eliminated from $t$ by interpreting every specific application of $\mathcal{P}$. This is done either by (2) or the functional interpretation of (1) in a way that retains the instance of comprehension. If this retained instance of comprehension is used for the next interpretation of $\mathcal{P}$ then an inductive treatment of every application of $\mathcal{P}$ yields that

(i) in the first case one instance of the functional interpretation of $\Pi^0_1$-CA suffices to prove to totality of $t$ and hence $\forall x \exists y A(x, y)$, see proposition 52,

(ii) in the second case one instance of $\Pi^0_1$-CA proves the totality of $t$ and hence $\forall x \exists y A(x, y)$, see proposition 35.

The instance of comprehension is then eliminated in favor of induction:

In the case (i) the solution to this functional interpreted instance of comprehension is provided by an instance of Spector’s bar recursion (in fact by an application of the rule of bar recursion). This usage of bar recursion is then eliminated using Howard’s ordinal analysis of bar recursion in favor of $\Sigma^0_2$-induction, see section 7.

In the case (ii) the instance of comprehension is eliminated through elimination of Skolem functions for monotone formulas (section 5) yielding that $\forall x \exists y A(x, y)$ is provable in primitive recursive arithmetic. For this it is crucial that $\mathcal{P}$ is proofwise low over a system that does not contain $\Sigma^0_2$-induction, for instance $\text{WKL}_0^∗$.

These techniques of elimination of instances of comprehension can be viewed as a proof-theoretic refinement of the arithmetical conservativity of $\text{ACA}_0$ over $\text{PA}$, see [4], [11], [50] and [45, IX.1.6].

Comparison to conservation results by syntactic forcing. Syntactic forcing is a method to prove conservativity result. It is commonly used in reverse mathematics.

To show that a second order principle $\mathcal{P}$ is conservative over $\mathcal{T}$ it proceeds by first taking an arbitrary countable model of $\mathcal{T}$. This model is then extended through a forcing argument to include sets solving all instances of $\mathcal{P}$ without altering the first order part. The conservativity then follows by Gödel’s completeness theorem. For details and further information see [2].

The elimination of monotone Skolem functions and Howard’s elimination of bar recursion are constructive: a careful analysis of the proofs would yield a uniform method to obtain a term of $\mathcal{T}$ for each function provable total using $\mathcal{P}$. Whereas
the forcing argument essentially uses a reductio ad absurdum argument (if \( P \) would not be conservative then by the completeness theorem there would be a model that could not be extended).

Forcing yields in many cases full \( \Pi^1_1 \)-conservativity whereas the functional interpretation usually stops at \( \Pi^3_0 \)-conservativity. This is a consequence of the way the functional interpretation works: it transforms every statement in a functional, where for every additional quantifier alternation the type-level rises, making it more complex to analyze. For instance, \( \Pi^0_1 \)-statements correspond to type 2 functionals (i.e. functionals essentially of the form \( N^N \rightarrow N \)).

This makes it easier to handle principles implying the \( \Pi^0_1 \)-bounded collection principle (\( \Pi^0_1 \)-CP). Due to the well-known fact that \( \Pi^0_1 \)-CP is \( \Pi^3_0 \)-conservative over \( \Sigma^0_1 \)-IA the base theory for the functional interpretation does not change. This circumvents the problems forcing experiences when proving conservativity over \( \Pi^0_1 \)-CP, see [15, §6].

The original proof that \( RT^2_2 \) or \( COH \) is \( \Pi^1_1 \)-conservative over \( \Sigma^0_0 \)-induction uses syntactic forcing, also the proof that \( COH \) is \( \Pi^1_1 \)-conservative over \( \Pi^2_0 \)-CP is done using a complicated double forcing, see [9]. Our proof of the fact that \( COH + \Pi^0_1 \)-CP is \( \Pi^1_1 \)-conservative over \( \Sigma^0_1 \)-IA is similar to the proof of [7] since we show conservativity over \( RCA_0 \) (without \( \Pi^0_1 \)-CP) and therefore do not face the problems forcing experiences with \( \Pi^0_1 \)-CP and that Chong, Slaman and Yang in [9] deal with. Additionally, our proof is open for proof mining, which means it provides a method for program extraction.

2. Logical systems

We will work in a setting based on fragments of Heyting and Peano arithmetic in all finite types introduced in [51], for details see also [32].

In general, theories will be written with a superscript \( \omega \) which indicates that this is a finite type theory. Axioms and rules will not have an \( \omega \). The only exceptions to this are the theories of reverse mathematics (\( RCA_0, WKL_0, ACA_0, RCA^*_0, WKL^*_0 \)) and PRA.

2.1. Finite types. The set of all finite types \( T \) is inductively defined as

\[
0 \in T, \quad \rho, \tau \in T \Rightarrow \tau(\rho) \in T,
\]

where 0 denotes the type of natural numbers and \( \tau(\rho) \) the type of functions from \( \rho \) to \( \tau \). The set of pure types \( P \subset T \) is defined as

\[
0 \in P, \quad \rho \in P \Rightarrow 0(\rho) \in P.
\]

They will often be denoted by natural numbers:

\[
0(n) := n + 1,
\]

e.g. \( 0(0) = 1 \). The degree \( deg(\rho) \) of a type \( \rho \) is inductively defined as

\[
deg(0) := 0, \quad deg(\tau(\rho)) := \max(deg(\tau), deg(\rho) + 1).
\]

We will often denote the type of a term or variable by a superscribed index. For two types \( \rho, \tau \) we will write \( \rho \leq \tau \) if \( deg(\rho) \leq deg(\tau) \).

Equality \( =_0 \) for type 0 objects will be added as primitive notion to the systems. Higher type equality \( =_{\tau^\rho} \) will be treated as abbreviation:

\[
x^{\tau^\rho} =_{\tau^\rho} y^{\tau^\rho} :\equiv \forall z^\rho \, x^z =_\tau y^z.
\]
2.2. Gödel’s system $T$. Define the $\lambda$-combinators $\Pi_{\rho,\sigma}$, $\Sigma_{\rho,\sigma,T}$ to be the functionals satisfying

$$\Pi_{\rho,\sigma,T} x^\rho y^\rho =_{\rho} x, \quad \Sigma_{\rho,\sigma,T} x^\sigma y^\rho z^\rho =_{\tau} x z(yz).$$

Similar define the recursor $R_\rho$ of type $\rho$ to be the functional satisfying

$$R_\rho 0 y z =_{\rho} y, \quad R_\rho(Sx^0)y z =_{\rho} z(R_\rho xy z).$$

Let Gödel’s system $T$ be the $T$-sorted set of closed terms that can be build up from $0^0$, the successor function $S^1$, the $\lambda$-combinators and, the recursors $R_\rho$ for all finite types $\rho$. Using the $\lambda$-combinators one easily sees that $T$ is closed under $\lambda$-abstraction, see [51].

$T_n$ denotes the subsystem of Gödel’s system $T$, where primitive recursion is restricted to recursors $R_\rho$ with $\deg(\rho) \leq n$. The system $T_0$ corresponds to the extension of Kleene’s primitive recursive functionals to mixed types, see [24], whereas full system $T$ corresponds to Gödel’s primitive recursive functionals, see [13].

2.3. Heyting and Peano arithmetic. Define the neutral Heyting/Peano arithmetic (N-HA$^\omega$, N-PA$^\omega$) to be the extension of the term system $T$ to a $T$-sorted intuitionistic resp. classical logical system plus the schema of full induction and the equality axioms for type 0, i.e.

- $x =_0 x$, $x =_0 y \rightarrow y =_0 x$, $x =_0 y \land y =_0 z \rightarrow x =_0 z$,
- $x_1 =_0 y_1 \land \cdots \land x_n =_0 y_n \rightarrow t(x_1, \ldots, x_n) =_0 t(y_1, \ldots, y_n)$ for any $n$-ary term $t$ of suitable type,

and substitution schemata for $\lambda$-combinators and the recursors, i.e.

$$(\text{SUB}): \begin{cases}
    t[\Pi xy] =_0 t[x] \\
    t[\Sigma x y z] =_0 t[xz(yz)] \\
    t[R 0 y z] =_0 t[y] \\
    t[R(Sx)y z] =_0 t[z(Rxyz)x]
\end{cases}$$

For a formal definition see [52, I.1.6.15] (there N-HA$^\omega$ is called HA$^\omega$).

These theories are neutral with respect to an intensional or an extensional interpretation of higher type objects. However, for type 0 objects the usual equality axioms hold. Higher type equality is of no effect except for the SUB-rule. Later we will add functionals yielding cohesive and homogeneous set which are not extensional (in the presence of extensionality they would prove full arithmetical comprehension, see [31]) and therefore can only be analyzed in a neutral context.

Let weakly extensional Heyting/Peano arithmetic (WE-HA$^\omega$, WE-PA$^\omega$) be N-HA$^\omega$ resp. N-PA$^\omega$ plus the quantifier-free rule of extensionality, i.e.

$$(\text{QF-ER}): A_\rho \rightarrow s \equiv_\rho t \overset{A_\rho \rightarrow R[s/x^\rho]}{\Rightarrow} r[t/y^\rho].$$

where $A_\rho$ is quantifier-free and $s^\rho$, $t^\rho$, $r^\tau$ are terms of WE-HA$^\omega$. Note that the addition of SUB here is redundant, since QF-ER together with the axioms for $\Pi$, $\Sigma$, $R$ proves it. The systems with full extensionality, i.e. N-HA$^\omega$, N-PA$^\omega$ plus the extensionality axioms

$$(E_{\rho,\tau}): \forall z^\tau, x^\rho, y^\rho (x =_{\rho} y \rightarrow zx =_{\tau} zy)$$

for all $\tau, \rho \in T$, will be denoted by E-HA$^\omega$ and E-PA$^\omega$. For a detailed definition of these systems, see [32, section 3].

The weakly extensional and neutral theories allow functional interpretation in themselves, which is not possible in the presence of full extensionality. Later we will eliminate the usage of extensionality (see proposition 3 below), hence neither the interpretation of constants yielding cohesive/homogeneous sets nor the functional
interpretation will lead to problems. For a discussion of these systems and the connection to functional interpretation we refer to [51].

It is also important to note that in presence of only QF-ER the deduction theorem in general fails, see [32, theorem 9.11]. To overcome this we will restrict the use of principles in premises of QF-ER. This will be denote by the \(\oplus\)-sign, e.g. WE-PA\(\omega\) \(\oplus\) WKL denote the system WE-PA\(\omega\) + WKL, where WKL may not be used in the premise of QF-ER. The weak extensional systems satisfy the deduction theorem with respect to \(\oplus\).

We now introduce fragments of neutral and (weakly) extensional Heyting/Peano arithmetic corresponding to \(T_n\):
Define \(\text{N-HA}_n^\omega\)\(\|\) to be the logical system extending \(T_n\) plus \(\Sigma_0^{n+1}\)-IA and plus the case-distinction functionals \((\text{Cond}_\rho)_{\rho \in T}\) and its substitution axioms

\[
\text{(SUBCond):} \quad \begin{cases} 
  t[\text{Cond}_\rho(t^0, x^0, y^0)] =_0 t[x] \\
  t[\text{Cond}_\rho(Su, x^0, y^0)] =_0 t[y] 
\end{cases} \quad \text{for all } t \text{ of type 0.}
\]

These case distinction functionals are needed for the functional interpretation and cannot be defined in these fragments of \(\text{N-HA}_n^\omega\), see [41, 3]. In the full system \(T\) they can be simulated by the recursors. Instead of \(\text{N-HA}_n^\omega\)\(\|\) we also write \(\hat{\text{N-HA}}_n^\omega\). The classical systems \(\text{N-PA}_n^\omega\)\(\|\), \(\hat{\text{N-PA}}_n^\omega\)\(\|\) are defined similarly. In the same way also the (weakly) extensional systems \((\text{W})\text{E-HA}_n^\omega\)\(\|\), \((\text{W})\text{E-HA}_n^\omega\)\(\|\), \((\text{W})\text{E-PA}_n^\omega\)\(\|\), \((\text{W})\text{E-PA}_n^\omega\)\(\|\) are defined.\(^4\) However for the classical systems defined here one does not need to add Cond to the system since it is provably definable with the \(\lambda\)-combinators and \(R_n\), see [41]. Note that \(\Sigma_0^{n+1}\)-induction is provable with the recursor \(R_n\) and quantifier-free induction and full QF-AC in all types (definition below) over the classical systems defined here. Hence the addition of it to the classical systems is actually superfluous. This follows from [41] and Kreisel’s characterization theorem, see [32, proposition 10.13].

2.4. Grzegorczyk arithmetic. We moreover need weaker fragments of Heyting and Peano arithmetic containing only quantifier-free induction.

Let weakly extensional Grzegorczyk arithmetic of level \(n\) in all finite types \(G_nA_n^\omega\)\(\|\) be the (intuitionistic) system containing \(=_0\)-axioms, QF-ER, \(\lambda\)-abstraction, the \(n\)-th branch of the Ackermann-function, bounded search and bounded primitive recursion. For a detailed definition see [26].\(^5\) The neutral variant will be denoted by \(\text{N-G}_nA_n^\omega\)\(\|\), the extensional one by \(\text{E-G}_nA_n^\omega\)\(\|\).

Let \(G_nA_n^\omega\)\(\|\) be the union of all these systems. This system contains all primitive recursive functions but not all primitive recursive functionals (in the sense of Kleene). For instance \(R_0\) is not contained in \(G_nA_n^\omega\)\(\|\). Thus it also contains no \(\Sigma_0^1\)-induction. The set of all closed terms of \(G_nA_n^\omega\)\(\|\) is called \(G_nR_n\). See [26] and [32, Chapter 3] for all of this.

2.5. Quantifier-free axiom of choice. Let QF-AC be the schema

\[
\forall x \exists y A_{gf}(x, y) \rightarrow \exists f \forall x A_{gf}(x, f(x)),
\]
where \(A_{gf}\) is a quantifier-free formula. If the types of \(x, y\) are restricted to \(\alpha, \beta\) we write QF-AC\(\alpha, \beta\).

\(^4\)For a formal definition let \((\text{W})\text{E-HA}_n^\omega\)\(\|\) be defined as in [32, section 3.4] and define \((\text{W})\text{E-HA}_n^\omega\)\(\|\) to be \((\text{W})\text{E-HA}_n^\omega\)\(\|\) plus \(\Sigma_0^{n+1}\)-IA and the defining axioms and constants for the recursors \(R_n\) with \(\text{deg}(\rho) \leq n\). The neutral variants are defined in the same way but without the rule of extensionality.

\(^5\)In [32] the system \(G_nA_n^\omega\)\(\|\) is defined to include all \(N, N^n, N^{n^n}\)-true \(\forall\)-sentences. In a pure proof-mining context these sentences do not matter because they have no impact on the provable primitive recursive functions in the system. We only add quantifier-free induction (QF-IA), to be able to later establish conservativity over PRA.
The scheme QF-AC\(^{0,0}\) corresponds to recursive comprehension (\(\Delta^{0}_1\)-CA) in a second order context. Thus \(\text{WE}-\text{PA}^{\omega\omega}\) + QF-AC\(^{1,0}\) and RCA\(_0\) share the same proof theoretic strength. RCA\(_0\) can easily be embedded into \(\text{WE}-\text{PA}^{\omega}\) + QF-AC\(^{1,0}\) and \(\text{WE}-\text{PA}^{\omega\omega}\) + QF-AC\(^{1,0}\) is conservative over RCA\(_0\) modulo this embedding, see [31].

For this reason \(\text{WE}-\text{PA}^{\omega\omega}\) + QF-AC\(^{1,0}\) is called RCA\(_0^\omega\).

The system RCA\(_0^\omega\) is RCA\(_0\), where \(\Sigma^{0}_{1}\)-induction is replaced by quantifier-free-induction and the exponential function, see [45, X.4.1]. This system can be embedded into \(\text{G}_2\text{A}^{\omega\omega}\) + QF-AC\(^{1,0}\) and both systems are \(\Pi^{0}_2\)-conservative over Kalmar elementary arithmetic.

In ordinary mathematics higher types usually do not occur and second order arithmetic is sufficient to formalize most of it. We require here a system containing all finite types to be able to carry out a functional interpretation and thus cannot use a second order system.

2.6. The quantifier-free subsystems. In order to exploit the full subtlety of the functional interpretation we will also need the quantifier-free subsystems of N-G\(_n\)A\(^\omega\) and N-HA\(_n^\omega\)\(^{\|}\). The quantifier-free subsystems are denoted by qf-N-G\(_n\)A\(^\omega\) resp. qf-N-PA\(_n^\omega\)\(^{\|}\). (The quantifier free subsystems satisfy the law of excluded middle and are therefore classical.)

