Some Results in Logic and Ergodic Theory

Henry Towsner

May 2008

Thesis Committee:
Jeremy Avigad (Chair)
James Cummings
Ernest Schimmerling
Stephen Simpson

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Acknowledgments

Firstly, I am grateful to the many people who took the time to read various parts of this thesis and provide helpful suggestions and corrections: Tim Austin, James Cummings, Philipp Gerhardy, John Griesmer, Ulrich Kohlenbach, Bryna Kra, and Terence Tao. Thanks also to my thesis committee members, James Cummings, Ernest Schimmerling, and Stephen Simpson.

Additionally, I am grateful to the mathematics department at CMU, including Bill Hrusa, as well as Stella, Ferna, and Nancy, for all their help over the last five years.

I also wish to thank Grisha Mints, without whose support I would never have begun this project. Most of all, I wish to thank my advisor, Jeremy Avigad, who has been a source of endless patience and wisdom.

Contents

Chapter 1

Introduction

One of the traditional goals of proof theory has been the development of an "ordinal analysis" for second order arithmetic. While the name suggests that ordinals are central to this project, this is somewhat deceptive: ordinals are a side product of an ordinal analysis, which is a rigorous description of the computational principles which can be proven to terminate in a particular system. The standard technique for producing an ordinal analysis is the method of cut-elimination invented by Gentzen in his ordinal analysis of first order arithmetic [37]. The most sophisticated developments of this method, for the subsystem of second order arithmetic known as $\Pi^1_2 - CA$, are extremely complicated [3, 75].

The main feature which makes cut-elimination arguments so unwieldy when applied to strong subsystems of second order arithmetic is that they become highly dependent on formal ordinal notations to describe the cut-elimination procedure. In Chapter 2, I propose an approach to cut-elimination of using new infinitary rules which do not need explicit ordinal notations to describe them. I apply this to the system μ_2, which is weaker than $\Pi^1_2 - CA$ but stronger than iterated forms of $\Pi^1_1 - CA$, which are the strongest systems which can already be analyzed without ordinal notations.

Theorem 1. *Given any proof of a Π_2 arithmetic formula in the system of inductive definitions μ_2, there is a cut-free proof of this formula.*

Another significant area of proof theoretic research is the development of interpretations of theories in systems of constructive functionals. A realizability interpretation of a theory is an assignment of λ-terms (from an appropriate theory of λ-terms) to proofs in the theory in such a way that proofs giving computational information—that is, proofs of Π_2 formulas—are witnessed by functions with the same computational content. Realizability interpretations for constructive theories are well-developed (see [92] for a detailed development of the standard varieties), but Avigad [6] observed that, by combining realizability interpretations of constructive theories with double-negation style embeddings of classical theories in constructive ones, it is possible to obtain realizability interpretations for classical theories as well.

Avigad gave a realizability interpretation for classical first order arithmetic, and in Chapter 3 I give a realizability interpretation for classical second order arithmetic. The interpretation is given in the system HRO^2 of second order hereditarily recursive functionals.

Theorem 2. *If f is a function, A is a primitive recursive relation representing the graph of f, and second order arithmetic proves that $\forall y \exists x A(y, x)$ then there is a term t of HRO^2 such that $f = \lambda x.t$.*

An alternative method for interpreting proofs by functionals is Gödel's Dialectica translation. Variations of the Dialectica translation are already known for second order arithmetic [39, 84], but these are unwieldy for practical application—that is, for the task of actually taking a particular proof of a Π_2 statement and extracting the corresponding functional. Fortunately, most actual theorems can be proven in much weaker systems [83], so it is helpful to have simpler Dialectica translations corresponding to these weaker systems. In Chapter 4 I give a translation for the particular theory ID_1 of non-iterated inductive definitions, and show how to generalize the interpretation to the theories ID_n of finitely iterated inductive definitions. (This chapter is joint work with Jeremy Avigad.)

Theorem 3 (Avigad-Towsner)**.** *If $\forall x \exists y \phi(x, y)$ is provable in ID_n then there is a term t in the system*

$QT^i_{\Omega_n}$ with at most the free variables x and those appearing in ϕ such that $QT^i_{\Omega_n}$ proves $\phi(x,t)$.

One particular application of such functional interpretations is in the field of ergodic Ramsey theory, in which combinatorial statements are proven using the measure-theoretic techniques of ergodic theory. Such proofs typically give no bounds, but it is possible to use functional interpretations to extract from the non-constructive ergodic proof a combinatorial proof which does give bounds. A necessary first step is giving the constructive combinatorial arguments which correspond to the basic theorems used in ergodic theory. In Chapter 5, I do so for the mean ergodic theorem. (This chapter is joint work with Jeremy Avigad and Philipp Gerhardy.)

The mean ergodic theorem states that if T is a nonexpansive linear operator on a Hilbert space and the averages $A_n f$ are given by $A_n f := \frac{1}{n}\sum_{i=0}^{n-1} T^i f$ for each f in the space, then the sequence $\langle A_n f\rangle$ converges (in the Hilbert space norm) to some element f^*. Since this is not a Π_2 statement, there is no guarantee that there is a constructive version of the full statement (and indeed, it can be shown that there is none [11]). However the Dialectica interpretation gives a Π_2 statement with quantifiers over functionals which is classically equivalent to the original one, and converts the non-constructive proof of the original form into a constructive proof of this modified form.

Theorem 4 (Avigad-Gerhardy-Towsner)**.** *Let T be a nonexpansive linear operator on a Hilbert space, and let f be any nonzero element of that space. Let K be any nondecreasing function satisfying $K(n)\geq n$ for every n. Then there is a function $\overline{K}$, computed explicitly in K, such that for every $\epsilon>0$ there is an n satisfying $1\leq n\leq \overline{K}(\epsilon)$ such that for every m in $[n,K(n)]$, $||A_m f - A_n f||\leq\epsilon$.*

Given a small amount of additional information, namely the norm of the limit $||f^*||$, the theorem does become constructive.

Theorem 5 (Avigad-Gerhardy-Towsner)**.** *Let T be a nonexpansive operator on a separable Hilbert space and let f be an element of that space. Let $f^* = \lim_n A_n f$. Then f^*, and a bound on the rate of convergence of $(A_n f)$ in the Hilbert space norm, can be computed from f, T, and $||f^*||$.*

This includes the special case where the Hilbert space is the L^2 space of a measure space and T is (the lift to the L^2 space of) a measure-preserving measurable ergodic operator on the measure space.

This constructive form of the mean ergodic theorem was independently given by Tao [89], who then used it to settle an open question in ergodic theory involving a more general class of averages. Tao's proof uses a finitary constructive setting in an essential way, even though the final statement proven is about the infinitary ergodic setting. The same ideas used in Chapter 5 to "unwind" the non-constructive proof of the mean ergodic theorem can be used in reverse to "wind" Tao's proof into an ergodic proof.

It turns out that the Furstenberg correspondence, the technique used to relate the infinitary and finitary settings, is insufficient for this proof. The needed generalization joins others recent variations [29, 88] in pointing towards a more general form of the correspondence. In Chapter 6, I give a single form which subsumes all these generalizations.

Theorem 6. *Let S be a countable set and let G be a semigroup acting on S. Let X be a second countable compact space. Let $E: S\rightarrow X$ be given, and let $\{I_n\}$ be a Følner sequence of subsets of G.*

Then there are a dynamical system $(Y,\mathcal{B},\nu,(T_g)_{g\in G})$ and functions $\tilde{E}_s : Y\rightarrow X$ for each $s\in S$ such that the following hold:

- *For any g, s, $\tilde{E}_{gs} = \tilde{E}_s\circ T_g$*
- *For any integer k, any continuous function $u: X^k\rightarrow\mathbb{R}$, and any finite sequence $s_1,\dots,s_k$,*

$$\begin{aligned}\liminf_{n\rightarrow\infty}\frac{1}{|I_n|}\sum_{g\in I_n} u(E(gs_1),\dots,E(gs_k)) &\leq \int u(\tilde{E}_{s_1},\dots,\tilde{E}_{s_k})d\nu\\ &\leq \limsup_{n\rightarrow\infty}\frac{1}{|I_n|}\sum_{g\in I_n} u(E(gs_1),\dots,E(gs_k))\end{aligned}$$

Using this theorem, in Chapter 7 I give an ergodic proof of Tao's theorem.

Theorem 7. *Let $l\geq 1$ be an integer. Assume $T_1,\dots,T_l : X\rightarrow X$ are commuting, invertible, measure-preserving transformations of a measure space $(X,\mathcal{B},\mu)$. Then for any $f_1,\dots,f_l\in L^\infty(X,\mathcal{B},\mu)$, the* *averages*

$$A_N(f_1, \ldots, f_l) := \frac{1}{N} \sum_{n=1}^{N} f_1(T_1^n x) \cdots f_l(T_l^n x)$$

converge in $L^2(X, \mathcal{B}, \mu)$.

Chapter 2 will appear in the Annals of Pure and Applied Logic. Chapter 3 has appeared as [90]. Chapter 4 is extracted from an article written with Jeremy Avigad [12] and Chapter 5 is extracted from an article written with Jeremy Avigad and Philipp Gerhardy [13].

Chapter 2

Cut-Elimination for μ_2

Infinitary inference rules have been a key tool in ordinal analysis since their introduction by Schütte [76]. The appropriate infinitary rule for Peano Arithmetic, the ω rule, is reasonably straightforward—it simply branches over the natural numbers—but suitable infinitary rules for stronger systems are less clear.

The first type proposed, Buchholz's Ω_μ rule [20], branches not over numbers, but over a particular class of derivations. Subsequently, Pohlers proposed the method of local predicativity [71], in which infinitary rules branch over infinite ordinals. Rules branching over ordinals have almost entirely replaced the Ω_μ rule, in large part because they led to productive generalizations, culminating in an analysis of Π^1_2-comprehension [75], while the Ω_μ rule seemed limited to iterated systems of Π^1_1-comprehension.

In the method of local predicativity, ordinals are built directly into the system, since they are necessary to even describe the system cut-elimination will take place in. This integration with ordinals is different from earlier analyses, in which the cut-elimination process came first and the ordinals could be "read off" from the reduction procedure; in local predicativity, the crucial collapsing step is justified by reference to the properties of the ordinals, which have, naturally, been defined just so as to make this possible. Unfortunately, as the systems get more complex, this leads to the appearance that the proof proceeds by "magic", obscuring the underlying structure of the argument. This problem isn't intrinsic to infinitary techniques—the most advanced finitary methods, as in [4],[5], and [3], also require systems defined in terms of ordinals, and face the same problems as a result.

Möllerfeld [68] proposed that an alternate approach could come from "game-theoretic" infinitary rules, based on his analysis of Π^1_2 comprehension in terms of a game quantifier. Möllerfeld's analysis shows that Π^1_2 comprehension is the union of systems μ_n for finite n, where μ_1 is (essentially) the system $ID_{<\omega}$, and the higher levels are novel systems based on complicated monotone inductive definitions.

2.1 Outline

Before launching into the technical details of the proof, we outline the general method as inspired by Buchholz's Ω_μ rule. Suppose that we can prove cut-elimination for some arbitrary theory T (say, Peano Arithmetic or ID_n) using an infinitary system T^∞. We may extend T to a theory T' by adding a least fixed point predicate

$$\mu x X.A(x, X)$$

where A is a formula of T and X appears positively, along with closure and induction axioms. We may then extend T^∞ by a closure rule

$$\frac{A(n, \mu x X.A(x, X))}{n \in \mu x X.A(x, X)}$$

and call proofs in this extended system "small". The full infinitary version of T' adds an Ω rule which branches over small proofs of $n \in \mu x X.A(x, X)$ and gives the conclusion $n \notin \mu x X.A(x, X)$. Cut-elimination is quite easy to prove, and the heart of the resulting argument is the demonstration that the induction axiom in the finitary system can be embedded as an Ω rule in the infinitary system. The proof that this is possible involves showing that, given any "small" proof of $n \in \mu x X.A(x, X)$, the predicate $y \in \mu x X.A(x, X)$ can be systematically replaced by any formula $F[y]$.

This method breaks down when we attempt to add the predicate

$$\mu x X.A(x, X, \mu y Y.B(x, X, y, Y))$$

where X appears negatively in B, and therefore $\mu y Y.B(x, X, y, Y)$ must appear negatively in A. (For convenience, we abbreviate $\mu y Y.B(x, X, y, Y)$ by $\mu_B(x, X)$ and $\mu x X.A(x, X, \mu_B(x, X))$ by μ_A.) If we attempt the same technique, a "small" proof of $n \in \mu_A$ must contain negative occurrences of $\mu_B(m, \mu_A)$, which must be introduced by an Ω rule for $\mu_B(m, \mu_A)$, which must in turn branch over proofs containing negative occurrences of μ_A, which gives a vicious cycle.

To find a way out of this dilemma, we can consider what we expect to happen when we attempt to embed an induction axiom for μ_A as a hypothetical Ω rule. We would expect to replace μ_A with some formula F, and therefore whatever rule introduces $\neg\mu_B(m, \mu_A)$ must be easily converted to a proof of $\neg\mu_B(m, F)$. This would not be true if we used an ordinary Ω rule, which would face the obstacle that the Ω rule for $\neg\mu_B(m, \mu_A)$ does not even necessarily branch over the right domain to become an Ω rule for $\neg\mu_B(m, F)$.

We resolve both these problems at once by introducing a new type of Ω rule to be the "small" rule introducing $\neg\mu_B(m, \mu_A)$; this rule will branch over proofs of $\mu_B(m, F)$ for any F. The difficulty is that such derivations may contain inference rules which more widely than is permitted in a small proof (for instance, the introduction of $\neg\mu_A$ when F is μ_A).

Such inferences will be converted into non-branching inference rules. We will call these inference rules *truncated inferences*, since rather than encoding the manner in which the original proof derived $\neg F[n]$, they merely note where such a derivation occurred. The Ω rule will then provide, to each derivation of $n \in \mu_B(m, F)$, a derivation of some $G[n, F]$ from instances of these truncated inferences, as well as an indication, for each truncated inference in the resulting derivation, a source inference in the original derivation.

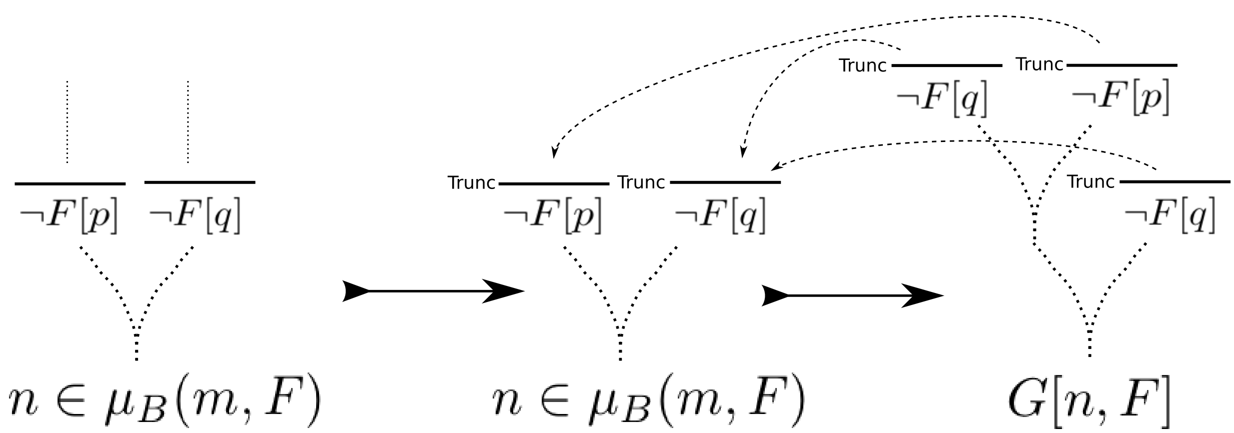

Figure 2.1: Example Transformation

We cannot be finished, because we have thrown away everything above a widely branching inference in the original derivation. In order to recover it, we must provide, for each truncated inference appearing in our derivation of $G[n, F]$, not only a truncated inference from the source derivation, but also a new Ω rule which will provide, for each possible premise d_i, a new derivation $\mathcal{F}(d_i)$ in such a way that $\{\mathcal{F}(d_i)\}_\iota$ are valid premises for the widely branching inference.

In order to keep all this information in one place, our Ω rule for $n \in \mu_B(m, F)$ will branch, not over derivations, but over *sequences* of derivations. Given a derivation d of $n \in \mu_B(m, F)$, we divide this derivation into pieces by chopping it at each introduction of some $\neg F[k]$. We then build up a new derivation coinductively; the bottommost piece, d_0, is replaced by some $\mathcal{F}(\langle d_0 \rangle)$. Each truncated inference θ appearing in $\mathcal{F}(d_0)$ is traced to some truncated inference in d_0, which in turn is traced to some introduction of $\neg F[k]$

in d using an inference rule $\mathcal{I}$. This introduction rule, whatever it is, has some list of premises $\{d_\iota\}$; for each d_ι there is an inference $\mathcal{F}(\langle d_0, d_\iota\rangle)$ which extends $\mathcal{F}(\langle d_0\rangle)$ at θ. By replacing θ in $\mathcal{F}(\langle d_0\rangle)$ with $\mathcal{I}$, taking for each premises ι the extension in $\mathcal{F}(\langle d_0, d_\iota\rangle)$, we obtain a new valid derivation. We then have new truncated inferences which first appeared in $\mathcal{F}(\langle d_0, d_\iota\rangle)$, and the process repeats.

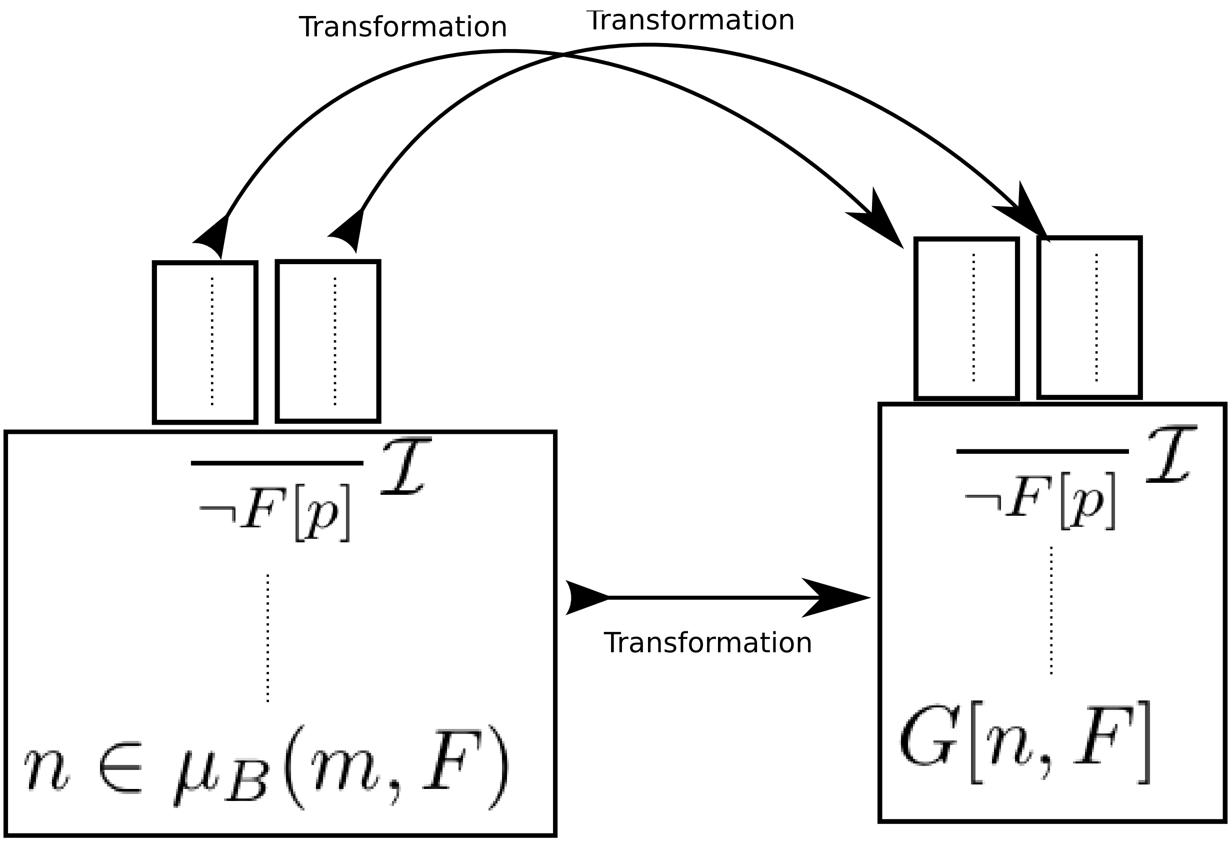

Figure 2.2: A derivation is divided into segments, and (the corresponding portion of) the transformation is applied to each segment in turn.

We may formulate this procedure as a game with two players, a Prover and a Transformer. Prover plays first, and must play a derivation of from our system of small proofs augmented by truncated inferences (which we wall call a truncated proof system). Transformer must play a derivation with appropriate endsequent from the same system (actually, transformer is given a bit more flexibility, for instance, being allowed to use the cut rule), with the additional property that, for each truncated inference in this derivation, transformer must name a source callback inference in Prover's play. Prover then chooses some truncated inference in Transformer's play, and plays this truncated inference together with a new derivation. From here, play continues alternating these last two steps. Transformer wins as long as it is possible to provide derivations with the appropriate endsequent relative to what Prover offers (and an additional condition to be described shortly). The Ω rule is simply an encoding of a winning strategy for Transformer. (The ordinary Ω rule may be viewed as the two step version of this game, where Prover is not permitted an additional play after Transformer has gone once.) Any derivation gives a collection of strategies for Prover, and applying the transformation to some derivation is the result of knitting together the results given by Transformer against all the strategies for Prover offered by the derivation.

Two points must be made about this procedure. First, it is convenient in the description of cut-elimination to take the view that Transformer's plays (that is, the premises of the Ω rule) are not merely the portion of the derivation to be placed above truncated rules, but the entire derivation below that point as well. That is, $\mathcal{F}(\sigma^\frown\langle d_n\rangle, \tau^\frown\langle\theta\rangle)$ should be an extension of $\mathcal{F}(\sigma, \tau)$ in which the only change is that the truncated inference θ, which had no premise in $\mathcal{F}(\sigma, \tau)$, is required to have a single premise with appropriate endsequent (based on d_n) in $\mathcal{F}(\sigma^\frown\langle d_n\rangle, \tau^\frown\langle\theta\rangle)$. These truncated rules with an additional premise will be called callback inferences, since they represent the point at which Transformer's play has to make reference to the content omitted in Prover's play.

The second point is that truncated inferences appearing in $\mathcal{F}(\sigma, \tau)$ may have their source in any inference in σ, not just the most recent one. This is necessary, since the cut-elimination process will cause this situation to occur. However this introduces a concern about well-foundedness; we wish to have the property that whenever d is a well-founded derivation, the result of applying the transformation to it is also well-founded. In order to preserve this, we must specify additional conditions on infinite play; if Prover's plays are given by the infinite sequences σ, τ and the τ_i are all selected from the newly extended part of Transformer's play, Transformer loses if there is some σ_i such that infinitely many τ_j belong to σ_i. In any other infinite play, Prover loses. (A well-foundedness criterion of some sort is to be expected, since we are producing an analysis of a system stronger than Π^1_1-comprehension. It is not hard to show that a transformation with this property maps well-founded derivations to well-founded ones.) Our Ω rule must remain a winning strategy for this clarified version of the game.

Given this Ω rule, the remainder of our proof is not so difficult. Such Ω rules are considered an additional type of "small" inference, and may appear in derivations of $n \in \mu_A$, which then has an ordinary Ω rule.

2.2 Transformations

2.2.1 Proof System

We first need a general notion of a proof system, which we take almost verbatim from [19]. In the following, we assume we have already fixed some suitable language, and are working with the formulas of this language.

Definition 2.1. *A* sequent *is a finite set of formulas.*

A proof system *consists of a set of formal* inference symbols *(generally denoted by the variable $\mathcal{I}$), and, for each inference symbol:*

- *A (possibly infinite) set $|\mathcal{I}|$ called its arity*
- *A sequent $\Delta(\mathcal{I})$*
- *For each $\iota \in |\mathcal{I}|$, a sequent $\Delta_\iota(\mathcal{I})$*
- *A set $Eig(\mathcal{I})$ which is either empty or a singleton $\{x\}$ where x is a variable not in $FV(\Delta(\mathcal{I}))$ (in this case we call x the* eigenvariable *of $\mathcal{I}$)*

When we say that a proof system contains an inference rule

$$\mathcal{I}\ \frac{\cdots \Delta_\iota \cdots (\iota \in I)}{\Delta}\ !u!$$

we are declaring $\mathcal{I}$ to be an inference symbol with arity I, $\Delta(\mathcal{I}) = \Delta$, $\Delta_\iota(\mathcal{I}) = \Delta_\iota$, and $Eig(\mathcal{I}) = \{u\}$ (or $\emptyset$ if u is omitted). When the arity is finite, we typically list all the premises explicitly.

Definition 2.2. *The* derivations *d of a proof system and the* end sequent *$\Gamma(d)$ are defined inductively. If, for each $\iota \in |\mathcal{I}|$, d_ι is a derivation and setting $\Gamma := \Delta(\mathcal{I}) \cup \bigcup_{\iota\in|\mathcal{I}|} \Gamma(d_\iota) \setminus \Delta_\iota(\mathcal{I})$, $Eig(\mathcal{I}) \cap FV(\Gamma) = \emptyset$ then $d := \mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}$ is a derivation with $\Gamma(d) := \Gamma$.*

If d is a derivation and $\Gamma(d) \subseteq \Gamma$ then we write $d \vdash \Gamma$.

Definition 2.3. *An expression of the form $\lambda x.F$ is called a predicate, and denoted $\mathcal{F}$. We write $\mathcal{F}[t] := F(x/t)$.*

2.2.2 Augmented and Truncated Derivations

We define proof systems with additional rules which serve to mark places where a derivation has been cut off. The rule $Trunc_{\Gamma\mapsto\Gamma,\Delta}$ indicates a point where the derivation has been truncated below an inference rule $\mathcal{I}$ with $\Delta(\mathcal{I}) = \Delta$ and $\bigcup_{\iota\in|\mathcal{I}|} \Gamma(d_\iota) \setminus \Delta_\iota(\mathcal{I}) = \Gamma$.

A $CB_{\Upsilon\mapsto\Delta}$ inference indicates a point where every branch besides the branch ι of some inference rule $\mathcal{I}$ has been cut off, $\Gamma(d_\iota)=\Upsilon$ and $\Delta(\mathcal{I})\cup\Gamma(d_\iota)\setminus\Delta_\iota(\mathcal{I})=\Delta$.

Definition 2.4. *Let $\mathcal{P}$ be a proof system. We define* truncated $\mathcal{P}$ *to consist of $\mathcal{P}$ together with inference rules*

$Trunc_{\Gamma\mapsto\Gamma,\Delta}\ \dfrac{}{\Gamma,\Delta}$

We define augmented $\mathcal{P}$ *to consist of truncated $\mathcal{P}$ together with inference rules*

$CB_{\Upsilon\mapsto\Delta}\ \dfrac{\Upsilon}{\Delta}$

We define $\Theta(d)$ to be the set of instances of $Trunc$ inferences appearing in d.

If θ is a truncated inference $Trunc_{\Gamma\mapsto\Gamma,\Delta}$, we set $In(\theta):=\Gamma$ and $Out(\theta):=\Delta$.

Note that Θ picks out instances, so it distinguishes two occurrences of the inference rule in different places, even if they have identical parameters.

We will want to be able to talk about systems such as truncated $\mathcal{P}$ where $\mathcal{P}$ is itself augmented $\mathcal{Q}$; when we speak of truncated inferences in a derivation in augmented $\mathcal{P}$, or refer to $\Theta(d)$, we mean to include only those inferences not belonging to $\mathcal{P}$. That is, augmenting and truncating give *disjoint* unions.

Definition 2.5. *We define the* exploded derivations *of $\mathcal{P}$ over $\mathcal{Q}$ by induction:*

- *If d is a derivation in truncated $\mathcal{Q}$ and $\mathcal{I},E$ are functions on $\Theta(d)$ such that $\mathcal{I}(\theta)$ is an inference rule from $\mathcal{P}$, $In(\theta)=\bigcup_\iota\Gamma(E(\theta,\iota))\setminus\Delta_\iota(\mathcal{I}(\theta))$, $Out(\theta)=\Delta(\mathcal{I}(\theta))$, and each $E(\theta,\iota)$ is an exploded derivation then $\langle d,\mathcal{I},E\rangle$ is an exploded derivation with endsequent $\Gamma(d)$*

We denote the endsequent of an exploded derivation E by $\Gamma(E)$. If $E=\langle d,\mathcal{I},E\rangle$ is an exploded derivation, we set $E_0:=d$ and call this the main part *of the exploded derivation.*

Definition 2.6. *If $\langle d,\mathcal{I},E\rangle$ is an exploded derivation, the* unexplosion $\mathbb{U}(\langle d,\mathcal{I},E\rangle)$ *is given by main induction on E and a side induction on d:*

- *If d is a $Trunc$ inference,*

$$\mathbb{U}(\langle Trunc,\mathcal{I},E\rangle):=\mathcal{I}(\theta)\{\mathbb{U}(E(\theta,\iota))\}_{\iota\in|\mathcal{I}(\theta)|}$$

- *Otherwise,*

$$\mathbb{U}(\langle\mathcal{J}\{d_\iota\},\mathcal{I},E\rangle):=\mathcal{J}\{\mathbb{U}(\langle d_\iota,\mathcal{I}\restriction\Theta(d_\iota),E\restriction\Theta(d_\iota)\rangle)\}_{\iota\in|\mathcal{J}|}$$

Definition 2.7. *If $\mathcal{P},\mathcal{Q}$ are proof systems, we define the* explosion $\mathbb{E}_\mathcal{Q}(d)$ *of a derivation d in $\mathcal{P}$ by:*

- *If $d=\mathcal{I}\{d_\iota\}$ where $\mathcal{I}$ is not an inference of $\mathcal{Q}$,*

$$\mathbb{E}_\mathcal{Q}(d):=\langle Trunc_{\Gamma(d)\setminus\Delta(\mathcal{I})\to\Gamma(d)},\theta\mapsto\mathcal{I},(\theta,\iota)\mapsto\mathbb{E}_\mathcal{Q}(d_\iota)\rangle$$

- *Otherwise $d=\mathcal{I}\{d_\iota\}$ where $\mathcal{I}$ is an inference of $\mathcal{Q}$ and set, for each $\iota\in|\mathcal{I}|$, $\langle d'_\iota,\mathcal{I}_\iota,E_\iota\rangle:=\mathbb{E}_\mathcal{Q}(d)$, and then*

$$\mathbb{E}_\mathcal{Q}(d):=\langle\mathcal{I}\{d'_\iota\},\bigcup\mathcal{I}_\iota,\bigcup E_\iota\rangle$$

Lemma 2.8. *For any $\mathcal{Q}$, $\mathbb{U}(\mathbb{E}_\mathcal{Q}(d))=d$.*

Definition 2.9. *Let d,d' be derivations in augmented $\mathcal{P}$ such that d and d' are identical except that there exist some $Trunc_{\Gamma\mapsto\Gamma,\Delta}$ inference in d, but at the corresponding place in d', there is a $\theta=CB_{\Upsilon\mapsto\Gamma,\Delta}$ inference. We say d'* narrowly extends *d, and write $d'\setminus d$ for the derivation which is the premise of the callback inference θ in d'. We call Υ the* key sequent *of this extension.*

Definition 2.10. *Let $\mathcal{P},\mathcal{Q}$ be proof systems, and let sequences of equal length σ_0,τ_0 be given. We say $\{d_{\sigma,\tau}\}_{\sigma\supseteq\sigma_0,\tau\supseteq_0\tau}$ together with supplementary functions $\Lambda_{\sigma,\tau}$ is a* transformation *from Γ out of $\mathcal{P}$ over some restricted set of formulas $\mathcal{F}$ (in a proof system $\mathcal{Q}$) with endsequent Σ and root σ_0,τ_0 if the following hold:*

- *For every derivation d of Γ,Υ in truncated $\mathcal{P}$ with $\Upsilon\subseteq\mathcal{F}$, $d_{\sigma_0^\frown\langle d\rangle,\tau_0}$ is a proof of Σ,Υ in truncated $\mathcal{Q}$, $\Lambda_{\sigma_0^\frown\langle d\rangle,\tau_0}:\Theta(d_{\sigma_0^\frown\langle d\rangle,\tau_0})\to\Theta(d)\cup\bigcup_{i<length(\sigma)}\Theta(\sigma_i)$, and for each θ, $Out(\theta)=Out(\Lambda_{\sigma_0^\frown\langle d\rangle,\tau_0}(\theta))$*
- *If $d_{\sigma,\tau}$ is defined, $\theta\in\Theta(d_{\sigma,\tau})$, and $d\vDash In(\Lambda_{\sigma,\tau}(\theta)),\Upsilon$ in truncated $\mathcal{P}$ with $\Upsilon\subseteq\mathcal{F}$ then $d_{\sigma^\frown\langle d\rangle,\tau^\frown\langle\theta\rangle}$ is a proof in augmented $\mathcal{Q}$ narrowly extending $d_{\sigma,\tau}$ at θ and $d_{\sigma^\frown\langle d\rangle,\tau^\frown\langle\theta\rangle}$ is a proof in truncated $\mathcal{Q}$ with key sequent $In(\theta),\Upsilon$. Furthermore, $\Lambda_{\sigma^\frown\langle d\rangle,\tau^\frown\langle\theta\rangle}$ has range in $\bigcup\Theta(\sigma_i)\cup\Theta(d)$, agrees with $\Lambda_{\sigma^\frown\langle d\rangle,\tau}$ on elements in their shared domain, and for each θ', $Out(\theta')=Out(\Lambda_{\sigma^\frown\langle d\rangle,\tau^\frown\langle\theta\rangle}(\theta'))$.*

If $T = \{d_{\sigma,\tau}\}_{\sigma\supseteq\sigma_0,\tau\supseteq\tau_0}$ *is a transformation and* $\sigma' \supseteq \sigma_0, \tau' \supseteq \tau_0$ *are such that* $d_{\sigma',\tau'}$ *is defined, we write* $T \upharpoonright \sigma', \tau'$ *for the transformation* $\{d_{\sigma,\tau}\}_{\sigma\supseteq\sigma',\tau\supseteq\tau'}$.