They are obtained from the full systems as follows:

- The quantifier-rules and -axioms are dropped from logic.
- For all axioms of the form \(A(x^{\omega}_1, \ldots, x^{\omega}_n)\), where \(A\) is quantifier-free, the following axioms are added to the system:
  \[
  A(t^{\omega}_1, \ldots, t^{\omega}_n),
  \]
  where \(t_i\) are arbitrary terms.
- The induction schema is replaced by the (quantifier-free) induction rule:
  \[
  \frac{A(0^{\omega}), \quad A(x^{\omega}) \rightarrow A(Sx^{\omega})}{A(t^{\omega})},
  \]
  where \(A\) is quantifier-free, \(x\) does not occur free in the assumption and \(t\) is an arbitrary term.

These quantifier-free systems contain only prime formulas of the form

\[
\Phi_0 =_0 \Phi_1,
\]

where \(\Phi_0, \Phi_1\) are terms in N-G\(_n\)A\(^\omega\) resp. N-HA\(_n^\omega\)\(^{\|}\). Formulas are logical combinations of these predicates. Obviously, qf-N-G\(_n\)A\(^\omega\) and qf-N-PA\(_n^\omega\)\(^{\|}\) are subsystems of N-G\(_n\)A\(^\omega\) resp. N-HA\(_n^\omega\)\(^{\|}\). For a detailed discussion of these systems we also refer the reader to [51, 1.6.5]. For technical reason we use here the variant of the systems described remark 1.5.8.)

Observe, that in these system we can only instantiate type 0 variables (via the induction rule) and not higher type variables, hence we immediately obtain the following lemma:

**Lemma 2.** Let \(A\) be a sentence and

\[
\mathcal{T} \vdash A,
\]

where \(\mathcal{T} = \text{qf-N-G}_n\text{A}^{\omega}, \text{qf-N-PA}_n^{\omega\|}\).

Then there exists a derivation of \(A\) in \(\mathcal{T}\) that contains only the variables of \(A\) plus some fresh variables of type 0.

**Proof.** In a derivation of \(A\) in \(\mathcal{T}\) replace every variable not of type 0 and not occurring in \(A\) by constant \(0^\Phi\) of suitable type. Since higher type variables cannot be instantiated the derivation remains valid. \(\square\)
2.7. **Functional interpretation.** Functional interpretation will denote in this paper a negative translation followed by Gödel’s Dialectica translation.

Gödel’s Dialectica translation is a proof interpretation that translates proofs from (a fragment of) \( \text{WE-HA}^{\omega} \) or \( \text{N-HA}^{\omega} \) into its quantifier-free subsystem, see [13].

Let \( T \) be such a system. The Dialectica translation associates to each formula \( A \) of \( T \) a \( \exists \forall \)-formula

\[
A^{D} := \exists x \forall y A_{D}(x, y),
\]

where \( A_{D} \) is quantifier-free. In particular, for a \( \Sigma_{1}^{0} \) sentence \( A \) the formula \( A_{D} \) is the quantifier-free matrix of \( A \).

From a proof of \( A \) one then can extract a term \( t \), such that

\[
\text{qf-} T \vdash A^{D}(t, x).
\]

A negative translation is a proof translation that translates classical proofs into intuitionistic proofs. It also proceeds by associating each formula \( A \) a formula \( A^{N} \) such that

\[
S \vdash A \iff A^{N} \quad \text{and} \quad S \vdash A \implies S_{i} \vdash A^{N}.
\]

Here \( S \) is any of \( (W)\text{E-PA}^{\omega} \), \( (\hat{W})\text{E-PA}^{\omega} \), \( \text{G}_{n}A^{\omega} \) or its neutral variants and \( S_{i} \) is its intuitionistic counterpart. (To be specific, Kuroda’s negative translation \( A^{N} \) is obtained from \( A \) by inserting \( \neg \neg \) after each \( \forall \) and in front of the whole formula.)

Thus we denote by functional interpretation a proof translation from (a fragment of) \( \text{WE-PA}^{\omega} \) or \( \text{N-PA}^{\omega} \) into its quantifier-free part. We abbreviate the functional interpretation by \( \text{ND} \). The ND-translation of a formula \( A \) will be denoted by \( A^{\text{ND}} \) and the quantifier-free matrix of it by \( A_{\text{qf}}^{\text{ND}} \).

The functional interpretation in particular has the property to extract a term for each provable recursive function, i.e. from a proof of a \( \forall \exists \)-statement (in \( \text{WE-PA}^{\omega} \) or any other fragment for which the functional interpretation holds)

\[
\text{WE-PA}^{\omega} \vdash \forall u \exists v A_{qf}(u, v)
\]

it extracts a term \( t \) such that

\[
\text{qf-WE-HA}^{\omega} \vdash A_{qf}(t, u) \equiv A^{\text{ND}}(t, u).
\]

For an introduction to the functional interpretation see [32, 3, 51].

Since the functional interpretation does not interpret full extensionality it is often combined with the elimination of extensionality.

**Proposition 3** (Elimination of extensionality, [38]). Let \( A \) be a formula containing only free variables and quantification of type \( \leq 1 \). If

\[
\text{E-PA}^{\omega} + \text{QF-AC}^{0,1} + \text{QF-AC}^{1,0} \vdash A
\]

then

\[
\text{N-PA}^{\omega} + \text{QF-AC}^{0,1} + \text{QF-AC}^{1,0} \vdash A.
\]

The same holds also for the fragments \( \hat{\text{N-PA}}^{\omega} \) and \( \text{N-G}_{n}A^{\omega} \).

**Proof.** Proposition 10.45 and lemma 10.41 of [32]. These lemma and proposition actually do not make use of weak extensionality and therefore show conservativity even over a neutral theory. \( \square \)
2.8. Additional notation and definitions. We denote sets by capital letters. Unless otherwise noted they are represented by characteristic functions. Sometimes capital letters also denote higher type functionals. It will be clear from the context what is meant.

It is important to note that in systems not containing $\Sigma^0_1$-induction it is in general not provable that every infinite set — that is a set $X$ satisfying $\forall k \exists n > k \ n \in X$ — can be strictly increasingly enumerated, i.e. there exists a strictly monotone function $f$ such that $\text{rng}(f) = X$. The system $\text{WE-HA}_ω^\omega + \text{AC}^{\omega,0}$ proves that the first statement implies the second. The converse — every strictly increasingly enumerable set is infinite — is already provable without $\Sigma^0_1$-induction, for instance $G_{1,Aω}$ suffices.

Sequence codes are denoted by $\langle x_0, \ldots, x_n \rangle$. The corresponding projection functions and length function are denoted by $(\cdot)_i$ and $lth(\cdot)$. We encode sequences using a bijective and monotone (in each component) sequence-coding based on the Cantor pairing, see [32, definition 3.30]. This coding is definable in every system containing qf-$G_{3,Aω}$.

We model in our systems $n$-colorings of $\mathbb{N}^2$ as functions $c: \mathbb{N} \times \mathbb{N} \to n$ with $c(x, y) = c(y, x)$.

Further we define the following notions:

- $\bar{f}$ denotes the course-of-value function of $f^1$, i.e. $\bar{f}(n) = \langle f(0), \ldots, f(n-1) \rangle$.
- $x \subseteq^\text{fin} X$ iff $x$ is an initial segment of a strictly monotone enumeration of $X$.
- $x \subseteq^\text{fin} X$ iff $x$ is an code for a finite subset of $X$.

**Definition 4** (Bounded type 1 recursor, $\check{R}_1$). The bounded type 1 recursor $\check{R}_1$ is defined as

\[
\check{R}_1(0)yzhu := 0 \min(y(u), h(0, u)) \\
\check{R}_1(x+1)yzhu := 0 \min(z(\check{R}_1xyzh)xu, h(x,u)).
\]

We will denote by $(\check{R}_1)$ the defining axioms. Note that they are purely universal and that $\check{R}_1$ can be trivially majorized.

**Definition 5** (Uniform weak König’s lemma, UWKL, [30]). Uniform weak König’s lemma is the statement

\[
\exists \Phi \leq_1 1(1) \forall f (T^{\infty}(f) \to \forall x^0 f(\Phi f x) = 0),
\]

where $T^{\infty}$ expresses that $f$ describes an infinite 0/1-tree.

We can modify (in $G_{3,Aω}$) every function $f$ such that it describes an infinite 0/1-tree and is not altered if it already described such a tree. We will write $\check{f}$ for this modification, see [25, 32].

With this we can restate UWKL equivalently as

\[
\exists \Phi \leq_1 1(1) \forall f^1 \forall x^0 \check{f}(\Phi \check{f} x) = 0.
\]

Note that the condition $\leq_1(1)$ is superfluous because the modified tree contains only 0/1-sequences.

By Skolemization we add a weak König’s Lemma functional constant $\mathcal{B}$ described by the (purely universal) axiom

\[
(3) \forall f \forall x^0 \check{f}(\mathcal{B} \check{f} x) = 0.
\]

This axiom will be denoted by $(\mathcal{B})$. Note that $\mathcal{B}$ can be trivially majorized.

In a system containing full extensionality UWKL implies $\Pi^1_1$-CA, see [30], hence it is too strong for our purpose. But in a weakly extensional system it often can be
handled like WKL, for instance it vanishes under a monotone functional interpretation like WKL and can be added to the elimination of monotone Skolem functions, see [30]. Note that proposition 3 does not cover UWKL.

3. Continuous functionals

Recall that a type 2 functional \( \varphi \) is continuous if

\[
\forall g \exists n \forall h \left( \bar{g}n = \bar{h}n \rightarrow \varphi(g) = \varphi(h) \right).
\]

**Definition 6** (Associate, [23, 33]). For every continuous type 2 functional \( \varphi \) we will denote by \( \alpha_\varphi \) an associate of \( \varphi \), i.e. a type 1 function with the properties

\[
\forall f \exists n \alpha_\varphi(fn) \neq 0,
\]

\[
\forall f, n \left( \alpha_\varphi(fn) \neq 0 \rightarrow \varphi(f) = \alpha_\varphi(fn) \right).
\]

The value of \( \varphi \) is uniquely determined through \( \alpha_\varphi \). For every continuous functional there exists an associate, though it is not uniquely determined. For details see also [39].

**Definition 7.** A functional given by a closed term \( \varphi^p \) of \( \mathcal{T} \) is called **provably continuous** if for some term \( \alpha_\varphi \) (containing at most the free variables of \( \varphi \)) of type 1 (if \( p > 0 \)) resp. 0 (if \( p = 0 \)), the following holds:

\[
\mathcal{T} \vdash \varphi \approx_\rho \alpha_\varphi.
\]

Here, for general \( x^p \) and \( \alpha^{0/1} \), the relation \( x \approx_\rho \alpha \) is defined by induction on \( \rho \):

\[
x \approx_0 \alpha : \equiv x =_0 \alpha,
\]

\[
x \approx_\rho \alpha : \equiv \alpha \in \text{ECF}_{\tau_\rho} \land \forall \bar{g}^p \forall \bar{\beta} \in \text{ECF}_\rho \left( \bar{y}^\rho \beta \rightarrow xy \approx_\tau \alpha \upharpoonright \beta \right),
\]

where ECF is the model of extensional hereditarily continuous functionals formalized in \( \mathcal{T} \) and \( \upharpoonright \) denotes the application in ECF. (See [24, 33, 51], for a definition see also [32, definitions 3.58, 3.59].)

Especially, in the case of \( \rho = 2 \) a functional \( \varphi \) is provably continuous in \( \mathcal{T} \) if it has an associate \( \alpha_\varphi \) in \( \mathcal{T} \) and (5) is provable.

**Proposition 8.** For every term \( t^2 \in \mathcal{G}_n R^\omega, T_0, T_1 \) there exists provably in \( \mathcal{G}_n A^\omega \) resp. \( \text{WE-PA}^{\omega} \lceil \), \( \text{WE-PA}^{\omega} \rceil \) a (primitive recursive) associate \( \alpha_t \). In other words \( t \) is provably continuous.

**Proof.** We first consider the case of \( \text{WE-PA}^{\omega} \lceil = \text{WE-PA}^{\omega} \rceil \) and \( \mathcal{G}_n A^\omega \). Here the only functional constants having no trivial associate are the \( \lambda \)-combinators and \( R_0 \) (in the case of \( \text{WE-PA}^{\omega} \lceil \) and the course-of-value functional (in the case of \( \mathcal{G}_n A^\omega \)). The associates of \( R_0 \) and the course-of-value functional can easily be computed and (5) be proven in the respective systems. By normalization one can find a term \( \tilde{t} =_2 t \) that does not include \( \lambda \)-abstraction of type \( \geq 1 \). The proposition for \( \text{WE-PA}^{\omega} \lceil \) and \( \mathcal{G}_n A^\omega \) follows from this.

In the case of \( \text{WE-PA}^{\omega} \lceil \) we prove by induction over the structure of \( t \) that \( t \) is provably continuous. For this it is sufficient to prove that every functional constant is provably continuous and to observe that this property is retained under composition. The associates for the \( \lambda \)-combinators are easily definable and provable in these systems, see [51].

Here we only show that the existence of an associate for \( R_1 \) is provable in \( \text{WE-PA}^{\omega} \lceil \), since we are only interested in this case. For the other recursors the
proof is similar. Let
\[ \alpha_R_1(0, y', z', u) := \begin{cases} (y')_u + 1 & \text{if } u < l h y', \\ 0 & \text{otherwise,} \end{cases} \]
\[ \alpha_R_1(x + 1, y', z', u) := \begin{cases} (z')_{(x, l h (x, y', z', k) + 1)} & \text{if } k < l h y', \text{such that } \alpha_R_1(x, y', z', k) > 0 \\ 0 & \text{otherwise.} \end{cases} \]

Using \( \Pi^0_2 \)-induction one shows that
\[ \forall x \ (\exists n \ \alpha_{R1}(x, y n, \alpha_x, \pi n, u) = R_1(x, y, z, u) + 1) \]
and hence that \( \alpha_{R1} \) is an associate of \( R_1 \). \( \square \)

4. Properties of instances of comprehension

Remark 9. A sequence of \( \Pi^1_0 \)-comprehension instances \( (\Pi^1_0-\text{CA}(f_i)) \) may be reduced to the single instance of \( \Pi^1_0-\text{CA}(f') \) with \( f'xy := f(x)_y(x, y) \), see [27, remark 3.8].

Lemma 10 ([27, 28]). For suitable terms \( \xi_i \) of \( G_3A^\omega \) we have
(i) \( G_3A^\omega + QF-\text{AC}^{0,0} \vdash \forall f (\Pi^1_0-\text{CA}(\xi_i f) \rightarrow \Pi^1_0-\text{AC}(f)) \),
(ii) \( G_3A^\omega + QF-\text{AC}^{0,0} \vdash \forall f (\Pi^1_0-\text{CA}(\xi_2 f) \rightarrow \Delta^0_2-\text{CA}(f)) \),
(iii) \( G_3A^\omega + QF-\text{AC}^{0,0} \vdash \forall f (\Pi^1_0-\text{CA}(\xi_3 f) \rightarrow \Delta^0_2-\text{IA}(f)) \),
(iv) \( G_3A^\omega + QF-\text{AC}^{0,0} \vdash \forall f (\Pi^1_0-\text{CA}(\xi_4 f) \rightarrow \Pi^1_0-\text{CP}(f)) \),
(v) \( G_3A^\omega + QF-\text{AC}^{0,0} + \text{WKL} \vdash \forall f (\Pi^1_0-\text{CA}(\xi_5 f) \rightarrow \Pi^1_0-\text{WKL}(f)) \).

Here the principle \( K-\text{AC} \) denotes the scheme of axiom of choice, where the base formula is of type \( K \). Similarly \( K-\text{WKL} \) denotes weak König’s lemma where the tree is given by a predicate of type \( K \). The principles \( K-\text{IA} \) and \( K-\text{CA} \) are defined likewise.

If \( K = \Pi^1_0, \Sigma^1_0 \) then an instance of those principles is given by a function \( f \) coding the quantifier-free part of the \( \Pi^1_0 \) resp. \( \Sigma^1_0 \) formula. For instance
\[ \Pi^1_0-\text{AC}(f) \equiv \forall x \exists y \forall z f(x, y, z) = 0 \rightarrow \exists Y \forall x \forall z f(x, Y(x), z) = 0. \]

Similar a \( \Delta^2_2 \)-formula is given by an \( f \) coding a function for a \( \Pi^1_0 \) and a function for a \( \Sigma^1_0 \) formula.

Proof of lemma 10. For (i), (ii) see [28, lemma 4.2]. The statements (iii), (iv) are immediate consequences of these. Note that we require here \( G_3A^\omega \) and not only \( G_3A^\omega \) as in the reference, since we do not add the true universal sentences to the system, see footnote 5.