Let d *be given and let* σ_0, τ_0 *be given with* d *an element of* σ. *We define the* d-well-founded transformations *inductively:*

- *If there is no* $\sigma \supseteq \sigma_0, \tau \supseteq \tau_0$, $\theta \in \Theta(d_{\sigma,\tau})$ *such that* $\Lambda_{\sigma,\tau}(\theta) \in \Theta(d)$ *then* $\{d_{\sigma,\tau}\}_{\sigma\supseteq\sigma_0,\tau\supseteq\tau_0}$ *is* d*-well-founded. We call such transformations* d-void.
- *If for every* d', θ, $T \upharpoonright \sigma_0^\frown\langle d'\rangle, \tau_0^\frown\langle\theta\rangle$ *is* d*-well-founded then so is* T

We say $T = \{d_{\sigma,\tau}\}$ *is* well-founded *if for every* σ, τ *and every* d *such that* $d_{\sigma^\frown\langle d\rangle,\tau}$ *is defined,* $T \upharpoonright \sigma^\frown\langle d\rangle, \tau$ *is* d*-well-founded.*

A transformation should, as the name suggests, give a way of transforming a derivation of Γ, Υ into a derivation of Σ, Υ. In order to get the right inductive hypothesis, we need to first show how to apply a transformation to an exploded derivation.

Lemma 2.11. *Let* $E = \langle d_0, \mathcal{I}, E_0\rangle$ *be an exploded derivation over* $\mathcal{P}$ *with endsequent* Γ, Υ *and let* $T = \{d_{\sigma,\tau}\}_{\sigma\supseteq\sigma_0,\tau\supseteq\tau_0}$ *be a well-founded transformation from* Γ *out of* $\mathcal{P}$ *over some* $\mathcal{F} \supseteq \Upsilon$ *with endsequent* Σ. *Then there is a derivation* d^* *with endsequent* Σ, Υ *and a function* $\Lambda : \Theta(d^*) \to \bigcup \Theta((\sigma_0)_i)$.

Proof. The proof is by main induction on E and side induction on T. Let $E = \langle d_0, \mathcal{I}, E_0\rangle$ be given. Then by induction, we produce from any d_0-well-founded transformation T a d_0-void transformation T'. If T is d_0-void then $T' = T$. Otherwise, for each d', θ, $T \upharpoonright \sigma_0^\frown\langle d'\rangle, \tau_0^\frown\langle\theta\rangle$ is d_0-well-founded, and by side IH there is a d_0-void transformation $\hat{T}_{d',\theta}$.

For each d, let $d'_{\sigma_0^\frown\langle d\rangle,\tau_0}$ be the result of replacing each $\theta \in \Theta(d_{\sigma_0^\frown\langle d\rangle,\tau_0})$ such that $\Lambda_{\sigma_0^\frown\langle d\rangle,\tau_0}(\theta) \in \Theta(d_0)$, with $\mathcal{I}(\Lambda_{\sigma_0^\frown\langle d\rangle,\tau_0}(\theta))$ and the premise ι given by T' applied to $E_0(\Lambda_{\sigma_0^\frown\langle d\rangle,\tau_0}(\theta), \iota)$; this application exists by the main IH.

Then $d^* := d'_{\sigma_0^\frown\langle d_0\rangle,\tau_0}$ and $\Lambda := \Lambda_{\sigma_0^\frown\langle d_0\rangle,\tau_0} \upharpoonright \Theta(d^*)$ witness the theorem. □

Theorem 2.12. *If* T *is a transformation out of* $\mathcal{Q}$ *from* Γ *over* $\mathcal{F}$ *with endsequent* Σ *and* $d \vDash \Gamma, \Upsilon$ *for some* $\Upsilon \subseteq \mathcal{F}$ *then there is a derivation* $T(d)$ *of* Σ, Υ.

Proof. Apply the preceding lemma to $\mathbb{E}_{\mathcal{Q}}(d)$. □

Definition 2.13. d' broadly extends d *if* d *can be derived from* d' *by replacing subderivations of* d' *ending in callback inferences with truncated inferences. If* S *is the set of such truncated inferences in* d, *we say* d' *broadly extends* d *at* S.

Lemma 2.14. *Let* T *be a wellfounded transformation, let* $\{\mathcal{O}_i\}$ *be a set of operators on derivations, all with the same domain, and for each* $\mathcal{O}_i$, *let* $\Lambda_{\mathcal{O}_i}$ *be a function with the properties that:*

- *Each* $\mathcal{O}_i$ *takes wellfounded derivations to wellfounded derivations*
- *Each* $\mathcal{O}_i$ *preserves extensions in the sense that if* d' *narrowly extends* d *at* θ *then* $\mathcal{O}_i(d')$ *broadly extends* $\mathcal{O}_i(d)$ *at* $\{\theta' \mid \theta = \Lambda_{\mathcal{O}_i}(d)(\theta')\}$
- *For every* d *in the domain of* $\mathcal{O}_i$, $\Lambda_{\mathcal{O}_i}(d) : \Theta(\mathcal{O}_i(d)) \to \Theta(d)$ *with the property that* $Out(\theta) = Out(\Lambda_{\mathcal{O}_i}(d)(\theta))$ *and if* $d' \vDash In(\Lambda_{\mathcal{O}}(d)(\theta)), \Upsilon$ *belongs to the domain then there is an operator* $\mathcal{O}_j$ *such that* $\mathcal{O}_j(d') \vDash In(\theta), \Upsilon$.

Then each $\mathcal{O}_i$ *extends to an operator on wellfounded transformations,* $T \mapsto \mathcal{O}_i \circ T$, *with appropriate domain and range with the property that*

$$(\mathcal{O}_i \circ T)(d) = \mathcal{O}_i(T(d))$$

for any derivation d.

Proof. Follows immediately by applying operators pointwise, using $\Lambda_{\mathcal{O}_i}$ to define $\mathcal{O}_i(\Lambda)$. □

We call such a system of such operators *uniform.*

2.3 The System μ_2

2.3.1 Language

Definition 2.15. *If $A(X,x)$ is a formula, we write $A(X)$ for $\{x \mid A(X,x)\}$; in particular, $A(X) \subseteq X$ means $\forall x(A(X,x) \to x \in X)$.*

As we define our system, we also assign depths to formulas. Depths will be ordinals $\leq \omega + \omega$, although we will immediately restrict ourselves to $\omega + 2$. (The use of the ordinal $\omega + \omega$ is somewhat artificial; we have ω levels corresponding to finitely many iterated inductive definitions, and then three levels above, corresponding to the inaccessible, the negated inaccessible, and an admissible above the inaccessible. The names $< I$, I, and $I + 1$ might convey this more clearly.)

Definition 2.16. *The language of $\mathcal{L}_{\mu_2}$ is defined as follows:*

- 0 *is a constant symbol*
- S *is a unary function constant symbol*
- *There are infinitely many symbols for variables*
- *For each n-ary primitive recursive relation, including $=$ and $\leq$, there is an n-ary predicate constant symbol R*
- *The logical symbols are $\neg, \wedge, \vee, \forall, \exists$*
- *If $A(x,X)$ contains no other free variables and contains X positively then $\mu xX.A(x,X)$ is a unary predicate symbol*
- *If $B(y,Y,Z)$ contains Y positively and Z negatively and $A(x,X,Z)$ contains X positively and Z negatively, and A and B have finite depth then $\mu xX.A(x,X,\mu yY.B(y,Y,X))$ is a unary predicate symbol; we call this a predicate of* inaccessible type

The terms are given by:

- 0 *is a term*
- *If t is a term then St is a term*
- *Each variable is a term*

The formulas are given by:

- *If R is a symbol for an n-ary primitive recursive relation and for each $i \leq n$, t_i is a term, then $Rt_1 \dots t_n$ is an atomic formula of depth n for any $n \geq 0$*
- *If $A(x,X)$ has depth n and t is a term then $t \in \mu xX.A(x,X)$ is an atomic formula of depth n*
- *If t is a term then $t \in \mu xX.A(x,X,\mu yY.B(y,Y,X))$ is an atomic formula of depth ω*
- *If A is an atomic formula of depth n, $\neg A$ is a formula of depth $n+1$*
- *If A_0 and A_1 are formulas of depth n then $A_0 \wedge A_1$ and $A_0 \vee A_1$ are formulas of depth n*
- *If x is a variable and A a formula of depth n then $\forall xA$ and $\exists xA$ are formulas of depth n*

If $n < \omega$ then $\mathcal{L}_{ID_n}$ is the restriction to formulas of depth n. The depth of a formula, $dp(A)$, is the least $n \geq 0$ such that A has depth n.

If $dp(A) \geq \omega + 1$ then we call $\mu xX.A(x,X)$, and any formula containing it, large.

Our theory will effectively restrict consideration to formulas of depth at most $\omega + 2$. Note that all formulas of higher depth are large. The restriction is somewhat artificial, since we have to "throttle" the formation rule for μ-expressions, but the alternative would be analyzing a stronger system corresponding to an inaccessible with infinitely many admissibles above it. (This phenomenon has been observed before, for instance in [78], where the addition of a constructor corresponding to an inaccessible immediately pushes the system up to infinitely many admissibles beyond it due to the presence of other constructors.)

Definition 2.17. *$FV(\phi)$ denotes the set of free variables of ϕ, and ϕ is closed if $FV(\phi) = \emptyset$. Here ϕ may be a formula, a term, or a sequent.*

If A is not atomic, $\neg A$ indicates the negation of A in negation normal form as given by de Morgan's laws.

The rank $rk(A)$ of a formula is defined by:

- $rk(A) := 0$ *if A is atomic*
- $rk(\neg A) := rk(A)$

- $rk(A \wedge B) = rk(A \vee B) := \max\{rk(A), rk(B)\} + 1$
- $rk(\forall x A) = rk(\exists x A) := rk(A) + 1$

$A(x/t)$ means the result of substituting t for every free occurrence of x in A (renaming bound variables if necessary). When x is clear, we just write $A(t)$.

Definition 2.18. *We define the true primitive recursive formulas to be those closed primitive recursive atomic formulas and negations of atomic formulas which are true in the standard interpretation.*

The system μ_2 contains the following inference symbols:

$$Ax_\Delta \ \frac{}{\Delta}$$

where Δ contains a true primitive recursive formula or a pair $t \in \mu x X.A(x, X), n \notin \mu x X.A(x, X)$

$$\bigwedge_{A_0 \wedge A_1} \frac{A_0 \qquad A_1}{A_0 \wedge A_1} \qquad \bigwedge^i_{A_0 \vee A_1} \frac{A_i}{A_0 \vee A_1}$$

$i \in \{0, 1\}$

$$\bigwedge^y_{\forall x A} \frac{A(y)}{\forall x A} \ !x! \qquad \bigvee^t_{\exists x A} \frac{A(t)}{\exists x A}$$

$$Cut_C \ \frac{C \qquad \neg C}{\emptyset} \qquad Ind^t_{\mathcal{F}} \ \frac{}{\neg \mathcal{F}[0], \neg \forall x(\mathcal{F}[x] \rightarrow \mathcal{F}[Sx]), \mathcal{F}[t]}$$

where C is not large

$$Cl_{t \in \mu x X.A(x,X)} \ \frac{A(t, \mu x X.A(x, X))}{t \in \mu x X.A(x, X)}$$

$$Ind^{\mu x X.A(x,X),t}_{\mathcal{F}} \ \frac{}{\neg(A(\mathcal{F}) \subseteq \mathcal{F}), t \notin \mu x X.A(x, X), \mathcal{F}[t]}$$

We say a derivation d belongs to ID_n if every formula in every endsequent in d belongs to $\mathcal{L}_{ID_n}$.

2.3.2 Infinitary Derivations

We define an infinitary system μ_2^∞; its language is the same language $\mathcal{L}_{\mu_2}$, but only closed formulas are permitted. This definition will require that a number of weaker systems be defined along the way.

The following, which we will call ID_0^∞, will be the basis for all the systems we need. Roughly, it is the standard infinitary system for Peano Arithmetic plus a closure rule—but not an induction rule—for $\mu x X.A(x, X)$ of depth 0.

Definition 2.19. $Ax_\Delta \ \frac{}{\Delta}$

where Δ contains a true primitive recursive formula

$$\bigwedge_{A_0 \wedge A_1} \frac{A_0 \qquad A_1}{A_0 \wedge A_1} \qquad \bigvee^i_{A_0 \vee A_1} \frac{A_i}{A_0 \vee A_1}$$

$i \in \{0, 1\}$

$$\bigwedge_{\forall x A} \frac{\cdots A(i) \cdots \qquad (i \in \mathbb{N})}{\forall x A} \qquad \bigvee^n_{\exists x A} \frac{A(n)}{\exists x A}$$

$n \in \mathbb{N}$

$$Cl_{n \in \mu x X.A(x,X)} \ \frac{A(n, \mu x X.A(x, X))}{n \in \mu x X.A(x, X)} \qquad Cut_C \ \frac{C \qquad \neg C}{\emptyset}$$

and all formulas have depth 0.

Definition 2.20. *If q is a proof and Γ a sequent, $\Delta^\Gamma_q := \Gamma(q) \setminus \Gamma$.*

The systems ID^∞_{n+1} are defined inductively; as the name suggests, they are essentially the infinitary systems from [19].

Definition 2.21. *Given ID^∞_n, the language of the system ID^∞_{n+1} is $\mathcal{L}_{ID_n}$—that is, formulas with depth $\leq n + 1$, and consists of the rules of ID^∞_n together with*

$$Ax_\Delta \ \frac{}{\Delta}$$

where Δ contains $n \in \mu x X.A(x, X), n \notin \mu x X.A(x, X)$ with $dp(\mu x X.A(x, X)) < n + 1$

$$\Omega_{k\notin\mu xX.A(x,X)}\ \frac{k\in\mu xX.A(x,X)\qquad \ldots\Delta_q^{k\in\mu xX.A(x,X)}\ldots\qquad (q\in|k\in\mu xX.A(x,X)|)}{\emptyset}$$

where $|k \in \mu xX.A(x,X)|$ is the set of cut-free proofs of $ID^{\infty}_{dp(k\in\mu xX.A(x,X))}$ and $dp(\mu xX.A(x,X)) \leq n$, and $\Delta_q(\Omega_{k\notin\mu xX.A(x,X)}) := \Upsilon$ where $q \vdash k \in \mu xX.A(x,X), \Upsilon$

Note that the premise of the Ω rule d defines a function taking proofs of $k \in \mu xX.A(x,X)$ to proofs of $\Gamma(d)$.

Definition 2.22. *Next we define a system μ^{∞}_{ω}, which extends the union of ID^{∞}_n over n with the closure rule for predicates of inaccessible type.*

Note that this doesn't add any derivations—there's no way to introduce $A(n,\mu_A)$ since there's no way to introduce $n \notin \mu_B(\mu_A)$. We're including the rule so that it will be present in the extensions we need.

Definition 2.23. *The system μ^{∞}_{I} extends μ^{∞}_{ω} by the rule*

$$\neg_{n\notin\mu xX.A(x,X)}\ \frac{n\in\mu xX.A(x,X,\mu_1,\ldots,\mu_k)\qquad \ldots d_{\sigma,\tau}\ldots}{\emptyset}$$

where $\mu_1,\ldots,\mu_k$ are predicates of inaccessible type appearing negatively in A, no other predicates of inaccessible type appear in A, and for every $\mathcal{F}_1,\ldots,\mathcal{F}_k$, the premises include a well-founded transformation from $n \in \mu xX.A(x,X,\mathcal{F}_1,\ldots,\mathcal{F}_k)$ out of the cut-free part of μ^{∞}_{ω} over $\mu xX.A(x,X,\mathcal{F}_1,\ldots,\mathcal{F}_k)$ positive formulas.

Now we can define our final system:

Definition 2.24. *The system μ^{∞} consists of μ^{∞}_{I} plus the rules*

$$Ax_{\Delta}\ \frac{}{\Delta}$$

where Δ contains $n \in \mu xX.A(x,X), n \notin \mu xX.A(x,X)$ and $\mu xX.A(x,X)$ has inaccessible type

$$\Omega_{n\notin\mu xX.A(x,X)}\ \frac{n\in\mu xX.A(x,X)\qquad \ldots d_{\sigma,\tau}\ldots}{\emptyset}$$

where the premises range over cut-free proofs of μ^{∞}_{I}

Note that none of these systems allow cut rules over large formulas.

Definition 2.25. *Given a system $\mathcal{P}$, the* augmentations *of $\mathcal{P}$ are given inductively: $\mathcal{P}$ is an augmentation of $\mathcal{P}$, and if $\mathcal{Q}$ is an augmentation of $\mathcal{P}$ then so is augmented $\mathcal{Q}$.*

Definition 2.26. *We define $c - rk(d)$, the cut-rank of d, inductively as follows:*

$$c - rk(d) = \max\{c - rk(d_\iota) \mid \iota \in |\mathcal{I}|\}$$

unless $\mathcal{I} = Cut_C$

$$c - rk(Cut_C(d_0, d_1)) = \max\{c - rk(d_0), c - rk(d_1), rk(C) + 1\}$$

2.4 Embedding

Definition 2.27. *A derivation in μ is closed if every number variable occurring free is the eigenvariable of an inference below that occurrence. In particular, $FV(\Gamma(h)) = \emptyset$ if h is closed.*

We will define a function taking closed proofs in μ_2 to proofs in μ^{∞}. The hard part will be the induction axioms, which will be embedded as Ω rules. Most of the work is defining the functions used to make these Ω rules.

Definition 2.28. *Let $\mathbf{d}_{\mathcal{F},\neg\mathcal{F}}$ be the canonical derivation of $\mathcal{F}, \neg\mathcal{F}$.*

If $d \vdash A(n,\mathcal{F})$ then $\mathbf{e}^{n}_{\mathcal{F},A}(d)$ is the derivation

$$\cfrac{\cfrac{\begin{matrix} d \\ \vdots \\ A(n,\mathcal{F}) \end{matrix}\qquad \begin{matrix} \mathbf{d}_{\mathcal{F}[n],\neg\mathcal{F}[n]} \\ \vdots \\ \mathcal{F}[n],\neg\mathcal{F}[n] \end{matrix}}{\mathcal{F}[n], A(n,\mathcal{F})\wedge\neg\mathcal{F}[n]}}{\mathcal{F}[n], \neg(A(\mathcal{F})\subseteq\mathcal{F})}$$

or symbolically

$$\bigvee_{\neg(A(\mathcal{F})\subseteq\mathcal{F})}^{n} \bigwedge_{A(n,\mathcal{F})\wedge\neg\mathcal{F}[n]} d\mathbf{d}_{\neg(\mathcal{F}[n]),\mathcal{F}[n]}$$

Lemma 2.29. *There is a function $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}$ such that if $dp(\mu xX.A(x,X))<\omega$ and $d\vdash\Pi(\mu xX.A(x,X)),\Sigma$ is a cut-free proof in $ID^{\infty}_{dp(\mu xX.A(x,X))}$ then*

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d)\vdash\Pi(\mathcal{F}),\neg(A(\mathcal{F})\subseteq\mathcal{F}),\Sigma$$

is a proof in μ^{∞}.

Proof. By induction on d. We simply proceed up through the proof, adding to Π as we encounter subformulas or new formulas produced by closures rules. A typical case is

$$\frac{B_0(\mu xX.A(x,X)) \qquad B_1(\mu xX.A(x,x))}{B_0(\mu xX.A(x,X))\wedge B_1(\mu xX.A(x,x))} \quad\mapsto\quad \frac{B_0(\mathcal{F}) \qquad B_1(\mathcal{F})}{B_0(\mathcal{F})\wedge B_1(\mathcal{F})}$$

where $B_0\wedge B_1$ belongs to Π.

The only difficult case is the closure rule, which we handle with the help of $\mathbf{e}$:

$$\frac{A(n,\mu xX.A(x,X))}{n\in\mu xX.A(x,X)} \quad\mapsto\quad \cfrac{\cfrac{A(n,\mathcal{F}) \qquad \begin{array}{c}\mathbf{d}_{\mathcal{F}[n],\neg\mathcal{F}[n]}\\ \vdots\\ \mathcal{F}[n],\neg\mathcal{F}[n]\end{array}}{\mathcal{F}[n],A(n,\mathcal{F})\wedge\neg\mathcal{F}[n]}}{\mathcal{F}[n],\neg(A(\mathcal{F})\subseteq\mathcal{F})}$$

Importantly, we never encounter $n\notin\mu xX.A(x,X)$ anywhere; in particular, we do not have to deal with the axiom $Ax_{n\in\mu xX.A(x,X),n\notin\mu xX.A(x,X)}$.

The full definition is given by

$$\mathcal{SUB}^{\Pi}_{n\in\mu xX.A(x,X),\mathcal{F}}(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}):=$$

$$\begin{cases} \mathbf{e}^{n}_{\mathcal{F},\mathcal{G}}(\mathcal{SUB}^{\Pi\cup\{\Delta_0(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_0)) & \text{if } \mathcal{I}=Cl_{n\in\mu xX.A(x,X)} \\ & \text{ and } n\in\mu xX.A(x,X)\in\Pi \\ \mathcal{I}_{A(\mathcal{F})}\left(\mathcal{SUB}^{\Pi\cup\{\Delta_\iota(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|}\right) & \text{if } \mathcal{I}=\mathcal{I}_{B(\mu xX.A(x,X))} \\ & \text{ and } B(\mu xX.A(x,X))\in\Pi \\ \mathcal{I}_A(\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|}) & \text{otherwise} \end{cases}$$

□

Lemma 2.30. *Let $A(x,X)$ be a formula. Then there is an operator $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}$ such that if $d\vdash\Pi(\mu xX.A(x,X)),\Sigma$ is a cut-free proof in an augmentation of $\mu^{\infty}_{<I}$ then*

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d)\vdash\Pi(\mathcal{F}),\neg(A(\mathcal{F})\subseteq\mathcal{F}),\Sigma$$

is a proof in the corresponding augmentation of μ^{∞}. Furthermore, this operator is uniform.

Proof. By induction on d. The proof is essentially the same as in the previous lemma, except that we add an additional case to handle $Trunc$ and CB inferences.

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}):=$$

$$\begin{cases} \mathbf{e}^{n}_{\mathcal{F}}(\mathcal{SUB}^{\Pi\cup\{\Delta_0(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_0)) & \text{if } \mathcal{I}=Cl_{n\in\mu xX.A(x,X)} \\ & \text{ and } n\in\mu xX.A(x,X)\in\Pi \\ Trunc_{F\mapsto\Pi(\mathcal{F}),\Sigma}\mathcal{SUB}^{\Pi}_{\mu xX.A(x,x),\mathcal{F}}(d_0) & \text{if } \mathcal{I}=Trunc_{F\mapsto\Pi(\mu xX.A(x,X)),\Sigma} \\ CB_{\Pi(\mathcal{F}),\Sigma}\mathcal{SUB}^{\Pi}_{\mu xX.A(x,x),\mathcal{F}}(d_0) & \text{if } \mathcal{I}=CB_{\Pi(\mu xX.A(x,X)),\Sigma} \\ \mathcal{I}_{A(\mathcal{F})}\left(\mathcal{SUB}^{\Pi\cup\{\Delta_\iota(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|}\right) & \text{if } \mathcal{I}=\mathcal{I}_{B(\mu xX.A(x,X))} \\ & \text{ and } B(\mu xX.A(x,X))\in\Pi \\ \mathcal{I}_A(\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|}) & \text{otherwise} \end{cases}$$

$\square$

Lemma 2.31. *Let $A(x,X)$ be a formula. Then there is an operator $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}$, and a companion $\Lambda_{\sigma,\tau}$, giving a well-founded transformation from $n \in \mu xX.A(x,X)$ out of $\mu^{\infty}_{<I}$ over $\mu xX.A(x,X)$ positive formulas with endsequent $\mathcal{F}[n], \neg(A(\mathcal{F}) \subseteq \mathcal{F})$.*

Proof. By induction on the length of σ. $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X,\mathcal{G}_1,\mathcal{G}_k),\mathcal{F}}(\langle d_0\rangle,\langle\rangle)$ is just $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d_0)$ as given by the previous lemma. Given $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\sigma,\tau)$, $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\sigma^\frown\langle d\rangle,\tau^\frown\langle\theta\rangle)$ is given by replacing θ in $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\sigma,\tau)$ with the derivation $\mathcal{SUB}^{\Pi'}_{\mu xX.A(x,X),\mathcal{F}}(d)$ where Π' is chosen to be the unique sequent such that θ was equal to $\mathcal{SUB}^{\Pi'}_{\mu xX.A(x,X),\mathcal{F}}(\Lambda_{\sigma,\tau}(\theta))$.

The function $\Lambda_{\sigma,\tau}$ is simply the association of each truncated inference in the range with the corresponding inference in the domain. $\square$

At last, we come to the key lemma:

Lemma 2.32. *If $\mu xX.A(x,X)$ has inaccessible type, there is an operator $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}$ such that whenever d is a proof of $\Pi(\mu xX.A(x,X)),\Gamma$ in μ^{∞}_{I} then*

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d) \vdash \Pi(\mathcal{F}), \Gamma, \neg(A(\mathcal{F}) \subseteq \mathcal{F})$$

Furthermore, $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}$ is uniform.

Proof. First, the simple cases are given by

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}, F\}) :=$$

$$\begin{cases} \mathbf{e}^{n}_{\mathcal{F}}(\mathcal{SUB}^{\Pi\cup\{\Delta_0(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_0)) & \text{if } \mathcal{I} = Cl_{n\in\mu xX.A(x,X)} \\ & \text{ and } n \in \mu xX.A(x,X) \in \Pi \\ \mathcal{I}_{A(\mathcal{F})}\left(\mathcal{SUB}^{\Pi\cup\{\Delta_\iota(\mathcal{I})\}}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|}\right) & \text{if } \mathcal{I} = \mathcal{I}_{A(\mu xX.A(x,X))} \\ & \text{ and } A(\mu xX.A(x,X)) \in \Pi \end{cases}$$

Next, consider $\neg_{n\notin\mu yY.B(y,Y,\mu xX.A(x,X),\mu_2,\ldots,\mu_k)}$ where $n \notin \mu yY.B(y,Y,\mu xX.A(x,X),\mu_2,\ldots,\mu_k) \in \Pi$. We use the abbreviations μ_A and $\mu_B(\mu_A)$ as in the introduction, and let T be the transformation formed by the premises. First, consider the simplest case, where $\mathcal{F}$ does not contain predicates of inaccessible type and $k=1$. Then we simply need to produce a function for an $\Omega_{n\notin\mu_B(\mathcal{F})}$ inference.

Since the premises give a transformation showing $n \in \mu_B(\mathcal{F}) \mapsto \Pi(\mu xX.A(x,X)), \Gamma\setminus\{n \notin \mu yY.B(y,Y,\mu xX.A(x,X))\}$, also $\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}} \circ d_{\sigma,\tau}$ gives a transformation T showing $n \in \mu_B(\mathcal{F}) \mapsto \Pi(\mathcal{F}), \Gamma\setminus\{n \notin \mu yY.B(y,Y,\mathcal{F})\}$. Then we may assign to each $q \vdash n \in \mu_B(\mathcal{F}), \Upsilon$ the derivation

$$d_q := \mathbb{U}(T_*(\mathbb{E}_{\mu^{\infty}_{I}}(q)))$$

More generally, if predicates of inaccessible type occur in $\mathcal{F}$ or $k > 1$ the same argument gives many transformations which collectively witness the corresponding $\neg$ inference.

In any other case, we do nothing:

$$\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(\mathcal{I}(d_\iota)) := \mathcal{I}_A(\mathcal{SUB}^{\Pi}_{\mu xX.A(x,X),\mathcal{F}}(d_\iota)_{\iota\in|\mathcal{I}|})$$

$\square$

Lemma 2.33. *If h is a closed μ_2 derivation of Δ with $dp(\Delta) \leq \omega + 2$ then there is a μ^{∞} derivation h^{∞} so that $h^{\infty} \vdash_m \Gamma(h)$ for some finite m.*

Proof. We define the $\cdot^{\infty}$ operation by induction on the proof h:

- $(\bigwedge^{y}_{\forall xA} d_0)^{\infty} := \bigwedge_{\forall xA}(d_0[n]^{\infty})_{n\in\mathbb{N}}$
- $(Ind^{0}_{\mathcal{F}})^{\infty} := \mathbf{d}_{\mathcal{F}[0],\neg\mathcal{F}[0]}$

- $(Ind^{n+1}_{\mathcal{F}})^\infty := \bigvee^n_{\exists x(\mathcal{F}[x]\wedge\neg\mathcal{F}[Sx])} \bigwedge_{\mathcal{F}[n]\wedge\neg\mathcal{F}[Sn]} (Ind^n_{\mathcal{F}})^\infty \mathbf{d}_{\neg\mathcal{F}[Sn],\mathcal{F}[Sn]}$
- $(Ind^{\mu_A,n}_{\mathcal{F}})^\infty := \Omega_{n\notin\mu_A} Ax_{n\in\mu_A,n\notin\mu_A}\{\mathcal{SUB}^{n\in\mu_A}_{\mu_A,\mathcal{F}}(q)\}$ if $dp(\mu_A) < \omega$ or has inaccessible type
- $(Ind^{\mu_A,n}_{\mathcal{F}})^\infty := \neg_{n\notin\mu_A} Ax_{n\in\mu_A,n\notin\mu_A}\{\mathcal{SUB}^{n\in\mu_A(\mathcal{F}_1,\ldots,\mathcal{F}_k)}_{\mu_A(\mathcal{F}_1,\ldots,\mathcal{F}_k)),\mathcal{F}}(\sigma,\tau)\}_{\sigma,\tau,\mathcal{F}_1,\ldots,\mathcal{F}_k}$ if $dp(\mu_A) \geq \omega$, $\mu_A(\mu_1,\ldots,\mu_k)$ does not have inaccessible type, and the μ_i are all predicates of inaccessible type appearing in A
- Otherwise $(\mathcal{I}h_0\ldots h_{n-1})^\infty := \mathcal{I}h_0^\infty \ldots h_{n-1}^\infty$

□

2.5 Cut-Elimination

Definition 2.34. *We say that A has $\bigwedge$-Form if it is either $A_0 \wedge A_1$ or $\forall x A_0$.*

We say that A has $\bigwedge^+$-Form if it has $\bigwedge$-Form, is a true primitive recursive formula, or has the form $\mu xX.A(x,X)n$. Define

$$C[k] := \begin{cases} C_k & \text{if } C = C_0 \wedge C_1 \text{ or } C = C_0 \vee C_1 \text{ where } k \in \{0,1\} \\ A(k) & \text{if } C = \forall x A \text{ or } C = \exists x A \text{ where } k \in \mathbb{N} \end{cases}$$

Lemma 2.35. *If C is a $\bigwedge$-Form then there is a uniform operator $\mathcal{J}^k_C$ such that whenever $d \vdash_m \Gamma, C$, $\mathcal{J}^k_C(d) \vdash_m \Gamma, C[k]$.*

Proof. By induction on d.

$$\mathcal{J}^k_C(d) := \begin{cases} \mathcal{J}^k_C(d_k) & \text{if } \mathcal{I} = \bigwedge_C \\ CB_{F\mapsto\Sigma,C[k]}(\mathcal{J}^k_C(d_0)) & \text{if } \mathcal{I} = CB_{F\mapsto\Sigma,C} \\ \neg\{\mathcal{J}^k_C \circ \mathcal{F}_q\}_q & \text{if } \mathcal{I} = \neg\{\mathcal{F}_q\}_q \\ \mathcal{I}(\mathcal{J}^k_C(d_\iota))_{\iota\in|\mathcal{I}|} & \text{otherwise} \end{cases}$$

□

Lemma 2.36. *Let $rk(C) \leq m$ with $\bigwedge^+$-Form and $e \vdash_m \Gamma, C$. Then there is an operator $\mathcal{R}_C(e,\cdot)$ such that whenever $d \vdash_m \Gamma, \neg C$, $\mathcal{R}_C(e,d) \vdash_m \Gamma$ and such that $\{\mathcal{R}_C\} \cup \{\mathcal{J}^k_D\}_{k,D}$ is uniform.*

Proof. By induction on d.

$$\mathcal{R}_C(e,d) := \begin{cases} Cut_{C[k]} \mathcal{J}^k_C(e) \mathcal{R}_C(e,d_k) & \text{if } \mathcal{I} = \bigvee^k_{\neg C} \\ e & \text{if } \mathcal{I} = Ax_{\neg C,C} \\ CB_{F\mapsto\Sigma} \mathcal{R}_C(e,d_0) & \text{if } \mathcal{I} = CB_{F\mapsto\Sigma,\neg C} \\ \neg\{\mathcal{R}_C \circ \mathcal{F}_q\}_q & \text{if } \mathcal{I} = \neg\{\mathcal{F}_q\}_q \\ \mathcal{I}(\mathcal{R}_C(e,d_\iota))_{\iota\in|\mathcal{I}|} & \text{otherwise} \end{cases}$$

□

Lemma 2.37. *For each m, there is an operator $\mathcal{E}_m$ so that whenever $d \vdash_{m+1} \Gamma$, $\mathcal{E}_m(d) \vdash_m \Gamma$ and $\{\mathcal{E}_m\} \cup \{\mathcal{R}_C\}_C \cup \{\mathcal{J}^k_D\}_{k,D}$ is uniform.*

Proof. By induction on d.

$$\mathcal{E}_m(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}) := \begin{cases} \mathcal{R}_C(\mathcal{E}_m(d_0), \mathcal{E}_m(d_1)) & \text{if } \mathcal{I} = Cut_C,\ rk(C) = m \\ & \text{ and } C \text{ has } \bigwedge^+\text{-Form} \\ \mathcal{R}_{\neg C}(\mathcal{E}_m(d_1), \mathcal{E}_m(d_0)) & \text{if } \mathcal{I} = Cut_C,\ rk(C) = m \\ & \text{ and } \neg C \text{ has } \bigwedge^+\text{-Form} \\ \neg\{\mathcal{E}_m \circ \mathcal{F}_q\}_q & \text{if } \mathcal{I} = \neg\{\mathcal{F}_q\}_q \\ \mathcal{I}(\mathcal{E}_m(d_\iota))_{\iota\in|\mathcal{I}|} & \text{otherwise} \end{cases}$$

□

Lemma 2.38. *There is a uniform operator $\mathcal{D}_I$ such that if Γ does not contain predicates of inaccessible type negatively and $d \vdash_0 \Gamma$ then $\mathcal{D}_I(d) \vdash \Gamma$ and $\mathcal{D}_I(d) \in \mu_I^\infty$.*

Proof. By induction on d.

$$\mathcal{D}_I(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}) := \begin{cases} \mathcal{D}_I \circ F & \text{if } \mathcal{I} = \Omega_{n\notin\mu xX.A(x,X)} \\ & \text{and } \mu xX.A(x,X) \text{ has inaccessible type} \\ \neg\{\mathcal{D}_I \circ \mathcal{F}_q\}_q & \text{if } \mathcal{I} = \neg\{\mathcal{F}_q\}_q \\ \mathcal{I}(\mathcal{D}_I(d_\iota))_{\iota\in|\mathcal{I}|} & \text{otherwise} \end{cases}$$

□

Lemma 2.39. *There is an operator $\mathcal{D}_n$ such that if $d \vdash_0 \Gamma$ and $dp(\Gamma) \leq n$ then $\mathcal{D}_n(d) \vdash_0 \Gamma$ and is a proof in ID_n^∞.*

Proof. By induction on d.

$$\mathcal{D}_n(\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}) := \begin{cases} \mathcal{D}_n(d_{\mathcal{D}_m(d_0)}) & \text{if } \mathcal{I} = \Omega_{n\notin\mu xX.A(x,X)} \\ & \text{and } dp(\mu xX.A(x,X)) = m \geq n \\ \mathcal{D}_n(d_{\mathcal{D}_I(d_0)}) & \text{if } \mathcal{I} = \Omega_{n\in\mu xX.A(x,X)} \\ & \text{and } \mu xX.A(x,X) \text{ has inaccessible type} \\ \mathcal{I}(\mathcal{D}_n(d_\iota))_{\iota\in|\mathcal{I}|} & \text{otherwise} \end{cases}$$

□

Theorem 2.40. *Let d be a proof in μ_2 of a sequent Γ of depth 0. Then there is a cut-free proof d^* of Γ in ID_0^∞. Furthermore, the existence may be shown in a constructive theory.*

Proof. Let $d^* := \mathcal{D}_0(\mathcal{E}_0(\cdots(\mathcal{E}_m(d^\infty))))$. Then d^* is a cut-free proof in ID_0^∞.