For (v) let \( \xi_5 \) be such that the instance of \( \Pi^1_0-\text{CA} \) yields the comprehension function for the innermost quantifier of the tree predicate reducing \( \Pi^1_0-\text{WKL} \) to \( \Pi^1_0-\text{WKL} \). This is equivalent to \( \text{WKL} \) and thus included in the system, see for instance [45]. \( \square \)

For the ordinal analysis of terms we will need the following abbreviation:
\[ \omega^\mu_0 = \mu \quad \text{and} \quad \omega^\mu_{k+1} = \omega^\mu_k, \]
where \( k \in \mathbb{N} \) and \( \mu \) is an ordinal.

Lemma 11. Let \( n \in \mathbb{N} \) and let \( t[g] \) be a type 1 term with the only free variable \( g \) such that \( \lambda g. t[g] \in T_n \). Then for every term \( \varphi \) in \( T_{n-1} \) or in \( G_\infty \mathbb{R}^\omega \) if \( n = 0 \) there exists a term \( \xi \) in the same system such that \( \text{WE-PA}_{n-1}[\varphi - \text{AC}] \lor \text{G_\infty A^\omega + QF-AC} \) in the case of \( n = 0 \) proves
\[ \forall g (\Pi^1_0-\text{CA}(\xi g) \rightarrow \exists f^1 (f \text{ satisfies the defining axioms of } t[g] \land \Pi^1_0-\text{CA}(\varphi fg)) ) \].
Defining axioms of $t[g]$ are a formula $A$, such that $\forall g, x, y \ (A(g, x, y) \leftrightarrow t[g]x = y)$. (Since $t^1$ can be defined by (unnested) ordinal recursion of order $< \omega^\omega_{n+1}$, one can take for $A$ the formula describing this recursion.)

Proof. First fix a suitable encoding for ordinals smaller than $\varepsilon_0$ in this system, see for instance [14].

Every term $t^1 \in T_n$ can be defined through (unnested) ordinal recursion of order $< \omega^\omega_{n+1}$; the totality of such a recursion can be proven using a suitable instance of $\Sigma^0_{n+1}$-CA, see [40] and theorem 17 below. Such an instance is included in the system because a suitable instance of $\Pi^0_1$-CA reduces it to $\Sigma^0_n$-IA. This proves the claim that there is a total function $f$ satisfying the definition of $t[g]$.

For the second part note that the defining axioms of unnested ordinal primitive recursion of order type $\alpha$ are given by

$$f(0) := f_0, \quad f(n) := h(n, f(l(n)),$$

where $l$ satisfies

$$l(n) < n \quad \text{for } n > 0$$

and $<$ defines a well-ordering on $\mathbb{N}$ of order type $\alpha$.

We say a finite sequence $s$ satisfies the defining axioms (6) up to $n$ if

$$(s)_0 = f_0, \quad (s)_i = h(i, (s)_{l(i)}) \quad \text{for all } i \in \bigcup_{n' \leq n} \{l^k(n')\} \setminus \{0\}$$

For notational ease we assume here that $l(0) = 0$. Note that because of (7) the set $\bigcup_{k} \{l^k(n')\}$ defines an $\prec$-descending chain and is therefore provably finite.

For the second part we have to prove a comprehension of the form

$$\exists H \forall k \ (k \in H \leftrightarrow \forall x \varphi(f, g, k, x) = 0).$$

We used the imposed instance of comprehension to prove the following comprehension

$$\exists H \forall k \ (k \in H \leftrightarrow \forall x \forall s, n \ (s \text{ satisfies the defining axioms of } t[g] \text{ up to } n)$$

$$\rightarrow \alpha_{\varphi(f, g, k, x)}(s) \leq 1).$$

Note that this comprehension is equivalent to (8) if $f$ is total.

The proof of the comprehension above is similar to the construction of a 1-generic set: If the statement

$$\forall x \varphi(f, g, k, x) = 0$$

for a fixed $k$ fails, then there is an $x$ such that $\varphi(f, g, k, x) \neq 0$. Since $\varphi$ is continuous this depends only on an initial segment of $f$. We express this by using associates, i.e. this statement is equivalent to

$$\exists n \alpha_{\varphi(f, g, k, x)}(fn) > 1.$$ 

Hence it suffices to consider only finite initial segments.

We will use this technique in most proofs of instances of comprehension in this paper. This is the reason why we require $\varphi$ to be provably continuous in the definition of proofwise low.

5. Elimination of monotone Skolem functions

Let $\Delta$ be a set of sentences of the form $\forall a \exists b < ra \forall c^0 B_{g}(a, b, c)$, where $r$ is a closed term and $B_{g}$ is quantifier-free and contains any further free variables than those shown. Let $\bar{\Delta}$ be the corresponding set of Skolem normal form of the sentence of $\Delta$, i.e. the corresponding formulas of the form $\exists B < r \forall a, c^0 B_{g}(a, Ba, c)$. 

Theorem 12 ([27, 3.8]). Let \( \gamma \) be an arbitrary type and let \( A_{\gamma} \) be a quantifier-free statement where only the shown variables are free and let \( s \) be a term in \( G_{\infty} \mathbb{R}^\omega \). If

\[
G_{\infty} \mathbb{A}^\omega + \text{QF-AC} \oplus \Delta \vdash \forall u \forall v \leq_\gamma su (\Pi^1_0 \text{-CA}(\xi uv) \rightarrow \exists w^0 A_{\gamma}(u, v, w))
\]

then one can extract from a proof a term \( t \in T_0 \) such that

\[
\text{WE-HA}^\omega \upharpoonright \oplus \Delta \vdash \forall u \forall v \leq_\gamma su \exists w \leq_0 tu A_{\gamma}(u, v, w).
\]

Especially, in case that \( A_{\gamma} \in \mathcal{L}(\text{PRA}) \), \( u \) of type 0, \( v \) absent and \( \Delta = \emptyset \) we have

\[
\text{PRA} \vdash \forall u^0 A_{\gamma}(u, tu).
\]

Corollary 13. Let \( \gamma, \xi, s, A_{\gamma} \) be as in Theorem 12. However \( \xi \) may contain \( B \) but \( s \) and \( A_{\gamma} \) must not. Then the following holds: If

\[
G_{\infty} \mathbb{A}^\omega + \text{QF-AC} \oplus (B) \vdash \forall u \forall v \leq_\gamma su (\Pi^1_0 \text{-CA}(\xi uv) \rightarrow \exists w^0 A_{\gamma}(u, v, w))
\]

then one can extract from a proof a term \( t \in T_0 \) such that

\[
\text{WE-HA}^\omega \upharpoonright \Delta \vdash \forall u \forall v \leq_\gamma su \exists w \leq_0 tu A_{\gamma}(u, v, w).
\]

Proof. First note that due to [27, remark 2.10] we may add the (majorizable) constant \( B \) to \( G_{\infty} \mathbb{A}^\omega \) in Theorem 12.

Apply this theorem to \( \Delta := \{ \forall f \exists x (Bf_x) = 0 \} \), cf. definition 5 and (3) on p. 10. The premise of the corollary implies that

\[
G_{\infty} \mathbb{A}^\omega + \text{QF-AC} \oplus \Delta \vdash \forall u \forall v \leq_\gamma su (\Pi^1_0 \text{-CA}(\xi uv) \rightarrow \exists w^0 A_{\gamma}(u, v, w)).
\]

Theorem 12 and noticing that \( \Delta \equiv \tilde{\Delta} \) yields

\[
\text{WE-HA}^\omega \upharpoonright \Delta \vdash \forall u \forall v \leq_\gamma su \exists w \leq_0 tu A_{\gamma}(u, v, w)
\]

and so

\[
\text{WE-HA}^\omega \upharpoonright \Delta \vdash \forall u \forall v \leq_\gamma su \exists w \leq_0 tu A_{\gamma}(u, v, w).
\]

Since the constant \( B \) only occurs in \( \Delta \), we may replace it with a new variable and so replace \( \Delta \) with \( \text{UWKL} \). The corollary now follows from [32, corollary 10.34]. \( \square \)

6. Bar recursor

With bar recursion (even of the lowest type) one can interpret the functional interpretation of (instances of) \( \Pi^1_0 \text{-CA} \). This will be discussed in detail in proposition 48 below. In this section we will show that the bar recursor \( B_{0,1} \) can be majorized provably in \( \text{WE-HA}^\omega \).

Definition 14 (bar induction of type 0). Let bar induction of type 0 be

\[
(Bl_0): \begin{cases}
\forall x^1 \exists n^0 \forall n \geq n_0 Q(x, n; n) \land \\
\forall x^1, n^0 (\forall d Q(x, x \ast d; n + 1) \rightarrow Q(x, n; n)) \land \\
\exists \forall x^1, n^0 Q(x, n; n),
\end{cases}
\]

where

\[
(x, x)k := \begin{cases}
x(k), & \text{if } k < n, \\
0, & \text{otherwise},
\end{cases} \quad (x, n \ast d)k := \begin{cases}
x(k), & \text{if } k < n, \\
d, & \text{if } k = n, \\
0, & \text{otherwise}.
\end{cases}
\]

If \( Q \) is restricted to formulas in \( \mathcal{K} \), we write \( \mathcal{K}-Bl_0 \).

Lemma 15.

\[
\text{WE-PA}^\omega \upharpoonright \text{QF-AC}^{0,0} \vdash \Pi^1_0 - Bl_0
\]
Proof. Let \( Q(x, n; n) \equiv \forall k Q_{Qf}(x, n; n; k) \). Suppose that \( \Pi^0_1\text{-}B_0 \) does not hold, i.e. the premises of \( \Pi^0_1\text{-}B_0 \) are true and
\[
\exists x_0, n_0, k_0 \neg Q_{Qf}(x_0, n_0; n_0; k_0),
\]
which is equivalent to
\[
\exists x_0, n_0, k_0 \neg Q_{Qf}(x_0, n_0; n_0; k_0).
\]

The second premise yields
\[
\forall x^1, n^0, k^0 \exists d, k' (-Q_{Qf}(x^1, n^0; n_0) \rightarrow -Q_{Qf}(x^1, n^0 \ast d; n_0 + 1; k')).
\]
Since the whole statement only depends on an initial segment of \( x^1 \), it can be coded in a type 0 object \( x^0 \). For instance let \( x' := \pi n \) then \( \lambda i. (x')_i, n = \pi n \).

Using QF-\( AC^0_0 \) we then obtain functions \( D(x, n, k), K(x, n, k) \) with
\[
\forall x^0, n, k (-Q_{Qf}(\lambda i. (x)_i, n; n_0) \rightarrow -Q_{Qf}(\lambda i. (x)_i, n \ast D(x, n, k); n_0 + 1; K(x, n, k))).
\]

Then define using simultaneous course-of-value recursion (\( n_0, x_0, k_0 \) are from (9)) the functions \( D_0, K_0 \):
\[
\begin{align*}
D_0(n) &:= x_0(n) \\
K_0(n) &:= k_0
\end{align*}
\]
for \( n \leq n_0 \),
\[
\begin{align*}
D_0(n) &:= D(D_0, n, n, K_0(n - 1)) \\
K_0(n) &:= K(D_0, n, n, K_0(n - 1))
\end{align*}
\]
for \( n > n_0 \).

The definition of \( D_0 \) and (9),(10) yield
\[
\forall n \geq n_0 \neg Q(D_0, n; n)
\]
and hence a contradiction to the first premise of \( \Pi^0_1\text{-}B_0 \).

\[ \square \]

Proposition 16. \( WE\text{-}\hat{\Pi}A^\omega \vdash QF\text{-}AC^0_0 \) proves that there exists a majorant \( B_{0,1} \) of \( B_{0,1} \).

Proof. Define \( B_{0,1} \) like in [32, proof of theorem 11.17]. By the cited proof it suffices to show \( \Pi^0_1\text{-}B_0 \). (Note that in that proof \( Q \) is a \( \Pi^0_1 \) formula in the case where \( \rho = 0 \).) Hence the proposition is an immediate consequence of lemma 15. See also [5].

\[ \square \]

7. Ordinal analysis of terms

7.1. Ordinal Peano/Heyting arithmetic. In this section we will investigate the strength of induction along ordinals the systems \( WE\text{-}\hat{\Pi}A^\omega \), \( WE\text{-}\hat{\Pi}A^\omega \).

We will code ordinals using the ordinal coding of [14, II.3.a]. (This coding uses the Cantor normal form for ordinals to define primitive recursive codes for ordinals.) For convenience we repeat the definition of \( \omega^\mu_k \):
\[
\omega^\mu_0 = \mu \quad \text{and} \quad \omega^\mu_{k+1} = \omega^\mu_k
\]
Here \( k \in \mathbb{N} \) and \( \mu \) is an arbitrary ordinal number.

Theorem 17 ([40], [53]). The functions and functionals of level 2 that are ordinal recursive (unnested) in an ordering \( \omega^\mu_k \) are exactly the functions and functionals in \( T_k \).

Theorem 18 ([14, II.3.18]).

\[
WE\text{-}\hat{\Pi}A^\omega \vdash \Sigma^0_{m+k-1}IA \vdash \Sigma^0_m\text{-}LNP(< \omega^\mu_k)
\]
for every \( m, k \in \mathbb{N} \), where LNP denotes the least number principle.
In particular, \( WE\text{-}\hat{\Pi}A^\omega \vdash \Sigma^0_{k+1}IA \) proves the totality of \( < \omega^\mu_k \) recursive functionals of type \( \leq 2 \).

Proof. See [14, II.3.18] and [40].

\[ \square \]
7.2. **Application to bar recursion.** Our goal is now to use the equivalences between ordinal induction and $\Sigma^0_n$-induction and an ordinal analysis of bar recursion to establish conservation results of bar recursion over induction along $\omega$.

**Definition 19** (Howard’s bar recursor). Define the bar recursor $B_{\rho,\tau}$ as

$$B_{\rho,\tau}AFGt := \tau \begin{cases} 
Gt, & \text{if } A[t] < \text{lth } t, \\
Ft(\lambda u^\rho. B_{\rho,\tau}AFG(t * u)), & \text{otherwise,}
\end{cases}$$

where $[t] := \lambda x.x(t)_x$.

**Definition 20** (restricted bar recursor).

$$\Phi'_{\rho}AFGt := \tau \begin{cases} 
Gt, & \text{if } A[t] < \text{lth } t, \\
Ft(\Phi'_{\rho}AFG(t * 0))(\Phi'_{\rho}AFG(t * 1)), & \text{otherwise.}
\end{cases}$$

The bar recursor $\Phi'_{\rho}$ can be used to solve the functional interpretation of WKL, see [18]. ($\Phi'_{\rho}$ is the restricted bar recursor schema 1 from there.)

We call a term **semi-closed** if it contains only variables of degree $\leq 1$ free. Howard introduced the notion of computational size for semi-closed terms, see [17, 18]. Roughly speaking the computation size of a semi-closed term of type 0 is an upper bound on the number of term reductions on has to apply to obtain a numeral. The computational size of a degree 1 term is the computational size of $t(H_0, \ldots, H_n)$, where $H_i$ are fresh variables such that the terms is of type 0.

**Theorem 21** ([18, 2.2, 2.3]). Let $\Phi'_{\rho}AFGc$ resp. $B_{0,1}AFGc$ be a semi-closed term and let $A, F, G$ have the computational sizes $a, f, g$ then

(i) $\Phi'_{\rho}AFGc$ has computational size $\sigma := (f + g + h)\omega + \omega(h + 1)$, where $h := wa + \omega$ and,

(ii) $B_{0,1}AFGc$ has computational size $\sigma := \omega^{a+f+2h}$, where $h := wa + \omega$.

This equivalence can be proven in $\Sigma^0_n$-LNP($\sigma$).

**Proof.** See the proofs in [18, 2.2, 2.3]. Note that these proofs actually define a counting function for the computation-tree through transfinite recursion. This recursion is essentially a transfinite primitive recursion over $\sigma$. Hence this proof may be carried out in $\Sigma^0_n$-LNP($\sigma$). $\square$

**Remark 22.** If we apply the rule of bar recursion to semi-closed, primitive recursive terms (in the sense of Kleene, i.e. terms of computation size $\omega^n$ for $n \in \omega$) we obtain a term with computation size $< \omega^{m+\omega}$ for an $m \in \omega$ and therefore a term that is provably definable already in WE-PA$^\omega$ for an $l \in \omega$ or in WE-PA$^\omega$ + $\Sigma^0_n$-lA. We can carried out the proof of the equivalence, theorem 21, in the same system, see theorem 17. Hence in each of these systems we can also proof the equivalence of both terms.