Constructivity follows via continuous cut-elimination carried in an appropriate constructive system; for specificity, intuitionistic $\Pi^1_2 - CA$ would be (more than) sufficient to formalize each instance of this argument. Although the derivations are nominally infinite, they can be replaced with finitary descriptions, with branches only produced when they are actually used. Since all our transformations are defined continuously, they remain well-defined in this context. □

Theorem 2.41. *μ_2 is consistent.*

Proof. If there is a proof of $0 = 1$ in μ_2 then there is a cut-free proof in ID_0^∞. But the cut-free proofs of primitive recursive formulas are also proofs in IS, so there is a cut-free proof of $0 = 1$ in μ_2. But this is impossible, since no inference rule other than cut can produce this as an end-sequent. □

Chapter 3

A Realizability Interpretation for Second Order Arithmetic

Although both classical and intuitionistic arithmetic prove the same Π_2 sentences, proofs in the intuitionistic version generally provide more information. The Curry-Howard isomorphism associates them with realizing λ terms, which associate numerical witnesses to existential quantifiers and appropriate functionals to strings of quantifiers.

Avigad [6] demonstrates a method of extending this realization to classical arithmetic to find numerical witnesses to Σ_1 sentences and type 1 functions witnessing Π_2 sentences. This method of witness extraction was derived from the composition of an embedding of classical logic in intuitionistic logic, the Friedman-Dragalin translation (first described in [33] and [28]), and the Curry-Howard isomorphism.

3.1 Preliminaries

A Tait style calculus based on the one in [77] will be used for PA^2. The primary difference is that $\neg$ is taken as a connective, rather than a shorthand for the negation-normal form. Atomic formulae will be either of the form $s = t$ or $Xt_1 \ldots t_n$ (where $s, t, t_1, \ldots, t_n$ are terms and X is an n-ary second order variable). The connectives will be $\neg$, $\vee$, $\exists$, and $\exists^2$. Other connectives can be defined in the usual way. We will write $\lambda\vec{y}.B$ for the predicate given by a formula $B(\vec{y})$.

The rules of this system will be:

1. Propositional Rules
   (a) $\Gamma, A, \neg A$ for any atomic A
   (b) From $\Gamma, \neg\phi$ and $\Gamma, \neg\psi$ conclude $\Gamma, \neg(\phi \vee \psi)$
   (c) From Γ, ϕ conclude $\Gamma, \phi \vee \psi$ and $\Gamma, \psi \vee \phi$
   (d) From Γ, ϕ and $\Gamma, \neg\phi$ conclude Γ
2. Quantifier rules
   (a) From $\Gamma, \neg\phi(y)$ conclude $\Gamma, \neg\exists x\phi(x)$ if y does not occur free in any formula of Γ
   (b) From $\Gamma, \neg\phi(Y)$ conclude $\Gamma, \neg\exists^2 X\phi(X)$ if Y does not occur free in any formula of Γ
   (c) From $\Gamma, \phi(t)$ conclude $\Gamma, \exists x\phi(x)$
   (d) From $\Gamma, \phi(\lambda\vec{y}.B)$ conclude $\Gamma, \exists^2 X\phi(X)$
3. Equality rules (quantifier free)
   - $\Gamma, t = t$ for any term t
   - From $\Gamma, t_1 = t_2$ conclude $\Gamma, t_2 = t_1$ for any terms t_1 and t_2

- From $\Gamma, t_1 = t_2$ and $\Gamma, \phi(t_1)$ conclude $\Gamma, \phi(t_2)$ for any terms t_1 and t_2

4. Arithmetical rules
   (a) Quantifier-free defining equations for all primitive recursive relations and functions
   (b) From $\Gamma, \neg\phi(0)$ and $\Gamma, \phi(y), \neg\phi(Sy)$ conclude $\Gamma, \neg\exists x\phi(x)$ if y does not occur free in Γ

All other normal rules of second order arithmetic can be derived from these, for example:

$$\frac{\Gamma, \phi \qquad \Gamma, \neg\phi, \neg\neg\phi}{\Gamma, \neg\neg\phi}$$

If $\Gamma = \{\phi_1, \ldots, \phi_k\}$ then $\neg\Gamma = \{\neg\phi_1, \ldots, \neg\phi_k\}$.

Intuitionistic logic and HA^2 will be given by a system of natural deduction with connectives $\forall$, $\exists$, $\exists^2$, $\vee$, and $\rightarrow$ ($\exists$ and $\vee$ are redundant, but it is more convenient to include them; $\forall^2$ and $\wedge$ will not be needed, so they are excluded).

3.2 Friedman-Dragalin Translation

A formula ϕ of PA^2 can be associated with a formula $\phi^{\neg\neg}$ of HA^2 such that $PA^2 \vdash \phi \Leftrightarrow HA^2 \vdash \phi^{\neg\neg}$. The embedding E used here is simpler, although the result proved will be correspondingly weaker:

- $\phi^E \equiv \phi$ for atomic ϕ
- $(\neg\phi)^E \equiv \phi^E \rightarrow \perp$
- $(\phi \vee \psi)^E \equiv \phi^E \vee \psi^E$
- $(\exists x\phi(x))^E \equiv \exists x\phi(x)^E$
- $(\exists X\phi(X))^E \equiv \exists X\phi(X)^E$

Given a fixed formula α of HA^2, a translation $FD(\alpha)$ of formulas within HA^2 can be defined so that $\alpha \rightarrow \phi^{FD(\alpha)}$ for every ϕ:

- $\phi^{FD(\alpha)} \equiv \phi$ (for $\phi = Xt_1 \ldots t_n$)
- $\phi^{FD(\alpha)} \equiv \phi \vee \alpha$ (for other atomic ϕ)
- $\perp^{FD(\alpha)} \equiv \alpha$
- $(\phi \rightarrow \psi)^{FD(\alpha)} \equiv \phi^{FD(\alpha)} \rightarrow \psi^{FD(\alpha)}$
- $(\phi \vee \psi)^{FD(\alpha)} \equiv \phi^{FD(\alpha)} \vee \psi^{FD(\alpha)}$
- $(\exists x\phi(x))^{FD(\alpha)} \equiv \exists x\phi(x)^{FD(\alpha)}$
- $(\exists X\phi(X))^{FD(\alpha)} \equiv \exists X\phi(X)^{FD(\alpha)}$

Note that $(Xt_1 \ldots t_n)^{FD(\alpha)} = Xt_1 \ldots t_n$ is not itself implied by α unless the range of X is restricted to the range of $FD(\alpha)$. This is necessary to ensure that $FD(\alpha)$ commutes with substitution.

When composed these operations give a transformation N from formulas of PA^2 to formulas of HA^2:

- $\phi^N \equiv \phi$ (for $\phi = Xt_1 \ldots t_n$)
- $\phi^N \equiv \phi \vee \alpha$ (for other atomic ϕ)
- $(\neg\phi)^N \equiv \phi^N \rightarrow \alpha$
- $(\phi \vee \psi)^N \equiv \phi^N \vee \psi^N$
- $(\exists x\phi(x))^N \equiv \exists x\phi(x)^N$
- $(\exists X\phi(X))^N \equiv \exists X\phi(X)^N$

Lemma 3.1. *The N-translation commutes with substitution:*

$$\phi(\lambda\vec{y}.B)^N = (\lambda Y.\phi(Y)^N)(\lambda\vec{y}.B^N)$$

or, equivalently:

$$(\phi[\lambda\vec{y}.B/Y])^N = \phi^N[\lambda\vec{y}.B^N/Y]$$

Proof. By induction on $\phi(Y)$. When $\phi(Y) \neq Yt_1 \dots t_n$, just apply the inductive hypothesis. When $\phi(Y) = Yt_1 \dots t_n$ then $\phi(\lambda\vec{y}.B)^N = (Bt_1 \dots t_n)^N = B^N t_1 \dots t_n$ while $(\lambda Y.\phi(Y)^N)(\lambda y.B^N) = (\lambda Y.Yt_1 \dots t_n)(\lambda\vec{y}.B^N) = B^N t_1 \dots t_n$. □

Lemma 3.2. *If $d : \Gamma$ is a proof in PA^2 then $(\neg\Gamma)^N \vdash \alpha$ is provable in HA^2.*

Proof. Proved by induction on the last step of d. The following two deductions will be used repeatedly:

$$\dfrac{\dfrac{\Gamma,\phi\Rightarrow\alpha}{\Gamma\Rightarrow\phi\to\alpha} \qquad (\phi\to\alpha)\to\alpha\Rightarrow(\phi\to\alpha)\to\alpha}{\Gamma,(\phi\to\alpha)\to\alpha\Rightarrow\alpha}$$

$$\dfrac{\dfrac{\Gamma,(\phi\to\alpha)\to\alpha\Rightarrow\alpha}{\Gamma\Rightarrow((\phi\to\alpha)\to\alpha)\to\alpha} \qquad \dfrac{\dfrac{\phi\Rightarrow\phi \qquad \phi\to\alpha\Rightarrow\phi\to\alpha}{\phi,\phi\to\alpha\Rightarrow\alpha}}{\phi\Rightarrow(\phi\to\alpha)\to\alpha}}{\Gamma,\phi\Rightarrow\alpha}$$

- If d is just the axiom $\Gamma, A, \neg A$ then either $(\neg A)^N = A\vee\alpha\to\alpha$ and $(\neg\neg A)^N = (A\vee\alpha\to\alpha)\to\alpha$ or $(\neg A)^N = A\to\alpha$ and $(\neg\neg A)^N = (A\to\alpha)\to\alpha$. In either case, α follows by $\to E$.
- If d concludes $\Gamma, \neg(\phi\vee\psi)$ from $\Gamma,\neg\phi$ and $\Gamma,\neg\psi$ then:

$$\dfrac{\dfrac{\dfrac{(\neg\Gamma)^N,(\phi^N\to\alpha)\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N,\phi^N\Rightarrow\alpha} \qquad \dfrac{(\neg\Gamma)^N,(\psi^N\to\alpha)\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N,\psi^N\Rightarrow\alpha}}{(\neg\Gamma)^N,\phi^N\vee\psi^N\Rightarrow\alpha}}{(\neg\Gamma)^N,(\phi^N\vee\psi^N\to\alpha)\to\alpha\Rightarrow\alpha}$$

- If d concludes $\Gamma, \phi\vee\psi$ from Γ,ϕ (the case for Γ,ψ is similar) then:

$$\dfrac{\dfrac{(\phi^N\vee\psi^N)\to\alpha\Rightarrow(\phi^N\vee\psi^N)\to\alpha \qquad \dfrac{\phi^N\Rightarrow\phi^N}{\phi^N\Rightarrow\phi^N\vee\psi^N}}{(\phi^N\vee\psi^N)\to\alpha,\phi^N\Rightarrow\alpha}}{(\phi^N\vee\psi^N)\to\alpha\Rightarrow\phi^N\to\alpha}$$

and

$$\dfrac{\dfrac{(\neg\Gamma)^N,\phi^N\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N\Rightarrow(\phi^N\to\alpha)\to\alpha} \qquad (\phi^N\vee\psi^N)\to\alpha\Rightarrow\phi^N\to\alpha}{(\neg\Gamma)^N,(\phi^N\vee\psi^N)\to\alpha\Rightarrow\alpha}$$

- If d concludes Γ from Γ,ϕ and $\Gamma,\neg\phi$ then:

$$\dfrac{\dfrac{(\neg\Gamma)^N,\phi^N\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N\Rightarrow(\phi^N\to\alpha)\to\alpha} \qquad \dfrac{(\neg\Gamma)^N,(\phi^N\to\alpha)\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N\Rightarrow((\phi^N\to\alpha)\to\alpha)\to\alpha}}{(\neg\Gamma)^N\Rightarrow\alpha}$$

- If d concludes $\Gamma,\neg\exists x\phi(x)$ from $\Gamma,\neg\phi(y)$ then:

$$\dfrac{\dfrac{\dfrac{(\neg\Gamma)^N,(\phi(y)^N\to\alpha)\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N,\phi(y)^N\Rightarrow\alpha} \qquad \exists x\phi(x)^N\Rightarrow\exists x\phi(x)^N}{(\neg\Gamma)^N,\exists x\phi(x)^N\Rightarrow\alpha}}{(\neg\Gamma)^N,(\exists x\phi(x)^N\to\alpha)\to\alpha\Rightarrow\alpha}$$

- If d concludes $\Gamma,\exists x\phi(x)$ from $\Gamma,\phi(t)$ then:

$$\dfrac{\dfrac{\dfrac{\phi(t)^N\Rightarrow\phi(t)^N}{\phi(t)^N\Rightarrow\exists x\phi(x)^N} \qquad \exists x\phi(x)^N\to\alpha\Rightarrow\exists x\phi(x)^N\to\alpha}{\exists x\phi(x)^N\to\alpha,\phi(t)^N\Rightarrow\alpha}}{\exists x\phi(x)^N\to\alpha\Rightarrow\phi(t)^N\to\alpha}$$

and

$$\dfrac{\dfrac{(\neg\Gamma)^N,\phi(t)^N\to\alpha\Rightarrow\alpha}{(\neg\Gamma)^N\Rightarrow(\phi(t)^N\to\alpha)\to\alpha} \qquad \exists x\phi(x)^N\to\alpha\Rightarrow\phi(t)^N\to\alpha}{(\neg\Gamma)^N,\exists x\phi(x)^N\to\alpha\Rightarrow\alpha}$$

- If d concludes $\Gamma, \neg\exists x\phi(x)$ from $\Gamma, \phi(0)$ and $\Gamma, \neg\phi(y), \phi(Sy)$ then:

$$\dfrac{\dfrac{(\neg\Gamma)^N, \phi(y)^N \to \alpha, (\phi(Sy)^N \to \alpha) \to \alpha \Rightarrow \alpha}{(\neg\Gamma)^N, \phi(y)^N \to \alpha, \phi(Sy)^N \Rightarrow \alpha}}{(\neg\Gamma)^N, \phi(y)^N \to \alpha \Rightarrow \phi(Sy)^N \to \alpha}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{(\neg\Gamma)^N, (\phi(0)^N \to \alpha) \to \alpha \Rightarrow \alpha}{(\neg\Gamma)^N, \phi(0)^N \Rightarrow \alpha}}{(\neg\Gamma)^N \Rightarrow \phi(0)^N \to \alpha} \quad (\neg\Gamma)^N, \phi(y)^N \to \alpha \Rightarrow \phi(Sy)^N \to \alpha}{(\neg\Gamma)^N \Rightarrow \forall x[\phi(x)^N \to \alpha]}}{(\neg\Gamma)^N \Rightarrow \phi(y)^N \to \alpha} \quad \phi(y)^N \Rightarrow \phi(y)^N}{(\neg\Gamma)^N, \phi(y)^N \Rightarrow \alpha}$$

and

$$\dfrac{\dfrac{(\neg\Gamma)^N, \phi(y)^N \Rightarrow \alpha \qquad \exists x\phi(x)^N \Rightarrow \exists x\phi(x)^N}{(\neg\Gamma)^N, \exists x\phi(x)^N \Rightarrow \alpha}}{(\neg\Gamma)^N, (\exists x\phi(x)^N \to \alpha) \to \alpha \Rightarrow \alpha}$$

- Suppose $d : \phi$. Then ϕ is also an axiom of HA^2, so:

$$\dfrac{\phi \to \alpha \Rightarrow \phi \to \alpha \qquad \Rightarrow \phi}{\phi \to \alpha \Rightarrow \alpha}$$

- If d concludes $\Gamma, \exists X\phi(X)$ from $\Gamma, \phi(\lambda\vec{y}.B)$ then:

$$\dfrac{\dfrac{\exists X\phi(X)^N \to \alpha \Rightarrow \exists X\phi(X)^N \to \alpha \qquad \dfrac{\phi(\lambda\vec{y}.B)^N \Rightarrow \phi(\lambda\vec{y}.B)^N}{\phi(\lambda\vec{y}.B)^N \Rightarrow \exists X\phi(X)^N}}{\exists X\phi(X)^N \to \alpha, \phi(\lambda\vec{y}.B)^N \Rightarrow \alpha}}{\exists X\phi(X)^N \to \alpha \Rightarrow \phi(\lambda\vec{y}.B)^N \to \alpha}$$

and

$$\dfrac{\dfrac{(\neg\Gamma)^N, \phi(\lambda\vec{y}.B)^N \to \alpha \Rightarrow \alpha}{\exists X\phi(X)^N \to \alpha \Rightarrow \phi(\lambda\vec{y}.B)^N \to \alpha} \qquad (\neg\Gamma)^N \Rightarrow (\phi(\lambda\vec{y}.B)^N \to \alpha) \to \alpha}{(\neg\Gamma)^N, \exists X\phi(X)^N \to \alpha \Rightarrow \alpha}$$

- If d concludes $\Gamma, \neg\exists X\phi(X)$ from $\Gamma, \neg\phi(Y)$ then:

$$\dfrac{\dfrac{\dfrac{(\neg\Gamma)^N, (\phi(Y)^N \to \alpha) \to \alpha \Rightarrow \alpha}{(\neg\Gamma)^N, \phi(Y)^N \Rightarrow \alpha} \qquad \exists X\phi(X)^N \Rightarrow \exists X\phi(X)^N}{(\neg\Gamma)^N, \exists X\phi(X)^N \Rightarrow \alpha}}{(\neg\Gamma)^N, (\exists X\phi(X)^N \to \alpha) \to \alpha \Rightarrow \alpha}$$

□

3.3 HRO^2

The language of HRO^2 is arithmetic augmented by definitions equating every hereditarily partially recursive function of finite type with a number. More precisely, each partially recursive function is associated with its Gödel number x, and $\{x\}(y)$ is used to denote the (possibly undefined) value of the function associated with x when applied to y; when $\{x\}(y)$ is defined, this is denoted $\{x\}(y) \downarrow$. For technical reasons, 0 should be the constantly 0 function.

The functionals in question are the second order functionals of system F; the set T of types of these functionals is given by:

- The type 0 of the natural numbers is in T
- If $\sigma, \tau \in T$ then $\sigma \to \tau \in T$
- For any n, a variable type $\alpha_n \in T$
- If $\sigma, \tau \in T$ then $\sigma \times \tau \in T$

- If $\sigma[\alpha_n] \in T$ then $\forall \alpha_n.\sigma[\alpha_n] \in T$
- If $\sigma[\alpha_n] \in T$ then $\exists \alpha_n.\sigma[\alpha_n] \in T$

HRO^2 is given by associating to each $\sigma \in T$ a set of numbers V_σ (representing the numbers denoting functions of that type) and to each type variable α a variable V_α ranging over the sets V_σ:

- All numbers are in V_0
- If $\alpha_n \in T$ is a type variable then there is a corresponding set variable V_{α_n}
- $x \in V_{\sigma \to \tau}$ if for any $y \in V_\sigma$, $\{x\}(y) \in V_\tau$
- $x \in V_{\sigma \times \tau}$ if $(x)_0 \in V_\sigma$ and $(x)_1 \in V_\tau$
- $x \in \forall \alpha_n.\sigma[\alpha_n]$ if for any $V \in T$, $x \in V_{\sigma[\alpha_n]}[V/V_{\alpha_n}]$
- $x \in \exists \alpha_n.\sigma[\alpha_n]$ if there is some $V \in T$ such that $x \in V_{\sigma[\alpha_n]}[V/V_{\alpha_n}]$

Full details of the construction are given in [92].

3.4 Realizability

The realizability here is the $HRO^2 - mr$ realizability given in [92] will be used, based on Kreisel's modified realizability presented in [61] and [59]. The modified realizability HRO^2-mr assigns a predicate, Realizes_ϕ from HRO^2, to each formula ϕ of HA^2. A number realizes a formula ϕ when the term it represents executes a computation which demonstrates the truth of the formula. It is then possible to assign a specific term to a deduction d which realizes the conclusion of d.

In order to define the realizability, it is first necessary to define a predicate which is satisfied when a number encodes a functional of the appropriate type to realize a formula. Following the notation in [92], a unary second order variable U^1_X of HRO^2 is uniquely associated to each second order variable X of HA^2. For technical reasons, the set denoted by U^1_X must contain 0, so $\exists U^1_X$ will represent quantification only over those formulae which are satisfied by 0. Then:

1. $\text{Type}_{s=t}(x) \equiv [x = x]$ where x is not free in s or t
2. $\text{Type}_{X\vec{t}}(x) \equiv U^1_X x$
3. $\text{Type}_{\phi\vee\psi}(x) \equiv ((x)_0 = 0 \to \text{Type}_\phi((x)_1)) \wedge ((x)_0 \neq 1 \to \text{Type}_\psi((x)_1))$
4. $\text{Type}_{\phi\to\psi}(x) \equiv \forall y(\text{Type}_\phi(y) \to \{x\}(y) \downarrow \wedge \text{Type}_\psi(\{x\}(y)))$
5. $\text{Type}_{\exists y \phi(y)}(x) \equiv \text{Type}_{\phi((x)_0)}((x)_1)$
6. $\text{Type}_{\exists X^n \phi(X)}(x) \equiv \exists U^1_X\, \text{Type}_{\phi(X)}(x)$

An $n+1$-ary second order variable of HRO^2, X^*, must be uniquely associated to each n-ary second order variable X of HA^2. Then the realizability is given by:

1. $\text{Realizes}_{s=t}(x) \equiv [s = t]$
2. $\text{Realizes}_{X\vec{t}}(x) \equiv X^*(x, \vec{t}) \wedge \text{Type}_{X\vec{t}}(x)$
3. $\begin{aligned}\text{Realizes}_{\phi\vee\psi}(x) \quad &\equiv ((x)_0 = 0 \to \text{Realizes}_\phi((x)_1)) \\ &\wedge((x)_0 \neq 0 \to \text{Realizes}_\psi((x)_1))\end{aligned}$
4. $\begin{aligned}\text{Realizes}_{\phi\to\psi}(x) \quad &\equiv \text{Type}_{\phi\to\psi}(x) \\ &\wedge\forall y(\text{Realizes}_\phi(y) \to \{x\}(y) \downarrow \wedge \text{Realizes}_\psi(\{x\}(y)))\end{aligned}$
5. $\text{Realizes}_{\exists y\phi(y)}(x) \equiv \text{Realizes}_{\phi((x)_0)}((x)_1)$
6. $\text{Realizes}_{\exists X^n\phi(X)}(x) \equiv \exists Y^* \exists U^1_Y\, \text{Realizes}_{\phi(Y)}(x)$

The rules of PA^2 are not sound for this realizability, but their N-translations are; for instance, there is no term corresponding to the axiom $\phi \vee \neg\phi$, but $\phi^N \to \alpha, \phi^N \to \alpha \to \alpha \vdash \alpha$ does correspond to a term. In particular, if $\alpha = \exists x A(x)$ where A is a primitive recursive relation then we say x PA^2-realizes a formula ϕ of PA^2 if $\text{Realizes}_{\phi^N}(x)$. Note that $\text{Realizes}_\alpha(x) \equiv A((x)_0)$, so

$$\text{Type}_{(\neg\phi)^N}(x) \equiv \forall y(\text{Type}_{\phi^N}(y) \to \{x\}(y) \downarrow)$$

$$\text{Realizes}_{(\neg\phi)^N} \equiv \text{Type}_{(\neg\phi)^N}(y) \wedge \forall y(\text{Realizes}_{\phi^N}(y) \rightarrow \{x\}(y)\downarrow \wedge A((\{x\}(y))_0))$$

α may have additional free variables so long as they are renamed to be different from the eigenvalues in any application of the induction or $\forall$ rules. Any free variables other than x will in general also be a free variable in Realizes_ϕ. In this case, $\text{Realizes}_\phi(t)$ means that t is a term (possibly with the same free variables as A) realizing ϕ for every value of those variables.

In general, we use α_ϕ for a first order variable intended to satisfy $\text{Type}_{\phi^N}(\alpha_\phi)$ and when $\Gamma = \{\phi_1, \ldots, \phi_k\}$ is a sequent, we intend $\alpha_\Gamma = (\alpha_{\phi_1}, \ldots, \alpha_{\phi_k})$ to be a sequence of variables such that $\text{Type}_{\phi_i^N}(\alpha_{\phi_i})$.

Lemma 3.3. *1. Write* $[*]$ *for* $[\lambda x\, \text{Type}_{B\vec{y}}(x)/U^1_X]$. *Then*

$$\text{Type}_{\phi(X\vec{t})}[*] = \text{Type}_{\phi(B\vec{t})}$$

2. Write $[\dagger]$ *for* $[\lambda x \lambda \vec{y}\, \text{Realizes}_{B\vec{y}}(x)/X^*]$. *Then*

$$\text{Realizes}_{\phi(X\vec{t})}(x)[*][\dagger] \leftrightarrow \text{Realizes}_{\phi(B\vec{t})}(x)$$

Proof. 1. Proved by a straightforward induction on $\phi(X)$. When $\phi(X) = X\vec{t}$ then

$$\text{Type}_{X\vec{t}}(x)[*] = U^1_X(x)[*] = \text{Type}_{B\vec{t}}(x)$$

The other cases just apply the inductive hypothesis.

2. Proved by induction on $\phi(X)$. When $\phi(X) = X\vec{t}$ then

$$\begin{aligned} \text{Realizes}_{X\vec{t}}(x)[*][\dagger] &= X^*(x, \vec{t})[\dagger] \wedge \text{Type}_{X\vec{t}}(x)[*] \\ &= \text{Realizes}_{B\vec{t}}(x) \wedge \text{Type}_{B\vec{t}}(x) \\ &\leftrightarrow \text{Realizes}_{B\vec{t}}(x) \end{aligned}$$

The other cases just apply the inductive hypothesis.

□

A deduction of $\Gamma \vdash \phi$ in HA^2 can be assigned a term of HRO^2-mr with free variables corresponding to the elements of Γ and which realizes ϕ whenever the free variables realize the corresponding elements of Γ. If Γ or ϕ has free variables, those will in general also be free variables of the term, and for any assignment of values to those variables, the term will realize ϕ. For axioms, the term is 0, and, for example, the deduction $\dfrac{d : \Gamma, \phi \Rightarrow \psi}{\Gamma \Rightarrow \phi \rightarrow \psi}$ becomes $\lambda a.t$ where t is the term correspond to d.

Free variables which appear in the premise but not conclusion of a proof rule can be eliminated in the corresponding terms. Specifically, if d applies $\forall I$, $\forall^2 I$, $\exists I$, $\vee E$, $\rightarrow E$, or $\exists E$ to d_0 (and d_1 and d_2 when appropriate) and x or X^n is a free variable appearing in d_0, d_1, or d_2 but not in d then if t_0 (t_1, t_2) are the corresponding terms, replace all occurrences of x with 0 and all occurrences of X^n with $\lambda\vec{y}.(\forall X^0)X$ before constructing t. For instance suppose $d_0 : \Gamma \Rightarrow \psi \rightarrow \phi(x, X)$ and $d_1 : \Sigma \Rightarrow \psi$ with x and X^n variable not appearing in ψ, Γ, or Σ. Then the corresponding term is $t_0[0/x][\lambda\vec{y}.(\forall X^0)X/X](t_1)$.

Theorem 3.4. *If d is a deduction of Γ in PA^2 then there is a term F_d with free variables among $\alpha_{\neg\Gamma} \cup \text{FV}(\Gamma) \cup \text{FV}(\alpha)$ such that if $\text{Type}_{(\neg\phi)^N}(\alpha_{\neg\phi})$ for each $\phi \in \Gamma$ then HA^2 proves $(\lambda\alpha_{\neg\Gamma}.F_d)(\alpha_{\neg\Gamma})\downarrow$ and if $\text{Realizes}_{(\neg\phi)^N}(\alpha_{\neg\phi})$ holds for each $\phi \in \Gamma$ then HA^2 proves $A(((\lambda\alpha_{\neg\Gamma}.F_d)(\alpha_{\neg\Gamma}))_0)$.*

Proof. Since d is a deduction of Γ in PA^2, there is a deduction D of $(\neg\Gamma)^N \vdash \exists x A(x)$ in HA^2. The theorem could be proved by simply appealing to the realization given in [92]. However this can also be proved directly by defining the term inductively on the last step of d; the appropriate can be easily found by taking the HA^2 deduction corresponding to an inference in PA^2 and applying the Curry-Howard isomorphism.

It will be necessary to remove extraneous free variables during this process. If d applies the cut rule or the first or second order $\exists$ rules, there may be free first or second order variables which appear in the premises but not the conclusion. If $d : \phi$ is an application of one of these three rules to $d_1 : \phi_1$ (and $d_2 : \phi_2$ in the case of cut) and x or X is a free variable in ϕ_1 (and ϕ_2 in the case of cut) which does not appear in ϕ then the inference

$$\frac{d_1[0/x][\lambda\vec{y}.(\forall X^0)X] \qquad \big(d_2[0/x][\lambda\vec{y}.(\forall X^0)X]\big)}{d}$$

is also a valid inference. The corresponding terms, $t_1[0/x][\lambda\vec{y}.(\forall X^0)X]$ and $t_2[0/x][\lambda\vec{y}.(\forall X^0)X]$ should be used in the inductive construction of t.

- d is any of the quantifier free axioms. Then $\Gamma = \{\phi_1, \ldots, \phi_k\}$ and at least one ϕ_i must be true, therefore it is never possible for $\alpha_{\neg\Gamma}$ to realize $\neg\Gamma$, so

$$F_d \equiv 0$$

- d is an axiom of the form $\Gamma, A, \neg A$. Then:

$$F_d \equiv \{\alpha_{\neg\neg A}\}(\alpha_{\neg A})$$

- d concludes $\Gamma, \phi \vee \psi$ from $d' : \Gamma, \phi$ (the case for $d' : \Gamma, \psi$ is similar). Then:

$$F_d \equiv (\lambda\alpha_{\neg\phi}.F_{d'})(\ulcorner\lambda\alpha_\phi.\{\alpha_{\neg(\phi\vee\psi)}\}(\langle 0, \alpha_\phi\rangle^\urcorner)$$

- d concludes $\Gamma, \neg(\phi \vee \psi)$ from $d_0 : \Gamma, \neg\phi$ and $d_1 : \Gamma, \neg\psi$. Then primitive recursion can be used to define by cases:

$$F' \equiv \ulcorner\begin{cases} (\lambda\alpha_{\neg\neg\phi}.F_0)(\ulcorner\lambda\alpha_{\neg\phi}.\{\alpha_{\neg\phi}\}((\alpha_{\phi\vee\psi})_1)^\urcorner) & \text{if } (\alpha_{\phi\vee\psi})_0 = 0 \\ (\lambda\alpha_{\neg\neg\psi}.F_1)(\ulcorner\lambda\alpha_{\neg\psi}.\{\alpha_{\neg\psi}\}((\alpha_{\phi\vee\psi})_1)^\urcorner) & \text{if } (\alpha_{\phi\vee\psi})_0 \neq 0 \end{cases}\urcorner$$

and define

$$F_d \equiv \{\alpha_{\neg\neg(\phi\vee\psi)}\}(\ulcorner\lambda\alpha_{\phi\vee\psi}.F'^\urcorner)$$

- d concludes $\Gamma, \exists x\phi(x)$ from $d' : \Gamma, \phi(t)$. If t has any free variables that do not occur in the conclusion the should be replaced with 0 in $F_{d'}$. Then:

$$F_d = (\lambda\alpha_{\neg\phi(t)}.F_{d'})(\ulcorner\lambda\alpha_{\phi(t)}.\{\alpha_{\neg\exists x\phi(x)}\}(\langle t, \alpha_{\phi(t)}\rangle)^\urcorner)$$

- d concludes $\Gamma, \neg\exists x\phi(x)$ from $d' : \Gamma, \neg\phi(y)$. Then $F_{d'}$ is a term which may contain y free and y does not occur free in Γ. So:

$$\begin{aligned} F_d \equiv \quad & \{\alpha_{\neg\neg\exists x\phi(x)}\}(\ulcorner\lambda\alpha_{\exists x\phi(x)}.(\lambda y\lambda\alpha_{\neg\neg\phi(y)}.F_{d'}) \\ & ((\alpha_{\exists x\phi(x)})_0)(\lambda\alpha_{\neg\phi(y)}.\{\alpha_{\neg\phi(y)}\}((\alpha_{\exists x\phi(x)})_1))^\urcorner) \end{aligned}$$

- d derives Γ from $d_0 : \Gamma, \neg\phi$ and $d_1 : \Gamma, \phi$. Replace any free variables which appear in d_0 and d_1 but not in d with 0 (for first order variables) and $(\forall X^0)X$ (for second order variables). Then:

$$F_d \equiv (\lambda\alpha_{\neg\neg\phi}.F_0)(\ulcorner\lambda\alpha_{\neg\phi}.F_1{}^\urcorner)$$

- d is a deduction of $\Gamma, \neg\exists x\phi(x)$ from $d_0 : \Gamma, \neg\phi(0)$ and $d_1 : \Gamma, \phi(y), \neg\phi(Sy)$. Then construct a function h by primitive recursion:

$$h(0) \equiv \ulcorner\lambda\alpha_{\phi(0)}.(\lambda\alpha_{\neg\neg\phi(0)}.F_{d_0})(\lambda\alpha_{\neg\phi(0)}.\{\alpha_{\neg\phi(0)}\}(\alpha_{\phi(0)}))^\urcorner$$

$$\begin{aligned} h(Sy) \equiv \quad & \ulcorner(\lambda\alpha_{\neg\phi(y)}.\lambda\alpha_{\phi(Sy)}.(\lambda\alpha_{\neg\neg\phi(Sy)}.F_{d_1}) \\ & (\lambda\alpha_{\neg\phi(Sy)}.\{\alpha_{\neg\phi(Sy)}\}(\alpha_{\phi(Sy)})))(h(y))^\urcorner \end{aligned}$$

Note that $\mathrm{Realizes}_{(\neg\phi(n))^N}(h(n))$ for every n.
Then:

$$F_d \equiv \{\alpha_{\neg\neg\exists x\phi(x)}\}(\lambda\alpha_{\exists x\phi(x)}.\{h((\alpha_{\exists x\phi(x)})_0)\}((\alpha_{\exists x\phi(x)})_1))$$

- d is a deduction of $\Gamma, \exists X\phi(X)$ from $d' : \Gamma, \phi(\lambda\vec{y}.B)$

$$F_d \equiv (\lambda\alpha_{\neg\phi(\lambda\vec{y}.B)}.F_{d'})(\ulcorner\lambda\alpha_{\phi(\lambda\vec{y}.B)}.\{\alpha_{\exists X\phi(X)}\}(\alpha_{\phi(\lambda\vec{y}.B)})^\urcorner)$$

Free variables appearing in d' but not d should be replaced.

- d is a deduction of $\Gamma, \neg\exists X\phi(X)$ from $d' : \Gamma, \neg\phi(Y)$ then:

$$
\begin{aligned}
F_d \equiv \quad & \{\alpha_{\neg\neg\exists X\phi(X)}\}(\ulcorner\lambda\alpha_{\exists X\phi(X)}.[(\lambda\alpha_{\phi(Y)}.(\lambda\alpha_{\neg\neg\phi(Y)}.F_{d'}) \\
& (\ulcorner\lambda\alpha_{\neg\phi(Y)}.\{\alpha_{\neg\phi(Y)}\}(\alpha_{\phi(Y)})\urcorner))(\alpha_{\exists X\phi(X)})]\urcorner)
\end{aligned}
$$

□

Theorem 3.5. *If d is a deduction of $\exists x A(x)$ where $A(x)$ is primitive recursive then it is possible to construct a term t of HRO^2 with the same free variables as $\exists x A(x)$ such that $A(t)$ holds for every value of those variables.*

Proof. Cut d with a hypothesis $h : \neg\exists x A(x)$; this gives a proof d' of the empty sequent. Let $F_h = \{\alpha_{\neg\neg\exists x A(x)}\}(\ulcorner\lambda\alpha_{\exists x A(x)}.\alpha_{\exists x A(x)}\urcorner)$. Then, applying the previous theorem, $t = F_{d'}$ is a term with no free variables, and therefore $A((t)_0)$. □

If A has free variables other than x, they will also, in general, be free variables in the corresponding term, so as an easy corollary we have:

Theorem 3.6. *If f is some function and A is primitive recursive relation symbol representing the graph of f and $PA^2 \vdash \forall y \exists x A(y, x)$ then there is a term t in HRO^2 with free variable y such that $f = \lambda y.t$.*

Proof. Since PA^2 proves $\forall y \exists x A(y, x)$, there is also a PA^2 deduction d of $\exists x A(y, x)$. Then the term $(F_d)_0$ given by the previous theorem suffices. □

Chapter 4

A Dialectica Interpretation for Inductive Definitions[1]

Let X be a set, and let Γ be a monotone operator from the power set of X to itself, so that $A \subseteq B$ implies $\Gamma(A) \subseteq \Gamma(B)$. Then the set

$$I = \bigcap \{A \mid \Gamma(A) \subseteq A\}$$

is a least fixed point of Γ; that is, $\Gamma(I) = I$, and I is a subset of any other set with this property. I can also be characterized as the limit of a sequence indexed by a sufficiently long segment of the ordinals, defined by $I_0 = \emptyset$, $I_{\alpha+1} = \Gamma(I_\alpha)$, and $I_\lambda = \bigcup_{\gamma<\lambda} I_\gamma$ for limit ordinals γ. Such inductive definitions are common in mathematics; they can be used, for example, to define substructures generated by sets of elements, the collection of Borel subsets of the real line, or the set of well-founded trees on the natural numbers.

From the point of view of proof theory and descriptive set theory, one is often interested in structures that are countably based, that is, can be coded so that X is a countable set. In that case, the sequence I_α stabilizes before the least uncountable ordinal. In many interesting situations, the operator Γ is given by a positive arithmetic formula $\varphi(x, P)$, in the sense that $\Gamma(A) = \{x \mid \varphi(x, A)\}$ and φ is an arithmetic formula in which the predicate P occurs only positively. (The positivity requirement can be expressed by saying that no occurrence of P is negated when φ is written in negation-normal form.)

The considerations above show that the least fixed point of a positive arithmetic inductive definition can be defined by a Π^1_1 formula. An analysis due to Stephen Kleene [51, 52] shows that, conversely, a positive arithmetic inductive definition can be used to define a complete Π^1_1 set. In the 1960's, Georg Kreisel presented axiomatic theories of such inductive definitions [23, 62]. In particular, the theory *ID₁* consists of first-order arithmetic augmented by additional predicates intended to denote least fixed-points of positive arithmetic operators. *ID₁* is known to have the same strength as the subsystem $\mathit{\Pi^1_1\text{-}CA}^-$ of second order arithmetic, which has a comprehension axiom asserting the existence of sets of numbers defined by Π^1_1 formulas without set parameters. It also has the same strength as Kripke Platek admissible set theory, $KP\omega$, with an axiom asserting the existence of an infinite set. (See [23, 46] for details.)

A Π_2 sentence is one of the form $\forall \bar{x}\ \exists \bar{y}\ R(\bar{x}, \bar{y})$, where $\bar{x}$ and $\bar{y}$ are tuples of variables ranging over the natural numbers, and R is a primitive recursive relation. Here we are concerned with the project of characterizing the Π_2 consequences of the theories *ID₁* in constructive or computational terms. This can be done in a number of ways. For example, every Π_2 theorem of *ID₁* is witnessed by a function that a can be defined in a language of higher-type functionals allowing primitive recursion on the natural numbers as well as a schema of recursion along well-founded trees, as described in Section 4.1 below. We are particularly interested in obtaining a translation from *ID₁* to a constructive theory of such functions that makes it possible to "read off" a description of the witnessing function from the proof of a Π_2 sentence in *ID₁*.

[1]This chapter is joint work and jointly written with Jeremy Avigad

There are currently two ways of obtaining this information. The first involves using ordinal analysis to reduce ID_1 to a constructive analogue [22, 72, 73], such as the theory $ID_1^{i,sp}$ discussed below, and then using either a realizability argument or a Dialectica interpretation of the latter [21, 45]. One can, alternatively, use a forcing interpretation due to Buchholz [2, 21] to reduce ID_1 to $ID_1^{i,sp}$.

Here we present a new method of carrying out this first step, based on a functional interpretation along the lines of Gödel's "Dialectica" interpretation of first-order arithmetic. Such functional interpretations have proved remarkably effective in "unwinding" computational and otherwise explicit information from classical arguments (see, for example, [53, 54, 56]). Howard [45] has provided a functional interpretation for a restricted version of the constructive theory $ID_1^{i,sp}$, but the problem of obtaining such an interpretation for classical theories of inductive definitions is more difficult, and was posed as an outstanding problem in [10, Section 9.8]. Feferman [30] used a Dialectica interpretation to obtain ordinal bounds on the strength of ID_1 (the details are sketched in [10, Section 9]), and Zucker [96] used a similar interpretation to bound the ordinal strength of ID_2. But these interpretations do not yield Π_2 reductions to constructive theories, and hence do not provide computational information; nor do the methods seem to extend extend to the theories beyond ID_2. Our interpretation bears similarities to those of Burr [24] and Ferreira and Oliva [31], but is not subsumed by either; some of the differences between the various approaches are indicated in Section 4.3.

The outline of the chapter is as follows. In Section 4.1, we describe the relevant theories and provide an overview our our results. Our interpretation of ID_1 is presented in three steps. In Section 4.2, we embed ID_1 in an intermediate theory, $OR_1 + (I)$, which makes the transfinite construction of the fixed-point explicit. In Section 4.3, we present a functional interpretation that reduces $OR_1 + (I)$ to a second intermediate theory, $Q_0 T_\Omega + (I)$. Finally, the latter theory is interpreted in a constructive theory, QT^i_Ω, using a cut elimination argument in Section 4.4. In Section 4.5, we show that our interpretation extends straightforwardly to cover theories of iterated inductive definitions as well.

We are grateful to Solomon Feferman, Philipp Gerhardy, and Wilfried Sieg for feedback on an earlier draft.

4.1 Background

In this chapter, we interpret classical theories of inductively defined sets in constructive theories of transfinite recursion on well-founded trees. In this section, we describe the relevant theories, and provide an overview of our results.

Take classical first-order Peano arithmetic, PA, to be formulated in a language with symbols for each primitive recursive function and relation. The axioms of PA consist of basic axioms defining these functions and relations, and the schema of induction,

$$\varphi(0) \wedge \forall x\ (\varphi(x) \to \varphi(x+1)) \to \forall x\ \varphi(x),$$

where φ is any formula in the language, possibly with free variables other than x. ID_1 is an extension of PA with additional predicates I_ψ intended to denote the least fixed point of the positive arithmetic operator given by ψ. Specifically, let $\psi(x, P)$ be an arithmetic formula with at most the free variable x, in which the predicate symbol P occurs only positively. We adopt the practice of writing $x \in I_\psi$ instead of $I_\psi(x)$. ID_1 then includes the following axioms:

- $\forall x\ (\psi(x, I_\psi) \to x \in I_\psi)$
- $\forall x\ (\psi(x, \theta/P) \to \theta(x)) \to \forall x \in I_\psi\ \theta(x)$, for each formula $\theta(x)$.

Here, the notation $\psi(\theta/P)$ denotes the result of replacing each atomic formula $P(t)$ with $\theta(t)$, renaming bound variables to prevent collisions. The first axiom asserts that I_ψ is closed with respect to Γ_ψ, while the second axiom schema expresses that I_ψ is the smallest such set, among those sets that can be defined in the language. Below we will use the fact that this schema, as well as the schema of induction, can be expressed as rules. For example, I_ψ-leastness is equivalent to the rule "from $\forall x\ (\psi(x, \theta'/P) \to \theta'(x))$ conclude $\forall x \in I_\psi\ \theta'(x)$." To see this, note that the rule is easily justified using the corresponding axiom; conversely, one obtains the axiom for $\theta(x)$ by taking $\theta'(x)$ to be the formula $(\forall z\ (\psi(z, \theta/P) \to \theta(z))) \to \theta(x)$ in the rule.

One can also design theories of inductive definitions based on intuitionistic logic. In order for these theories to be given a reasonable constructive interpretation, however, one needs to be more careful in specifying the positivity requirement on ψ. One option is to insist that P does not occur in the antecedent of any implication, where $\neg\eta$ is taken to abbreviate $\eta \to \bot$. Such a definition is said to be *strictly positive*, and we denote the corresponding axiomatic theory $ID_1^{i,sp}$. An even more restrictive requirement is to insist that $\psi(x)$ is of the form $\forall y \prec x\ P(y)$, where $\prec$ is a primitive recursive relation. These are called *accessibility* inductive definitions, and serve to pick out the well-founded part of the relation. In the case where $\prec$ is the "child-of" relation on a tree, the inductive definition picks out the well-founded part of that tree. We will denote the corresponding theory $ID_1^{i,acc}$.

The following conservation theorem can be obtained via an ordinal analysis [23] or the methods of Buchholz [21]:

Theorem 4.1. *Every Π_2 sentence provable in ID_1 is provable in $ID_1^{i,acc}$.*

The methods we introduce here provide another route to this result.

Using a primitive recursive coding of pairs and writing $x \in I_y$ for $(x, y) \in I$ allows us to code any finite or infinite sequence of sets as a single set. One can show that in any of the theories just described, any number of inductively defined sets can coded into a single one, and so, for expository convenience, we will assume that each theory uses only a single inductively defined set.

We now turn to theories of transfinite induction and recursion on well-founded trees. The starting point is a quantifier-free theory, T_Ω, of computable functionals over the natural numbers and the set of well-founded trees on the natural numbers. In particular, T_Ω extends Gödel's theory T of computable functionals over the natural numbers. We begin by reviewing the theory T. The set of *finite types* is defined inductively, as follows:

- N is a finite type; and
- assuming σ and τ are finite types, so are $\sigma \times \tau$ and $\sigma \to \tau$.

In the "full" set-theoretic interpretation, N denotes the set of natural numbers, $\sigma \times \tau$ denotes the set of ordered pairs consisting of an element of σ and an element of τ, and $\sigma \to \tau$ denotes the set of functions from σ to τ. But we can also view the finite types as nothing more than datatype specifications of computational objects. The set of *primitive recursive functionals of finite type* is a set of computable functionals obtained from the use of explicit definition, application, pairing, and projections, and a scheme allowing the definition of a new functional F by primitive recursion:

$$F(0) = a$$
$$F(x+1) = G(x, F(x))$$

Here, the range of F may be any finite type. The theory T includes defining equations for all the primitive recursive functionals, and a rule providing induction for quantifier-free formulas φ:

$$\frac{\varphi(0) \qquad \varphi(x) \to \varphi(S(x))}{\varphi(t)}$$

Gödel's *Dialectica* interpretation shows:

Theorem 4.2. *If PA proves a Π_2 theorem $\forall \bar{x}\ \exists \bar{y}\ R(\bar{x}, \bar{y})$, there is a sequence of function symbols $\bar{f}$ such that T proves $R(\bar{x}, \bar{f}(\bar{x}))$. In particular, every Π_2 theorem of PA is witnessed by sequence of primitive recursive functionals of type $N^k \to N$.*

See [10, 41, 91] for details. If (st) is used to denote the result of applying s to t, we adopt the usual conventions of writing, for example, $stuv$ for $(((st)u)v)$. To improve readability, however, we will also sometimes adopt conventional function notation, and write $s(t, u, v)$ for the same term.

In order to capture the Π_2 theorems of ID_1, we use an extension of T that is essentially due to Howard [45], and described in [10, Section 9.1]. Extend the finite types by adding a new base type, Ω, which is intended to denote the set of well-founded (full) trees on N. We add a constant, e, which denote the tree with just one node, and two new operations: sup, of type $(N \to \Omega) \to \Omega$, which forms a new tree from a sequence of subtrees, and $\sup^{-1}$, of type $\Omega \to (N \to \Omega)$, which returns the immediate subtrees of a nontrivial tree. We extend the schema of primitive recursion on N in T to the larger system, and add a principle of primitive recursion on Ω:

$$F(e) = a$$
$$F(\sup\ h) = G(\lambda n\ F(h(n))),$$

where the range of F can be any of the new types. We call the resulting theory T_Ω, and the resulting set of functionals the *primitive recursive tree functionals*. Below we will adopt the notation $\alpha[n]$ instead of $\sup^{-1}(\alpha, n)$ to denote the nth subtree of α. In that case definition by transfinite recursion can be expressed as follows:[2]

$$F(\alpha) = \begin{cases} a & \text{if } \alpha = e \\ G(\lambda n\ F(\alpha[n])) & \text{otherwise.} \end{cases}$$

A trick due to Kreisel (see [44, 45]) allows us to derive a quantifier-free rule of transfinite induction on Ω in T_Ω, using induction on N and transfinite recursion.

Proposition 4.3. *The following is a derived rule of T_Ω:*

$$\frac{\varphi(e, x) \qquad \alpha \neq e \wedge \varphi(\alpha[g(\alpha, x)], h(\alpha, x)) \to \varphi(\alpha, x)}{\varphi(s, t)}$$

for quantifier-free formulas φ.

For the sake of completeness, we sketch a proof in the Appendix.

We define QT_Ω to be the extension of T_Ω which allows quantifiers over all the types of the latter theory, replacing transfinite induction rule with the axiom schema,

$$\varphi(e) \wedge \forall \alpha\ (\alpha \neq e \wedge \forall n\ \varphi(\alpha[n]) \to \varphi(\alpha)) \to \forall \alpha\ \varphi(\alpha)$$

where φ is any formula in the expanded language; and an "ω bounding" axiom,

$$\forall x\ \exists \alpha\ \psi(x, \alpha) \to \exists \beta\ \forall x\ \exists i\ \psi(x, \beta[i]),$$

where x and y can have any type and ψ is any formula. Let QT^i_Ω denote the version of this theory based on intuitionistic logic. The following theorem shows that all of the intuitionistic theories described in this section are "morally equivalent," and reducible to T_Ω.

Theorem 4.4. *The following theories all prove the same Π_2 sentences:*

1. $ID_1^{i,sp}$
2. $ID_1^{i,acc}$
3. QT^i_Ω
4. T_Ω

Proof. Buchholz [21] presents a realizability interpretation of $ID_1^{i,sp}$ in the theory $ID_1^{i,acc}$. Howard [45] presents an embedding of $ID_1^{i,acc}$ in QT^i_Ω, and a functional interpretation of QT^i_Ω in T_Ω. Interpreting T_Ω in $ID_1^{i,sp}$ is straightforward, using the set O of Church-Kleene ordinal notations to interpret the type Ω, and interpreting the constants of T_Ω as hereditarily recursive operations over O (see [10, Sections 4.1, 9.5, and 9.6]). $\square$

We can now describe our main results. In Sections 4.2 to 4.4, we present the interpretation outlined in the introduction, which yields:

Theorem 4.5. *Every Π_2 sentence provable in ID_1 is provable in QT^i_Ω.*

[2]We are glossing over issues involving the treatment of equality in our descriptions of both T and T_Ω. All of the ways of dealing with equality in T described in [10, Section 2.5] carry over to T_Ω, and our interpretations work with even the most minimal version of equality axioms associated with the theory denoted T_0 there. In particular, our interpretations to not rely on extensionality, or the assumption $\forall n\ (\alpha[n] = \beta[n]) \to \alpha = \beta$. We do make use of the decidability of the atomic formula $\alpha =_\Omega e$, but this can be interpreted as the formula $f(\alpha) =_N 0$, where f is the function from Ω to N defined recursively by $f(e) = 0, f(\sup\ g) = 1$.

Our theory T_Ω is essentially the theory V of Howard [45]. Our theory QT^i_Ω is essentially a finite-type version of the theory U of [45], and contained in the theory V^* described there. One minor difference is that Howard takes the nodes of his trees to be labeled, with end-nodes labeled by a positive natural number, and internal nodes labeled 0.

In fact, if ID_1 proves a Π_2 theorem $\forall \bar{x}\ \exists \bar{y}\ R(\bar{x}, \bar{y})$, our proof yields a sequence of function symbols $\bar{f}$ such that QT^i_Ω proves $R(\bar{x}, \bar{f}(\bar{x}))$. By Theorem 4.4, this last assertion can even be proved in T_Ω. Thus we have:

Theorem 4.6. *If ID_1 proves a Π_2 theorem $\forall \bar{x}\ \exists \bar{y}\ R(\bar{x}, \bar{y})$, there is a sequence of function symbols $\bar{f}$ such that T_Ω proves $R(\bar{x}, \bar{f}(\bar{x}))$. In particular, every Π_2 theorem of ID_1 is witnessed by sequence of primitive recursive tree functionals of type $N^k \to N$.*

The reduction described by Sections 4.2 to 4.4 is thus analogous to the reduction of ID_1 given by Buchholz [21], but relies on a functional interpretation instead of forcing.

4.2 Embedding ID_1 in $OR_1 + (I)$

In this section, we introduce theories OR_1 and $OR_1 + (I)$, and show that ID_1 is easily interpreted in the latter. The theory $OR_1 + (I)$ is closely related to Feferman's theory OR^ω_1, as described in [30] and [10, Section 9], and the details of the embedding are essentially the ones described there.

The language of OR_1 is two-sorted, with variables $\alpha, \beta, \gamma, \ldots$ ranging over type Ω, and variables $i, j, k, n, x, \ldots$ ranging over N. We include symbols for the primitive recursive functions on N, and a function symbol $\sup^{-1}(\alpha, n)$ which returns an element of type Ω. As above, we write $\alpha[n]$ for $\sup^{-1}(\alpha, n)$. Recall that $\alpha[n]$ is intended to denote the nth subtree of α, or e if $\alpha = e$. The axioms of OR_1 are as follows:

1. defining axioms for the primitive recursive functions
2. induction on N
3. transfinite induction on Ω
4. a schema of ω bounding:
$$\forall x\ \exists \alpha\ \varphi(x, \alpha) \to \exists \beta\ \forall x\ \exists i\ \varphi(x, \beta[i]),$$
where φ has no quantifiers over type Ω.

We will often think of an element $\alpha \neq e$ of Ω as denoting a countable set $\{\alpha[i] \mid i \in \mathbb{N}\}$ of elements of Ω. We write $\alpha \sqsubseteq \beta$ for $\forall i\ \exists j\ (\alpha[i] = \beta[j])$, to express inclusion between the corresponding sets. Now let $t(i)$ be any term of type Ω, and let i_0 and i_1 denote the projections of i under a primitive recursive coding of pairs. Since trivially we have $\forall i\ \exists \beta\ (\beta = t(i_0)[i_1])$, ω bounding implies that for some γ we have $\forall i\ \exists j\ (\gamma[j] = t(i_0)[i_1])$. This implies $\forall k\ \forall l\ \exists j\ (\gamma[j] = t(k)[l])$, in other words, $\forall k\ (t(k) \sqsubseteq \gamma)$. In other words, if we think of $t(k)$ as a sequence of countable sets, ω bounding guarantees the existence of a countable set γ that includes their union.

Now fix any instance of ID_1 with inductively defined predicate I given by the positive arithmetic formula $\psi(x, P)$. To define the corresponding instance of $OR_1 + (I)$, we extend the language of OR_1 with a new binary predicate $I(\alpha, x)$, where α ranges over Ω and x ranges over N. We will write $x \in I_\alpha$ instead of $I(\alpha, x)$, and write $x \in I_{\prec\alpha}$ for $\exists i\ (x \in I_{\alpha[i]})$. The schemas of induction, transfinite induction, and ω bounding are extended to the new language. We also add the following defining axioms for the predicate I:

- $\forall x\ (x \notin I_e)$
- $\forall \alpha\ (\alpha \neq e \to \forall x\ (x \in I_\alpha \leftrightarrow \exists i\ \psi(x, I_{\alpha[i]})))$

For any formula φ of ID_1, let $\hat{\varphi}$ be the formula obtained by interpreting $t \in I$ as $\exists \alpha\ (t \in I_\alpha)$.

Theorem 4.7. *If ID_1 proves φ, then $OR_1 + (I)$ proves $\hat{\varphi}$.*

Before proving this, we need a lemma. Note that if $\eta(x, P)$ is any arithmetic formula involving a new predicate symbol P and $\theta(y)$ is any formula, applying the $\hat{\cdot}$-translation to $\eta(x, \theta/P)$ changes only the instances of θ. In particular, $\widehat{\eta(x, I)}$ is $\eta(x, \exists \alpha\ (y \in I_\alpha)/P)$.

Lemma 4.8. *Let $\eta(x, P)$ be a positive arithmetic formula. Then $OR_1 + (I)$ proves the following:*

1. *$\alpha \sqsubseteq \beta$ and $\eta(x, I_{\prec\alpha})$ implies $\eta(x, I_{\prec\beta})$*
2. *$\widehat{\eta(x, I)}$ implies $\exists \beta\ \eta(x, I_{\prec\beta})$.*

Proof. Both claims are proved by a straightforward induction on positive arithmetic formulas (expressed in negation normal form). To prove the second claim, in the base case, suppose we have $t \in I_\alpha$. Applying ω bounding with antecedent $\forall x\ \exists \gamma\ (\gamma = \alpha)$, we obtain a β such that $\beta[i] = \alpha$ for some i. Then we have $t \in I_{\prec \beta}$, as required.

In the case where the outermost connective is a universal quantifier, suppose $\varphi(x, y, I)$ implies $\exists \beta\ \varphi(x, y, I_{\prec\beta})$. Using ω bounding, $\forall y\ \varphi(x, y, I)$ then implies

$$\exists \gamma\ \forall y\ \exists i\ \varphi(x, y, I_{\prec \gamma[i]}).$$

Using ω bounding again, as described above, we obtain an α such that for every i, $\gamma[i] \subseteq \alpha$. Using (1), we have

$$\exists \alpha\ \forall y\ \varphi(x, y, I_{\prec\alpha}),$$

as required. The remaining cases are easy. $\square$

Proof of Theorem 4.7. The defining axioms for the primitive recursive functions and induction axioms of ID_1 are again axioms of $OR_1 + (I)$ under the translation, so we only have to deal with the defining axioms for I.

The translation of the closure axiom, $\forall x\ (\psi(x, I) \to x \in I)$, is immediate from Lemma 4.8 and the defining axiom for I_α.

This leaves only transfinite induction, which can be expressed as a rule, "from $\forall x\ (\psi(x, \theta/P) \to \theta(x))$, conclude $\forall x \in I\ \theta(x)$." To verify the translation in $OR_1 + (I)$, suppose $\forall x\ (\psi(x, \hat{\theta}/P) \to \hat{\theta}(x))$. It suffices to show that for every α, we have $\forall x \in I_\alpha\ \hat{\theta}(x)$. We use transfinite induction on α. In the base case, when $\alpha = e$, this is immediate from the defining axiom for I_e. In the inductive step, suppose we have $\forall i\ \forall x \in I_{\alpha[i]}\ \hat{\theta}(x)$. This is equivalent to $\forall x \in I_{\prec\alpha}\ \hat{\theta}(x)$. Using the positivity of P, we have $\forall x\ (\psi(x, I_{\prec\alpha}) \to \psi(x, \hat{\theta}/P))$. Using the definition of I_α, we then have $\forall x \in I_\alpha\ \hat{\theta}(x)$, as required. $\square$

4.3 A functional interpretation of $OR_1 + (I)$

Our next step is to interpret the theory $OR_1 + (I)$ in a second intermediate theory, $Q_0 T_\Omega + (I)$. First, we describe a fragment $Q_0 T_\Omega$ of QT_Ω, obtained by restricting the language of QT_Ω to allow quantification over the natural numbers only, though we continue to allow free variables and constants of all types. We also restrict the language so that the only atomic formulas are equalities $s = t$ between terms of type N. The axioms of $Q_0 T_\Omega$ are as follows:

1. any equality between terms of type N that can be derived in T_Ω
2. the schema of induction on N.
3. the schema of transfinite induction, given as a rule:

$$\frac{\theta(e) \qquad \alpha \neq e \wedge \forall n\ \theta(\alpha[n]) \to \theta(\alpha)}{\theta(t)}$$

   for any formula θ and term t of type Ω.

It is not hard to check that substitution is a derived rule in $Q_0 T_\Omega$, which is to say, if the theory proves $\varphi(x)$ where x is a variable of any type, it proves $\varphi(s)$ for any term s of that type. Similarly, if T_Ω proves $s = t$ for any terms s and t of that type, then $Q_0 T_\Omega$ proves $\varphi(s) \leftrightarrow \varphi(t)$.

The following lemma shows that in $Q_0 T_\Omega$ we can use instances of induction in which higher-type parameters are allowed to vary. For example, the first rule states that in order to prove $\theta(\alpha, x)$ for arbitrary α and x, it suffices to prove $\theta(e, x)$ for an arbitrary x, and then, in the induction step, prove that $\theta(\alpha, x)$ follows from $\theta(\alpha[n], a)$, as n ranges over the natural numbers and a ranges over a countable sequence of parameters depending on n and x.

Proposition 4.9. *The following are derived rules of $Q_0 T_\Omega$:*

$$\frac{\theta(e, x) \qquad \alpha \neq e \wedge \forall i\ \forall j\ \theta(\alpha[i], f(\alpha, x, i, j)) \to \theta(\alpha, x)}{\theta(\alpha, x)}$$

and

$$\frac{\psi(0,x) \qquad \forall j\ \psi(n, f(x,n,j)) \rightarrow \psi(n+1,x)}{\psi(n,x)}$$

A proof of Proposition 4.9 can be found in the Appendix.

The theories $Q_0 T_\Omega + (I)$ are now defined in analogy with $OR_1 + (I)$: we extend the language with a new binary predicate $I(\alpha, x)$, which is allowed to occur in the induction axioms and the transfinite induction rules, and add the same defining axioms. Proposition 4.9 extends to this new theory. In this section, we will use a functional interpretation to interpret $OR_1 + (I)$ in $Q_0 T_\Omega + (I)$.

Recall that we have defined $\alpha \sqsubseteq \beta$ by

$$\alpha \sqsubseteq \beta \equiv \forall i\ \exists j\ (\alpha[i] = \beta[j]),$$

thinking of elements α of Ω as coding countable sets $\{\alpha[i] \mid i \in \mathbb{N}\}$. Let $t(i)$ be any term of type Ω, where i is of type N. Then we can define the union of the sets $t(0), t(1), t(2), \ldots$ by

$$\sqcup_i t(i) = \sup_j t(j_0)[j_1],$$

where j_0 and j_1 denote the projections of j under a primitive recursive coding of pairs. In other words, $\sqcup_i t(i)$ represents the set $\{t(i)[k] \mid i \in \mathbb{N}, k \in \mathbb{N}\}$. In particular, we have that for every i, $t(i) \sqsubseteq \sqcup_i t(i)$, since for every k we have $t(i)[k] = (\sqcup_i t(i))[(i,k)]$.