If we apply the rule of restricted bar recursion to primitive recursive terms, which contain only free variable of type 0, we even end up with a primitive recursive term.

8. **Term-normalization**

Denote by $T_{(k)}[F_0, \ldots, F_{n-1}]$ the extension of the system $T_k$ resp. $T$ with the constants $F_0, \ldots, F_{n-1}$. Further we treat here $R_\rho$ as an unspecified constant (without $R_\rho$ axioms) in the case of qf-$G_\omega$A$^\omega$.

In the following we will call the reduction of

$$\text{Cond}_{\rho(\tau)}(x, y, z)u^\tau \text{ to } \text{Cond}_{\rho}(x, yu, zu)$$

a Cond-reduction. These Cond-reductions are provably valid in qf-$N$-$G_\omega$A$^\omega$. 
**Theorem 23** (term-normalization for type 2). Let \( F_i \) be constants of type \( \leq 2 \).

For every term \( t^i \in T_0[\text{Cond}_\rho]_T, F_0, \ldots, F_{n-1} \) there is provably in qf-N-\( \Gamma_0 \mathcal{A}^\omega \) a term \( \tilde{t} \in T_0[\text{Cond}_0, F_0, \ldots, F_{n-1}] \) for which
\[
\forall x \; tx =_0 \tilde{t}x
\]
and where every occurrence of an \( F_i \) is of the form
\[
F_i(\tilde{t}_0[y^0], \ldots, \tilde{t}_{k-1}[y^0]).
\]
Here \( k \) is the arity of \( F_i \), and \( \tilde{t}_j[y^0] \) are fixed terms whose only free variable is \( y^0 \).

**Proof.** Without loss of generality we take the system \( T_0[F] \) where \( F \) is of type 2. For notational simplication we assume that the recursor \( R_0 \) can be obtained from \( F \). This can always be achieved with coding.

Let \( t^i \) be a term in \( T_0[F] \). The term \( tx \), where \( x \) is a fresh variable, is \( =_0 \) equal to a term \( t'[x] \) where \( t' \) results from \( tx \) by carrying out all possible \( \Pi \)-, \( \Sigma \)-, and Cond-reductions. The outermost symbol of \( t' \) cannot be \( \Pi \), \( \Sigma \), or \( \text{Cond}_\rho \) with \( \rho \neq 0 \), since otherwise in \( t' \) either not all \( \Pi \)-, \( \Sigma \)-reductions had been carried out or \( t' \) would not be of type 0.

Hence one of the following holds:
1) \( t'[x] = 0^0 \)
2) \( t'[x] = \lambda x.t'[x] \)
3) \( t'[x] = F(t'_0[x]) \)
4) \( t'[x] = \text{Cond}_0(t'_0[x], t'_j[x], t'_k[x]) \)

In the first case we are done, \( \lambda x.t'[x] \) satisfies the theorem. In the second case we proceed the same way with the term \( t_0. \) In the third case we proceed with the term \( t_0 y^0 \) where \( y^0 \) is a new variable making \( t_0 \) to type 0 and in the fourth case we proceed with the terms \( t_0, t_0 y^0, t_0 y^0. \) Note that we can code the variables \( x \) and \( y \) in one type 0 variable. Also note that since we applied all Cond-reductions only \( \text{Cond}_0 \) occurs.

By the strong normalization theorem this process stops, yielding the desired term, see e.g. [12].

**Theorem 24** (term-normalization for type 3). Now let \( G_i \) be constants of type \( \leq 3 \).

For every term \( t^i \in T_0[\text{Cond}_\rho]_T, G_0, \ldots, G_{n-1} \) there is provably in qf-N-\( \Gamma_3 \mathcal{A}^\omega \) a term \( \tilde{t} \in T_0[\text{Cond}_0, G_0, \ldots, G_{n-1}] \) for which
\[
\forall x \; tx =_0 \tilde{t}x
\]
and where every occurrence of an \( G_i \) is of the form
\[
G_i(t_0[f^1], \ldots, t_{k-1}[f^1]).
\]
Here \( k \) is the arity of \( G_i \), and \( \tilde{t}_j[f^1] \) are fixed terms whose only free variable is \( f^1 \).

**Proof.** Analogous to proof of theorem 23. See also [29, proof of proposition 4.2].

Note that the equality between \( t, \tilde{t} \) is only pointwise. Therefore one needs (weak) extensionality to conclude that \( s[t] =_0 s[\tilde{t}] \) for an arbitrary term \( s \).

**Application to proofs in quantifier-free systems.** For a term \( t \) call the term where every maximal type 0 subterm (i.e. every subterm of type 0 which is not included in a different subterm of type 0) is replaced by a fresh type 0 variable skeleton. Obviously, \( t \) can be regained from its skeleton by substitution of type 0 terms.

**Lemma 25.** Let \( \mathcal{T} \) be qf-\( \Gamma_n \mathcal{A}^\omega \) with \( n \geq 3 \) or qf-\( \Gamma-\mathcal{PA}^\omega \) augmented with arbitrary constants \( H_0, H_1, \ldots \), let \( t_0, t_1 \in T_0[\text{Cond}_0, H_0, H_1, \ldots] \) and in \( t_0, t_1 \) all possible \( \Pi \)-, \( \Sigma \)-reductions have been carried out.

Then the following are equivalent:
(i) The terms $t_0, t_1$ are provable equal in every term context ($T \vdash s[t_0] =_0 s[t_1]$ for every term $s$).

(ii) $T \vdash P(t_0) =_0 P(t_1)$, where $P$ is a variable of suitable type.

(iii) The terms $t_0, t_1$ have the same skeleton (modulo renaming of type 0 variables) and $t_0, t_1$ are obtained from the skeleton by substitution of $=_0$-equal terms.

Proof. (i) $\Rightarrow$ (ii) is clear. (ii) $\Rightarrow$ (i) follows from the fact that one can replace $P$ by any term in the derivation and so in particular by $\lambda x.s[x]$. By definition of the axioms of a quantifier-free system the axioms of this new derivation are also in $T$.

For (ii) $\Rightarrow$ (iii) observe that the only way to prove the equality in (ii) are the $=_0$-axioms, $\text{SUB}_{\text{Cond}}$-axioms for $\text{Cond}_0$, and the $=_{\omega}$-axioms for $\text{R}_0$ are used. These axioms only change type 0 values and, therefore, the skeletons have to be the same. The lemma follows.

Proposition 26. Let $T$ be $\text{qf-N-G_nA}^\omega$ where $n \geq 3$ or $\text{qf-\overline{N-PA}^\omega}$ augmented by a type 2 constant $F$. Further let $A$ be a formula containing only type 0 variables free and satisfying $T \vdash A$.

Then there exists a formula $\tilde{A}$ such that the weakly extensional intuitionistic system $T_{\text{WE}}$ corresponding to $T$ (i.e. $G_nA^\omega$ or $\text{WE-HA}^\omega$) proves $A \iff \tilde{A}$ and that there is a derivation $\tilde{D}$ of $T \vdash \tilde{A}$ where every occurring term is normalized according to theorem 23, i.e. each occurrence of $F$ is of the form $F(t_i[x])$.

Moreover, these applications of $F$ can be chosen independently from each other in the sense that

$$T \not\vdash P[F(t')] =_0 P[F(t'')]$$

for all type 0 substitution instances $t', t''$ of $t_i$ resp. $t_j$ with $i \neq j$. (In other words, the theory $T$ does not see that the $F(t')$, $F(t'')$ are applications of $F$ and not just an arbitrary term of suitable type and with the same free variables. Hence they may be replaced independently.)

Using coding we may also allow finitely many constants $F_i$ of type $\leq 2$.

Proof. Let $D$ be a derivation of $T \vdash A$. By lemma 2 we may assume that only the variables of $A$ and some free type 0 variables occur in $D$. Hence every term showing up in $D$ satisfies the premise of theorem 23.

We obtain a new derivation $\tilde{D}$ by replacing every term in $D$ with its normal form as defined in the proof of theorem 23 (in particular all possible $\Pi$- and $\Sigma$-reductions have been carried out and only $\text{Cond}_0$ occurs in $\tilde{t}$). The derivation $\tilde{D}$ is still valid because the used logical axioms and rules, $\text{SUB}$-axioms for the recursor and $\text{Cond}$, $=_{\omega}$-axioms, and quantifier-free induction rule are translated into other instances of themselves. The used $\text{SUB}$-axioms for $\Pi$ and $\Sigma$ become trivial since in all terms all possible $\Pi$- and $\Sigma$-reductions have been carried out.

Let $\tilde{A}$ be the result of $\tilde{D}$. Each term occurring in $\tilde{A}$ is just the normal form of the term at the same position in $A$ and therefore weakly extensional equal to it. Hence

$$T_{\text{WE}} \vdash A \iff \tilde{A}.$$

Obviously, the derivation $\tilde{D}$ contains only finitely many applications $t_i$ of $F$. Each of the $t_i$ contains only type 0 variables free. However, these applications of $F$ are not independent from each other because there might be equalities between them provable in $T$. 

The infinite pigeonhole principle is equivalent to for each \( \forall n \ \Theta \rightarrow \exists l \\exists R \ \forall G \subseteq^\ast R_n \cup G \subseteq^\ast R_n \), i.e.
\[
\forall n \exists s \ (\forall j \geq s (j \in G \rightarrow j \in R_n) \lor \forall j \geq s (j \in G \rightarrow j \notin R_n)).
\]

A set \( G \) is strongly cohesive for \((R_n)_{n \in \mathbb{N}}\) if
\[
\forall n \exists s \forall i < n (\forall j \geq s (j \in G \rightarrow j \in R_i) \lor \forall j \geq s (j \in G \rightarrow j \notin R_i)).
\]

The cohesive principle (COH) is the statement that for every sequence of sets an infinite strongly cohesive set exists. Similarly the strong cohesive principle (StCOH) is the statement that for every sequence of sets an infinite strongly cohesive set exists. We denote by \((St)COH(r, G)\) the statement that \( G \) is a set that satisfies the (strong) cohesive principle for the sets given by the characteristic functions \((\lambda x.r(i, x))_i\), where \( r : \mathbb{N} \times \mathbb{N} \rightarrow 2 \).

**Proposition 28** ([15, 4.4]).
1. \( G_3^{-} \models StCOH \rightarrow COH \)
2. \( G_3^{-} \models StCOH \rightarrow \Pi^1_1\)-CP
3. \( G_3^{-} \models StCOH \leftrightarrow COH \land \Pi^1_1\)-CP

**Proof.** The first statement is clear and the third statement is an immediate consequence of the first and second.

For the second we prove the infinite pigeonhole principle \( RT^3_{\infty} \) from \( StCOH \). The infinite pigeonhole principle is equivalent to \( \Pi^1_1\)-CP, over \( \Sigma^0_1\)-induction. This was shown in [16]. The proof can even be carried out in \( G_3^\alpha \), see [36]. Let \( f : \mathbb{N} \rightarrow \mathbb{N} \) be a coloring. Define \( R_i := \{x \mid f(x) = i\} \). Let \( G \) be an infinite, strongly cohesive set for \( R_i \). By definition there is an \( s \) with
\[
\forall i < n \ (\forall j \geq s (j \in G \rightarrow j \in R_i) \lor \forall j \geq s (j \in G \rightarrow j \notin R_i)).
\]

By the totality of \( f \) there is exactly one \( i \) such that the first disjunction holds, i.e. the color \( i \) occurs infinitely often on \( G \) and thus on \( \mathbb{N} \).

**Lemma 29.** \( G_3^\alpha \) proves that a countable number of instances of \((St)COH\) is uniformly equivalent to a single instance of \((St)COH\).

**Proof.** Let \((R_{j,i})_{i \in \mathbb{N}}\) be a sequence of sequences of sets. A set which is (strongly) cohesive for all of these sets is obviously also (strongly) cohesive for the sets \((R_{j,i})_{i \in \mathbb{N}}\) for each \( j \). Hence a single application of \((St)COH\) is sufficient to solve the sequence of instance of \((St)COH\) given by \((R_{j,i})_{i \in \mathbb{N}}\) for each \( j \).

**Proposition 30.**
\[
G_3^\alpha \models QF-AC \oplus WKL \vdash \forall r : \mathbb{N} \times \mathbb{N} \rightarrow 2 \ (\Pi^1_1\)-CA \((\xi r) \rightarrow \exists G \ StCOH(r, G))
\]
where \( \xi \) is a suitable term.
Proof. Define

\[ R_n := \lambda x. r(n, x), \quad R^x := \bigcap_{i<l(x)} \left\{ R_i \mid \frac{R_i}{\overline{R}_i} \right\} \text{ if } x_i = 0, \text{ otherwise.} \]

Here and in the following let \( x \) be the code of the sequence \( \langle x_0, \ldots, x_{\text{th}(x)-1} \rangle \).

For every \( n \) the set (of sets) \( \{ R^x \mid x \in \mathbb{N} \} \) is a partition of \( \mathbb{N} \), i.e.

\[ \forall n \forall z \exists x \in \mathbb{N} \; \exists z \in R^x. \]

This statement can be proved with an instance of quantifier-free induction (the tuple \( \langle x_0, \ldots, x_{n-1} \rangle \) is bounded by \( T_n \) and \( z \) is a parameter).

We construct an infinite \( \Pi^0_1 \)-tree \( T \) deciding at level \( n \) whether for the solution set \( G \) either \( G \subseteq^* \overline{R}_n \) or \( G \subseteq^* R_n \) holds: Let

\[ T(\langle x_0, \ldots, x_n \rangle) \iff R^{\langle x_0, \ldots, x_n \rangle} \text{ is infinite.} \]

The statement “\( R^x \) is infinite” is \( \Pi^0_1 \). The predicate \( T \) clearly defines a tree. The tree is infinite because otherwise

\[ \exists n \forall x \in \mathbb{N} \exists y \forall z \; y > z \notin R^x \]

and this together with an instance of \( \Pi^0_1 \)-CP yields a contradiction to (11). \( x \) can be bounded by \( 1n \).

With an application of an instance of \( \Sigma^0_1 \)-induction we prove

\[ \forall x \left( R^x \text{ infinite} \rightarrow \forall n \exists \langle l_0, \ldots, l_{n-1} \rangle \left( \forall i < n - 1 \; l_i < l_{i+1} \land \forall i < n \; l_i \in R^x \right) \right) \]

and then conclude

\[ \forall n \forall x \left( \text{th}(x) = n \right. \]

\[ \left. \wedge R^x \text{ infinite} \rightarrow \exists \langle l_0, \ldots, l_{n-1} \rangle \forall i < n - 1 \; l_i < l_{i+1} \land \forall i < n \; l_i \in R^x \right). \]

An instance of \( \Pi^0_2 \)-WKL yields an infinite branch \( b \) of \( T \), i.e. \( \forall n \left( R^{x(n)}_n \text{ infinite} \right) \).

Using (12) we obtain

\[ \forall n \exists \langle l_0, \ldots, l_{n-1} \rangle \left( \forall i < n - 1 \; l_i < l_{i+1} \land \forall i < n \; l_i \in R^{x(n)} \subseteq R^{x(n)} \right). \]

An application of QF-\( \text{AC} \) yields an enumeration \( n \mapsto \langle l_0, \ldots, l_{n-1} \rangle \) of finite tuples. Searching for the least code of a tuple and the properties of (13) assure that every tuple is extended by the following. Hence we may diagonalize to obtain an the set \( G := \{ l_0, l_1, \ldots \} \). This set is strongly cohesive and solves the proposition.