We now extend these notions to higher types. Define the set of *pure* Ω*-types* to be the smallest set of types containing Ω and closed under the operation taking σ and τ to $\sigma \rightarrow \tau$. Note that every pure Ω-type τ has the form $\sigma_1 \rightarrow \sigma_2 \rightarrow \ldots \sigma_k \rightarrow \Omega$. We can therefore lift the notions above to a, b, and t of arbitrary pure type, as follows:

$$\begin{aligned} a[i] &= \lambda x\ ((ax)[i]) \\ a \sqsubseteq b &\equiv \forall i\ \exists j\ \forall x\ ((ax)[i] = (bx)[j]) \\ \sqcup_i t(i) &= \lambda x\ (\sup_j (tx)(j_0)[j_1]), \end{aligned}$$

where in each case x is a tuple of variables chosen so that the resulting term has type Ω. Thus, if a is of any pure type, we can think of a as representing the countable set $\{a[i] \mid i \in \mathbb{N}\}$, in which case $\sqsubseteq$ and $\sqcup$ have the expected behavior.

Note that the relation $a \sqsubseteq b$ can be expressed in the language of QT_Ω, but when τ is not Ω, it cannot be expressed in the language of $Q_0 T_\Omega$, which does not allow quantification over pure types. We will, however, be interested in situations where $Q_0 T_\Omega$ can prove $a \sqsubseteq_\tau b$ in the sense that there is an explicit function $j(i)$ such that it can prove $\forall i\ ((ax)[i] \sqsubseteq (bx)[j(i)])$, and hence also the the result of substituting any particular sequence of terms for x. In particular, $Q_0 T_\Omega$ can prove $t(i) \subseteq \sqcup_i t(i)$ in this sense.

As in Burr [24], we use a variant of Shoenfield's interpretation [79] which incorporates an idea due to Diller and Nahm [27]. The Shoenfield interpretation works for classical logic, based on the connectives $\forall$, $\vee$, and $\neg$. This has the virtue of cutting down on the number of axioms and rules that need to be verified, and keeping complexity down. Alternatively, we could have used a Diller-Nahm variant of the ordinary Gödel interpretation, combined with a double-negation interpretation. The relationship between the latter approach and the Shoenfield interpretation is now well understood (see [9, 86]).

To each formula φ in the language of $OR_1 + (I)$, we associate a formula φ^S of the form $\forall a\ \exists b\ \varphi_S(a,b)$, where a and b are tuples of variables of certain pure Ω-types (which are implicit in the definitions below), and φ_S is a formula in the language of $Q_0 T_\Omega + (I)$. The interpretation is defined, inductively, in such a way that the following monotonicity property is preserved: whenever $Q_0 T_\Omega + (I)$ proves $b \sqsubseteq b'$, it proves $\varphi_S(a,b) \rightarrow \varphi_S(a,b')$. In the base case, we define

$$\begin{aligned} I(\alpha, t)^S &\equiv I(\alpha, t) \\ (s = t)^S &\equiv s = t \end{aligned}$$

In the inductive step, suppose φ^S is $\forall a\, \exists b\, \varphi_S(a, b)$ and ψ^S is $\forall c\, \exists d\, \psi_S(c, d)$. Then we define

$$\begin{aligned}
(\varphi \lor \psi)^S &\equiv \forall a, c\, \exists b, d\, (\varphi_S(a, b) \lor \psi_S(c, d)) \\
(\forall x\, \varphi)^S &\equiv \forall a\, \exists b\, (\forall x\, \varphi_S(a, b)) \\
(\forall \alpha\, \varphi)^S &\equiv \forall \alpha, a\, \exists b\, \varphi_S(a, b) \\
(\neg\varphi)^S &\equiv \forall B\, \exists a\, (\exists i\, \neg\varphi_S(a[i], B(a[i]))).
\end{aligned}$$

Verifying the monotonicity claim above is straightforward; the inner existential quantifier in the clause for negation takes care of the only case that would otherwise have given us trouble. Note in particular the clause for universal quantification over the natural numbers. Our functional interpretation is concerned with bounds; because we can compute "countable unions" using the operator $\sqcup$, we can view quantification over the natural numbers as "small" and insist that the bound provided by b is independent of x. Note also that if φ is a purely arithmetic formula, φ^S is just φ.

The rest of this section is devoted to proving the following:

Theorem 4.10. *Suppose $OR_1 + (I)$ proves φ, and φ^S is the formula $\forall a\, \exists b\, \varphi_S(a, b)$. Then there are terms b of T_Ω involving at most the variables a and the free variables of φ of type Ω such that $Q_0\, T_\Omega + (I)$ proves $\varphi_S(a, b)$.*

Importantly, the terms b in the statement of the theorem do not depend on the free variables of φ of type N.

As usual, the proof is by induction on derivations. The details are similar to those in Burr [24]. As in Shoenfield [79], we can take the logical axioms and rules to be the following:

1. excluded middle: $\neg\varphi \lor \varphi$
2. substitution: $\forall x\, \varphi(x) \to \varphi(t)$, and $\forall \alpha\, \varphi(\alpha) \to \varphi(t)$
3. expansion: from φ conclude $\varphi \lor \psi$
4. contraction: from $\varphi \lor \varphi$ conclude φ
5. cut: from $\varphi \lor \psi$ and $\neg\varphi \lor \theta$, conclude $\psi \lor \theta$.
6. $\forall$-introduction: from $\varphi \lor \psi$ conclude $\forall x\, \varphi \lor \psi$, assuming x is not free in ψ; and similarly for variables of type Ω
7. equality axioms

The translation of excluded middle is

$$\forall B, a'\, \exists a, b'\, (\exists i\, \neg\varphi_S(a[i], B(a[i])) \lor \varphi_S(a', b')).$$

Given B and a', let $a = \sup_i a'$, so that $a[i] = a'$ for every i; in other words, a represents the singleton set $\{a'\}$. Let $b' = B(a')$. Then the matrix of the formula holds with $i = 0$.

The translation of substitution for the natural numbers is equivalent to

$$\forall B, a'\, \exists a, b'\, (\forall i, x\, \varphi_S(x, a[i], B(a[i])) \to \varphi_S(t, a', b')).$$

(In this context, "equivalent to" means that $Q_0\, T_\Omega + (I)$ proves that the $\cdot_S$ part of the translation is equivalent to the expression in parentheses.) Once again, given B and a', letting $a = \sup_i a'$ and $b' = B(a')$ works.

Handling substitution for Ω, expansion, and contraction is straightforward, and so we consider cut. By the inductive hypothesis we have terms $b = b(a, c)$ and $d = d(a, c)$ satisfying

$$\varphi_S(a, b(a, c)) \lor \psi_S(c, d(a, c)), \tag{4.1}$$

and terms $a' = a'(B, e)$ and $f = f(B, e)$ satisfying

$$\exists i\, \neg\varphi_S(a'(B, e)[i], B(a'(B, e))[i]) \lor \theta_S(e, f(B, e)). \tag{4.2}$$

We need terms $d' = d'(c', e')$ and $f' = f'(c', e')$ satisfying

$$\psi_S(c', d'(c', e')) \lor \theta_S(e', f'(c', e')).$$

Given c' and e', and the terms $b(a,c)$, $d(a,c)$, $a'(B,e)$, and $f(B,e)$, define $B' = \lambda a\ b(a,c')$, define $a'' = \sup_i a'(B',e')$, and then define $d' = d(a'',c')$ and $f' = f(B',e')$. Since $Q_0\,T_\Omega + (I)$ proves $B'(a'') = b(a'',c')$, from (4.1) we have

$$\varphi_S(a'', B'(a'')) \vee \psi_S(c', d').$$

Since $a''[i] = a'(B',e')$ for every i, from (4.2) we have

$$\neg\varphi_S(a'', B'(a'')) \vee \theta_S(e', f').$$

Applying cut in $Q_0\,T_\Omega + (I)$, we have $\psi_S(c',d') \vee \theta_S(e',f')$, as required.

The treatment of $\forall$-introduction over N and Ω is straightforward. We can take the equality axioms to be reflexivity, symmetry, transitivity, and congruence with respect to the basic function and relation symbols in the language. These, as well as the defining equations for primitive recursive function symbols in the language and the defining axioms for I, are verified by the fact that for formulas whose quantifiers ranging only over N, $\varphi^S = \varphi$.

Thus we only have to deal with the other axioms of $OR_1 + (I)$, namely, ω bounding, induction on N, and transfinite induction on Ω. Note that if φ has quantifiers ranging only over N, the definition of $\exists$ in terms of $\forall$ implies that $(\exists\alpha\ \varphi(\alpha))^S$ is equivalent to $\exists\alpha\ \exists i\ \varphi(\alpha[i])$. To interpret the translation of ω-bounding, we therefore need to define a term $\beta = \beta(\alpha)$ satisfying

$$\forall x\ \exists i\ \varphi_S(x, \alpha[i]) \rightarrow \exists j\ \forall x\ \exists k\ \varphi_S(x, (\beta[j])[k]).$$

Setting $\beta = \sup_j \alpha$ means that for every j we have $\beta[j] = \alpha$, so this β works.

We can take induction on the natural numbers to be given by the rule "from $\varphi(0)$ and $\varphi(x) \rightarrow \varphi(x+1)$ conclude $\varphi(t)$ for any term t." From a proof of the first hypothesis, we obtain a term $b = b(a)$ satisfying

$$\varphi_S(0, a, b). \tag{4.3}$$

From a proof of the second hypothesis, we obtain terms $a' = a'(B', a'')$ and $b'' = b''(B', a'')$ satisfying

$$\forall i\ \varphi_S(x, a'[i], B'(a'[i])) \rightarrow \varphi_S(x+1, a'', b''). \tag{4.4}$$

If suffices to define a function $f(x, \hat{a})$ and show that we can prove

$$\varphi_S(x, \hat{a}, f(x, \hat{a})), \tag{4.5}$$

since if we then define $\hat{b}(\hat{a}) = \sqcup_x f(x, \hat{a})$, we have $\varphi_S(x, \hat{a}, \hat{b})$ by the monotonicity property of our translation. Define f by

$$f(0, \hat{a}) = b(\hat{a})$$
$$f(x+1, \hat{a}) = b''(\lambda a\ f(x,a), \hat{a})$$

Let B' denote $\lambda a\ f(x,a)$, so $f(x+1,\hat{a}) = b''(B', \hat{a})$. Let $A(x, \hat{a})$ denote the formula (4.5). From (4.3), we have $A(0, \hat{a})$, and from (4.4) we have $\forall i\ A(x, a'(B', \hat{a})[i]) \rightarrow A(x+1, \hat{a})$. Using Proposition 4.9, we obtain $A(x, \hat{a})$, as required.

Transfinite induction, expressed as the rule "from $\varphi(e)$ and $\forall n\ \varphi(\alpha[n]) \rightarrow \varphi(\alpha)$ conclude $\varphi(\alpha)$," is handled in a similar way. From a proof of the first hypothesis we obtain a term $b = b(a)$ satisfying

$$\varphi_S(e, a, b). \tag{4.6}$$

From a proof of the second hypothesis we obtain terms $a' = a'(\alpha, B', a'')$ and $b'' = b''(\alpha, B', a'')$ satisfying

$$\forall i\ \forall n\ \varphi_S(\alpha[n], a'[i], B'(a'[i])) \rightarrow \varphi_S(\alpha, a'', b''). \tag{4.7}$$

It suffices to define a function f satisfying

$$\varphi_S(\alpha, \hat{a}, f(\alpha, \hat{a}))$$

for every α and $\hat{a}$, since then $\hat{b} = f(\alpha, \hat{a})$ is the desired term. Let $A(\alpha, \hat{a})$ be this last formula, and define f by recursion on α:

$$f(\alpha, \hat{a}) = \begin{cases} b(a) & \text{if } \alpha = e \\ b''(\alpha, \lambda a\ (\sqcup_j f(\alpha[j], a)), \hat{a}) & \text{otherwise.} \end{cases}$$

Write B' for the expression $\lambda a\ (\sqcup_j f(\alpha[j], a))$, so we have $f(\alpha, \hat{a}) = b''(\alpha, B', \hat{a})$ when $\alpha \neq e$. We will use the transfinite induction rule given by Proposition 4.9 to show that $A(\alpha, \hat{a})$ holds for every α and $\hat{a}$. From (4.6), we have $A(e, \hat{a})$, so it suffices to show

$$\alpha \neq e \wedge \forall n, i\ A(\alpha[n], a'[i]) \rightarrow A(\alpha, \hat{a}),$$

where a' is the term $a'(\alpha, B', \hat{a})$. Arguing in $Q_0\, T_\Omega + (I)$, assume $\alpha \neq e$ and $\forall n, i\ A(\alpha[n], a'[i])$, that is,

$$\forall n, i\ \varphi_S(\alpha[n], a'[i], f(\alpha[n], a'[i])).$$

By monotonicity, we have

$$\forall n, i\ \varphi_S(\alpha[n], a'[i], \sqcup_j f(\alpha[j], a'[i])).$$

By the definition of B', this is just

$$\forall n, i\ \varphi_S(\alpha[n], a'[i], B'(a'[i])).$$

By (4.7), this implies

$$\varphi_S(\alpha, \hat{a}, f(\alpha, \hat{a})),$$

which is $A(\alpha, \hat{a})$ as required. This concludes the proof of Theorem 4.10.

Our theory $Q_0\, T_\Omega + (I)$ is inspired by Feferman [30], and, in particular, the theory denoted $T_\Omega + (\mu)$ in [10, Section 9]. That theory, like $Q_0\, T_\Omega + (I)$, combines a classical treatment of quantification over the natural numbers with a constructive treatment of the finite types over Ω.

The principal novelty of our interpretation, however, is the use of the Diller-Nahm method in the clause for negation, and the resulting monotonicity property. This played a crucial rule in the interpretation of transfinite induction. The usual Dialectica interpretation would require us to choose a single candidate for the failure of an inductive hypothesis, something that cannot be done constructively. Instead, using the Diller-Nahm trick, we recursively "collect up" a countable sequence of possible counterexamples.

Similar uses of monotonicity can be found in functional interpretations developed by Kohlenbach [53, 56] and Ferreira and Oliva [31], as well as in the forcing interpretations described in Avigad [7]. The functional interpretations of Avigad [8], Burr [24], and Ferreira and Oliva [31] also make use of the Diller-Nahm trick. But Kohlenbach, Ferreira, and Oliva rely on majorizability relations, which cannot be represented in $Q_0\, T_\Omega$, due to the restricted uses of quantification in that theory. Our interpretation is perhaps closest to the one found in Burr [24], but a key difference is in our interpretation of universal quantification over the natural numbers; as noted above, because we are computing bounds and our functionals are closed under countable sequences, the universal quantifier is absorbed by the witnessing functional.

4.4 Interpreting $Q_0\, T_\Omega + (I)$ in QT^i_Ω

The hard part of the interpretation is now behind us. It is by now well known that one can embed infinitary proof systems for classical logic in the various constructive theories listed in Theorem 4.4. This idea was used by Tait [87], to provide a constructive consistency proof for the subsystem Σ^1_1-CA of second-order arithmetic. It was later used by Sieg [80, 81] to provide a direct reduction of the classical theory ID_1 to the constructive theory $ID^{i,sp}_2$, as well as the corresponding reductions for theories of transfinitely iterated inductive definitions (see Section 4.5). Here we show that, in particular, one can define an infinitary proof system in QT^i_Ω, and use it to interpret $Q_0\, T_\Omega + (I)$ in a way that preserves Π_2 formulas. The methods are essentially those of Sieg [80, 81], adapted to the theories at hand. In fact, our interpretation yields particular witnessing functions in T_Ω, yielding Theorem 4.5.

Let us define the set of infinitary *constant* propositional formulas, inductively, as follows:

- $\top$ and $\bot$ are formulas.
- If $\varphi_0, \varphi_1, \varphi_2, \ldots$ are formulas, so are $\bigvee_{i \in N} \varphi_i$ and $\bigwedge_{i \in N} \varphi_i$.

Take a *sequent* Γ to be a finite set of such formulas. As usual, we write Γ, Δ for $\Gamma \cup \Delta$ and Γ, φ for $\Gamma \cup \{\varphi\}$. We define a cut-free infinitary proof system for such formulas with the following rules:

- $\Gamma, \top$ is an axiom for each sequent Γ.
- From Γ, φ_i for some i conclude $\Gamma, \bigvee_{i \in N} \varphi_i$.
- From Γ, φ_i for every i conclude $\Gamma, \bigwedge_{i \in N} \varphi_i$.

We also define a mapping $\varphi \mapsto \neg\varphi$ recursively, as follows:

- $\neg\top = \bot$
- $\neg\bot = \top$
- $\neg \bigvee_{i \in N} \varphi_i = \bigwedge_{i \in N} \neg\varphi_i$.
- $\neg \bigwedge_{i \in N} \varphi_i = \bigvee_{i \in N} \neg\varphi_i$.

Note that the proof system does not include the cut rule, namely, "from Γ, φ and $\Gamma, \neg\varphi$ include Γ." In this section we will show that it is possible to represent propositional formulas and infinitary proofs in the language of QT^i_Ω in such a way that QT^i_Ω proves that the set of provable sequents is closed under cut. We will then show that this infinitary proof system makes it possible to interpret $Q_0 T_\Omega + (I)$ in a way that preserves Π_2 sentences. This will yield Theorem 4.5. In fact, our interpretation will yield explicit functions witnessing the truth of the Π_2 from the proof in $Q_0 T_\Omega + (I)$.

We can represent formulas in QT^i_Ω as well-founded trees whose end nodes are labeled either $\top$ or $\bot$ and whose internal nodes are labeled either $\bigvee$ or $\bigwedge$. A well-founded tree is simply an element of Ω. As in the Appendix, if α is an element of Ω, then one can assign to each node of α a unique "address," σ, where σ is a finite sequence of natural numbers. Since these can be coded as natural numbers, a labeling of α from the set $\{\top, \bot, \bigvee, \bigwedge\}$ is a function l from N to N. The assertion that α, l is a proof, i.e. that the labeling has the requisite properties, is given by a universal formula in QT^i_Ω. Using λ-abstraction we can define functions F with recursion of the following form:

$$F(\alpha, l) = \begin{cases} G(l(\emptyset)) & \text{if } \alpha = e \\ H(\lambda n\, G(\alpha[n], \lambda\sigma\, l((n)\hat{}\sigma))) & \text{otherwise,} \end{cases}$$

where $\emptyset$ denotes the sequence of length 0. This yields a principle of recursive definition on formulas, which can be used, for example, to define the map $\varphi \to \neg\varphi$. (This particular function can be defined more simply by just switching $\top$ with $\bot$ and $\bigwedge$ with $\bigvee$ in the labeling.) A principle of induction on formulas is obtained in a similar way. We can now represent proofs as well-founded trees labeled by finite sets of formulas and rules of inference, yielding principles of induction and recursion on proofs as well.

We will write $\vdash \Gamma$ for the assertion that Γ has an infinitary proof, and we will write $\vdash \varphi$ instead of $\vdash \{\varphi\}$. The proofs of the following in QT^i_Ω are now standard and straightforward (see, for example, [77, 81]).

Lemma 4.11 (Weakening)**.** *If $\vdash \Gamma$ and $\Gamma' \supseteq \Gamma$ then $\vdash \Gamma'$.*

Lemma 4.12 (Excluded middle)**.** *For every formula φ, $\vdash \{\varphi, \neg\varphi\}$.*

Lemma 4.13 (Inversion)**.**

- *If $\vdash \Gamma, \bot$, then $\vdash \Gamma$.*
- *If $\vdash \Gamma, \bigwedge_{i \in N} \varphi_i$, then $\vdash \Gamma, \varphi_i$ for every i.*

The first and third of these is proved using induction on proofs in QT^i_Ω. The second is proved using induction on formulas.

Lemma 4.14 (Admissibility of cut)**.** *If $\vdash \Gamma, \varphi$ and $\vdash \Gamma, \neg\varphi$, then $\vdash \Gamma$.*

Proof. We show how to cast the usual proof as a proof by induction on formulas, with a secondary induction on proofs. For any formula φ, define

$$\varphi^\vee = \begin{cases} \varphi & \text{if } \varphi \text{ is } \top \text{ or of the form } \bigvee_{i \in N} \psi_i \\ \neg\varphi & \text{otherwise.} \end{cases}$$

We express the claim to be proved as follows:

For every formula φ, for every proof d, the following holds: if d is a proof of a sequent of the form $\Gamma, \varphi^{\vee}$, then $\vdash \Gamma, \neg(\varphi^{\vee})$ implies $\vdash \Gamma$.

The most interesting case occurs when $\varphi = \varphi^{\vee}$ is of the form $\bigvee_{i \in N} \psi_i$, and the last inference of d is of the form

$$\frac{\Gamma, \bigvee_{i \in N} \psi_i, \psi_j}{\Gamma, \bigvee_{i \in N} \psi_i}$$

Given a proof of $\Gamma, \bigwedge_{i \in N} \neg\psi_i$, apply weakening and the inner inductive hypothesis for the immediate subproof of d to obtain a proof of Γ, ψ_j, apply inversion to obtain a proof of $\Gamma, \neg\psi_j$, and then apply the outer inductive hypothesis to the subformula $\neg\psi_j$ of φ. $\square$

We now assign, to each formula $\varphi(\bar{x})$ in the language of $Q_0\, T^i_\Omega + (I)$, an infinitary formula $\hat{\varphi}(\bar{x})$. More precisely, to each formula $\varphi(\bar{x})$ we assign a function $F_\varphi(\bar{x})$ of T_Ω, in such a way that QT^i_Ω proves "for every $\bar{x}$, $F_\varphi(\bar{x})$ is an infinitary propositional formula." We may as well take $\vee$, $\neg$, and $\forall$ to be the logical connectives of $Q_0\, T^i_\Omega + (I)$, and use the Shoenfield axiomatization of predicate logic given in the last section. For formulas not involving I_α, the assignment is defined inductively as follows:

- $\widehat{s = t}$ is equal to $\top$ if $s = t$, and $\perp$ otherwise.
- $\widehat{\varphi \vee \psi}$ is equal to $\bigvee_j \widehat{\theta}_j$, where $\theta_0 = \varphi$ and $\theta_j = \psi$ for $j > 0$.
- $\widehat{\forall x\, \varphi(x)}$ is $\bigwedge_j \widehat{\varphi}(j)$.
- $\widehat{\neg\varphi}$ is $\neg\widehat{\varphi}$.

If I corresponds to the inductive definition $\psi(x, P)$, the interpretation of $x \in I_\alpha$ is defined recursively:

$$x \in I_\alpha = \begin{cases} \perp & \text{if } \alpha = e \\ \widehat{\psi(x, I_{\prec\alpha})} & \text{otherwise.} \end{cases}$$

The following lemma asserts that this interpretation is sound.

Lemma 4.15. *If $Q_0\, T_\Omega + (I)$ proves $\varphi(\bar{x})$, then QT^i_Ω proves that for every $\bar{x}$, $\vdash \hat{\varphi}(\bar{x})$.*

Proof. We simply run through the axioms and rules of inference in $Q_0\, T_\Omega + (I)$. If $s = t$ is a theorem of T_Ω, it is also a theorem of QT^i_Ω. Hence QT^i_Ω proves $\widehat{s = t} = \top$, and so $\vdash \widehat{s = t}$.

The interpretation of the logical axioms and rules are easily validated in the infinitary propositional calculus augmented with the cut rule, and the interpretation of the defining axioms for I_α are trivially verified given the translation of $\widehat{t \in I_\alpha}$. This leaves only induction on N and transfinite induction on Ω. We will consider transfinite induction on Ω; the treatment of induction on N is similar.

We take transfinite induction to be given by the rule "from $\varphi(e)$ and $\alpha \neq e \wedge \forall n\ \varphi(\alpha[n]) \to \varphi(\alpha)$ conclude $\varphi(\alpha)$." Arguing in $Q_0\, T_\Omega + (I)$, suppose for every instantiation of α and the parameters of φ there is an infinitary derivation of the $\widehat{\cdot}$ translation of these hypothesis. Fixing the other parameters, use transfinite induction to show that for every α there is an infinitary proof of $\widehat{\varphi}(\alpha)$. When $\alpha = e$, this is immediate. In the inductive step we have infinitary proofs of $\widehat{\varphi}(\alpha[n])$ for every n. Applying the $\bigwedge$-rule, we obtain an infinitary proof of $\forall n\ \widehat{\varphi(\alpha[n])}$, and hence, using ordinary logical operations in the calculus with cut, a proof of $\widehat{\varphi}(\alpha)$. $\square$

We note that with a little more care, one can obtain cut-free proofs of the induction and transfinite induction axioms; see, for example, [21].

Lemma 4.16. *Let φ be a formula of the form $\forall\bar{x}\ \exists\bar{y}\ R(\bar{x}, \bar{y})$, where R is primitive recursive. Then QT^i_Ω proves that $\vdash \hat{\varphi}$ implies φ.*

Proof. Using a primitive recursive coding of tuples we can assume, without loss of generality, that each of $\bar{x}$ and $\bar{y}$ is a single variable. Using the inversion lemma, it suffices to prove the statement for Σ_1 formulas, which we can take to be of the form $\exists y\ S(y)$ for some primitive recursive S. Use induction on proofs to prove the slightly more general claim that given any proof of either $\{\widehat{\exists y\ S(y)}\}$ or $\{\widehat{\exists y\ S(y)}, \perp\}$ there is a

j satisfying $S(j)$. In a proof of either sequent, the last rule rule can only have been a $\bigvee$ rule, applied to a sequent of the form $\{\widehat{S(j)}\}$ or $\{\widehat{\exists y\, S(y)}, \widehat{S(j)}\}$. If $\widehat{S(j)}$ equals $\top$, j is the desired witness; otherwise, apply the inductive hypothesis. □

Putting the pieces together, we have shown:

Theorem 4.17. *Every* Π_2 *theorem of* $Q_0\, T_\Omega + (I)$ *is a theorem of* QT^i_Ω.

Together with Theorems 4.7 and 4.10, this yields Theorem 4.5. Note that every time we used induction on formulas or proofs in the lemmas above, the arguments give explicit constructions that are represented by terms of T_Ω. So we actually obtain, from an ID_1 proof of a Π_2 sentence, a T_Ω term witnessing the conclusion and a proof that this is the case in QT^i_Ω. By Theorem 4.4, this can be converting to a proof in T_Ω, if desired.

Our reduction of ID_1 to a constructive theory has been carried out in three steps, amounting, essentially, to a functional interpretation on top of a straightforward cut elimination argument. A similar setup is implicit in the interpretation of ID_1 due to Buchholz [21], where a forcing interpretation is used in conjunction with an infinitary calculus akin to the one we have used here. We have also considered alternative reductions of $Q_0\, T_\Omega + (I)$ that involve either a transfinite version of the Friedman A-translation [33] or a transfinite version of the Dialectica interpretation. These yield interpretations of $Q_0\, T_\Omega + (I)$ not in QT^i_Ω, however, but in a Martin-Löf type theory $ML_1\, V$ with a universe and a type of well-founded sets [1]. $ML_1\, V$ is known to have the same strength as ID_1, but although many consider $ML_1\, V$ to be a legitimate constructive theory in its own right, we do not know of any reduction of $ML_1\, V$ to one of the other constructive theories listed in Theorem 4.4 that does not subsume an reduction of ID_1. Thus the methods described in this section seem to provide an easier route to a stronger result.

4.5 Iterating the interpretation

In this section, we consider theories ID_n of finitely iterated inductive definitions. These are defined in the expected way: ID_{n+1} bears the same relationship to ID_n that ID_1 bears to PA. In other words, in ID_{n+1} one can introduce a inductive definitions given by formulas $\psi(x, P)$, where ψ is a formula in the language of ID_n together with the new predicate P, in which P occurs only positively.

In a similar way, taking T_{Ω_1} to be T_Ω, we can define a sequence of theories T_{Ω_n}. For each $n \geq 1$ take $T_{\Omega_{n+1}}$ to add to T_{Ω_n} a type Ω_{n+1} of trees branching over Ω_n, with corresponding constant e and functionals $\sup : (\Omega_n \to \Omega_{n+1}) \to \Omega_{n+1}$ and $\sup^{-1} : \Omega_{n+1} \to (\Omega_n \to \Omega_{n+1})$. Once again, we extend primitive recursion in T_{Ω_n} to the larger system and add a principle of primitive recursion on Ω_{n+1}. The theories $QT^i_{\Omega_{n+1}}$ are defined analogously. It will be convenient to act as though for each $i < j$, Ω_j is closed under unions indexed by N or Ω_i; this can arranged by fixing injections of N and each Ω_i into Ω_{j-1}.

In this section, we show that our interpretation extends straightforwardly to this more general setting, yielding the following generalization of Theorem 4.5:

Theorem 4.18. *Every* Π_2 *sentence provable in* ID_n *is provable in* $QT^i_{\Omega_n}$.

The interpretation can be further extended to theories of transfinitely iterated inductive definitions, as described in [23]. We do not, however, know of any ordinary mathematical arguments that are naturally represented in such theories.

First, we extend the theories $OR_1 + (I)$ to theories $OR_n + (I)$ in the expected way. In addition to the schema of ω bounding, we add a schema of Ω_j bounding for each $j < n$: for each $i < j$ and formula φ with quantifiers ranging over the types $N, \Omega_1, \ldots, \Omega_i$, we add the axiom

$$\forall \alpha^{\Omega_i}\, \exists \beta^{\Omega_j}\, \varphi(\alpha, \beta) \to \exists \beta\, \forall \alpha\, \exists \gamma^{\Omega_i}\, \varphi(\alpha, \beta[\gamma]). \tag{4.8}$$

The fixed points $I_1, \ldots, I_n$ of ID_n are interpreted iteratively according to the recipe in Section 4.2. In particular, if $\psi_j(x, P)$ is gives the definition of the jth inductively defined predicate I_j, the translation of ψ_j has quantifiers ranging over at most Ω_{j-1}; $t \in I_j$ is interpreted as $\exists \alpha^{\Omega_j}\, (t \in I_{j,\alpha})$, where the predicates $I_j(\alpha, x)$ are defined in analogy to $I(\alpha, x)$. This yields:

Theorem 4.19. *If* ID_n *proves* φ, *then* $OR_n + (I)$ *proves* $\hat{\varphi}$.

Next, we define theories $Q_n\, T_{\Omega_n} + (I)$ in analogy to the theory $Q_0\, T_\Omega + (I)$ of Section 4.3. Now it is quantification over the types $N, \Omega_1, \ldots, \Omega_{n-1}$ that is considered "small," and absorbed into the target theory. In particular, the bounding axioms for these types are unchanged by the functional interpretation. Ω_n bounding, induction on N, and transfinite induction on Ω_n are interpreted as before. In other words, with the corresponding modifications to φ^S, we have the analogue to Theorem 4.10:

Theorem 4.20. *Suppose $OR_n + (I)$ proves φ, and φ^S is the formula $\forall a\, \exists b\, \varphi_S(a, b)$. Then there are terms b of T_{Ω_n} involving at most the variables a and the free variables of φ of type Ω_n such that $Q_n\, T_{\Omega_n} + (I)$ proves $\varphi_S(a, b)$.*

In the last step, we have to embed $Q_n\, T_{\Omega_n} + (I)$ into an infinitary proof system in $QT^i_{\Omega_n}$. The method of doing this is once again found in [80, 81], and an extension of the argument described in Section 4.4. We extend the definition of the infinitary propositional formulas so that when, for each $\alpha \in \Omega_j$ with $j < n$, φ_α is a formula, so are $\bigvee_{\alpha \in \Omega_j} \varphi_\alpha$ and $\bigwedge_{\alpha \in \Omega_j} \varphi_\alpha$. The proof of cut elimination, and the verification of transfinite induction and the defining axioms for the predicates $I_j(\alpha, x)$, are essentially unchanged. The only additional work that is required is to handle the bounding axioms; this is taken care of using a style of bounding argument that is fundamental to the ordinal analysis of such infinitary systems (see [72, 73, 80, 81]).

Lemma 4.21. *For every $i < j < n$ and formula $\varphi(\alpha, \beta)$ with quantifiers ranging only over $N, \Omega_1, \ldots, \Omega_{j-1}$, $QT^i_{\Omega_n}$ proves the translation of the bounding axiom* (4.8).

Proof (sketch). Since $QT^i_{\Omega_n}$ establishes the provability of the law of the excluded middle in the infinitary language, it suffices to show that for every sequent Γ with quantifiers ranging over at most Ω_i, if $\vdash \Gamma, \forall \alpha^{\Omega_i}\, \exists \beta^{\Omega_j}\, \varphi(\alpha, \beta)$, then there is a β in Ω^j such that for every α in Ω_i, $\vdash \exists \gamma^{\Omega_i}\, \varphi(\alpha, \beta[\gamma])$. But this is essentially a consequence of the "Boundedness lemma for Σ" in Sieg [81, page 182]. $\square$

This gives us the proper analogue of Theorem 4.17, and hence Theorem 4.18.

Theorem 4.22. *Every Π_2 theorem of $Q_n\, T_{\Omega_n} + (I)$ is a theorem of $QT^i_{\Omega_n}$.*

Appendix: Kreisel's trick and induction with parameters

For completeness, we sketch a proof of Proposition 4.3. Full details can be found in [44, 45].

Proposition 4.3. *The following is a derived rule of T_Ω:*

$$\frac{\varphi(e, x) \qquad \alpha \neq e \wedge \varphi(\alpha[g(\alpha, x)], h(\alpha, x)) \to \varphi(\alpha, x)}{\varphi(s, t)}$$

for quantifier-free formulas φ.

Proof. We associate to each node of an element of Ω a finite sequence σ of natural numbers, where the ith child of the node corresponding to σ is assigned $\sigma\hat{}(i)$. Then the subtree α_σ of α rooted at σ (or e if σ is not a node of α) can be defined by recursion on Ω as follows:

$$e_\sigma = e$$

$$(\sup f)_\sigma = \begin{cases} \sup f & \text{if } \sigma = \emptyset \\ (f(i))_\tau & \text{if } \sigma = (i)\hat{}\tau \end{cases}$$

Here $\emptyset$ denotes the sequence of length 0.

Now, given φ, g, and h as in the statement of the lemma, we define a function $k(\alpha, x, n)$ by primitive recursion on n. The function k uses the the second clause of the rule to compute a sequence of pairs (σ, y) with the property that $\varphi(\alpha_\sigma, y)$ implies $\varphi(\alpha, x)$. For readability, we fix α and x and write $k(n)$ instead of $k(\alpha, x, n)$. We also write $k_0(n)$ for $(k(n))_0$ and $k_1(n)$ for $(k(n))_1$.

$$k(0) = (\emptyset, x)$$

$$k(n+1) = \begin{cases} (k_0(n)\hat{}(g(\alpha_{k_0(n)}, k_1(n))), & \\ \quad h(\alpha_{k_0(n)}, k_1(n))) & \text{if } \alpha_{k_0(n)} \neq e \\ k(n) & \text{otherwise.} \end{cases}$$

Ordinary induction on the natural numbers shows that for every n, $\varphi(\alpha_{k_0(n)}, k_1(n))$ implies $\varphi(\alpha, x)$. So, it suffices to show that for some n, $\alpha_{k_0(n)} = e$.

Since $k_0(0) \subseteq k_0(1) \subseteq k_0(2) \subseteq \ldots$ is an increasing sequence of sequences, it suffices to establish the more general claim that for every α and every function f from N to N, there is an n such that $\alpha_{(f(0),\ldots,f(n-1))} = e$. To that end, by recursion on Ω, define

$$g(\alpha, f) = \begin{cases} 1 + g(\alpha[f(0)], \lambda n\ f(n+1)) & \text{if } \alpha \neq e \\ 0 & \text{otherwise} \end{cases}$$

Let $h(m) = g(\alpha_{(f(0),\ldots,f(m-1))}, \lambda n\ f(n+m))$. By induction on m we have $h(0) = m + h(m)$ as long as $\alpha_{(f(0),\ldots,f(m-1))} \neq e$. In particular, setting $m = h(0)$, we have $h(h(0)) = 0$, which implies $\alpha_{(f(0),\ldots,f(h(0)-1))} = e$, as required. □

The following principles of induction and recursion were used in Section 4.3.

Proposition 4.9. *The following are derived rules of $Q_0 T_\Omega$:*

$$\frac{\theta(e, x) \qquad \alpha \neq e \wedge \forall i\ \forall j\ \theta(\alpha[i], f(\alpha, x, i, j)) \to \theta(\alpha, x)}{\theta(\alpha, x)}$$

and

$$\frac{\psi(0, x) \qquad \forall j\ \psi(n, f(x, n, j)) \to \psi(n+1, x)}{\psi(n, x)}$$

Proof. Consider the first rule. For any element α of Ω and finite sequence of natural numbers σ (coded as a natural number), once again we let α_σ denote the subtree of α rooted at σ. Let τ be the type of x. We will define a function $h(\alpha, g, \sigma)$ by recursion on α, which returns a function of type $N \to \tau$, with the property that $h(\alpha, g, \emptyset) = g$, and for every σ, $\theta(\alpha_\sigma, x)$ holds for every x in the range of $h(\alpha, g, \sigma)$. Applying the conclusion to $h(\alpha, \lambda i\ x, \emptyset)$ will yield the desired result.

The function h is defined as follows:

$$h(\alpha, g, \sigma) = \begin{cases} g & \text{if } \alpha = e \text{ or } \sigma = \emptyset \\ h(\alpha[i], \lambda l\ f(\alpha, g(l_0), i, l_1), \sigma') & \text{if } \alpha \neq e \text{ and } \sigma = \sigma'\hat{\ }(i) \end{cases}$$

Using transfinite induction on α, we have

$$\forall \sigma\ \forall v\ \theta(\alpha_\sigma, h(\alpha, g, \sigma)(v))$$

for every α, and hence and hence $\theta(\alpha, h(\alpha, \lambda i\ x, \emptyset)(0))$. Since $h(\alpha, \lambda i\ x, \emptyset)(0) = (\lambda i\ x)(0) = x$, we have the desired conclusion.

The second rule is handled in a similar way. □

Chapter 5

A Constructive Proof of the Mean Ergodic Theorem[1]

Let T be a nonexpansive linear operator on a Hilbert space $\mathcal{H}$, that is, a linear operator satisfying $\|Tf\| \leq \|f\|$ for all $f \in \mathcal{H}$. For each $n \geq 1$, let $S_n f = f + Tf + \ldots + T^{n-1}f$ denote the sum of first n iterates of T on f, and let $A_n f = S_n f/n$ denote their average. The von Neumann ergodic theorem asserts that the sequence $A_n f$ converges in the Hilbert space norm. The most important example occurs when T is a Koopman operator $Tf = f \circ \tau$ on $L^2(\mathcal{X})$, where τ is a measure preserving transformation of a probability space $\mathcal{X} = (X, \mathcal{B}, \mu)$. In that setting, the Birkhoff pointwise ergodic theorem asserts that the sequence $A_n f$ converges pointwise almost everywhere, and in the L^1 norm, for any f in $L^1(\mathcal{X})$.

It is known that, in general, the sequence $(A_n f)$ can converge very slowly. For example, Krengel [63] has shown that for any ergodic automorphism of the unit interval under Lebesgue measure, and any sequence (a_n) of positive reals converging to 0, no matter how slowly, there is a subset A of the interval such that

$$\lim_{n\to\infty} \frac{1}{a_n} |A_n(\chi_A) - \mu(A)| = \infty$$

almost everywhere, and

$$\lim_{n\to\infty} \frac{1}{a_n} \|A_n(\chi_A) - \mu(A)\|_p = \infty$$

for every $p \in [1, \infty)$. For related results and references, see [47, Section 0.2] and [64, notes to Section 1.2] for related results and references.) Here, however, we will be concerned with the extent to which a bound on the rate of convergence can be computed from the initial data. That is, given $\mathcal{H}$, T, and f in the statement of the von Neumann ergodic theorem, we can ask whether it is possible to compute, for each rational $\varepsilon > 0$, a value $r(\varepsilon)$, such that for every n greater or equal to $r(\varepsilon)$, we have $\|A_{r(\varepsilon)} f - A_n f\| < \varepsilon$.

Determining whether such an r is computable from the initial data is not the same as determining its rate of growth. For example, if $(a_n)_{n\in\mathbb{N}}$ is any computable sequence of rational numbers that decreases monotonically to 0, then a rate of convergence can be computed trivially from the sequence: given ε, one need only run through the elements of the sequence and until one of them drops below ε. On the other hand, it is relatively easy to construct a computable sequence (b_n) of rational numbers that converge to 0, for which there is *no* computable bound on the rate of convergence. It is also relatively easy to construct a computable, monotone, bounded sequence (c_n) of rationals that does not have a computable limit, which implies that there is no computable bound on the rate of convergence of this sequence, either. These examples are discussed in [11, 13].

In situations where the rate of convergence of the ergodic averages is not computable from T and f, is there any useful information to be had? The logical form of a statement of convergence provides some guidance. The assertion that the sequence $(A_n f)$ converges can be represented as follows:

$$\forall \varepsilon > 0 \, \exists n \, \forall m \geq n \, (\|A_m f - A_n f\| < \varepsilon). \tag{5.1}$$

[1] This chapter is joint work and jointly written with Jeremy Avigad and Philipp Gerhardy

A bound on the rate of convergence is a function $r(\varepsilon)$ that returns a witness to the existential quantifier for each $\varepsilon > 0$. It is the second universal quantifier that leads to noncomputability, since, in general, there is no finite test that can determine whether a particular value of n is large enough. But, classically, the statement of convergence is equivalent to the following:

$$\forall\varepsilon > 0, M : \mathbb{N} \to \mathbb{N}\ \exists n\ (M(n) \geq n \to \|A_{M(n)}f - A_n f\| < \varepsilon). \tag{5.2}$$

To see this, note that if, for some $\varepsilon > 0$, the existential assertion in (5.1) is false, then for every n there is an $m \geq n$ such that $\|A_m f - A_n f\| \geq \varepsilon$. In that case, ε together with any function $M(n)$ that returns such an m for each n represents a counterexample to (5.2). Assertion (5.1) is therefore equivalent to the statement that there is no such counterexample, i.e. assertion (5.2).

But now notice that if (5.2) is true, then for each $\varepsilon > 0$ and M one can compute a witness to the existential quantifier in (5.2) simply by trying values of n until one satisfying $\|A_{M(n)} - A_n\|$ is found. Thus, (5.2) has an inherent computational interpretation. In particular, given any function $K(n)$, suppose we apply (5.2) to a function $M(n)$ which, for each n, returns a value m in the interval $[n, K(n)]$ maximizing $\|A_m f - A_n f\|$. In that case, (5.2) asserts

$$\forall\varepsilon > 0\ \exists n\ \forall m \in [n, K(n)]\ \|A_m f - A_n f\| < \varepsilon.$$

In other words, if $r(\varepsilon)$ is a function producing a witness to the existential quantifier, then, rather than computing an absolute rate of convergence, $r(\varepsilon)$ provides, for each $\varepsilon > 0$, a value n such that the ergodic averages $A_m f$ are stable to within ε on the interval $[n, K(n)]$.

It is now reasonable to ask for an explicit bound on $r(\varepsilon)$, expressed in terms of in terms of K, T, f, and ε. In Section 5.1, we obtain bounds on $r(\varepsilon)$ that, in fact, depend only on K and $\rho = \lceil \|f\|/\varepsilon \rceil$. Since the bound on the rate of convergence is clearly monotone with ρ, our results show that, for fixed K, the bounds are uniform on any bounded region of the Hilbert space and independent of T. As special cases, we have the following:

- If $K = n^{O(1)}$, then $r(f, \varepsilon) = 2^{2^{O(\rho^2 \log\log\rho)}}$.
- If $K = 2^{O(n)}$, then $r(f, \varepsilon) = 2^1_{O(\rho^2)}$, where 2^x_n denotes the nth iterate of $y \mapsto 2^y$ starting with x.
- If $K = O(n)$ and T is an isometry, then $r(f, \varepsilon) = 2^{O(\rho^2 \log\rho)}$.

Fixing ρ and a parameterized class of functions K, one similarly obtains information on the dependence of the bounds on the parameters defining K.

The techniques given here can also be extended to the pointwise ergodic theorem, as is done in [13].

Our constructive versions of the mean and pointwise ergodic theorems are examples of Kreisel's no-counterexample interpretation [60, 61]. Our extractions of bounds can be viewed as applications of a body of proof theoretic results that fall under the heading "proof mining" (see, for example, [54, 57, 58]). What makes it difficult to obtain explicit information from the usual proofs of the mean ergodic theorem is their reliance on a nonconstructive principle, namely, the assertion that any bounded increasing sequence of real numbers converges. Qualitative features of our bounds—specifically, the dependence only on $\|f\|$, K, and ε—are predicted by the general metamathematical results of Gerhardy and Kohlenbach [38]. Moreover, methods due to Kohlenbach make it possible to extract useful bounds from proofs that make use of nonconstructive principles like the one just mentioned. These connections are explained in Section 5.4.

In the field of constructive mathematics, one is generally interested in obtaining constructive analogues of nonconstructive mathematical theorems. Other constructive versions of the ergodic theorems, due to Bishop [15, 16, 17], Nuber [70], and Spitters [85], are discussed in Sections 5.2 and 5.3. Connections to the field of reverse mathematics are also discussed in Section 5.3.

5.1 A constructive mean ergodic theorem

Given any operator T on a Hilbert space and $n \geq 1$, let $S_n f = \sum_{i<n} T^i f$, and let $A_n f = \frac{1}{n} S_n f$. The Riesz version of the mean ergodic theorem is as follows.

Theorem 5.1. *If T is any nonexpansive linear operator on a Hilbert space and f is any element, then the sequence $(A_n f)$ converges.*

We present a proof in a form that will be amenable to extracting a constructive version.

Proof. Let $M = \{f \mid Tf = f\}$ be the subspace consisting of fixed-points of T, and let N be the subspace generated by vectors of the form $u - Tu$ (that is, N is the closure of the set of linear combinations of such vectors).

For any g of the form $u - Tu$ we have $\|A_n g\| = \frac{1}{n}\|u - T^n u\| \leq 2\|u\|/n$, which converges to 0. Passing to limits (using the fact that A_n satisfies $\|A_n v\| \leq \|v\|$ for any v), we have that $A_n g$ converges to 0 for every $g \in N$.

On the other hand, clearly $A_n h = h$ for every $h \in M$. For arbitrary f, write $f = g + h$, where g is the projection of f on N, and $h = f - g$. It suffices to show that h is in M. But we have

$$\begin{aligned}
\|Th - h\|^2 &= \|Th\|^2 - 2\langle Th, h\rangle + \|h\|^2 \\
&\leq \|h\|^2 - 2\langle Th, h\rangle + \|h\|^2 \\
&= 2\langle h, h\rangle - 2\langle Th, h\rangle \\
&= 2\langle h - Th, h\rangle,
\end{aligned} \tag{5.3}$$

and the right-hand side is equal to 0, since h is orthogonal to N. So $Th = h$. □

The last paragraph of proof shows that $N^\perp \subseteq M$, and moreover that $A_n f$ converges to h. It is also possible to show that $M^\perp \subseteq N$, and hence $M = N^\perp$, which implies that h is the projection of f on M. We will not, however, make use of this additional information below.

As indicated in the introduction, the mean ergodic theorem is classically equivalent to the following:

Theorem 5.2. *Let T and f be as above and let $M : \mathbb{N} \to \mathbb{N}$ be any function satisfying $M(n) \geq n$ for every n. Then for every $\varepsilon > 0$ there is an $n \geq 1$ such that $\|A_{M(n)} f - A_n f\| \leq \varepsilon$.*

Our goal here is to provide a constructive proof of this theorem. We will, in particular, provide explicit bounds on n solely in terms of M and $\|f\|/\varepsilon$.

For the rest of this section, we fix a nonexpansive map T and an element f of the Hilbert space. A moment's reflection shows that $A_n f$ lies in the cyclic subspace $\mathcal{H}_f$ spanned by $\{f, Tf, T^2 f, \ldots\}$, and so it suffices to consider the subspace N_f spanned by vectors of the form $T^i f - T^{i+1} f$. Let g be the projection of f onto N_f. Then g is the limit of the sequence $(g_i)_{i \in \mathbb{N}}$, where, for each i, g_i is the projection of f onto the finite dimensional subspace spanned by

$$f - Tf, Tf - T^2 f, \ldots, T^i f - T^{i+1} f.$$

The sequence (g_i) can be defined explicitly by

$$g_0 = \frac{\langle f, f - Tf\rangle}{\|f - Tf\|^2}(f - Tf),$$

and

$$g_{i+1} = g_i + \frac{\langle f - g_i, T^i f - T^{i+1} f\rangle}{\|T^i f - T^{i+1} f\|^2}(T^i f - T^{i+1} f).$$

For each i, we can write $g_i = u_i - Tu_i$, where the sequence $(u_i)_{i \in \mathbb{N}}$ is defined by

$$u_0 = \frac{\langle f, f - Tf\rangle}{\|f - Tf\|^2} f,$$

and

$$u_{i+1} = u_i + \frac{\langle f - g_i, T^i f - T^{i+1} f\rangle}{\|T^i f - T^{i+1} f\|^2} T^i f.$$

Note that this representation of g_i as an element of the form $u - Tu$ is not unique, since if u and u' differ by any fixed point of T, $u - Tu = u' - Tu'$.

Finally, if we define the sequence $(a_i)_{i\in\mathbb{N}}$ by $a_i = \|g_i\|$, then (a_i) is nondecreasing and converges to $\|g\|$. We will see in Section 5.3 that a bound on the rate of convergence of (a_i) might not be computable from T and f. Our strategy here will be to show that, given a fixed "counterexample" function M as in the statement of Theorem 5.2, the fact that the sequence (a_i) is bounded and increasing allows us to bound the number of times that M can foil our attempts to provide a witness to the conclusion of the theorem.

First, let us record some easy but useful facts:

Lemma 5.3. *1. For every n and f, $\|A_n f\| \leq \|f\|$.*

2. For every n and u, $A_n(u - Tu) = (u - T^n u)/n$, and $\|A_n(u - Tu)\| \leq 2\|u\|/n$.

3. For every f, g, and $\varepsilon > 0$, if $\|f - g\| \leq \varepsilon$, then $\|A_n f - A_n g\| \leq \varepsilon$ for any n.

4. For every f, if $\langle f, f - Tf\rangle \leq \varepsilon$, then $\|Tf - f\| \leq \sqrt{2\varepsilon}$.

Proof. The first two are straightforward calculations, the third follows from the first by the linearity of A_n, and the fourth follows from inequality (5.3) in the proof of Theorem 5.1. □

Lemma 5.4. *For every f, if $\|Tf - f\| \leq \varepsilon$, then for every $m \geq n \geq 1$ we have $\|A_m f - A_n f\| \leq (m-n)\varepsilon/2$. In particular, if $\|Tf - f\| \leq \varepsilon$ and $m \geq 1$, then $\|A_m f - f\| \leq m\varepsilon/2$.*

Proof. Suppose $m \geq n \geq 1$. Then

$$\begin{aligned}\|A_m f - A_n f\| &= \|\frac{1}{m}\sum_{i=0}^{m-1} T^i f - \frac{1}{n}\sum_{j=0}^{n-1} T^j f\| \\ &= \frac{1}{mn}\|n\sum_{i=0}^{m-1} T^i f - m\sum_{j=0}^{n-1} T^j f\| \\ &= \frac{1}{mn}\|n\sum_{i=n}^{m-1} T^i f - (m-n)\sum_{j=0}^{n-1} T^j f\|\end{aligned}$$

There are now $n\cdot(m-n)$ instances of $T^i f$ in the first term and $n\cdot(m-n)$ instances of $T^j f$ in the second term. Pairing them off and using that $\|T^i f - T^j f\| \leq (i-j)\cdot\varepsilon$ for each such pair, we have

$$\begin{aligned}\ldots &\leq \frac{1}{mn}\left(n\sum_{i=n}^{m-1} i - (m-n)\sum_{j=0}^{n-1} j\right)\varepsilon \\ &= \frac{1}{mn}\left(n\left(\frac{m(m-1)}{2} - \frac{n(n-1)}{2}\right) - (m-n)\left(\frac{n(n-1)}{2}\right)\right)\varepsilon \\ &= \frac{1}{mn}\left(n\left(\frac{m(m-1)}{2}\right) - m\left(\frac{n(n-1)}{2}\right)\right)\varepsilon \\ &= (m-n)\varepsilon/2\end{aligned}$$

as required. □

We now turn to the proof of the constructive mean ergodic theorem proper. The first lemma relates changes in g_i to changes in a_i.

Lemma 5.5. *Suppose $|a_j - a_i| \leq \varepsilon^2/(2\|f\|)$. Then $\|g_j - g_i\| \leq \varepsilon$.*

Proof. Assume, without loss of generality, $j > i$. Since g_j is the projection of f onto a bigger subspace,

$g_j - g_i$ is orthogonal to g_i. Thus, by the Pythagorean theorem, we have

$$\begin{aligned}\|g_j - g_i\|^2 &= \|g_j\|^2 - \|g_i\|^2 \\ &= |a_j^2 - a_i^2| \\ &= |a_j - a_i| \cdot |a_j + a_i| \\ &\le \frac{\varepsilon^2}{2\|f\|} \cdot 2\|f\| \\ &= \varepsilon^2,\end{aligned}$$

as required. □

The next lemma introduces a strategy that we will exploit a number of times. Namely, we define an increasing function F such that if, for some j, $\|g_{F(j)} - g_j\|$ is sufficiently small, we have a desired conclusion; and then argue that because the sequence (a_i) is nondecreasing and bounded, sufficiently many iterations of F will necessarily produce such a j. (In the next lemma, we use $F(j) = j + 1$.)

Lemma 5.6. *Let $\varepsilon > 0$, let $d = d(\varepsilon) = \lceil 32\|f\|^4/\varepsilon^4 \rceil$. Then for every i there is a j in the interval $[i, i+d)$ such that $\|T(f - g_j) - (f - g_j)\| \le \varepsilon$.*

Proof. By Lemma 5.3.4, to obtain the conclusion, it suffices to ensure $\langle f - g_j, f - g_j - T(f - g_j)\rangle \le \varepsilon^2/2$. We have

$$\begin{aligned}\langle f - g_j, f - g_j - T(f - g_j)\rangle &= \langle f - g_j, f - Tf\rangle + \langle f - g_j, Tg_j - g_j\rangle \\ &= \langle f - g_j, Tg_j - g_j\rangle\end{aligned}$$

because g_j is the projection of f on a space that includes $f - Tf$, and $f - g_j$ is orthogonal to that space. Recall that g_j is a linear combination of vectors of the form $T^k f - T^{k+1} f$ for $k \le j$, and g_{j+1} is the projection of f onto a space that includes $Tg_j - g_j$. Thus, continuing the calculation, we have

$$\begin{aligned}\ldots &= \langle f - g_{j+1}, Tg_j - g_j\rangle + \langle g_{j+1} - g_j, Tg_j - g_j\rangle \\ &= \langle g_{j+1} - g_j, g_j - Tg_j\rangle \\ &\le \|g_{j+1} - g_j\| \cdot \|Tg_j - g_j\| \\ &\le \|g_{j+1} - g_j\|(\|Tg_j\| + \|g_j\|) \\ &\le 2\|g_{j+1} - g_j\| \cdot \|f\|\end{aligned}$$

Thus, if $\|g_j - g_{j+1}\| \le \frac{\varepsilon^2}{4\|f\|}$, we have the desired conclusion.

Consider the sequence $a_i, a_{i+1}, a_{i+2}, \ldots, a_{i+d-1}$. Since the a_j's are increasing and bounded by $\|f\|$, for some $j \in [i, i+d)$ we have $|a_{j+1} - a_j| \le \frac{\|f\|}{d} \le \frac{\varepsilon^4}{32\|f\|^3}$. By Lemma 5.5, this implies $\|g_j - g_{j+1}\| \le \frac{\varepsilon^2}{4\|f\|}$, as required. □

Lemma 5.7. *Let $\varepsilon > 0$, let $n \ge 1$, and let $d' = d'(n, \varepsilon) = d(2\varepsilon/n) = \lceil 2n^4\|f\|^4/\varepsilon^4 \rceil$. Then for any i, there is an j in the interval $[i, i+d')$ satisfying $\|A_n(f - g_j) - (f - g_j)\| \le \varepsilon$.*

Proof. By the previous lemma, there is some j in the interval $[i, i+d')$ such that $\|T(f - g_j) - (f - g_j)\| \le 2\varepsilon/n$. By Lemma 5.4 this implies $\|A_n(f - g_j) - (f - g_j)\| \le \varepsilon$. □

Lemma 5.8. *Let $\varepsilon > 0$, let $m \ge 1$, let $d'' = d''(m, \varepsilon) = d'(m, \varepsilon/2) = \lceil 32m^4\|f\|^4/\varepsilon^4 \rceil$. Further suppose $\|g_i - g_{i+d''}\| \le \varepsilon/4$. Then for any $n \le m$, $\|A_n(f - g_i) - (f - g_i)\| \le \varepsilon$.*

Proof. By the previous lemma, for any $n \leq m$, there is some j in the interval $[i, i + d'')$ such that $\|A_n(f - g_j) - (f - g_j)\| \leq \varepsilon/2$. This implies

$$\begin{aligned}
\|A_n(f - g_i) - (f - g_i)\| &\leq \|A_n(f - g_i) - A_n(f - g_j)\| \\
&\qquad + \|A_n(f - g_j) - (f - g_j)\| + \|(f - g_j) - (f - g_i)\| \\
&= \|A_n(g_j - g_i)\| + \|A_n(f - g_j) - (f - g_j)\| + \|g_i - g_j\| \\
&\leq \|A_n(f - g_j) - (f - g_j)\| + 2\|g_j - g_i\| \\
&\leq \varepsilon
\end{aligned}$$

as required. □

Lemma 5.9. *Let $\varepsilon > 0$, let $m \geq 1$, let $d''' = d'''(m, \varepsilon) = d''(m, \varepsilon/2) = \lceil 2^9 m^4 \|f\|^4 / \varepsilon^4 \rceil$. Further suppose $\|g_i - g_{i+d'''}\| \leq \varepsilon/8$. Then for any $n \leq m$, $\|A_m(f - g_i) - A_n(f - g_i)\| \leq \varepsilon$.*

Proof. Apply the previous lemma with $\varepsilon/2$ in place of ε. Then for every $n \leq m$,

$$\begin{aligned}
\|A_m(f - g_i) - A_n(f - g_i)\| &\leq \|A_m(f - g_i) - (f - g_i)\| + \\
&\qquad \|A_n(f - g_i) - (f - g_i)\| \\
&\leq \varepsilon/2 + \varepsilon/2 = \varepsilon,
\end{aligned}$$

as required. □

Let us consider where we stand. Given $\varepsilon > 0$ and a function M satisfying $M(n) \geq n$ for every n, our goal is to find an n such that $\|A_{M(n)} f - A_n f\| \leq \varepsilon$. Now, for any n and i, we have

$$\begin{aligned}
\|A_{M(n)} f - A_n f\| &= \|A_{M(n)}(f - g_i) + A_{M(n)} g_i - (A_n(f - g_i) + A_n g_i)\| \\
&\leq \|A_{M(n)}(f - g_i) - A_n(f - g_i)\| + \|A_{M(n)} g_i\| + \|A_n g_i\|.
\end{aligned}$$

Lemma 5.9 tells us how to ensure that the first term on the right-hand side is small: we need only find an i such that $\|g_{i+d'''} - g_i\|$ is small, for some d''', depending on $M(n)$, that is sufficiently large. On the other hand, by Lemma 5.3 and $M(n) \geq n$, we have $\|A_n g_i\| \leq \|u_i\|/(2n)$ and $\|A_{M(n)} g_i\| \leq \|u_i\|/(2M(n)) \leq \|u_i\|/(2n)$. Thus, to guarantee that the remaining two terms are small, it suffices to ensure that n is sufficiently large, in terms of u_i.

There is some circularity here: our choice of i depends on $M(n)$, and hence n, whereas our choice of n depends on u_i, and hence i. The solution is to define sequences $(i_k)_{k \in \mathbb{N}}$ and $(n_k)_{k \in \mathbb{N}}$ recursively, as follows. Set $i_0 = 1$, and, assuming i_k has been defined, set

$$n_k = \max\left(\left\lceil \frac{2\|u_{i_k}\|}{\varepsilon} \right\rceil, 1\right) \tag{5.4}$$

and

$$i_{k+1} = i_k + d'''(\varepsilon/2, M(n_k)) = i_k + \left\lceil \frac{2^{13} M(n_k)^4 \|f\|^4}{\varepsilon^4} \right\rceil \tag{5.5}$$

Let $e = \lceil 2^9 \|f\|^2 / \varepsilon^2 \rceil$, and consider the sequence $a_{i_0}, a_{i_1}, \ldots, a_{i_{e-1}}$. Once again, since this is increasing and bounded by $\|f\|$, for some $k < e$ we have $|a_{i_{k+1}} - a_{i_k}| \leq \varepsilon^2 / 2^9 \|f\|$. Lemma 5.5 implies $\|g_{i_{k+1}} - g_{i_k}\| \leq \varepsilon/16$. Write $i = i_k$ and $n = n_k$, so that $i_{k+1} = i + d'''(M(n), \varepsilon/2)$. Applying Lemma 5.9, we have

$$\|A_{M(n)}(f - g_i) - A_n(f - g_i)\| \leq \varepsilon/2.$$

On the other hand, from the definition of $n = n_k$, we have

$$\|A_n g_i\| \leq \|u_i\|/(2n) \leq \varepsilon/4$$

and

$$\|A_{M(n)} g_i\| \leq \|u_i\|/(2n) \leq \varepsilon/4,$$

so $\|A_M(n) f - A_n f\| \leq \varepsilon$, as required. Notice that the argument also goes through for any sequences (i_k) and (n_k) that grow faster than the ones we have defined, that is, satisfy (5.4) and (5.5) with "=" replaced by "≥." In sum, we have proved the following:

Lemma 5.10. *Given T, f, ε, and M, sequences (i_k) and (n_k) as above, and the value e as above, there is an n satisfying $1 \leq n \leq n_{e-1}$ and $\|A_{M(n)}f - A_n f\| \leq \varepsilon$.*

This is almost the constructive version of the ergodic theorem that we have promised. The problem is that the bound, i_e, is expressed in terms of sequence of values $\|u_i\|$ as well as the parameters M, f, and ε. The fact that the term $\|T^i f - T^{i+1} f\|$ appears in the denominator of a fraction in the definition of the sequence (u_i) makes it impossible to obtain an upper bound in terms of the other parameters. But we can show that if, for any i, $\|T^i f - T^{i+1} f\|$ is sufficiently small (so $T^i f$ is almost a fixed point of T), we can find alternative bounds on an n satisfying the conclusion of our constructive mean ergodic theorem. Thus we can obtain the desired bounds on n by reasoning by cases: if $T^i f - T^{i+1} f$ is sufficiently small for some i, we are done; otherwise, we can bound $\|u_i\|$.

The analysis is somewhat simpler in the case where T is an isometry, since then $\|T^i f - T^{i+1} f\| = \|f - Tf\|$ for every i. Let us deal with that case first.

Lemma 5.11. *If T is an isometry, then for any $m \geq 1$ and $\varepsilon > 0$, one of the following holds:*

1. $\|A_m f - f\| \leq \varepsilon$, *or*
2. $\|u_i\| \leq \frac{(i+1)m\|f\|^2}{2\varepsilon}$ *for every i.*

Proof. By the Cauchy-Schwartz inequality we have $\|u_0\| \leq \|f\|^2/\|f - Tf\|$. By Lemma 5.4, if $\|f - Tf\| \leq 2\varepsilon/m$ then $\|A_m f - f\| \leq \varepsilon$.

Otherwise, $2\varepsilon/m < \|f - Tf\| = \|T^i f - T^{i+1} f\|$ for every i. In that case, we have $\|u_0\| \leq \frac{m\|f\|^2}{2\varepsilon}$, and, since $\|f - g_i\| \leq \|f\|$, we obtain

$$\|u_{i+1}\| \leq \|u_i\| + \frac{m\|f\|\|f - g_i\|}{2\varepsilon} \leq \|u_i\| + \frac{m\|f\|^2}{2\varepsilon}$$

for every i. The result follows by induction on i. □

We can now obtain the desired bounds. If $n = 1$ does not satisfy $\|A_{M(n)}f - A_n f\| \leq \varepsilon$, we have $n_k \leq \lceil \frac{(i_k+1)M(1)\|f\|^2}{\varepsilon^2} \rceil$ for each k. Otherwise, let K be any nondecreasing function satisfying $K(n) \geq M(n) \geq n$ for every n. From the definition of the sequence (i_k), we can extract a function $\widehat{K}(i)$ such that for every k, $\widehat{K}^k(1) \geq i_k$:

- $\rho = \lceil \|f\|/\varepsilon \rceil$
- $\widehat{K}(i) = i + 2^{13}\rho^4 K((i+1)K(1)\rho^2)$
- $e = 2^9\rho^2$

As long as f is nonzero, we have $\rho \geq 1$, which ensures that $\widehat{K}^e(1) \geq n_{e-1}$ and $\widehat{K}^e(1) \geq 1$. Thus, we have $\|A_{M(n)}f - A_n f\| \leq \varepsilon$ for some $n \leq \widehat{K}^e(1)$.

On the other hand, given a nondecreasing function K to serve as a bound for M, the best a counterexample function $M(n)$ can do is return any m in the interval $[n, K(n)]$ satisfying $\|A_m f - A_n f\| > \varepsilon$, if there is one. Thus, we have the following:

Theorem 5.12. *Let T be an isometry on a Hilbert space, and let f be any nonzero element of that space. Let K be any nondecreasing function satisfying $K(n) \geq n$ for every n, and let $\widehat{K}$ be as defined above. Then for every $\varepsilon > 0$, there is an n satisfying $1 \leq n \leq \widehat{K}^e(1)$, such that for every m in $[n, K(n)]$, $\|A_m f - A_n f\| \leq \varepsilon$.*

This is our explicit, constructive version of the mean ergodic theorem, for the case where T is an isometry. If T is merely nonexpansive instead of an isometry, the argument is more complicated and requires a more general version of Lemma 5.4.