Note that the instances of \( \Sigma^0_1 \)-IA, \( \Pi^0_1 \)-CP and \( \Pi^0_2 \)-WKL can be reduced to an instance of \( \Pi^0_1 \)-CA using lemma 10 and remark 9 yielding a suitable term \( \xi \). □

We now strengthen this proposition to

**Proposition 31.** For every closed term \( \varphi \) one can construct a closed term \( \xi \) such that

\[ G_{\infty}A^\omega + \text{QF-AC} \oplus \text{WKL} \vdash \forall r: \mathbb{N} \times \mathbb{N} \rightarrow 2 \left( \Pi^0_1 \text{-CA}(\xi r) \rightarrow \exists G \left( \text{StCOH}(r, G) \land \Pi^0_1 \text{-CA}(\varphi r G) \right) \right). \]

**Proof.** We construct an infinite \( \Pi^0_1 \)-0/1-tree, in which we decide at level

- \( 2n \) whether \( G \subseteq^* \overline{R}_n \) or \( G \subseteq^* R_n \) and at level,
- \( 2n + 1 \) the \( n \)-th value of the instance of \( \Pi^0_1 \)-comprehension, i.e. whether \( \forall k (\varphi r G)nk = 0 \) is true.
We assign to every element of the tree a finite (potential) initial segment $L^x$ of $G$. At level $2n$ we add — as in the previous proposition — the next element of $R^x$; at level $2n + 1$ we only add the smallest counterexample (extending our old initial segment of $G$ with elements from $R^x$) to the statement $\forall k (\neg \varphi r G)nk = 0$ if it is false and nothing otherwise. Define:

\[
T((x_0, \ldots, x_{2n})) \text{ iff } R^{(x_0, x_2, \ldots, x_{2n})} \text{ is infinite},
\]

\[
T((x_0, \ldots, x_{2n}, 0)) \text{ iff } \forall l \subseteq \text{fin } R^{(x_0, x_2, \ldots, x_{2n})} \exists k \alpha_\varphi(L^{(x_0, \ldots, x_{2n})} * l, n, k) \leq 1,
\]

\[
T((x_0, \ldots, x_{2n}, 1)) \text{ iff } \exists l \subseteq \text{fin } R^{(x_0, x_2, \ldots, x_{2n})} \exists k \alpha_\varphi(L^{(x_0, \ldots, x_{2n})} * l, n, k) > 1,
\]

\[
L_0 := \langle \rangle,
\]

\[
L^{(x_0, \ldots, x_{2n})} := L^{(x_0, \ldots, x_{2n-1})} * \left\{ \min \left\{ x \in R^{(x_0, x_2, \ldots, x_{2n})} \mid x > \max L^{(x_0, \ldots, x_{2n-1})} \right\} \right\},
\]

\[
L^{(x_0, \ldots, x_{2n}, 0)} := L^{(x_0, \ldots, x_{2n})},
\]

\[
L^{(x_0, \ldots, x_{2n}, 1)} := L^{(x_0, \ldots, x_{2n})} * l,
\]

\[
k^{(x_0, \ldots, x_{2n}, 1)} := k,
\]

\[
k^x := 0 \text{ for all } x \text{ not of this form},
\]

where $(l, k)$ minimal with

\[
l \sqsubset R^{(x_0, x_2, \ldots, x_{2n})} \land \alpha_\varphi(L^{(x_0, \ldots, x_{2n})} * l, n, k) > 1.
\]

For notational simplification we omitted the requirements to make $T$ closed under prefix, but we can simply add the conditions of the previous levels to the definition of $T$ making it a tree.

$L^x$ and $k^x$ is clearly defined if $T(x)$ is true (use an instance of $\Sigma^0_1$-induction to show this — weaken the $\Pi^0_2$-statement “$R^x$ is infinite” in the definition of $T$ to $\exists z \in R^x$).

Using the same argument as in the previous proposition we see that the tree is infinite. But we cannot apply $\Sigma^0_1$-WKL$(\xi r)$, because this instance contains $L$, which is in general not computable in $r$ (in the sense of $G_\infty A^\infty$).

The graph of $x \mapsto (L^x, k^x)$ is definable and $\Delta^1_1$. For notational ease we define the graph of its course-of-value function:

\[
((x_0, \ldots, x_n), (L_0, \ldots, L_n), (k_0, \ldots, k_n)) \in G_{L,k} \quad \text{iff}
\]

\[
n = 0: \ L_n = \langle \rangle, \ k = 0,
\]

\[
n \text{ even: } \ L_n = L_{n-1} * \langle y \rangle, \ k_n = 0
\]

where $y$ minimal with $y \in R^{(x_0, \ldots, x_{2n-1})} \land y > \max (L_{n-1}),$

\[
n \text{ odd and } x_n = 0: \ L_n = L_{n-1}, \ k_n = 0,
\]

\[
n \text{ odd and } x_n = 1: \ L_n = L_{n-1} * l \text{ and } (l, k_n) \text{ minimal with } l \sqsubset R^{(x_0, x_2, \ldots, x_{2n})} \land \alpha_\varphi(L_n * l, (n-1)/2, k_n) > 2.
\]

(Note that equations like $L_n = L_{n-1} * l$ we omitted for notational ease the bounded quantifier $\exists l < L_n$ for $l$.) So we can replace every reference to $L^x$ in the definition of $T$ by

\[
\exists k, y (x, (y, k)) \in G_{L,k} \quad \text{or} \quad \forall k, y (x, (y, k)) \in G_{L,k}.
\]

The resulting tree is still $\Pi^0_2$ so we may apply an instance of $\Pi^0_2$-WKL and obtain an infinite branch $b$.

Setting $G := \bigcup_n L^{(n)}$ now enumerates an infinite strongly cohesive set and from $b$ we can decide $\forall k (\neg \varphi r G)nk = 0$ for every $n$. \(\square\)
Corollary 32 (to the proof). For every system \( \mathcal{T} \) containing \( G_\infty \mathrm{A}_\omega \) and every provably continuous term \( \varphi \) there exists a term \( \xi \), such that
\[
\mathcal{T} + \mathrm{QF-AC} \oplus \mathrm{WKL} \vdash
\forall r : \mathbb{N} \times \mathbb{N} \rightarrow 2 \left( \Pi^1_1\mathrm{-CA}(\xi r) \rightarrow \exists G \left( \mathrm{StCOH}(r, G) \land \Pi^1_1\mathrm{-CA}(\varphi r G) \right) \right).
\]

Corollary 33. \((\mathrm{St})\mathrm{COH}\) is proofwise low in sequence over \( G_\infty \mathrm{A}_\omega + \mathrm{QF-AC} \oplus \mathrm{WKL}\).

**Proof.** Lemma 29 and proposition 31 (with corollary 32). \(\square\)

Our goal is now to interpret consequences (of the form \( \forall x^1 \exists y^0 A_g(x, y) \)) of a principle \( \mathcal{P} \) that is proofwise low in sequence. For this we will strengthen \( \mathcal{P} \) to the statement that there exists a uniform solution functional \( \mathcal{P} \) for \( \mathcal{P} \). The functional \( \mathcal{P} \) must be of type \( \leq 2 \), such that after extracting terms using the functional interpretation one can normalizing them with the tools of Section 8. With this we will see that \( \mathcal{P} \) is only used finitely many times and can be replaced using the lowness property in favor of an instance of \( \Pi^1_1\mathrm{-CA} \).

The properties of the solution functional \( \mathcal{P} \) must be axiomatizable universally, since they will become an implicative assumption. After prenexation they will become purely existential and the functional interpretation will extract terms witnessing them. Existential quantifier in the axiomitation of \( \mathcal{P} \) would become universal after prenexation and therefore would need to be presented afterward.

If \( \mathcal{P} \) is of the form
\[
\forall S \exists G \forall x P_g(S, G, x),
\]
where \( P_g \) is quantifier-free. Then one can take for \( \mathcal{P} \) the Skolem functional for \( G \), i.e. a functional \( \mathcal{P} \) satisfying
\[
\forall S \forall x P_g(S, \mathcal{P}(S), x).
\]

With the help of the following lemma we can obtain a functional for \( \mathcal{P} \) where \( P \) is a \( \Pi^0_1 \) formula. This is sufficient for \( \mathrm{StCOH} \).

**Lemma 34.** Let \( \mathcal{P} \) be a principle proofwise low in sequence over \( G_\infty \mathrm{A}_\omega \) + QF-AC + WKL, that has the form
\[
(\mathcal{P}) : \forall S \exists G \forall x \exists y \forall z P_g(S, G, x, y, z),
\]
where \( P_g \) is quantifier-free.

Then the principle
\[
(\mathcal{P}') : \forall S \exists G, Y \forall Z^1 \forall x P_g(S, G, x, Y(x, Z), Z(Y(x, Z))
\]
is proofwise low in sequence, in the sense that for every closed term \( \varphi \) a closed term \( \xi \) exists, such that \( \Pi^1_1\mathrm{-CA}(\xi(S_i)_1(Z_i)_i) \) proves
\[
\exists(G_i)_i, (Y_i)_i (\forall i, Z', x P_g(S_i, G_i, x, (Y)_i(x, Z'), Z'(Y_i(x, Z')) \land 
\Pi^1_1\mathrm{-CA}(\varphi(S_i)(Z_i)(G_i)(\lambda x. Y_i(x, Z_i)_i))).$
\]

**Proof.** The lowness of \( \mathcal{P} \) provides that for every term \( \varphi' \) an instance of \( \Pi^1_1\mathrm{-comprehension} \( \Pi^1_1\mathrm{-CA}(\xi SZ) \) proves
\[
\exists G \left( \forall x^0 \exists y^0 \forall z^0 P_g(S, G, x, y, z) \land \Pi^1_1\mathrm{-CA}(\varphi' SZG) \right).
\]
Hence it also proves
\[
\exists G \left( \forall x, Z \exists y P_g(S, G, x, y, Z(y)) \land \Pi^1_1\mathrm{-CA}(\varphi' SZG) \right).
\]
By searching for the least \( y \) we may assume that there exists a unique \( y \) for each \( x, Z \). Let \( Y(x, Z) \) be the choice function for \( y \) obtained using \( \text{QF-AC} \). To show that \( \mathcal{P}' \) is provewise low it suffices to show for every closed \( \varphi \) that there is a closed \( \varphi' \) (and thus a closed \( \xi \)) such that \( \Pi^0_1\text{CA}(\varphi S Z G(\lambda x. Y(x, Z))) \) is provable from \( \Pi^0_1\text{CA}(\varphi S Z) \).

Since \( Y \) is computable in \( S, G \) a suitable \( \varphi \) can easily be constructed with the same generic construction used in the proof of lemma 11.

One also easily verifies that the whole argumentation is stable under sequences and hence that \( \mathcal{P}' \) is provewise low in sequence.

\[ (18) \quad \text{qf-N-G} = \text{A}^{\omega} \vdash (P_5(\mathcal{P}, t_\mathcal{P}(x, \mathcal{P}, R_0, B)) \land (R_0)_{qf}(R_0, t_{R_0}(x, \mathcal{P}, R_0, B)) \land (B)_{qf}(B, t_B(x, \mathcal{P}, R_0, B)) \Rightarrow A_{qf}(x, t_{qf}(x, \mathcal{P}, R_0, B))), \]

see [51, 26].

The terms \( t_\mathcal{P}(x, \mathcal{P}, R_0, B), t_{R_0}(x, \mathcal{P}, R_0, B), t_B(x, \mathcal{P}, R_0, B), t_{qf}(x, \mathcal{P}, R_0, B) \) have type \( \leq 1 \). By proposition 26 we obtain a new derivation in \( \text{qf-N-G} = \text{A}^{\omega} \) of a sentence which is equivalent to (18) over \( \text{qf-G} = \text{A}^{\omega} \) and where each application of \( \mathcal{P} \) is of the form \( \mathcal{P}(t_i[z^n_0]) \) or a substitution instance of \( \mathcal{P}(t_i[z^n_0]) \) and \( \mathcal{P}(t_j[z^n_0]) \) and \( \mathcal{P}(t_j[z^n_0]) \) are independent (in the sense of proposition 26). Same for \( R_0, B \).

Our goal is now to replace these occurrences of \( \mathcal{P}, R_0, \) and \( B \) in the normalized derivation of (18) by a low solution to those principles, such that the premise of (18) becomes provable.

We proceed by inductively over the nesting-depth of \( \mathcal{P}, R_0, \) and \( B \) replacing the applications (and their substitution instances) with low solutions retaining the
instance of comprehension. This operation leaves the derivation valid since the
different applications are independent. Concretely we replace \( P, R_0, B \) by the fol-
lowing:

- \( R_0(t_1[z]) \) simply defines a primitive recursive function, which is provably
total using an instance of \( \Sigma^0_1 \)-induction. This instance can be obtained from
\( \text{QF-I} \) and an instance of \( \Pi^0_1 \)-comprehension. Then lemma 11 yields a new
instance of comprehension (which allows \( R_0(t_1[z]) \) as parameter).
- \( P(t_1[z]) \) can be handled using the assumption that \( P \) is proofwise low in
sequence (lemma 34)
- \( B(t_1[z]) \) can trivially be handled because it is present in the verifying sys-
tem.

For the construction of these replacements we work in the system \( G_\omega A^\omega \), i.e. with
weak extensionality and quantifiers. After this the premise of (18) becomes prov-
able. Quantifying over all \( x \) and coding \( x, z \) together into a new variable \( x \), yields
the proposition without \( \Pi^0_1 \)-CP.

To prove the full proposition note that we can add \( \text{StCOH} \) to the system since it
is proofwise low in sequence, see corollary 33, and that \( \text{StCOH} \) implies \( \Pi^0_1 \)-CP, see
proposition 28. This completes the proof. □

**Theorem 36** (Conservation for proofwise low in sequence). Let \( P \) be a principle
of the form (15) that is proofwise low in sequence over \( G_\omega A^\omega + \text{QF-AC} \oplus \text{WKL} \). In
particular, this includes all principles of this form proofwise low in sequence over
\( \text{WKL}^*_\omega \). If

\[
\text{WE-HA}^\omega \models \forall x^1 A_\psi(x,z) \quad \text{and} \quad x \text{ is of type 0 we have PRA} \models \forall x A_\psi(x,tx).
\]

then one can extract a primitive recursive term \( t \) such that

\[
\text{WE-HA}^\omega \models \forall x^1 A_\psi(x,tx).
\]

In particular, if \( A_\psi \in \mathcal{L}(\text{PRA}) \) and \( x \) is of type 0 we have PRA \( \vdash \forall x A_\psi(x,tx) \).

Proof. We may assume that \( A_\psi \in \mathcal{L}(G_\omega A^\omega) \). Otherwise it would contain \( R_0 \). If this
is the case we normalize every term occurring in \( A_\psi \) and replace every occurrence of
\( R_0uvw \) by a fresh variable that will be \( \exists \)-quantified. There are no other occurrence
of \( R_0 \) in \( A_\psi \) since it contains (beside \( \Pi, \Sigma \)) no constant of type \( > 2 \). These fresh
variables will hold the value of \( R_0uvw \). This values exists provably with \( \Sigma^0_1 \)-IA
and can be expressed in a quantifier-free way.

Apply now elimination of Skolem function for monotone formulas (corollary 13)
to the result of proposition 35. □

**Corollary 37.** Especially from a proof of

\[
\text{WE-HA}^\omega \models \forall x^1 A_\psi(x,z) \quad \text{one can extract a primitive recursive term} \ t \ \text{such that}
\]

\[
\text{WE-HA}^\omega \models \forall x^1 A_\psi(x,tx).
\]

Proof. Theorem 36 and corollary 33. □

**Corollary 38.** The system \( \text{WKL}^*_\omega + \Pi^0_1 \text{-CP} + \text{COH} \) is \( \Pi^0_1 \)-conservative over \( \text{PRA} \).
Additionally, for every \( \Pi^0_2 \)-sentence one can extract uniformly a primitive recursive
(provably) realizing term.

Further \( \text{WKL}^*_\omega + \Pi^0_1 \text{-CP} + \text{COH} \) is conservative over \( \text{RCA}^\omega_0 \) for sentences of the form
\( \forall x^1 \exists y^1 \forall z^0 A_\psi(x,y,z) \).

As consequence we also obtain that \( \text{WKL}^*_\omega + \Pi^0_1 \text{-CP} + \text{COH} \) is conservative over
\( \text{RCA}^\omega_0 \) for sentences of the form \( \forall X \exists y \forall z A(X,y,z) \), where \( A \) is \( \Delta^0_1 \), and thus in
particular is \( \Pi^0_2 \)-conservative.
Lemma 41. Stable Ramsey’s theorem for pairs (10.1. principle).

The point convergence but possibly without a computable rate of convergence, i.e.

This in some sense is the best possible result since RCA0 + Π10-CP is not Σ30-conservative over a theory containing only Σ20-induction, see [1].

Remark 39. Recall that BW is the statement that every bounded sequence \((y_i)_{i \in \mathbb{N}}\) of reals contains a subsequence \((y_{f(i)})_{i \in \mathbb{N}}\) converging with the rate \(2^{-n}\), i.e. \(\forall n \forall i, j \geq n \ |y_{f(i)} - y_{f(j)}| < 2^{-n}\), see [45]. It turns out that StCOH is equivalent to a natural variant of this principle, namely the statement that each bounded sequence \((y_i)_{i \in \mathbb{N}}\) of reals contains a Cauchy subsequence \((y_{f(i)})_{i \in \mathbb{N}}\). This means a sequence which converges but possibly without a computable rate of convergence, i.e. \(\forall n \exists k \forall i, j \geq k \ |y_{f(i)} - y_{f(j)}| < 2^{-n}\), see [37]. Hence the term extraction results we obtain below for StCOH also apply to this variant of the Bolzano-Weierstraß principle.

10. RAMSEY’S THEOREM FOR PAIRS

10.1. Stable Ramsey’s theorem for pairs (SRT2 1). An \(n\)-coloring \(c : [\mathbb{N}]^2 \to n\) is called stable if

The point \(k\) is called a stability point for \(x\).