Lemma 5.13. *Assume T is a nonexpansive mapping on a Hilbert space, f is any element, $m \geq n \geq 1$, and $\varepsilon > 0$. Then for any k, if $n \geq 2k\|f\|/\varepsilon > k$, then either $\|T^k f - T^{k+1} f\| > \varepsilon/(2m)$ or $\|A_m f - A_n f\| \leq \varepsilon$.*

Proof. We have

$$\begin{aligned}
\|A_m f - A_n f\| &= \frac{1}{mn}\|n\sum_{i=0}^{m-1} T^i f - m\sum_{j=0}^{n-1} T^j f\| \\
&\leq \frac{1}{mn}\|n\sum_{i=0}^{k-1} T^i f - m\sum_{j=0}^{k-1} T^j f\| + \frac{1}{mn}\|n\sum_{i=k}^{m-1} T^i f - m\sum_{j=k}^{n-1} T^j f\| \\
&\leq \frac{1}{mn}\|(n-m)\sum_{j=0}^{k-1} T^j f\| + \frac{1}{mn}\|n\sum_{i=0}^{m-k-1} T^i(T^k f) - m\sum_{j=0}^{n-k-1} T^j(T^k f)\|.
\end{aligned}$$

The first term is less than or equal to

$$\frac{(m-n)}{mn}\|\sum_{i=0}^{k-1} T^j f\| \leq \frac{k\|f\|}{n} \leq \varepsilon/2.$$

Using an argument similar to the one used in the proof of Lemma 5.4, we have

$$\frac{1}{nm}\|n\sum_{i=0}^{m-k-1} T^i(T^k f) - m\sum_{j=0}^{n-k-1} T^j(T^k f)\| \leq (m-n)\|T^k f - T^{k+1} f\| \leq m\|T^k f - T^{k+1} f\|$$

If $\|T^k f - T^{k+1} f\| \leq \frac{\varepsilon}{2m}$, the second term in the last expression is also less than or equal to $\varepsilon/2$, in which case $\|A_m f - A_n f\| \leq \varepsilon$. □

We now have an analogue to Lemma 5.11 for the nonexpansive case.

Lemma 5.14. *For any $i \geq 0$, $n \geq 1$, and $\varepsilon > 0$, either*

1. *there is an $n \leq 2i\lceil\frac{\|f\|}{\varepsilon}\rceil$ such that $\|A_{M(n)} f - A_n f\| \leq \varepsilon$, or*
2. $\|u_i\| \leq \frac{\|f\|^2}{2\varepsilon}\sum_{j=0}^{i} M(2j\lceil\frac{\|f\|}{\varepsilon}\rceil)$

Proof. Use induction on i. At stage $i+1$, if clause 1 doesn't hold, we have $\|A_{M(i+1)} f - A_{i+1} f\| > \varepsilon$, in which case we can use the inductive hypothesis, the previous lemma, and the definition of u_{i+1} to obtain clause 2. □

The definition of the sequences (i_k) and (n_k) remain valid. What has changed is that we now have a more complex expression for the bounds on n_k in the case where case 2 of Lemma 5.14 holds for each i_k. In other words, we have that for every k,

$$n_k \leq \left\lceil \frac{\|f\|^2}{\varepsilon^2}\sum_{l=0}^{i_k} M(2l\lceil\|f\|/\varepsilon\rceil)\right\rceil.$$

unless there is an $n \leq 2i_k\lceil\|f\|/\varepsilon\rceil$ such that $\|A_{M(n)} f - A_n f\| \leq \varepsilon$. Assuming K is a nondecreasing function satisfying $K(n) \geq M(n)$, we can replace this last bound by $\left\lceil\frac{\|f\|^2}{\varepsilon^2}(i_k+1)K(2i_k\lceil\|f\|/\varepsilon\rceil)\right\rceil$. Define

- $\rho = \lceil\|f\|/\varepsilon\rceil$
- $\overline{K}(i) = i + 2^{13}\rho^4 K((i+1)K(2i\rho)\rho^2)$
- $e = 2^9\rho^2$

Then we have:

Theorem 5.15. *Let T be a nonexpansive linear operator on a Hilbert space, and let f be any nonzero element of that space. Let K be any nondecreasing function satisfying $K(n) \geq n$ for every n, and let $\overline{K}$ be as defined above. Then for every $\varepsilon > 0$, there is an n satisfying $1 \leq n \leq \overline{K}^e(1)$, such that for every m in $[n, K(n)]$, $\|A_m f - A_n f\| \leq \varepsilon$.*

Direct calculation yields the following asymptotic bounds.

Theorem 5.16. *Let T be any nonexpansive map on a Hilbert space, let K be any nondecreasing function satisfying $K(n) \geq n$ for every n, and for every nonzero f and $\varepsilon > 0$, let $r_K(f, \varepsilon)$ be the least $n \geq 1$ such that $\|A_m f - A_n f\| \leq \varepsilon$ for every m in $[n, K(n)]$.*

- *If $K = n^{O(1)}$, then $r_K(f, \varepsilon) = 2^{2^{O(\rho^2 \log\log\rho)}}$.*
- *If $K = 2^{O(n)}$, then $r_K(f, \varepsilon) = 2^{1}_{O(\rho^2)}$.*
- *If $K = O(n)$ and T is an isometry, then $r_K(f, \varepsilon) = 2^{O(\rho^2 \log \rho)}$.*

In these expressions, ρ abbreviates $\lceil \|f\|/\varepsilon \rceil$ and 2^x_n denotes the nth iterate of $y \mapsto 2^y$ starting with x.

Alternatively, we can fix ρ and consider the dependence on K. Here are two special cases.

Theorem 5.17. *Let T be an isometry on a Hilbert space, and let K be as above. Fix $\rho = \lceil \|f\|/\varepsilon \rceil$.*

- *If $K(x) = x + c$, then, as a function of c, $r_K(f, \varepsilon) = O(c)$.*
- *If $K(x) = cx + d$, then, as a function of c, $r_K(f, \varepsilon) = c^{O(1)}$.*

5.2 Results from upcrossing inequalities

We are not the first to develop constructive versions of the ergodic theorems. A different type of constructive ergodic theorem, upcrossing inequalities, can be used, indirectly, to obtain bounds on our constructive mean ergodic theorem, Theorem 5.2, in the specific case where the operator in question is the Koopman operator corresponding to a measure preserving transformation. Of course, the upcrossing inequalities characterize the overall oscillatory behavior of a sequence, and thus provide a more information. On the other hand, our results in Section 5.1 apply to any nonexpansive mapping on a Hilbert space, and so are more general. There are further differences: because we obtain our pointwise results from our constructive version of the mean ergodic theorem, the L^2 norm $\|f\|_2$ of f plays a central role. In contrast, results obtained using upcrossing techniques are more naturally expressed in terms of $\|f\|_1$ and $\|f\|_\infty$. In this section, we will see that when the two methods yield analogous results, they provide qualitatively different bounds.

A sequence of real numbers a_n is said to admit k ε-fluctuations if there is a sequence

$$m_1 < n_1 \leq m_2 < n_2 \leq \ldots \leq m_k < n_k$$

such that for every i, $1 \leq i \leq k$, $|a_{m_i} - a_{n_i}| \geq \varepsilon$. Let T be the Koopman operator arising from a measure preserving transformation on $\mathcal{X}$. By the mean ergodic theorem, for every $\varepsilon > 0$, the number k_ε of ε-fluctuations is finite. Kachurovskii [47, Theorem 29] shows:

Theorem 5.18. *Let f be any element of $L^\infty(\mathcal{X})$. Then for every $\varepsilon > 0$,*

$$k_\varepsilon \leq C \left(\frac{\|f\|_\infty}{\varepsilon} \right)^4 \left(1 + \ln \left(\frac{\|f\|_\infty}{\varepsilon} \right) \right)$$

for some constant C.

Now, given any counterexample function M satisfying $M(n) \geq n$ for every n, consider the sequence

$$A_1 f, A_{M(1)} f, A_{M(M(1))} f, \ldots, A_{M^{k_\varepsilon + 1}(1)} f.$$

At least one step must change by less than ε. Thus, we have the following analogue to our Theorem 5.15:

Theorem 5.19. *Let T be a Koopman operator corresponding to a measure preserving transformation of a space $\mathcal{X}$ and let f be any element of $L^\infty(\mathcal{X})$. Let K be any function satisfying $K(n) \geq n$ for every n. Let $k(f, \varepsilon)$ be the bound on k_ε given in the preceding theorem. Then for every $\varepsilon > 0$, there is an n, $1 \leq n \leq K^{k(f,\varepsilon)}(1)$, satisfying $\|A_m f - A_n f\| \leq \varepsilon$ for every $m \in [n, K(n)]$.*

In other words, we can bound a witness to the conclusion of the constructive mean ergodic theorem with $k(f, \varepsilon)$ iterates of K. In contrast, Theorem 5.15 required $e(f, \varepsilon) = C \lceil \|f\|_2/\varepsilon \rceil$ iterates of a faster-growing function $\overline{K}$.

5.3 Computability of rates of convergences

Suppose $(a_n)_{n\in\mathbb{N}}$ is any sequence of rational numbers that decreases monotonically to 0. No matter how slowly the sequence converges, if one is allowed to query the values of the sequence, one can compute a function $r(\varepsilon)$ with the property that for every rational $\varepsilon > 0$ and every $m > r(\varepsilon)$, $|a_m - a_{r(\varepsilon)}| < \varepsilon$. The algorithm is simple: on input ε, just search for an m such that $a_m < \varepsilon$.

On the other hand, it is not hard to construct a computable sequence $(a_n)_{n\in\mathbb{N}}$ of rational numbers that converges to 0, with the property that no computable function $r(\varepsilon)$ meets the specification above. This is an easy consequence of the unsolvability of the halting problem. Let $(M_i)_{i>0}$ be an enumeration of Turing machines, and let j_i be an enumeration of the natural numbers with the property that every natural number appears infinitely often in the enumeration. For every i, let $a_i = 1/j_i$ if Turing machine M_{j_i}, when started input 0, halts in less than i steps, but not in i' steps for any $i' < i$ such that $j_{i'} = j_i$; and let $a_i = 0$ otherwise. Then (a_j) converges to 0, since once we have recognized all the machines among $M_1, \ldots, M_n$ that eventually halt, a_i remains below $1/n$. But any any value $r(1/n)$ meeting the specification above tells us how long we have to wait to determine whether M_n halts, and so any such such r would enable us to solve the halting problem.

In a similar way, one can construct a computable sequence $(a_n)_{n\in\mathbb{N}}$ of rational numbers that is monotone and bounded, but converges to a noncomputable real number. This, too, implies that no computable function $r(\varepsilon)$ meets the specification above. Such a sequence is known as a *Specker sequence*, and an example is given in the proof of Theorem 5.20, below. Thus neither monotonicity nor the existence of a computable limit alone is enough to guarantee the effective convergence of a sequence of rationals.

What these examples show is that the question as to whether it is possible to compute a bound on a rate of convergence of a sequence from some initial data is not a question about the speed of the sequence's convergence, but, rather, its predictability. In this section, we show that in general, one cannot compute a bound on the rate of convergence of ergodic averages from the initial data, although one can do so when dealing when dealing with an ergodic transformation of a (finite) measure space.

The results in this section presuppose notions of computability for various objects of analysis. There are a number of natural, and equivalent, frameworks for defining such notions. For complete detail, the reader should consult Pour el and Richards [74] or Weihrauch [93]. To make sense of the results below, however, the following sketchy overview should suffice.

The general strategy is to focus on infinitary objects that can be represented with a countable set of data. For example, a real number can be taken to be represented by a sequence of rational numbers together with a bound on its rate of convergence; the corresponding real number is said to be computable if it has a computable representation. In other words, a computable real number is given by computable functions $a : \mathbb{N} \to \mathbb{Q}$ and $r : \mathbb{Q} \to \mathbb{N}$ with the property that for every rational $\varepsilon > 0$, $|a_m - a_{r(\varepsilon)}| < \varepsilon$ for every $m \geq r(\varepsilon)$. A function taking infinitary objects as arguments is said to be computable if the output can be computed by a procedure that queries any legitimate representation of the input. For example, a computable function $f(x)$ from $\mathbb{R}$ to $\mathbb{R}$ is given by an algorithm which, given the ability to request arbitrarily good rational approximations to x, produces arbitrarily good rational approximations to $y = f(x)$, in the sense above. In other words, f is given by algorithms that compute functions a_y and r_y representing y, given the ability to query "oracles" a_x and r_x representing x.

Similar considerations apply to separable Hilbert spaces, where are assumed to come with a fixed choice of basis. An element of the space can be represented by a sequence of finite linear combinations of basis elements together with a bound on their rate of rate of convergence in the Hilbert space norm; once again, such an element is said to be computable if it has a computable representation. The inner product and norm are then computable operations on the entire space. A bounded linear operator can be represented by the sequence of values on elements of the basis, and is computable if that sequence is. In general, a computable bounded linear operator need not have a computable norm (see [11, 18]).

Computability with respect to a measure space can be understood in similar ways. A measurable function is represented by a sequence of suitably simple functions that approximate it in the L^1 norm, together with a rate of convergence. Note that this means that a measurable function is represented only up to a.e. equivalence. One can associate to any measure μ the bounded linear operator $Tf = \int f \, d\mu$ on L^1, and take μ to be represented by any representative of the associated T.

The following theorem shows that it is not always possible to compute a bound on the rate of convergence of a sequence of ergodic averages from the initial data.

Theorem 5.20. *There are a computable measure-preserving transformation of* $[0,1]$ *under Lebesgue measure and a computable characteristic function* $f = \chi_A$ *such that if* $f^* = \lim_n A_n f$*, then* $\|f^*\|_2$ *is not a computable real number.*

In particular, f^* *is not a computable element of the Hilbert space, and there is no computable bound on the rate of convergence of* $(A_n f)$ *in either the* L^2 *or* L^1 *norm.*

Proof. First, observe that it suffices to prove the assertion in the first sentence. If f^* were computable, $\|f^*\|_2$ would be computable, and if there were a computable bound on the rate of convergence of $(A_n f)$ in the L^2 norm, then f^* would be a computable element of $L^2([0,1])$. Computable bounds on the rate of convergence in either of the other senses mentioned in the second sentence would imply a computable bound on the rate of convergence in the L^2 norm.

We use a variant of constructions described in [11, 82]. First, suppose f is the characteristic function of the interval $[0,1/2)$, and τ is the rotation $\tau x = (x+a) \mod 1$, where a is either 0 or $1/2^j$ for some $j \geq 1$. If $a = 0$, then $f^* = f$ and $\|f^*\|_2^2 = 1/2$. If $a = 1/2^j$ for any $j \geq 1$, then f^* is the constant function equal to $1/2$, and $\|f^*\|_2^2 = 1/4$. Thus knowing f^* allows us to determine whether $a = 0$.

Our strategy will be to divide $[0,1)$ into intervals $[1-2^i, 1-2^{i+1})$, and let T rotate each interval by a computable real number a_i that depends on whether the ith Turing machine halts. With a suitable choice of f, the limit f^* of the sequence $(A_n f)$ will then encode information as to which Turing machines halt on input 0.

The details are as follows. Let $T(e,x,s)$ be Kleene's T predicate, which asserts that s is a code for a halting computation sequence of Turing machine e on input x. The predicate T is computable, but the set $\{e \mid \exists s\, T(e,0,s)\}$ is not. Without loss of generality, we can assume that for any e and x there is at most one s such that $T(e,x,s)$ holds. We will prove the theorem by constructing computable τ and f such that $\{e \mid \exists s\, T(e,0,s)\}$ is computable from $\|f^*\|$.

Define the computable sequence (a_i) of computable reals by setting

$$a_i = \begin{cases} 1/2^{i+j+1} & \text{if } j \text{ is the unique } j \text{ such that } T(i,0,j), \text{ if such } j \text{ exists} \\ 0 & \text{otherwise} \end{cases}$$

Let τ be the measure preserving transformation that rotates each interval $[1-2^i, 1-2^{i+1})$ by a_i. To see that the sequence (a_i) is computable, remember that we only need to by able to compute approximations to the a_i's uniformly; we can do this by testing $T(i,0,j)$ up to a sufficiently large value of j. To see that τ is computable, remember that it is sufficient to be able to compute approximations to the value of T applied to any simple function, given rational approximations to the the a_i's.

Let f be the characteristic function of the set $\bigcup_i [1-2^i, 1-3\cdot 2^{i+2})$, so that f is equal to 1 on the left half of each interval $[1-2^i, 1-2^{i+1})$ and 0 on the right half. Let $f^* = \lim_n A_n f$. Then

$$\|f^*\|_2^2 = \sum_{\{i \mid \exists j\, T(i,0,j)\}} \frac{1}{4}\cdot\frac{1}{2^{i+1}} + \sum_{\{i \mid \neg\exists j\, T(i,0,j)\}} \frac{1}{2}\cdot\frac{1}{2^{i+1}}$$

and

$$\frac{1}{2} - \|f^*\|_2^2 = \sum_{i\in\mathbb{N}} \frac{1}{2}\cdot\frac{1}{2^{i+1}} - \|f^*\|_2^2 = \sum_{\{i \mid \exists j\, T(i,0,j)\}} \frac{1}{2^{i+3}}.$$

Calling this last expression r, it suffices to show that $\{i \mid \exists j\, T(i,0,j)\}$ is computable from r. But the argument is now standard (see [74, Section 0.2, Corollary 2a] or [83, Theorem III.2.2]). For each n, let

$$r_n = \sum_{\{i \mid \exists j \leq n\, T(i,0,j)\}} \frac{1}{2^{i+3}}.$$

Then the sequence (r_n) is computable and increases monotonically to r. To determine whether Turing machine i halts on input 0, it suffices to search for an n and an approximation to r sufficiently good to ensure $|r - r_n| < 1/2^{i+3}$. Then we only need to check if there is a $j < n$ such that $T(i,0,j)$ holds; if there isn't, $T(i,0,j)$ is false for every j. ☐

The proof of Theorem 5.20 relied on the fact that the system we constructed is not ergodic; we used the behavior of the system on each ergodic component to encode the behavior of a Turing machine. The next theorem and its corollary show that if, on the other hand, the space in question is ergodic, then one always has a computable rate of convergence.

Theorem 5.21. *Let T be a nonexpansive operator on a separable Hilbert space and let f be an element of that space. Let $f^* = \lim_n A_n f$. Then f^*, and a bound on the rate of convergence of $(A_n f)$ in the Hilbert space norm, can be computed from f, T, and $\|f^*\|$.*

Proof. It suffices to show that one can compute a bound on the rate of convergence of $(A_n f)$ from the given data. Assuming f is not already a fixed point of T, write $f = f^* + g$, and let the sequences (g_i), (u_i), and (a_i) be defined as in Section 5.1. Then $g = \lim_i g_i$, and $g_i = u_i - Tu_i$ and $a_i = \|g_i\|$ for every i. Let $a = \lim_i a_i$. Then $a = \|g\|_2 = \sqrt{\|f\|^2 - \|f^*\|^2}$ can be computed from f and $\|f^*\|$. For any m, $n \geq m$, and i, we have

$$\begin{aligned} \|A_m f - A_n f\| &= \|A_m g - A_n g\| \\ &\leq \|A_m g_i - A_n g_i\|_2 + \|A_m(g - g_i)\| + \|A_n(g - g_i)\| \\ &\leq \|A_m g_i\| + \|A_n g_i\| + 2\|g - g_i\| \\ &\leq \|A_m g_i\| + \|A_n g_i\| + 4(a - a_i)\|f\| \end{aligned}$$

as in the proof of Lemma 5.4. Given ε, using the given data we can now find an i such that the last term on the right hand side is less than $\varepsilon/2$, compute u_i, and then, using Lemma 5.3.2, determine an m large enough so that for any $n \geq m$, $\|A_m g_i\| + \|A_n g_i\| < \varepsilon/2$. □

Corollary 5.22. *Let $\mathcal{X} = (X, \mathcal{B}, \mu)$ be a separable measure space, let τ be a an ergodic transformation of $\mathcal{X}$, and let T be the associated Koopman operator. Then for any f in $L^2(\mathcal{X})$, bounds on the rate of convergence in the L^2 norm can be computed from f, T, and μ.*

Proof. If the space is ergodic, f^* is any constant function equal to $\int f \, d\mu$ a.e., and so $\|f^*\|_2 = |\int f \, d\mu|$, which is computable from f and μ. Thus, Theorem 5.21 gives the result for convergence in the L^2 norm. □

The issues raised here can be considered from a spectral standpoint as well. If T is a unitary transformation of a Hilbert space, then the spectral measure σ_f associated to f can be described in the following way. For each $k \in \mathbb{Z}$, let $b_k = \langle T^k f, f \rangle$ be the kth autocorrelation coefficient of f. Let $\mathbb{T}$ be the circle with radius 1, identified with the interval $[0, 2\pi)$. Let I be the linear operator on the complex Hilbert space $L_2^{\mathbb{C}}(\mathbb{T})$ defined with respect to the basis $\langle e^{ik\theta} \rangle_{k \in \mathbb{Z}}$ by $I(e^{ik\theta}) = b_k$. The sequence b_k is a positive definite sequence, and so by Bochner's theorem (see [48, 69]), there is a positive measure σ_f on $\mathbb{T}$ such that $I(g) = \int g \, d\sigma_f$. It is well known that $\|f^*\|_2^2 = \sigma_f(\{0\})$, and Kachurovskii [47, page 670] shows that if $f^* = 0$, then for every n and $\delta \in (0, \pi)$,

$$\|A_n f\|_2 \leq \sqrt{\sigma_f(-\delta, \delta)} + \frac{4\|f\|_2}{n \sin(\delta/2)}.$$

This last expression shows that, in the case where $f^* = 0$, one can compute a bound the rate of convergence of $(A_n f)$ from a bound on the rate of convergence of $\sigma_f(-\delta, \delta)$ as δ approaches 0. The problem is that I is not necessarily a bounded linear transformation, and so σ_f is not generally computable from f. Theorem 5.21 above shows that for any f it is nonetheless possible to compute f^* from $\sigma_f(\{0\})$, f, and T.

For any set of natural numbers X, let X' denote the halting problem relative to X. The proof of Theorem 5.20 shows, more generally, the following:

Theorem 5.23. *For any set of natural numbers X, there are a Lebesgue-measure preserving transformation τ of $[0, 1]$, computable from X, and a computable element f of $L^2([0, 1])$, such that X' is computable from $\|f^*\|_2$.*

The results in this section can be adapted to yield information with respect to provability in restricted axiomatic frameworks. Constructive mathematics, for example, aims to use only principles that can be given a direct computational interpretation (see, for example, [16, 18]). There is also a long tradition of developing mathematics in classical theories that are significantly weaker than set theory. In the field of reverse mathematics, this is done with an eye towards calibrating the degree of nonconstructivity of various theorems of mathematics (see [83]); in the field of proof mining, this is done with an eye towards mining proofs for additional information (see Section 5.4, below).

When a theorem of modern mathematics is not constructively valid, one can search for an "equal hypothesis" substitute, i.e. a constructive theorem with the same hypotheses, and with a conclusion that is easily seen to be classically equivalent to the original theorem. Bishop's upcrossing inequalities, as well as the results of Spitters [85], are of this form. The results of Section 5.1 are also of this form, and are provable both constructively and in the weak base theory RCA_0 of reverse mathematics. One can also look for "equal conclusion" substitutes, by seeking classically equivalent but constructively stronger hypotheses. Theorem 5.21 has this flavor, but it is hard to see how one can turn it into a constructive theorem, because it is not clear how one can refer to $\|f^*\|_2$ without presupposing that $(A_n f)$ converges. One can show, constructively and in RCA_0, that if the projection of f on the subspace N described in the proof of Theorem 5.1 exists then $(A_n f)$ converges; but the assumption that the projection of f on M exists is not sufficient (see [11, 85] and the corrigendum to the latter). An interesting equal conclusion constructive version of the pointwise ergodic theorem can be found in Nuber [70].

Theorem 5.20 shows that the mean and pointwise ergodic theorems do not have constructive proofs. In fact, in the setting of reverse mathematics, they are equivalent to a set-existence principle known as arithmetic comprehension over RCA_0. For stronger results, see [11, 82].

5.4 Proof-theoretic techniques

The methods we have used belong to a branch of mathematical logic called "proof mining," where the aim is to develop general techniques that allow one to extract additional information from nonconstructive or ineffective mathematical proofs. The program is based on two simple observations: first, ordinary mathematical proofs can typically be represented in formal systems that are much weaker then axiomatic set theory; and, second, proof theory provides general methods of analyzing formal proofs in such theories, with an eye towards locating their constructive content. Traditional research has aimed to show that many classical theories can be reduced to constructive theories, at least in principle, and has developed a variety of techniques for establishing such reductions. These include double-negation translations, cut-elimination, Herbrand's theorem, realizability, and functional interpretations. (The *Handbook of Proof Theory* [25] provides an overview of the range of methods.) Proof mining involves adapting and specializing these techniques to specific mathematical domains where additional information can fruitfully be sought.

Our constructive versions of the mean and pointwise ergodic theorems are examples of Kreisel's no-counterexample interpretation [60, 61]. Effective proofs of such translated statements can often be obtained using variants of Gödel's functional ("Dialectica") interpretation [41] (see also [10]). Ulrich Kohlenbach has shown that the Dialectica interpretation can be used as an effective tool; see, for example, [54, 58]. For example, our constructive mean ergodic theorem, Theorem 5.15, provides bounds that depend only on K and $\|f\|/\varepsilon$. In fact, the usual proofs of the mean ergodic theorem can be carried out in axiomatic frameworks for which the general metamathematical results of Gerhardy and Kohlenbach [38] guarantee such uniformity.

While the methods of the paper just cited do show how one can find an explicit expression for the requisite bound, the resulting expression would not yield, a priori, useful bounds. For that, a more refined analysis, due to Kohlenbach [55], can be used. The nonconstructive content of the Riesz proof of the mean ergodic theorem can be traced to the use of the principle of convergence for bounded monotone sequences of real numbers. In formal symbolic terms, the fact that every bounded increasing sequence of real numbers converges can be expressed as follows:

$$\forall a : \mathbb{N} \to \mathbb{R}, c \in \mathbb{R} (\forall i \quad (a_i \leq a_{i+1} \leq c) \to \forall \varepsilon > 0 \; \exists n \; \forall m \geq n \; (|a_m - a_n| \leq \varepsilon)).$$

Using a principle known as "arithmetic comprehension," we can conclude that there is a function, r, bounding the rate of convergence:

$$\forall a : \mathbb{N} \to \mathbb{R}, c \in \mathbb{R} \ (\forall i \ (a_i \leq a_{i+1} \leq c) \to \exists r \ \forall \varepsilon > 0 \ \forall m \geq r(\varepsilon) \ (|a_m - a_{r(\varepsilon)}| \leq \varepsilon)). \quad (5.6)$$

In general, r cannot be computed from the sequence (a_i). On the other hand, the proof of Theorem 5.21 shows that witnesses to the mean ergodic theorem can be computed from a bound r on the rate of convergence, for a sequence (a_i) that is explicitly computed from T and f. Moreover, the proof of this fact can be carried out in a weak theory. Kohlenbach's results show that, in such situations, one can compute explicit witnesses to the Dialectica translation to the theorem in question from a weaker version of principle (5.6):

$$\forall a : \mathbb{N} \to \mathbb{R}, c \in \mathbb{R} \ (\forall i \ (a_i \leq a_{i+1} \leq c) \to \forall \varepsilon > 0, M \exists n \ (M(n) \geq n \to (|a_{M(n)} - a_n| \leq \varepsilon)).$$

This last principle can be given a clear computational interpretation: given ε and M, one can iteratively compute $0, M(0), M(M(0)), \ldots$ until one finds a value of n such that $|a_{M(n)} - a_n| \leq \varepsilon$. This information can then be used to witness the Dialectica translation of the conclusion, that is, our constructive mean ergodic theorem.

This strategy is clearly in evidence in Section 5.1. In practice, it is both infeasible and unnecessary to express the initial proof in completely formal terms. Rather, one undertakes a good deal of heuristic manipulation of the original proof, using the translation to determine what form intermediate lemmas should have, and how they should be combined. The metamathematical results are therefore used as a guide, providing both guarantees as to what results can be achieved, and the strategies needed to achieve them.

Chapter 6

Generalizing the Furstenberg Correspondence

The Furstenberg correspondence [35] was originally developed in order to use ergodic methods to prove Szemerédi's Theorem, that every set of integers with positive upper Banach density contains arbitrarily long arithmetic progressions. The correspondence is based on the following theorem:

Theorem 6.1. *Let $E \subseteq \mathbb{Z}$ with positive upper Banach density be given. Then there is a dynamical system $(Y, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$ such that for any finite set of integers U, the upper Banach density of $\bigcap_{n\in U}(E - n)$ is at least $\mu(\bigcap_{n\in U} T^n A)$.*

Bergelson [14] generalizes this to countable amenable groups:

Definition 6.2. *If $\{I_n\}_{n\in\mathbb{N}}$ is a left Følner sequence of G, for any $E \subseteq G$ define*

$$\overline{d}_{\{I_n\}}(E) = \limsup_{n\to\infty} \frac{|E \cap I_n|}{|I_n|}$$

Say E has positive upper density *with respect to $\{I_n\}$ if $\overline{d}_{\{I_n\}}(E) > 0$.*

Theorem 6.3. *Let G be a countable amenable group and assume that a set $E \subseteq G$ has positive upper density with respect to a left Følner sequence $\{I_n\}_{n\in\mathbb{N}}$. Then there exists a dynamical system $(X, \mathcal{B}, \mu, (T_g)_{g\in G})$ and a set $A \in \mathcal{B}$ with $\mu(A) = \overline{d}_{\{I_n\}}(E)$ such that for any $k \in \mathbb{N}$ and $g_1, \ldots, g_k \in G$, one has*

$$\overline{d}_{\{I_n\}}(E \cap g_1^{-1}E \cap \cdots \cap g_k^{-1}E) \geq \mu(A \cap T_{g_1}^{-1}A \cdots \cap T_{g_k}^{-1}A)$$

In this chapter, we state a generalized version of this theorem and give two proofs. A proof using combinatorial techniques based on those of Furstenberg's will be given in Section 6.2, while a proof using nonstandard analysis will be given in Section 6.4.

6.1 A Generalized Correspondence

Definition 6.4. *Let G be a semigroup, and let $\{I_n\}$ be a sequence of finite subsets of G. The sequence $\{I_n\}$ is a (left) Følner sequence if for each $g \in G$ and each $\epsilon > 0$, there is some $M_{g,\epsilon}$ such that whenever $n \geq M_{g,\epsilon}$,*

$$\frac{|I_n \triangle g \cdot I_n|}{|I_n|} < \epsilon$$

Semigroups for which Følner sequences exist are precisely the countable amenable semigroups.

Theorem 6.5. *Let S be a countable set and let G be a semigroup acting on S. Let X be a second countable compact space. Let $E : S \to X$ be given, and let $\{I_n\}$ be a Følner sequence of subsets of G.*

Then there are a dynamical system $(Y, \mathcal{B}, \nu, (T_g)_{g\in G})$ and measurable functions (with respect to the Borel sets generated by the topology on X) $\tilde{E}_s : Y \to X$ for each $s \in S$ such that the following hold:

- *For any g, s, $\tilde{E}_{gs} = \tilde{E}_s \circ T_g$*
- *For any integer k, any continuous function $u : X^k \to \mathbb{R}$, and any finite sequence $s_1, \ldots, s_k$,*

$$\liminf_{n\to\infty} \frac{1}{|I_n|} \sum_{g\in I_n} u(E(gs_1), \ldots, E(gs_k)) \leq \int u(\tilde{E}_{s_1}, \ldots, \tilde{E}_{s_k}) d\nu$$
$$\leq \limsup_{n\to\infty} \frac{1}{|I_n|} \sum_{g\in I_n} u(E(gs_1), \ldots, E(gs_k))$$

To illustrate this, we give six special cases; the first five have been proven separately, while the sixth, as far as we know, is novel. The first is Theorem 6.1 above:

Corollary 6.6 ([34],[35])**.** *Let $\hat{E} \subseteq \mathbb{Z}$ with positive upper Banach density be given. Then there are a dynamical system $(Y, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$ such that for any finite set of integers U, the upper Banach density of $\bigcap_{n\in U}(E - n)$ is at least $\mu(\bigcap_{n\in U} T^n A)$.*

Proof. Apply Theorem 7.7 by letting $G = S = \mathbb{Z}$ and $E := \chi_{\hat{E}} : \mathbb{Z} \to \{0, 1\}$. Let the sequence $\{I_n\}$ be a sequence of intervals witnessing the positive upper Banach density of E.

This gives a dynamical system $(Y, \mathcal{B}, \mu, \mathbb{Z})$. Let T be the action of 1 on Y. Set $\tilde{E} := \tilde{E}_1$, so $\tilde{E}_n = \tilde{E} \circ T^n$ for each n. Then for any k, the function $u : \{0, 1\}^k \to \mathbb{R}$ given by

$$u(b_1, \ldots, b_k) := \prod_{i\leq k} b_i$$

is continuous, so the upper Banach density of $\bigcap_{n\in U}(E - n)$ is bounded below by $\int \prod_{n\in U} \tilde{E}_n d\mu = \int \prod_{n\in U} \tilde{E} \circ T^n d\mu$. □

By similar arguments:

Corollary 6.7 ([36],[35])**.** *Let $\hat{E} \subseteq \mathbb{Z}^k$ with positive upper Banach density be given. Then there are a dynamical system $(Y, \mathcal{B}, \mu, T_1, \ldots, T_k)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$ such that for any finite set of tuples U, the upper Banach density of $\bigcap_{\vec{n}\in U}(E - \vec{n})$ is at least $\mu(\bigcap_{\vec{n}\in U} T_1^{n_1} \cdots T_k^{n_k} A)$.*

Corollary 6.8. *Let $f : \mathbb{Z} \to [-1, 1]$ be given. Then there are a dynamical system $(Y, \mathcal{B}, \mu, T)$ and a function $F \in L^\infty(Y)$ such that for any finite set of integers U,*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{k\in U} f(n + k) \leq \int \prod_{k\in U} F \circ T^k d\mu \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{k\in U} f(n + k)$$

Corollary 6.9. *Let $(Z, \mathcal{B}, \mu)$ be a separable measure space, and for each $s \leq k$, let $(Z_s, \mathcal{B}_s, \mu_s)$ be a factor. Let a real number b be given, and for $s \leq k$, let $\{f_{s,n}\}_{i\in\mathbb{N}}$ be a sequence of $L^\infty(Z_s, \mathcal{B}_s, \mu_s)$ functions almost everywhere bounded by b. Let g be a weak limit point of the sequence $\frac{1}{N} \sum_{n=0}^{N-1} \prod_{s\leq k} f_{s,n}$ as N goes to infinity.*

Then there are a measure space $(Y, \mathcal{C}, \nu)$ and functions $\tilde{f}_s \in L^\infty(Z_s \times Y)$ such that $\int \prod \tilde{f}_s d\nu$ is g.