We call an \(n\)-coloring strongly stable if

Over \(\Pi1\)-CP strongly stable and stable coincide. Even an instance of the collection principle of the form \(\Pi1\)-CP(\(\xi(c)\)) where \(\xi\) is a suitable term and \(c\) the coloring suffices to prove this equivalence.

Let \(\text{SRT}_n^2\) be the statement expressing that every stable \(n\)-coloring of pairs has an infinite homogeneous set and let \(\text{SRT}^2_{\infty} := \forall n \text{SRT}_n^2\). For a stable \(n\)-coloring \(c\) the statement \(\text{SRT}_n^2(c, H)\) denotes that \(H\) is a homogeneous set for \(c\).

The principle \(\text{SRT}^2_2\) is over \(\Sigma_1\)-induction equivalent to the statement that for every \(\Delta_2\)-set \(X\) there exists an infinite set \(Y\) with \(Y \subseteq X\) or \(Y \subseteq X^c\), see [7, 8].

Before we go on with the main result we need some auxiliary lemmata:

Lemma 40 ([7, lemma 4.2]). For every fixed \(n\), let \((\xi_{k, i})_{k < n, i \in \mathbb{N}}\) be a sequence of \(\Pi1\)-sentences of the form \(\forall x A(k, i, x)\) for a quantifier-free \(A\) such that \(\forall i \exists k < n \xi_{k, i}\). Then \(\text{WKL}\) proves that there exists a choice function \(g : \mathbb{N} \to n\) satisfying \(\forall i \xi_{g(i), i}\).

If \(\text{WKL}\) is replaced by \(\Sigma1\)-\(\text{WKL}\) the same holds for \(\Pi1\)-sentences.

Proof. Define

The function \(f\) clearly defines a \(\Pi1\)-0-\(n\)-tree and is by assumption infinite.

Via the equivalence of 0-\(n\)-trees and 0/1-trees and of \(\Pi1\)-\(\text{WKL}\) and \(\text{WKL}\) (see [45]), weak König’s lemma yields an infinite branch \(g\) solving the lemma.

Lemma 41 (and definition, \(\Pi1\)-class, [22]). A \(\Pi1\)-class \(A\) of \(2^\omega\) is a set of functions of the form

\(A = \{ f \in 2^\omega \mid \forall n A(\tilde{f}n)\},\)
where $A$ is a quantifier-free formula.

WKL proves that a $\Pi^0_1$-class $A$ is not empty if
\begin{equation}
\forall n^0 \exists s \in 2^n \forall s' \subseteq s \ A(s').
\end{equation}
(The definition of $\Pi^0_1$-class induces an infinite tree in which every $f \in A$ codes an infinite path through it.) The statement (19) is equivalent to a $\Pi^0_1$-statement.

Note that one may also allow $A$ to be a $\Pi^0_1$-formula as the $\forall$-quantifier can be coded into the quantification over $n$ (see for instance [45]).

**Remark 42** (Treatment of $\Pi^0_1$-0/1-trees). Let $T(w) := (\forall k T_{qf}(w, k) = 0)$ be a $\Pi^0_1$-predicate. Using the UWWKL functional $B$ one can define the functional
\[B_{\Pi^0_1}(T_q) := B \left( \min_{w' \in w, k \leq \text{th} w} T_{qf}(w', k) \right)\]
that yields an infinite branch of $T$, if $T$ defines an infinite 0/1-tree.

Furthermore, an instance of $\Pi^0_1$-CA decides whether the tree $T$ is infinite, since
\[\forall n \exists w \in 2^n \forall k T_{qf}(w, k)\]
is equivalent a $\Pi^0_1$-statement (over $G_{\infty} A^\omega + \text{QF-AC}$).

Hence one can treat $\Pi^0_1$-0/1-trees mostly like quantifier free trees.

**Proposition 43.**
\[G_{\infty} A^\omega + \text{QF-AC} \vdash \forall c : \mathbb{N} \times \mathbb{N} \rightarrow 2 \left( \Pi^0_1 \text{-CA}(\xi c) \rightarrow \exists H \ SRT^2_3(c, H) \right),\]
where $\xi$ is a suitable term.

**Proof.** Assume that the coloring $c$ is stable. Define for $i < 2$
\[A_i := \{ x \mid \forall k \exists y \geq k \ c(x, y) = i \}.\]
By stability $A_i = \{ x \mid \exists k \forall y \geq k \ c(x, y) = i \}$. Hence each $A_i$ is a $\Delta^0_2$-set.

At least for one $i$ the set $A_i$ is infinite (by RT$^1_2$). Fix such an $i$. With an instance of $\Pi^0_1$-CP we obtain strong stability, i.e.
\[\forall x \exists k \forall y > k \forall x' \leq x \ c(x', k) = c(x', y).\]
This instance of $\Pi^0_1$-CP follows from a suitable instance of $\Pi^0_1$-CA, see lemma 10.(iv).

Together with the infinity of $A_i$ we get
\[\forall x \exists k \in A_i \forall x' \leq x \ (x' \in A_i \rightarrow c(x', k) = i).\]
Define the set $H$ inductively by $x \in H$ iff $x \in A_i$ and $c(x', x) = i$ for all $x' < x$ with $x' \in H$.

This definition only uses bounded course-of-value recursion in the characteristic function of $A_i$ which can be obtained from a suitable instance of $\Pi^0_1$-CA, see lemma 10.(ii). (The characteristic function $\chi_H$ of $H$ is clearly bounded and hence also its course-of-value function $\chi_H$, which is actually defined in the recursion.)

The set $H$ is clearly infinite and homogeneous. (The two instances of $\Pi^0_1$-CA can be coded into one term $\xi$, see remark 9.)

**Proposition 44.** Let $\varphi c H$ be a term that is provably continuous in $H$, where $\alpha_{\varphi c}(\cdot, n, k)$ is an associate for $\lambda H. \varphi(c, H, n, k)$. Then there exists a term $\xi$, such that
\[\text{WE-PA}^\omega + \text{QF-AC} \oplus (B) \oplus (R_1) \vdash \forall c : \mathbb{N} \times \mathbb{N} \rightarrow 2 \left( \Pi^0_1 \text{-CA}(\xi c) \rightarrow \exists H \ SRT^2_3(c, H) \land \Pi^0_1 \text{-CA}(\varphi c H) \right).\]

If $\varphi c H$ is moreover provably continuous in $c$ the term $\xi$ can be chosen such that it is provably continuous.
Case ii) No partition satisfying (23) exists.

Case i) A partition and at least for one Sketch of proof.

and for some \(i < 2\) the comprehension \(\Pi^0_1\text{-CA}(\varphi_c(G \cap A_i))\) is decided. The set \(H := G \cap A_i\) then solves this proposition.

We will construct the set \(G\) in steps such that at each step \(n\) we will assure that 

\[|G \cap A_i| \geq n \quad \text{for every } i < 2\]

and for some \(i < 2\) the comprehension for \(G \cap A_i\) at the position \((n)\), will be decided, i.e. whether the statement

\[\forall k \, (\varphi_c(G \cap A_i)(n), k) = 0\]

holds. More precisely, we will construct functions \(I, J : \mathbb{N} \rightarrow 2\), such that

\[\exists I, J \forall n \, (\forall k \, (\varphi_c(G \cap A_I(n))(n), k) = 0 \leftrightarrow J(n) = 0).\]

With these functions we can then obtain a comprehension function for one of the sets \(G \cap A_i\), because either

\[\forall m \exists n \, (m = (n)_{I(n)} \land I(n) = 0)\]

and then \(J(N(m))\), where \(N(m)\) is some choice function for \(n\) obtained by QF-AC, decides the comprehension for \(G \cap A_0\) or

\[\exists m \forall n \, (m \neq (n)_{I(n)} \lor I(n) = 1).\]

By choosing \(n = (m, m')\) we obtain \(\forall m' I((m, m')) = 1\) and therefore the function \(\lambda n. J((m, m'))\) decides the comprehension for \(G \cap A_i\).

The set \(G\) and the functions \(I, J\) will be constructed by recursion. We will first give a sketch of the argument and later show that \(R_0\) and the imposed comprehension suffice for the construction.

By induction we construct \((d_n, L_n)\), such that the sequence \((d_n)\) is an ascending sequence of finite sets and \((L_n)\) is a descending sequence of infinite sets of possible candidates to extend \(d_n\) (i.e. \(d_{n+1} \setminus d_n \subset L_n\) and \(\min(L_n)\) is greater than the stability point of \(d_n\)). Each set \(L_n\) is low, in the sense that it can be described by a term containing \(B\) and \(R_1\). The set \(G\) will be given by \(\bigcup_n d_n\).

We start with \((\emptyset, \mathbb{N})\). Assume \((d_n, L_n)\) is already defined. We distinguish two cases:

**Case i)** A partition \(Z_0\) and \(Z_1\) of \(L_n\) exists such that

\[\forall z \subseteq^{fin} Z_i \, (z \text{ is } i\text{-homogeneous} \rightarrow \forall k \, \alpha_{\varphi_c}(d_n \cup z)(n), k \leq 1),\]

where \(d_n = d_n \cap A_i\), holds for all \(i < 2\). (If we extend the initial segment \(d_n\) with elements from \(Z_i\) the comprehension remains true.)

At least one of \(Z_0\) and \(Z_1\) is infinite because \(L_n\) is infinite. We take this set as \(L_{n+1}\), forcing (20) to be true for this \(i\) on all further extensions and let \(d_{n+1} := d_n\).

**Case ii)** No partition satisfying (23) exists.

We know then that especially \(L_n \cap A_0\) and \(L_n \cap A_1\) is no such partition. So we can find for one \(i\) a finite \(i\text{-homogeneous} \, \text{set } d' \subseteq^{fin} A_i\) such that

\[\exists k \, \alpha_{\varphi_c}(d_n \cup d')(n), k > 1.\]

Setting \(d_{n+1} := d_n \cup d'\) and \(L_{n+1} := \{x \in L_n \mid x > \max d'\}\) forces the comprehension function to be \(\neq 0\) at \((n)\).

Note that (23) defines a \(\Pi^0_1\)-class of \(2^\omega\). (We view here a partition of \(\mathbb{N}\) into two sets \(Z_0, Z_1\) as a function \(f \in 2^\omega\) with \(f(n) = i\) iff \(n \in Z_i\).) Thus we may assume that the \(Z_i\) are low and we can decide which case holds by asking if a certain \(0/1\)-tree is infinite (this is a \(\Pi^0_1\)-statement).

The size requirements are met by extending \(d_{n+1}\) with suitable elements of \(L_n\).
The set $G := \bigcup_n d_n$ then satisfies the proposition.

**Proof.** Define

\[
L^0(w) := 0,
\]

\[
L^{(x_0,\ldots,x_{n-1},(d,k,y))}(w) := \begin{cases} 1 & \text{if } w \leq y, \\ sg \left[ B_{\Pi^0_1}(\theta(L^{(x_0,\ldots,x_{n-1}),d}) - (k-1) \right] & \text{if } k \geq 1 \land w > y, \\ L^{(x_0,\ldots,x_{n-1})}w & \text{if } k = 0 \land w > y.
\end{cases}
\]

($d$ is just an auxiliary parameter used to build the tree, it will be set to $d_{n-1}$ defined below; $k$ denotes the case, $k = 0$ for case ii), $k \geq 1$ for case i) and $Z_{k-1}$ infinite in the sketch; $y$ is a lower bound for $L$.)

Here $\theta(B, d^0, d^1)wk$ will be the characteristic function of the predicate

\[
\forall i < 2 \forall y \leq^* \in \mathbb{N} B \cap \{x < \text{lth}(w) \mid (w)_x = i\}
\]

\[
(y \text{ is } i\text{-homogeneous } \rightarrow \alpha_{\varphi}(d^0 \cup y)(n), k \leq 1),
\]

where the variables $w, y$ are numerals coding finite sets. The statement

\[
T_{B,d^0,d^1}(w) := \forall k \theta(B, d^0, d^1)wk = 0
\]

defines the $\Pi^0_1$-1-tree build in (23) in the sketch.

We will write $T_{B,d}$ and $\theta(B, d)wk$ for $T_{B,d^0,d^1}$ resp. $\theta(B, d \cap A_0, d \cap A_1)wk$. This will not lead to problems because $d \cap A_i$ is just a number computable from $d$ relative to the imposed instance of comprehension. Note that $L^x$ can be defined in $B$ and $\theta$ using the bounded iterator $\tilde{R}_1$. Thus the function $L^x$ can be described by a term in this system.

We assume that for all $x$ and $i$ the set $L^x \cap A_i$ is infinite if $L^x$ is infinite. Otherwise the set $L^x \cap [k, \infty]$ for a suitable $k$ would be an infinite subset of $A_{1-i}$ and therefore solve the proposition.

Using this and an instance of $\Delta^0_2$-comprehension (over $L$) we generate functions $g_i$ such that

\[
g_i(x) := \min(L^x \cap A_i).
\]

With an application of $\Pi^0_1$-AC and taking a maximum we obtain a function $h((x_1, \ldots, x_n))$ giving a common stability point of $x_1, \ldots, x_n$.

We now define $(d_n, l_n)$ by recursion. ($L^{l_n}$ should match $L_n$ from sketch above.)

We use primitive recursion in the sense of Kleene, i.e. the recursion can be defined with the recursor $R_0$.

Let $d_0 := ()$ and $l_0 := ()$. For the recursion step we distinguish the cases:

**Case i)** The tree $T_{L^{l_n},d_n}(w)$ is infinite, i.e.

\[
\forall m \exists w \in 2^m \forall k \theta(L^{l_n},d_n)wk = 0.
\]

By RT$_1$ there is at least one $j < 2$ such that $\{x \in \mathbb{N} \mid B_{\Pi^0_1}(\theta(L^{l_n},d_n)|x = j\}$ is infinite. An index $j$ can be chosen constructively relative to $\Sigma^0_1$-WKL, see lemma 40. Set $d_{n+1} := d_n$ and $l_{n+1} := j + 1$.

**Case ii)** The tree $T_{L^{l_n},d_n}(w)$ is finite, i.e.

\[
\exists m \forall w \in 2^m \exists k \theta(L^{l_n},d_n)wk \neq 0.
\]

Then especially the set $A_0$ does not code a path through the tree, i.e. for this $m$

\[
\exists k \theta(L^{l_n},d_n)(\chi_{A_0}m)k \neq 0,
\]
where \( \chi_{A_i} \) is the characteristic function of \( A_i \). So there is an \( i \) and a finite \( i \)-homogeneous set \( y \subseteq^{fin} A_i \cap \{0, \ldots, m-1\} \cap L^y \) such that

\[
\exists k \alpha_{\varphi c}(d^y \cup y)(n), k > 1.
\]

Set

\[
d'_{n+1} := d \cup y \quad \text{and} \quad k'_{n+1} := 0.
\]

Note that this case distinction is constructive relative to the given instance of comprehension (the second quantifier of the formula is bounded).

Now we extend \( d'_{n+1} \) with suitable elements, such that the size requirements are met:

\[
d_{n+1} := d_n \cup \bigcup_{i<2} \{ g_i(l_n \ast (d_n, l', h(d'_{n+1}) + 1)) \}
\]

\[
l_{n+1} := l_n \ast (d_n, k'_{n+1}, h(d_{n+1}) + 1)
\]

Applying \( RT^1_2 \) yields an \( i \) such that all comprehension instances are decided. From the \( d_n \) and the given comprehension one can easily obtain a enumeration of the set \( G \cap A_i := H \).

This solves the proposition. The term \( \xi c \) is continuous in \( c \) because the only discontinuous functional in this system is \( B \) but it is only used to define \( L^c \) and to prove \( WKL \). Hence \( \xi \) can be chosen such that \( c \) does not occur as a parameter to \( B \). More precisely \( \xi c \) is of the form \( \xi[l_1c, \lambda x.L^x] \) with \( \xi, t \in T_0 \) and therefore continuous.

\[ \square \]

**Proposition 45.** Let \( \varphi cH \) be a term that is provably continuous in \( H \) and let \( \alpha_{\varphi c} \) be as in proposition 44. Then there exists a term \( \xi \) such that

\[
\mathbf{WPE-PA}^w + \Sigma^0_2\text{-IA} + \mathbf{QF-AC} \oplus (B) \oplus (\tilde{R}_1) \vdash \\
\forall c : N \times N \rightarrow n \ (\Pi^0_2\text{-CA}(\xi c) \rightarrow \exists H \ SRT^0_{<\infty}(c, H) \land \Pi^0_2\text{-CA}(\varphi cH)) .
\]

If \( \varphi \) is moreover provably continuous in \( c \) the term \( \xi \) can be chosen such that it is provably continuous in \( c \).

**Proof.** Analogous to Proposition 44.

The applications of \( RT^1_2 \) become applications of \( RT^1_{<\infty} \), which is equivalent to \( \Pi^0_2\text{-CP} \) and thus provable using \( \Sigma^0_2\text{-IA} \). The \( 0/1 \)-trees will become \( 0/1 \)-trees; but these trees can be constructively transformed into \( 0/1 \)-trees, see [45].

The only difficult part is adopt the assumption that

\[(27) \quad \forall x \forall i < n \ (L^x \text{ infinite} \rightarrow L^x \cap A_i \text{ infinite}) ,\]

which leads to the definition of \( g_i \) in (26) because we cannot simply deduce the existence of a solution from the failure of (27).

First note that (27) due to the minimal element parameter (\( y \) in (24)) is equivalent to

\[(28) \quad \forall x \forall i < n \ (L^x \text{ infinite} \rightarrow L^x \cap A_i \text{ not empty}) .\]

If (27) resp. (28) does not hold, our goal is to find a set \( L^x \) on which — provided we neglect colors that do not occur — the assumption holds. This can be done by finding a maximal set \( K \subseteq n \), such that there is an \( x \) with \( L^x \cap \bigcup_{k \in K} A_k \) is empty. Then for all \( x' \equiv x \) and \( i \notin K \) the sets \( A_i \cup L^{x'} \) are not empty. Thus if we relativize our argumentation to \( L^x \) and the colors \( n \setminus K \) the condition (27) holds.

To find such a \( K \) and \( x \) define

\[ \eta(s_0, \ldots, s_{n-1}) := \exists x \left( L^x \text{ infinite} \land \bigwedge_i (s_i = 0 \rightarrow L^x \cap A_i = \emptyset) \right) .\]
\( \eta \) is clearly \( \Sigma^0_3 \). Finding a minimal tuple \( \langle s_0, \ldots, s_{n-1} \rangle \) satisfying \( \eta \) yields a suitable solution. A minimal tuple can be obtained using an instance of \( \Sigma^0_3 \)-induction, which is provable from \( \Sigma^0_3 \)-IA and an instance of \( \Pi^0_1 \)-comprehension. \( \Box \)

**Corollary 46.** Let \( \varphi \in H \) be a term that is provably continuous in \( H \). Then there exists a term \( \xi \) such that

\[
\forall c: [N] \times N \to N \quad (\forall c \in [N] \quad (\exists c \in [N] \to \exists H \quad \operatorname{RT}_2^\xi(c, H) \land \Pi^1_1-CA(\varphi c H)).
\]

The term \( \xi \) can be chosen such that \( c \) does not occur as a subterm of a parameter of \( B \).

If \( \Sigma^0_3 \)-IA is added to the system, \( \operatorname{RT}_2^\xi \) may be replaced by \( \operatorname{RT}_{2<\infty}^\xi \).

Hence \( \operatorname{RT}_2^\xi \) is provewise low over \( \text{WE-PA}^+ \vdash \text{QF-AC} \oplus (B) \oplus (\hat{R}_1) \vdash \forall c: [N] \times N \to N \quad (\forall c \in [N] \quad (\exists c \in [N] \to \exists H \quad \operatorname{RT}_2^\xi(c, H) \land \Pi^1_1-CA(\varphi c H)).
\]

Proof. Let \( R_i = \{ x \in N \mid c(i, x) = 0 \} \) and let \( g \) be a strictly increasing enumeration of a cohesive set for \( R_i \). The coloring \( c'(x, y) := c(gx, gy) \) is stable and for each homogeneous set \( H' \) of \( c' \) the set \( gH' \) is homogeneous for \( c \). See [7].

Hence the corollary follows from corollary 32 and proposition 44 resp. proposition 45. \( \Box \)

By the proposition below \( \operatorname{RT}_2^\xi \) implies \( \Pi^0_1\text{-LEM} \). Therefore, sequences of instances of \( \operatorname{RT}_2^\xi \) imply \( \Pi^0_1\text{-CA} \). Hence it is not possible to show that \( \operatorname{RT}_2^\xi \) is provewise low in sequence.

**Proposition 47.** \( \text{iRCA}^*_0 \vdash \text{RT}_2^\xi \to \Pi^0_1\text{-LEM} \), where \( \text{iRCA}^*_0 \) is the intuitionistic system corresponding to \( \text{RCA}^*_0 \) and \( \Pi^0_1\text{-LEM} \) is the \( \Pi^0_2 \)-law of excluded middle.

More precisely, for every \( \Pi^0_2 \)-statement \( \forall x \exists y \ A_{xy}(x, y) \) there is a coloring such that one can decide constructively from a homogeneous set whether the \( \Pi^0_2 \)-statement is true or not.

Proof. We show for an arbitrary quantifier-free formula \( A_{xy} \) that \( \forall x \exists y \ A_{xy}(x, y) \lor \exists x \forall y \neg A_{xy}(x, y) \). First note that \( \text{iRCA}^*_0 \) proves that \( \forall x \exists y A_{xy}(x, y) \leftrightarrow \forall x \exists y y' \leq x \exists y' \leq y A_{xy}(x', y') \).

Hence we may assume that \( A_{xy} \) is monotone in the sense that \( A_{xy}(x, y) \to \forall u x \forall y \geq y A_{xy}(u, v) \).

Now color each pair \( \{ x, y \} \) with \( x < y \) red if \( A_{xy}(x, y) \) holds and blue otherwise. It is easy to see that there exists an infinite red homogeneous set iff \( \forall x \exists y A_{xy}(x, y) \) is true. \( \Box \)

To overcome this problem we switch to the functional interpretation (i.e. ND-interpretation) where the need for \( \Pi^0_1\text{-LEM} \) vanishes.

### 10.2. ND-Interpretation of \( \text{RT}_2^\xi \)

We now formulate the ND-interpretation of \( \text{RT}_2^\xi \) and of corollary 46. For notational simplification we sometimes will not apply the last application of QF-AC to the ND-interpretation. This corresponds to the so-called Shoenfield translation, see [49]. For \( \text{RT}_2^\xi \) we use the formalization

\[
\text{RT}_2^\xi \equiv \forall c: [N] \to 2 \exists H \forall u < v c(H(u, H)) = c(H(0, H1)).
\]

The ND-interpretation then yields

\[
\text{RT}_2^{\text{ND}} \equiv \forall c: [N] \to 2 \forall U < V \exists H c(H(UH), H(VH)) = c(H(0, H1)) \equiv \text{RT}_2^{\text{ND}}(H,c,U,V)
\]
Here the set $H$ is given as an enumeration, i.e. $H$ is strictly monotone and $Hn$ is the $n$-th element of $H$, and $U < V$ is defined pointwise.\footnote{Officially, quantification over functions like $c: \mathbb{N}^2 \to 2$ or strictly monotone increasing functions like $H$ are not included in our system as primitive notions, but we can enforce the same behavior by quantifying over $c: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ and $H: \mathbb{N} \to \mathbb{N}$ and replacing every occurrence of $c, H$ with}

For the ND-interpretation of $\Pi^0_1$-comprehension we use an $\varepsilon$-calculus like formulation:

$$(31) \quad \Pi^0_1-\mathcal{CA}(\varphi) : \equiv \exists f \forall x, y (\varphi(x, f(x)) = 0 \lor \varphi(x, y) \neq 0) =: (\Pi^0_1-\mathcal{CA}(\varphi))_{QF}(f, x, y)$$

This leads to following ND-interpretation (modulo a last application of QF-AC)

$$(\Pi^0_1-\mathcal{CA}(\varphi))_{ND} : \equiv \forall X, Y \exists f (\varphi(X, f(X)) = 0) \lor \varphi(X, Y f) \neq 0).$$

Because $\text{RT}^2_2$ and $\Pi^0_1-\mathcal{CA}(\varphi)$ are only $\forall \exists \forall$-statements, the ND-interpretation coincides with the no-counterexample interpretation. So one might view a solution to $\text{RT}^2_2$, i.e. a term $(c, U, V)$ that yields for every $U, V$ a set $H$ that may not be homogeneous in total but for which $c(H0, H1) = c(H(UH), H(VH))$ holds, as a procedure that disproves every possible counterexample to $\text{RT}^2_2$. Same for $\Pi^0_1-\mathcal{CA}(\varphi)$.

**Proposition 48** ([48, 32, 42]). The solution to $(\Pi^0_1-\mathcal{CA}(\varphi))_{ND}$ can be defined with a single use of $\Phi_0$, this is Spector’s bar recursor for type $0$:

$$t_f := \Phi_0 X u 0(\lambda k^0.0), \quad \text{unv} := \begin{cases} 1 & \text{if } \varphi(n, Y(u1)), \\ Y(u1) & \text{otherwise}. \end{cases}$$

The bar recursor $\Phi_0$ is defined as in [32]. It is primitive recursively and instance-wise definable in the bar recursor $B_{01}$, see definition 19 below.

The statement from corollary 46 spelled out is

$$\text{WE-PA}^\omega \vdash \text{QF-AC} \oplus (B) \oplus (\tilde{R}_1) \vdash \forall \varphi \exists f^0_\varphi \forall x, y (\Pi^0_1-\mathcal{CA}(\varphi c))_{QF}(f^0_\varphi, x, y) \rightarrow \\ \exists H (\forall u < v c(Hu, Hv) = c(H0, H1) \land \exists f^0_\varphi \forall x, y (\Pi^0_1-\mathcal{CA}(\varphi c H))_{QF}(f^0_\varphi, x, y, y)).$$

An ND-interpretation leads then to

**Theorem 49** (ND-interpretation of corollary 46). For every provably continuous (in $c, H$) term $\varphi \in T_0[B, \tilde{R}_1]$ a term $\xi \in T_0[B, \tilde{R}_1]$ (that is continuous in $c$) exists such that

$$\text{WE-HA}^\omega \vdash (B) \oplus (\tilde{R}_1) \vdash \forall \varphi \exists f^0_\varphi \forall U < V \forall X, Y, Y, \exists x, y (\exists H \exists f^0_\varphi \\
(\Pi^0_1-\mathcal{CA}(\xi c))_{QF}(f^0_\varphi, x, y) \rightarrow (c(H(UH f^0_\varphi), H(VH f^0_\varphi)) = c(H0, H1) \\
\land \Pi^0_1-\mathcal{CA}(\varphi c H))_{QF}(f^0_\varphi, X, Y, H f^0_\varphi, Y H f^0_\varphi)).$$

Moreover, there exist terms $t_\varphi, t_H, t_f, t_\tilde{H}, t_{\tilde{R}_1} \in T_0[B, \tilde{R}_1]$ (with the given parameters) satisfying this formula.
Proof. The system \( \mathsf{WE-FA} \vdash + \mathsf{QF-AC} \) has an ND-interpretation into \( \mathsf{WE-HA} \vdash \). This also extends to additions of new constants and universal axioms. See e.g. [3, 32].

The term \( t_H \) and \( t_{f_x} \) can be seen as procedures transforming the no-counterexample interpretation of the premise to the no-counterexample interpretation of the conclusion; the terms \( t_x \) and \( t_y \) yield which instance of the premise is needed to prove the conclusion.

Note that the counter-functions of \( \mathsf{RT}_2 \) and \( \Pi^0_1 \mathsf{CA} \) have access to both \( t_H \) and \( t_{f_x} \). The proof of proposition 50 bellow will use this.

To show that the no-counterexample interpretation of the conclusion (and hence the conclusion) holds we have to provide an \( f_x \) that satisfies \( (\Pi^0_1 \mathsf{CA} (\bar{x}))_{\mathsf{QF}} (f_x,t_x,t_y) \).

This can be done using \( B_0,1 \), see proposition 48.

Note that here the application of \( \Pi^0_1 \mathsf{CA} (\bar{\varphi}) \) in the premise is not fully interpreted. We obtain this form by applying logical simplifications after the negative translation. This leads to fixed terms in the second and third parameter of the premise and will reduce the need for the bar recursor \( B_0.1 \) to the rule of \( B_0.1 \).

10.3. Application to Ramsey's theorem.

Proposition 50. Let \( t^1[g] \) be a term such that \( \lambda g.t^1[g] \in \mathcal{T}_0[\mathcal{R}] \), where \( \mathcal{R} \) is a functional solving \( \mathsf{RT}_2^{\mathsf{ND}} \), and every occurrence of \( \mathcal{R} \) is of the form

\[ \mathcal{R}(t_c[g], t_u[g], t_v[g]). \]

Then there exist terms \( t_x, t_y, t_z \in \mathcal{T}_0[\mathcal{R}, \mathcal{B}] \), such that one can inductively replace every occurrence of \( \mathcal{R} \) in \( t \) with a new term

\[ r(f, g; t_c[g], t_u[g], t_v[g]) \]

(Here \( r \) is a term and \( t_c[g], t_u[g], t_v[g] \) are the results of replacing \( \mathcal{R} \) in \( t_c[g], t_u[g], t_v[g] \), such that

\[ \mathsf{WE-HA} \vdash + \mathsf{QF-AC} \oplus (\mathcal{B}) \oplus (\bar{\mathcal{R}}_1) \vdash \forall g^1, f \left( (\Pi^0_1 \mathsf{CA}(\bar{x}))_{\mathsf{QF}} (f, t_x, t_y) \right) \rightarrow \mathsf{RT}_2^{\mathsf{ND}}(r(f, g; t_c[g], t_u[g], t_v[g])); t_c[g], t_u[g], t_v[g]). \]

The formula \( \mathsf{RT}_2^{\mathsf{ND}} \) denotes the quantifier-free part of \( \mathsf{RT}_2^{\mathsf{ND}} \), see (30) on p. 30.

Proof. We use theorem 49 to inductively interpret the term \( t \). For convenience we repeat (32), the existential quantified variables are replaced by their realizing terms constructed in that theorem:

\begin{align*}
(\Pi^0_1 \mathsf{CA}(\bar{\varphi}))_{\mathsf{QF}} (f, t_x, t_y) & \rightarrow c(t_H(U t_H f_x), t_H(V t_H f_x)) = c(t_H0, t_H1) \\
\wedge (\Pi^0_1 \mathsf{CA}(\varphi t_H))_{\mathsf{QF}} (f_c, t_c, t_x, t_y) & \rightarrow c(t_H(U t_H f_x), t_H(V t_H f_x)) = c(t_H0, t_H1)
\end{align*}

It is clear that in case of \( t_c, t_u, t_v \in \mathcal{T}_0 \), i.e. there are no nested applications of \( \mathcal{R} \), every application of \( \mathcal{R} \) in the term \( t \) can be interpreted using (33). (Just set \( c = t_c, U = \lambda f_x, t_u, V = \lambda f_x, t_v \) and the others variable to 0.) Using contraction of \( \Pi^0_1 \)-comprehension, see remark 9, a term containing multiple such occurrence of \( \mathcal{R} \) can be interpreted.

If the term \( t_c \) contains a single occurrence of \( \mathcal{R} \) then we first interpret this inner \( \mathcal{R} \) but now we will take advantage of \( \varphi \) and set \( \varphi, X_\varphi, Y_\varphi \), so that the resulting instance of ND-comprehension suffices to interpret the outer occurrence of \( \mathcal{R} \) in \( t \).

Iterating this process allows us to interpret all terms \( t \in \mathcal{T}_0[\mathcal{R}] \) where every occurrence of \( \mathcal{R} \) is of the form \( \mathcal{R}(t_c[g], t_u[g], t_v[g]) \) with \( t_u, t_v \in \mathcal{T}_0 \).
Now inductively assume that \( t_u, t_v \) are terms for which this proposition holds, i.e.
there exists terms \( \hat{t}_u, \hat{t}_v \) equal to \( t_u, t_v \) modulo a given instance of ND-comprehension with the parameter \( H \). The problem is now that the instances of comprehension cannot be generated parallel to \( t_v \) because they include the parameter \( H \). But we take advantage of the argument \( \ell_{f,\varphi} \) of \( U \) and \( V \). Coding the instances of ND-comprehension together (ND-interpretation of remark 9) we can find \( \varphi', X'_\varphi, Y'_\varphi \) such that
\[
(\Pi_1^0-\mathcal{CA}(\varphi'cH))_{\mathcal{QF}}(f_\varphi, X'_\varphi(Hf_\varphi), Y'_\varphi(Hf_\varphi))
\]
proves the original ND-instance of \( \Pi_1^0-\mathcal{CA} \) for \( \varphi \) and those needed for \( t_u, t_v \).

This proves the proposition. \( \square \)

**Corollary 51**  (Extension to \( R_1, \Phi_0^1 \)). The statement of proposition 50 also holds for terms \( t^1[g] \) with \( \lambda g . t[g] \) \( \in_0 [\mathcal{R}, R_1, \Phi_0^1] \to T_1[\mathcal{R}, \Phi_0^1] \), where every occurrence of \( \mathcal{R} \) is of the form required in proposition 50 and every occurrence of \( R_1 \) or \( \Phi_0' \) is of the form
\[
R_1(t_1[g], t_2[g], t_3[g]) \quad \text{resp.} \quad \Phi_0'(t_1[g], t_2[g], t_3[g]).
\]

**Proof.** The proof proceeds like in proposition 50:

To interpret \( R_1 \) while retaining the instance of ND-comprehension, we will essentially use a functional interpretation of the proof of lemma 11 (for \( n = 1 \)). First note that \( s := R_1(t_1[g], t_2[g], t_3[g]) \) defines a type 1 function in \( T_1[g] \). Arguing as in lemma 11, it is clear that over \( \mathcal{WE-PA}^0 \) a suitable instance of \( \Pi_1^0-\mathcal{CA} \) with the parameter \( g \) proves that \( s \) is total \( (\forall x \exists y \langle x, y \rangle \in \mathcal{G}_s[g], \mathcal{G}_s \) is the graph of \( s \)). An ND-interpretation of this statement yields that even an instance of the ND-interpretation of \( \Pi_1^0-\mathcal{CA} \) is sufficient to prove that \( s \) is total. Another instance of ND-comprehension proves the ND-interpretation of the \( \Pi_1^0-\mathcal{CA} \)-instance in (8) on p. 13. This instance is modulo the totality of \( s \) equivalent to an instance of ND-comprehension with the parameter \( s \). The two instances of ND-comprehension used can be coded together, see remark 9.

The functional \( \Phi_0' \) can be replaced by a function in \( T_1[g] \), see theorem 21 and remark 22, and hence can also be interpreted. \( \square \)

**Proposition 52.** Let \( A_{yf} \) be a quantifier-free formula that contains only the shown variables free. If
\begin{equation}
\text{(34)} \quad \mathcal{N-PA}^0 + \text{QF-AC} + \Sigma^0_f-\mathcal{IA} + \text{RT}^2 \cup \text{WKL} \vdash \forall x_1 \exists y^0 A_{yf}(x, y)
\end{equation}

then one can find a terms \( t_y, t_u, t_v, \xi \in T_0[\mathcal{B}, \check{R}_1] \) such that
\[
\mathcal{WE-HA}^0 \uplus (\mathcal{B}) \uplus (\check{R}_1)
\vdash \forall x_1 \forall f \left( (\Pi_1^0-\mathcal{CA}(\xi x))_{\mathcal{QF}}(f, t_u fx, t_v fx) \rightarrow A_{yf}(x, t_y fx) \right).
\]

**Proof.** A functional interpretation of the statement (34) yields closed terms resp. term tuples \( t_y, t_R, t_\xi, t_\varphi, t_\Phi \in T_0 \), such that
\[
\text{qf-\mathcal{N-PA}^0} \vdash \left( (R_1)_\text{ND}(R_1, t_R, R_1; \Phi_0 x) \wedge \text{RT}^2_{\text{ND}}(R, t_R, R_1; \Phi_0 x) \right)
\wedge \text{WKL}_{\text{ND}}(\Phi_0, t_\Phi, R_1; \Phi_0 x) \rightarrow A_{yf}(x, t_y R_1; \Phi_0 x).
\]

Here we use that \( (\Sigma^0_f-\mathcal{IA})^\text{ND} \) can be solved by \( R_1 \), see [41].

Apply now proposition 26 and remark 27 to this derivation to normalize it such that only finitely many independent applications of \( \mathcal{R}, R_1, \Phi_0 \) occur, where each of them is of the form
\[
\mathcal{R}^*(t_1[g], t_2[g], t_3[g]) \quad \text{resp.} \quad R_1(t_1[g], t_2[g], t_3[g]), \quad \Phi_0'(t_1[g], t_2[g], t_3[g])
\]
and \( t_1, t_2, t_3 \) are semi-closed.
The terms occurring in this normalized derivation can be interpreted using corollary 51. (Applications to literally equal terms are replaced by the same interpretation.)

The instances of ND-comprehension needed for corollary 51 can be coded together in one instance using remark 9.

The application of $\Pi^0_1$-$\text{CA}$ can be interpreted by a non-iterated use $\text{R-}(B_{0,1})$ of the rule of bar-recursion — this means we substitute $f$ with a solution $t_f$ to $(\Pi^0_1$-$\text{CA})^{\text{ND}}$:

$$\text{WE-HA}^\omega | \oplus (\mathcal{B}) \oplus (\bar{R}_1) \oplus \text{R-}(B_{0,1}) \vdash \forall x^1 \left( (\Pi^0_1$-$\text{CA}(\xi x))_{\text{QP}}(t_f[x], t_u t_f[x] x, t_v t_f[x] x) \rightarrow A_{gf}(x, t_g t_f[x] x) \right)$$

The term $t_f \in T_0[\mathcal{B}, \bar{R}_1, B_{0,1}]$ is defined as in proposition 48. Note that $t_f$ depends on $\xi, t_u, t_v$ and that it is of type 2 containing only one occurrence of $B_{0,1}$ to semi-closed terms defining a type 2 object.

Since $t_f$ solves the instance of comprehension we obtain:

$$\text{WE-HA}^\omega | \oplus (\mathcal{B}) \oplus (\bar{R}_1) \oplus \text{R-}(B_{0,1}) \vdash \forall x^1 A_{gf}(x, t_g t_f[x] x).$$

The term $t := \lambda x. t_g t_f[x] x \in T_0[\mathcal{B}, \bar{R}_1, B_{0,1}]$, contains only majorizable constants; the majorants to $\mathcal{B}, \bar{R}_1$ are trivial and $B_{0,1}$ is essentially majorized by itself, see proposition 16, hence we can find a majorant $t^* \in T_0[\mathcal{B}_{0,1}]$ to $t$ containing also only one application of $B_{0,1}$ to semi-closed terms. Now we can apply bounded search to obtain a new realizer $t'$ for $y$ not containing $\mathcal{B}$ or $\bar{R}_1$:

$$t' := \begin{cases} 
\text{minimal } y \leq t^* x \text{ with } A_{gf}(x, y), & \text{if such a } y \text{ exists,} \\
0, & \text{otherwise.}
\end{cases}$$

Since $t'$ now does not contain $\mathcal{B}$ anymore we may weaken $(\mathcal{B})$ to $\text{UWKL}$ and then eliminate it from the system using a monotone functional interpretation, see [25, 32]. Hence we obtain a term $t'' \in T_0[\mathcal{B}_{0,1}]$ containing after normalization only one occurrence of $B_{0,1}$ defining a type 2 object, such that with the rule $\text{R-}(B_{0,1})$ of $B_{0,1}$

$$\text{WE-HA}^\omega | \oplus (\bar{R}_1) \oplus \text{R-}(B_{0,1}) \vdash \forall x^1 A_{gf}(x, t'' x).$$

Using ordinal analysis of the $\text{B}_{0,1}$-rule (cf. theorem 21 and remark 22) yields a new term $t'''$ definable with ordinal primitive recursion up to $\omega_2''$ such that

$$\text{WE-HA}^\omega | \omega_2 \oplus (\bar{R}_1) \vdash \forall x^1 A_{gf}(x, t''' x).$$

Combining this with theorem 17 and noting that $\bar{R}_1$ is included in $\text{WE-HA}^\omega _I$ and that $\text{WE-PA}^\omega + \Sigma^0_2$-IA has an ND-interpretation in $\text{WE-HA}^\omega _I$ we obtain the following theorem:

**Theorem 53** (Conservation for $\text{RT}^2_2$). If

$$\text{N-PA}^\omega + \text{QF-AC} + \Sigma^0_2$-IA + $\text{RT}^2_2 + \text{WKL} \vdash \forall x^1 \exists y^0 A_{gf}(x, y)$$

then one can extract a term $t \in T_1$ such that

$$\text{WE-HA}^\omega _I \vdash \forall x^1 A_{gf}(x, tx).$$
10.3.1. Extension to $RT^2_{<\infty}$. Proposition 50 holds analogously for $RT^2_{<\infty}$ if one adds $R_1$ and $\Sigma^0_3$-IA to the verifying system; corollary 51 holds if one replaces $R_1$ by $R_2$.

But in contrast to the previous the technique used in remark 27 to extract terms that meet the requirements of these propositions can only be applied to terms in $T^2_1[R_\infty]$ and not to terms $T^2_2[R_\infty]$, because $deg(R_2) = 4$ and therefore we could not apply the term normalization. The mathematical reason is that $R_2$ is strong enough to iterate $B_{0,1}$ and $R_\infty$.

This will hinder us to achieve full conservativity for full $\Sigma^0_3$-IA over a system in all finite types but a restricted variant of $\Sigma^0_3$-induction can be handled. Define the rule of $\Sigma^0_3$-induction $\Sigma^0_3$-IR as

\[
\forall n \left( \exists x \forall y \exists z A_{gf}(n,x,y,z,q) \right) \\
\forall n \exists x \forall y \exists z A_{gf}(n,x,y,z,q),
\]

where $A_{gf}$ is quantifier-free and contains only the variables shown, $u,v,w,x,y,z,n$ are type 0 variables and $q$ denotes an arbitrary tuple of parameters. Let $\Sigma^0_3$-IR$_2$ be the restriction of $\Sigma^0_3$-IR to parameters $q$ of type $\leq 2$.

**Theorem 54** (Conservation for $RT^2_{<\infty}$). If

\[
N\text{-PA}^\omega_1 \vdash QF\text{-AC} + \Sigma^0_3\text{-IR}_2 + RT^2_{<\infty} + WKL \vdash \forall x^1 \exists y^0 A_{gf}(x,y)
\]

then one can extract a term $t \in T_2$ such that

$$\text{WE-HA}^\omega_1 \vdash \forall x^1 A_{gf}(x,tx).$$

**Proof.** The ND-interpretation of the conclusion of $\Sigma^0_3$-IR$_2$ is given by

$$\forall n^0 \forall y^2 \exists x^0, Z^1 A_{gf}(n,x,YxZ, Z(YxZ), y^2).$$

One immediately see that $\Sigma^0_3$-IR$_2$ introduces only type 3 terms ($t_Z, t_x$ ranging over $n^0, Y^2, y^2$). Hence we can ND-interpret (35) in

$$q\in N\text{-PA}^\omega_1 \vdash (G_1) + \cdots + (G_n)$$

where $(G_i)$ are defining axioms and constants of type $\leq 3$ introduced by the rule $\Sigma^0_3$-IR$_2$. The terms occurring in the derivation can be viewed as terms in $T^2_1[R_\infty, \Phi^0_0, G_1, \ldots, G_n]$. The requirements of theorem 24 in remark 27 are met and we obtain a normalized derivation.

By [41], $(\Sigma^0_3\text{-IA})^{ND}$ can be solved by $R_2$. Since $\Sigma^0_3\text{-IA}$ implies $\Sigma^0_3\text{-IR}_2$ the constants $G_i$ may be chosen to be in $T^2_2[R_\infty, \Phi^0_0]$. These terms can be handled like in proposition 52.

This completes the proof. \(\square\)

**Corollary 55.** If

$$\text{E-PAC}^\omega + QF\text{-AC}^{0,1} + QF\text{-AC}^{1,0} + \Sigma^0_2\text{-IA} + RT^2 + WKL \vdash \forall x^1 \exists y^0 A_{gf}(x,y)$$

one can extract a term $t \in T_2$ such that

$$\text{WE-HA}^\omega_1 \vdash \forall x^1 A_{gf}(x,tx).$$

If $RT^2_{<\infty} + \Sigma^0_3$-IR$_2$ is added to the above system then one can extract a term $t \in T_2$ realizing $y$ provably in $\text{WE-HA}^\omega_1$ instead of $\text{WE-HA}^\omega_1$.

**Proof.** Apply elimination of extensionality (proposition 3) and use theorem 53.

For the second statement use theorem 54. To be able to use the elimination of extensionality the induction rule $\Sigma^0_3$-IR$_2$ has to be altered to include the premise that the parameters are extensional. Since this is a formula of the form $\forall u^1 \exists v^0 B_{gf}(u,v)$, the functional interpretation does not introduce terms of type $> 3$ and the rule which still follows from $\Sigma^0_3$-IA can be interpreted like in the proof of theorem 54. \(\square\)
Corollary 56.

- $\text{WKL}_0^\omega + \Sigma^0_3\text{-IA} + \text{RT}^2_2$ is conservative over $\text{RCA}_0^\omega + \Sigma^0_3\text{-IA}$ for sentences of the form $\forall x \exists y \forall z A(y, x, y, z)$.

As a consequence, $\text{WKL}_0 + \Sigma^0_3\text{-IA} + \text{RT}^2_2$ is conservative over $\text{RCA}_0 + \Sigma^0_3\text{-IA}$ for sentences of the form $\forall x \forall y \forall z A(x, x, y, z)$, where $A$ is $\Delta^0_2$, and thus, in particular, $\Pi^0_3$-conservative.

- $\text{WKL}_0^\omega + \Sigma^0_3\text{-IA} + \Sigma^0_3\text{-IR}_2 + \text{RT}^2_{<\infty}$ is conservative over $\text{RCA}_0^\omega + \Sigma^0_3\text{-IA}$ for sentences of the form $\forall x \exists y \forall z A(y, x, y, z)$.

Hence, $\text{WKL}_0 + \Sigma^0_3\text{-IA} + \Sigma^0_3\text{-IR} + \text{RT}^2_{<\infty}$ is conservative over $\text{RCA}_0 + \Sigma^0_3\text{-IA}$ for sentences of the form $\forall x \forall y \forall z A(x, x, y, z)$, where $A$ is $\Delta^0_2$, and thus, in particular, $\Pi^0_3$-conservative.

Moreover from of $\forall x \exists y \forall z A(y, x, y, z)$ the above theories one can extract terms in $T_1$ resp. $T_2$ realizing $y$.

Proof. The former statements follow from the previous theorem with the fact every sentence of the form $\forall x \exists y \forall z A(y, x, y, z)$ is over $\text{QF-AC}^{0,0}$ equivalent to a sentence of the form $\forall x \exists y \forall z A(y, x, y, z)$.

The conservativity over $\text{RCA}_0$ follows from the fact that $\text{RCA}_0^\omega$ is conservative over its second order fragment, which can be simulated in $\text{RCA}_0$, see [31]. The quantification over $X$ and $x$ can be coded into a quantification over a function. The restrict on the rule of $\Sigma^0_3$-induction is automatically met in a second order system.

11. Possible extensions

The question arises whether $\text{RT}^2_2$ also is proofwise low in sequence over $\text{WKL}_0^\omega$ (or $\Sigma^0_\infty \text{A}^\omega - \text{QF-AC}^{0,0} + (B)$) and hence does not imply $\Sigma^0_3$-induction.

The first obstacle to show this is that the proof of the lowness-property crucially depends on full $\Sigma^0_3$-induction which renders $\Sigma^0_\infty \text{A}^\omega$ or equivalently $\text{RCA}_0^\omega$ insufficent. The other obstacle is that $\text{RT}^2_2$ implies $\Pi^0_3\text{-LEM}$ so that sequences of solutions would imply $\Pi^0_3\text{-CA}$. Thus $\text{RT}^2_2$ cannot be proofwise low in sequence over a theory which does not include $\Pi^0_3\text{-CA}$, see proposition 47. Actually even the so called stable chain-antichain principle (SCAC) implies $\Pi^0_3\text{-LEM}$ (for a definition see [15]).

In [35] we refined the method based on the bar recursion (section 10.3) and could show that the type 2 functionals that are provable from principles which are proofwise low over $\text{WKL}_0^\omega$ are primitive recursive. We also show that $\text{CAC}$ is proofwise low in sequence and thus that this theorem applies to it, see also [9]. However, we were not able to show that $\text{RT}^2_2$ is proofwise low in sequence over $\text{WKL}_0^\omega$. (In other words we could overcome the second obstacle but not the first one.) Still the question remains whether one could do the same with $\text{RT}^2_2$ or any other principle which is stronger than $\text{CAC}$. The principle $\text{RT}^2_2$ splits into the so called Erdős-Moser principle (EM) and $\text{CAC}$ (actually even $\text{ADS}$), see [6]. Therefore EM seems to be a good candidate for further investigations.

References


20. ______, *Correction to: “A cohesive set which is not high”*, Math. Logic Quart. 43 (1997), no. 4, 569. MR 1477624

Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstrasse 7, 64289 Darmstadt, Germany
E-mail address: akreuzer@mathematik.tu-darmstadt.de
E-mail address: kohlenbach@mathematik.tu-darmstadt.de