Proof. Let S be $[0, k] \times \mathbb{N}$, let X be the set of $L^\infty(Z, \mathcal{B}, \mu)$ functions with norm at most b, under the weak* topology. Let G be $\mathbb{N}$, acting on S by $n(s, m) := (s, m + n)$, let N_t be a sequence of integers witnessing that g is a weak limit point, and let $I_t := [0, N_t]$. Let $E(i, m) := f_{i,m}$.

Fix some orthonormal basis $\{g_j\}$ for $L^2(Z)$. Observe that for each j, the function S_j given by $S_j(h) := \sum_{i\leq j} \int h g_j d\mu$ is continuous, and since each $f_{s,n}$ is almost everywhere bounded by b, it follows that for every j, $||S_j(f_{s,0})||_{L^2(Y)} \leq b$. Therefore $|| \sum_{i\leq j} g_i(z) \int \tilde{E}_{s,0}(y) g_j d\mu|| \leq b$ for every j. Then the infinite sum $\sum_i g_i(z) \int \tilde{E}_{s,0}(y) g_j d\mu$ is a convergent sum of functions measurable in $L^2(Z \times Y)$, and is therefore measurable. We may then take this function to be represented by $\tilde{f}_s(z, y) := \tilde{E}_{s,0}(y)(z)$.

Next, observe that for each g_j, the function $\int \prod \tilde{f}_s g_j d\mu \times \nu$ is equal to

$$\lim_{t\to\infty} \int \frac{1}{N_t} \sum_{n=0}^{N_t-1} \prod_{i\leq s} f_{i,n} g_j d\mu$$

(The limit exists since we have chosen the sequence N_t to witness a particular limit point of the sequence.)

Finally, for each s, if h is orthonormal to $(Z_s, \mathcal{B}_s, \mu_s)$, the set of y such that $\int y_{s,0} h d\mu \neq 0$ has measure 0, and so $\tilde{f}_s$ is measurable with respect to $(Z_s, \mathcal{B}_s, \mu_s)$. $\square$

Definition 6.10 ([67])**.** *If K, H are finite graphs. define $t(K,H)$ to be the fraction of embeddings $\pi : |K| \to |H|$ such that π is a graph embedding.*

Corollary 6.11 ([29],[88])**.** *Let $H_n := (V_n, E_n)$ be a sequence of finite graphs. Then there are a measure space $(Y, \mathcal{B}, \nu)$ and, for every finite graph $K := (V, E)$, a function $\tilde{V}$ on Y such that*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} t(K, H_n) \leq \int \tilde{V} d\nu \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} t(K, H_n)$$

Proof. Let $S = G = \mathbb{N}$. Let X be the space of finite graphs, viewed as functions from finite subsets of $\mathbb{N}$ to $\{0,1\}$. For any graph K, the function $u_K(H) := t(K,H)$ is continuous, so the result follows from Theorem 7.7. $\square$

Note that a the sequence (V_n, E_n) is convergent, in the sense of Lovász and Szegedy, just if $t(K, H_n)$ converges for each K, in which case the lim inf and lim sup of the averages will converge to the same value.

Corollary 6.12. *Let $(\mathbb{N}, E)$ be a countable graph. Then there is a measure space $(Y, \mathcal{B}, \mu)$ such that for any finite graph V, there is a measurable function $\tilde{V}$ such that*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} t(V, ([0,n], E \upharpoonright [0,n])) \leq \int \tilde{V} d\mu \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} t(V, ([1,n], E \upharpoonright [1,n]))$$

Proof. Let S be the space $\mathbb{N}^{[2]}$ (that is, the set of pairs of distinct elements n, m), and G the permutations on finite subsets of $\mathbb{N}$, acting on S by $s \cdot \{a,b\} = \{s(a), s(b)\}$. Take I_n to be the set of permutations on $[1,n]$, and note that this is a Følner sequence. Let $X = \{0,1\}$; then the characteristic function of E is a function from S to X. The measure space $(Y, \mathcal{B}, \mu)$ and functions $\tilde{V}$ exist by the main theorem. $\square$

6.2 Furstenberg-Style Proof

Let $S, G, X, E : S \to X$, and $\{I_n\}$ be given as in the statement of Theorem 7.7. Let Y be the space of functions from S to X. Let $\mathcal{O}$ be a countable subbasis for X. The product topology is on Y is compact by the Tychonoff theorem, and is generated by sets of the form $\{y \mid y(s) \in U\}$ for elements $U \in \mathcal{O}$. Call sets of this form and complements of such sets *simple*.

Let $\mathcal{C}$ be the algebra generated from the simple sets by finite unions and intersections (the simple sets are already closed under complements). This algebra is countable, so by diagonalizing, choose a subsequence $n_t \to \infty$ such that for every $C \in \mathcal{C}$, the limit

$$\rho(C) := \lim_{t\to\infty} \frac{1}{|I_{n_t}|} \sum_{g \in I_{n_t}} \chi_C(E \circ g)$$

is defined.

Lemma 6.13. *ρ is finitely additive.*

Proof. Immediate by expanding the definition, since finite sums distribute over limits and multiplication by constants. $\square$

Since ρ is non-negative, it is also monotonic. Define a G action on Y by $T_g(y)(s) := y(gs)$. For any g and large enough n_t, $\frac{|I_{n_t} \triangle g \cdot I_{n_t}|}{|I_{n_t}|} \to 0$, so

$$|\rho(C) - \rho(T_g C)| = \lim_{t\to\infty} \frac{1}{|I_{n_t}|} \sum_{g \in I_{n_t} \triangle g \cdot I_{n_t}} \chi_C(E \circ g) = 0$$

For an open set $C \in \mathcal{C}$, define $\nu(C)$ to be the supremum of $\rho(C')$ where C' ranges over closed elements of $\mathcal{C}$ contained in C:

$$\nu(C) = \sup_{D \subseteq C, D \text{ closed}} \rho(D)$$

ν is finitely additive on open sets, since if A and B are disjoint, $\nu(A \cup B) = \nu(A) + \nu(B)$ by the definition. Therefore there is a unique finitely additive extension of ν to all of $\mathcal{C}$.

Lemma 6.14. *If C is open then $\nu(C) \leq \rho(C)$.*

Proof. If $D \subseteq C$ then $\rho(D) \leq \rho(C)$. Since

$$\nu(C) = \sup_{D \subseteq C, D \text{ closed}} \rho(D)$$

also $\nu(C) \leq \rho(C)$. □

Lemma 6.15. *For any $C \in \mathcal{C}$ and any $\epsilon > 0$, there is a $D \subseteq C$ such that $D \in \mathcal{C}$, D is closed, and $\nu(C) - \nu(D) < \epsilon$.*

Proof. First, suppose C is simple. If C is closed, C itself suffices, so suppose C is open. Then there is a closed $D \subseteq C$ such that $\rho(D) \geq \nu(C) - \epsilon$. Since $\rho(D) \leq \nu(D)$, also $\nu(D) \geq \nu(C) - \epsilon$.

Any C may be written as a finite union of finite intersections $C = \bigcup_{i \leq k} \bigcap_{j \leq m_i} C_{i,j}$ where each $C_{i,j}$ is simple; by adding additional terms, $\bigcap_j C_{i,j}$ may be assumed to be disjoint from $\bigcap_j C_{i',j}$ for $i' \neq i$. Then it suffices to show the lemma for each $i \leq k$ separately. Suppose that whenever $C = \bigcap_{j \leq m} C_j \cap D$ and D is open then we may find a closed $D' \subseteq D$ such that $\nu(\bigcap C_j \cap D) - \nu(\bigcap C_j \cap D') < \epsilon$. Then we may define C'_j inductively so that if C_j is closed, $C'_j := C_j$, and if C_j is open, $C'_j \subseteq C_j$, C'_j is closed, and

$$\nu(\bigcap_{j'<j} C'_{j'} \cap \bigcap_{j'>j} C_{j'} \cap C_j) - \nu(\bigcap_{j'<j} C'_{j'} \cap \bigcap_{j'>j} C_{j'} \cap C'_j) < \epsilon/j$$

Then $\bigcap C'_j$ is closed and $\nu(\bigcap C_j) - \nu(\bigcap C'_j) < \epsilon$.

To show the assumption, choose $D' \subseteq D$ so that $\nu(D) - \nu(D') < \epsilon$ and write

$$C = \left(\bigcap C_i \cap D'\right) \cup \left(\bigcap C_i \cap (D \setminus D')\right)$$

Since $\bigcap C_i \cap (D \setminus D') \subseteq D \setminus D'$, it follows that $\nu(\bigcap C_i \cap (D \setminus D')) < \epsilon$, and therefore $\nu(C) - \nu(\bigcap C_i \cap D') < \epsilon$. □

By taking complements, for any $C \in \mathcal{C}$ and any $\epsilon > 0$, there is a $D \supseteq C$ such that $D \in \mathcal{C}$, D is open, and $\nu(D) - \nu(C) < \epsilon$.

Lemma 6.16. *ν is σ-additive.*

Proof. Suppose $\bigcup_{i \in \mathbb{N}} C_i = C$. For any $\epsilon > 0$, we may choose $C' \subseteq C$ such that $\nu(C) - \nu(C') < \epsilon/2$ and C' is closed, and sets $C'_i \supseteq C_i$ such that C'_i is open and $\sum_i (\nu(C'_i) - \nu(C_i)) < \epsilon/2$. $C' \subseteq \bigcup_i C'_i$, so by the compactness of Y, there is a finite subcover of C', $C' \subseteq \bigcup_{j \leq k} C'_{i_j}$, so by finite additivity $\nu(C') \leq \sum_j \nu(C'_{i_j})$. But then $\nu(C) \leq \sum_j \nu(C_{i_j}) + \epsilon$. Since we may choose ϵ arbitrarily small, it follows that $\nu(C) \leq \sum_i \nu(C_i)$.

Conversely, since for each k, $\bigcup_{i \leq k} C_i \subseteq C$, finite additivity gives $\sum_{i \leq k} \nu(C_i) \leq \nu(C)$, and therefore $\sum_i \nu(C_i) \leq \nu(C)$. □

For each $s \in S$, define $\tilde{E}_s$ to be the function $y \mapsto y(s)$. These functions are measurable since for any Borel set B on X, $\{y \mid y(s) \in B\}$ belongs to $\mathcal{C}$; indeed, replacing B by an arbitrary open set U and applying the same argument shows that these functions are continuous. Then by definition, $\tilde{E}_{gs} = \tilde{E}_s \circ T_g$.

Let some $u, s_1, \ldots, s_k$ be given. Since X^k is compact and u is continuous, and a continuous function with compact support is integrable, it follows that the function $\tilde{u}$ given by $\tilde{u}(y) := u(\tilde{E}_{s_1}(x), \ldots, \tilde{E}_{s_k}(x))$ is integrable. In particular, $|\tilde{u}(y)|$ has a compact range, and is therefore bounded by some B.

Therefore there is a sequence of functions u_n of the form

$$u_n = \sum_{i=0}^{N_n} v_{n,i}\chi_{S_{n,i}}$$

for elements $v_{n,i} \in V$ and measurable sets $S_{n,i}$ such that

$$\lim_{N\to\infty} |u_n(y) - \tilde{u}(y)| = 0$$

almost everywhere. Since $|v_{n,i}|$ is bounded by B, for any $\epsilon > 0$ we may choose n so that

$$u_n = \sum_{i=0}^{N} v_i\chi_{S_i}$$

and there is a set C so that $C \cup \bigcup_i S_i = Y$, $C \cap \bigcup_i S_i = \emptyset$, the S_i are pairwise disjoint, $\nu(C) < \epsilon/2B$, and $|u_n(y) - \tilde{u}(y)| < \epsilon/2B$ whenever $y \notin C$.

For each i, we may choose an open set S'_i containing S_i such that $\nu(S'_i) - \nu(S_i) < \epsilon/2|v_i|$; since the set of $v' \in V$ within $\epsilon/|v_i|$ of v_i is open and $\tilde{u}$ is continuous, we may further require that for every $y \in S'_i$, $|\tilde{u}(y) - v_i| < \epsilon/2|v_i|$.

We may then choose a closed $S''_i \subseteq S'_i$ such that $\nu(S'_i) - \rho(S''_i) < \epsilon/2N|v_i|$. Then

$$|\sum_i v_i\nu(S_i) - \sum_i v_i\rho(S''_i)| < \epsilon/2$$

But also

$$\sum_i v_i\rho(S''_i) = \lim_{t\to\infty}\sum_i v_i\frac{|\{g \in I_{n_t} \mid E \circ g \in S''_i\}|}{|I_{n_t}|} = \lim_{t\to\infty}\frac{1}{|I_{n_t}|}\sum_i\sum_{g\in I_{n_t}} v_i \cdot \chi_{S''_i}(E \circ g)$$

Whenever $\chi_{S''_i}(E \circ g) = 1$, it follows that

$$|u(E(gs_1), \ldots, E(gs_k)) - v_i| < \epsilon/2|v_i|$$

Therefore this limit is within $\epsilon/2$ of

$$\lim_{t\to\infty}\frac{1}{|I_{n_t}|}\sum_i\sum_{g\in I_{n_t}} u(E(gs_1), \ldots, E(gs_k)) \cdot \chi_{S''_i}(E \circ g)$$

But since $\lim_{t\to\infty}\frac{1}{|I_{n_t}|}\sum_i\sum_{g\in I_{n_t}}\chi_{S''_i}(E \circ g) \geq 1 - \epsilon/2B$, it follows that the limit is within ϵ of

$$\lim_{t\to\infty}\frac{1}{|I_{n_t}|}\sum_{g\in I_{n_t}} u(E(gs_1), \ldots, E(gs_k))$$

Putting this together,

$$|\int \tilde{u}d\nu - \lim_{t\to\infty}\frac{1}{|I_{n_t}|}\sum_{g\in I_{n_t}} u(E(gs_1), \ldots, E(gs_k))| < 2\epsilon$$

for all ϵ. Since $\{I_{n_t}\}$ is a subsequence of $\{I_t\}$, the result follows.

6.3 Minimality

We may further prove that the construction given in the previous section is the smallest such system up to the choices made in the construction.

Theorem 6.17. *Let $(Y, \mathcal{C}, \nu)$ and $\{\tilde{E}_s\}$ satisfy the conclusion of the main theorem. Let $(X, \mathcal{B}, \mu)$ and $\{\tilde{D}_s\}$ be given by the Furstenberg-style proof in the previous section so that for every $u, s_1, \ldots, s_k$, and α,*

$$\nu(\{y \mid u(\tilde{E}_{s_1}(y), \ldots, \tilde{E}_{s_k}(y)) > \alpha\}) = \mu(\{x \mid u(\tilde{D}_{s_1}(x), \ldots, \tilde{D}_{s_k}(x)) > \alpha\})$$

Then there is a measurable measure-preserving function $\pi : Y \to X$ such that $\tilde{E}_s = \tilde{D}_s \circ \pi$.

Proof. The second condition uniquely defines $\pi(y) \in S \to X$ by $\pi(y)(s) := \tilde{E}_s(y)$. Observe that the inverse image of each simple set on X is measurable since each $\tilde{E}_s$ is measurable.

The measure space $(X, \mathcal{B}, \mu)$ is generated by sets of the form

$$B_{u,s_1,\ldots,s_k,\alpha} := \{x \mid u(\tilde{D}_{s_1}(x), \ldots, \tilde{D}_{s_k}(x)) > \alpha\}$$

But then

$$\nu(\{y \mid u(\tilde{D}_{s_1}(\pi(y)), \ldots, \tilde{D}_{s_k}(\pi(y))) > \alpha\} = \nu(\{y \mid u(\tilde{E}_{s_1}(y), \ldots, \tilde{E}_{s_k}(y)) > \alpha\})$$

which is equal to $\mu(B_{u,s_1,\ldots,s_k,\alpha})$ by assumption.

□

6.4 Nonstandard Proof

For a general reference on nonstandard analysis, see, for example, [42].

Fix an ultrapower extension of a universe containing all objects given in the premises and their powersets. The sequence $\{I_n\}$ is a sequence of subsets of G, so let $Y := I_m$ for some nonstandard m be a subset of G^*. For any internal $B \subseteq Y$, let $\mu(B)$ be the standard part of $\frac{|B|}{|Y|}$. By the Loeb measure construction, this extends to a true measure on the σ-algebra extending the set of internal subsets of Y.

For each $s \in S$, define

$$E_s(g) := E^*(gs^*)$$

for every $g \in G^*$. These functions are internal, and therefore measurable. Define an action of G on Y by $T_g g' := g' g^*$. Then $gE_s(g') = E_s(g'g^*) = E^*(g'g^*s^*) = E_{gs}(g')$, as required. Then E_s is a function from Y to X^*.

For every ϵ and large enough t, $\frac{|I_t \triangle g \cdot I_t|}{|I_t|} < \epsilon$, so in particular, $\frac{|Y \triangle g \cdot Y|}{|Y|}$ is infinitesimal, and therefore $\mu(Y \setminus T_g Y) = 0$. Then, for any internal $B \subseteq Y$, $T_g B \subseteq T_g Y$, and so $\mu(T_g B) = \mu(B)$.

Letting $\mathcal{O}$ be the open sets of X. For each $x \in X^*$, consider

$$\{U \in \mathcal{O} \mid x \notin U^*\}$$

Then the starred versions of the complements of these U have the finite intersection property, and therefore the complements of these U have the finite intersection property (for any finite set of these U, since there exists an element in X^* in all of them, by transfer, there also exists an element in X in all of them). By the compactness of X, it follows that there is an element $st(x)$ contained in all of these sets.

Now let $\tilde{E}_s := st \circ E_s$. If u is a continuous function from $X^k \to \mathbb{R}$, $\{s_1, \ldots, s_k\}$ is a finite set then for every ϵ and all but finitely many N,

$$\frac{1}{|I_N|} \sum_{g \in I_N} u(E(g \cdot s_1), \ldots, E(g \cdot s_k)) + \epsilon \geq \liminf_{N \to \infty} \frac{1}{|I_N|} \sum_{g \in I_N} u(E(g \cdot s_1), \ldots, E(g \cdot s_k))$$

and so, by transfer,

$$\frac{1}{|Y|} \sum_{g \in Y} u^*(E^*(g \cdot s_1), \ldots, E^*(g \cdot s_k)) \geq \liminf_{N \to \infty} \frac{1}{|I_N|} \sum_{g \in I_N} u(E(g \cdot s_1), \ldots, E(g \cdot s_k))$$

Note that for any x and any open set U of X containing $st(x)$, U^* contains x, since otherwise the complement of U^* would be a closed set containing x, and therefore also $st(x)$. Since X is compact and u is continuous, for each $\epsilon > 0$ and each $x_1, \ldots, x_k$, there is an open subset U of X^k containing $st(x_1), \ldots, st(x_k)$ so that $|u(x'_1, \ldots, x'_k) - u(st(x_1), \ldots, st(x_k))| < \alpha$ for each $x'_1, \ldots, x_k \in U$. But then U^* contains $x_1, \ldots, x_k$, which means that $|u^*(x_1, \ldots, x_k) - u(st(x_1), \ldots, st(x_k))| < \alpha$. Since this holds for every α, it follows that $st(u^*(x_1, \ldots, x_k)) = u(st(x_1), \ldots, st(x_k))$. Therefore

$$\frac{1}{|Y|}\sum_{g\in Y} u(\tilde{E}_{s_1}(g), \ldots, \tilde{E}_{s_k}(g)) \geq \liminf_{N\to\infty} \frac{1}{|I_N|}\sum_{g\in I_N} u(E(g\cdot s_1), \ldots, E(g\cdot s_k))$$

and also

$$\int u(\tilde{E}_{s_1}(g), \ldots, \tilde{E}_{s_k}(g))d\mu = st(\frac{1}{|Y|}\sum_{g\in Y} u(\tilde{E}(g\cdot s_1), \ldots, \tilde{E}(g\cdot s_k))) \geq \liminf_{N\to\infty} \frac{1}{|I_N|}\sum_{g\in I_N} u(E(g\cdot s_1), \ldots, E(g\cdot s_k))$$

Since the dual claim holds for the $\limsup$, the result follows.

Chapter 7

Norm Convergence for Diagonal Ergodic Averages

Tao has proven [89]

Theorem 7.1. *Let $l \geq 1$ be an integer. Assume $T_1, \ldots, T_l : X \to X$ are commuting, invertible, measure-preserving transformations of a measure space $(X, \mathcal{B}, \mu)$. Then for any $f_1, \ldots, f_l \in L^\infty(X, \mathcal{B}, \mu)$, the averages*

$$A_N(f_1, \ldots, f_l) := \frac{1}{N} \sum_{n=1}^{N} f_1(T_1^n x) \cdots f_l(T_l^n x)$$

converge in $L^2(X, \mathcal{B}, \mu)$.

This can be viewed as a generalization of the mean ergodic theorem, which is the special case when $l = 1$. Several special cases were proven before Tao's result: the $l = 2$ case was proven by Conze and Lesigne [26], and various special cases for higher l have been shown by Zhang [94], Frantzikinakis and Kra [32], Lesigne [65], Host and Kra [43], and Ziegler [95].

Tao's argument is unusual, in that he uses the Furstenberg correspondence principle, which is traditionally used to obtain combinatorial results via ergodic proofs, in reverse: he takes the ergodic system and produces a sequence of finite structures. He then proves a related result for these finitary systems and shows that a counterexample in the ergodic setting would give rise to a counterexample in the finite setting.

7.1 Extensions of Product Spaces

We wish to reduce convergence of the expression in Theorem 1.1 in arbitrary spaces to convergence in spaces where the transformations have been, in some sense, disentangled. The useful location turns out to be extensions of product spaces—that is, given an ergodic dynamical system $\mathbb{Y} = (Y, \mathcal{C}, \nu, U_1, \ldots, U_l)$, we would like to find a system $\mathbb{X} = (\prod_{i \leq l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ where each T_i acts only on the i-th coordinate, but which preserves enough properties of the original system that proving convergence for all $L^\infty(\mathbb{X})$ functions is sufficient to give convergence for all $L^\infty(\mathbb{Y})$ functions.

$\mathbb{X}$ naturally gives rise to a product space, taking $\mathcal{B}_i$ to be the restriction of $\mathcal{B}$ to those sets depending only on the i-th coordinate, but we do not require that $\mathcal{B}$ be the product of the $\mathcal{B}_i$; in general, $\mathcal{B}$ may properly extend the product.

Given any such system, there is a maximal factor $\mathbb{X}' = (X', \mathcal{B}', \mu \upharpoonright \mathcal{B}')$ in which all sets are $T_i T_j^{-1}$ invariant for each $i, j \leq l$. We must either accept poor pointwise behavior, since, for example, this factor does not separate x from $T_i T_j^{-1} x$, or, as will do here, take X' to be a different set. Formally, we will want the property that if γ is the projection of $\prod X_i$ onto X', then for every $i \leq l$ and almost every $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_l$, the function $x_i \mapsto \gamma(x_1, \ldots, x_l)$ is an isomorphism from $(X_i, \mathcal{B}_i, \mu \upharpoonright \mathcal{B}_i)$ to $\mathbb{X}'$.

This obviously requires that all the $\mathbb{X}_i$ be pairwise isomorphic themselves (and further, that $\mathcal{B}$ be symmetric under any change of coordinates).

This requirement is derived from the behavior exhibited by Tao's finitary setting. Here the product space is the finite measure space on $\mathbb{Z}_N^l$ and $\mathbb{X}'$ is the finite measure space on $\mathbb{Z}_N$. The map $\gamma : \mathbb{Z}_N^l \to \mathbb{Z}_N$ is just the map $x_1, \ldots, x_l \mapsto \sum_i x_i$, which has the property that if we fix x_i for $i \neq k$, the map $x_k \mapsto \sum_{i \neq k} x_i + x_k$ is an isomorphism.

Since $(\prod X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ is not a true product space, we cannot rely on Fubini's Theorem. Since we nonetheless wish to integrate over coordinates, we have to rely on the use of certain invariant subsets to produce an analogous property. If $e \subseteq [1, l]$, we will write x_e for an element of $\prod_{i \in e} X_i$; we also write $\overline{e}$ for the complement of e, and, when e is the singleton $\{i\}$, write $\overline{i}$ for the complement of $\{i\}$. Given some x_e, if $i \in e$ then x_i denotes the corresponding element of the sequence x_e. Given two such variables, say, $x_e, x_{\overline{e}}$, will write x for the combination of these two vectors, that is

$$(x_e, x_{\overline{e}})_i := \begin{cases} (x_e)_i & \text{if } i \in e \\ (x_{\overline{e}})_i & \text{otherwise} \end{cases}$$

Note that this is not simply concatenation. For instance, if f is a function on $\prod X_i$, we will often write $f(x_{\overline{k}}, z)$ as an abbreviation for $f(x_1, \ldots, x_{k-1}, z, x_{k+1}, \ldots, x_l)$.

Definition 7.2. *Given a measure space $(\prod_{i \leq l} X_i, \mathcal{B}, \mu)$, for $k \leq l$, let $\mathcal{B}_{\overline{k}}$ be the restriction of $\mathcal{B}$ to those sets of the form $\prod_{i \neq k} B_i \times X_k$ where $B_i \subseteq X_i$ (or having symmetric difference of measure* 0 *with such a set).*

With respect to $\mathcal{B}_{\overline{k}}$, we may identify elements of $\prod_{i \leq l} X_i$ with elements of $\prod_{i \neq k} X_i$ by discarding the k-th coordinate.

Definition 7.3. *Let $\mathbb{Z}$ a dynamical system. We say a* measure disintegration *exists for some factor $\pi : \mathbb{Z} \to \mathbb{Z}'$ if there is a map $z' \mapsto \mu_{z'}$ from $\mathbb{Z}'$ to the space of measures on $\mathbb{Z}$ so that $\mu_{z'}$ is supported on $\pi^{-1}(z')$, the map commutes with the group action (so $\mu_{T_g z'} = T_g \mu_{z'}$ for each g and almost every z'), and for any $f \in L^2(\mathbb{Z})$,*

$$\int f d\mu = \iint f d\mu_{z'} d\mu'$$

where in particular, the right side is defined.

This disintegration always exists given certain conditions on $\mathbb{Z}$ [35], but in our case it is easier to prove that one exists outright than to arrange for those conditions to hold. We will be dealing with a dynamical system $\mathbb{X} = (\prod_{i \leq l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ where each T_i acts only on the i-th coordinate, and where the measure algebra is a (possibly proper) extension of the product measure $\prod_{i \leq l} (X_i, \mathcal{B}_i, \nu_i)$. Furthermore, taking $\mathcal{B}_e$ to consist of those sets depending only on coordinates in e, we will assume the measure disintegration onto $(\prod_{i \neq k} X_i, \mathcal{B}_{\overline{k}}, \mu \upharpoonright \mathcal{B}_{\overline{k}})$ exists (we denote the corresponding measures in the disintegration $\mu_{k, x_{\overline{k}}}$). We want to be able to exchange coordinates, and further, to have an additional, $l+1$-st, coordinate which can "stand in" for any of the other coordinates.

This extra coordinate will be a factor $\mathbb{X}' = (X', \mathcal{B}', \mu', T_1', \ldots, T_l')$ of $\mathbb{X}$ such that a measure disintegration exists and the projection $\gamma : \prod_{i \leq l} X_i \to X'$ is $T_i T_j^{-1}$ invariant almost everywhere (that is, for almost every x, $\gamma(x) = \gamma(T_i T_j^{-1} x)$ for each i, j). If we fix all but one coordinate of $\mathbb{X}$, we obtain a function $\gamma_{x_{\overline{k}}} : X_k \to X'$ by setting $\gamma_{x_{\overline{k}}}(x_k) := \gamma(x_{\overline{k}}, x_k)$.

Definition 7.4. *Let $\mathbb{X}$ have the form $(\prod_{i \leq l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$, and let $\mathbb{X}' = (X', \mathcal{B}', \mu', T_1', \ldots, T_l')$ be a factor of $\mathbb{X}$ with γ the corresponding factor map. We say $\mathbb{X}'$* cleanly factors *$\mathbb{X}$ if:*

- *γ is $T_i T_j^{-1}$ invariant for each i, j*
- *For each $k \leq l$ and almost every $x_{\overline{k}} \in \prod_{i \neq k} X_i$, the function $\gamma_{x_{\overline{k}}}$ is an isomorphism from $(X_k, \mathcal{B}, \mu_{k, x_{\overline{k}}})$ onto $\mathbb{X}'$*

Recall that $\mu_{k, x_{\overline{k}}}$ is the measure given by the measure disintegration of $(\prod_{i \neq k} X_i, \mathcal{B}_{\overline{k}}, \mu \upharpoonright \mathcal{B}_{\overline{k}})$ evaluated at $x_{\overline{k}}$.

It is instructive to consider the finite case, where each X_i is just the discrete measure space on $[1, N]$, as is X', and the function $\gamma(x)$ is simply $\sum_{i \leq l} x_i \mod N$. The clean factoring property here just asserts that if we fix all but one coordinate (and therefore $y := \sum_{i \neq k} x_i \mod N$), the map $\gamma_{x_{\overline{k}}}$, which is given by $\gamma_{x_{\overline{k}}}(x_k) = y + x \mod N$, is a one-to-one mapping.

Theorem 7.5. *If $\mathbb{Y} = (Y, \mathcal{C}, \nu, T_1, \ldots, T_l)$ is an ergodic dynamical system with the T_i commuting, invertible, measure-preserving transformations and $f_1, \ldots, f_l \in L^\infty(\mathbb{Y})$ then there is a dynamical system $\mathbb{X} := (\prod_{i \leq l} X_i, \mathcal{B}, \mu, \tilde{T}_1, \ldots, \tilde{T}_l)$ and functions $\tilde{f}_1, \ldots, \tilde{f}_l \in L^\infty(\mathbb{X})$ such that:*

- *For each of the factors $(\prod_{i \neq k} X_i, \mathcal{B}_{\overline{k}}, \mu \upharpoonright \mathcal{B}_{\overline{k}}, \tilde{T}_1, \ldots, \tilde{T}_l)$ a measure disintegration exists*
- *An $\mathbb{X}'$ exists which cleanly factors $\mathbb{X}$*
- *For each i there is an S_i such that $\tilde{T}_i$ has the form $\tilde{T}_i(x_1, \ldots, x_i, \ldots, x_l) = (x_1, \ldots, S_i x_i, \ldots x_l)$*
- *If $A_N(\tilde{f}_1, \ldots, \tilde{f}_l)$ converges in the L^2 norm then $A_N(f_1, \ldots, f_l)$ does as well*

Note that in the first A_N above, the transformations in question are the $\tilde{T}_i$, while in the latter, the transformations are the T_i. The proof depends on arguments from nonstandard analysis and the Loeb measure construction; see, for instance, [42] for a reference on these topics.

Proof. If $\vec{v} \in [1, P]^l$, write $T^{\vec{v}}$ for $T_1^{v_1} \cdots T_l^{v_l}$. By a multidimensional version of the pointwise ergodic theorem (for instance, the general version of the theorem for amenable groups of transformations [66]), for any function g and almost every x,

$$\int g d\nu = \lim_{P \to \infty} \frac{1}{P^l} \sum_{\vec{v} \in [1,P]^l} g(T^{\vec{v}} x)$$

A point with this property is called *generic* for g. Let $\mathcal{G}$ be the set of polynomial combinations of shifts of the functions f_i with rational coefficients. Since this is a countable set, we may choose a single point x_0 which is generic for every element of $\mathcal{G}$. For each $g \in \mathcal{G}$, define

$$\hat{g}(\vec{n}) := g(T^{\vec{n}} x_0)$$

Since the f_i are L^∞ functions, we may replace them with functions uniformly bounded by some M_{f_i} only changing them on a set of measure 0, and we may therefore assume that each $\hat{g}$ is bounded.

Working in an $\aleph_1$-saturated nonstandard extension, choose some nonstandard c. Using the Loeb measure construction, we may extend the internal counting measure on $[1, c]^l$ to a true external measure μ on the σ-algebra generated by the internal subsets of $[1, c]^l$. The functions $\tilde{g} := \hat{g}^* \upharpoonright [1, c]^l$, the restriction of the nonstandard extension of $\hat{g}$, are internal, and therefore measurable, and bounded since each $\hat{g}$ is.

For each $g \in \mathcal{G}$, by the definition of μ

$$\int \tilde{g} d\mu = st \left(\frac{1}{c^l} \sum_{\vec{n} \in [1,c]^l} \hat{g}^*(\vec{n}) \right)$$

where st is the standard part of a bounded nonstandard real. Furthermore

$$st \left(\frac{1}{c^l} \sum_{\vec{n} \in [1,c]^l} \hat{g}^*(\vec{n}) \right) = \lim_{P \to \infty} \frac{1}{P^l} \sum_{\vec{v} \in [1,P]^l} g(T^{\vec{v}} x_0)$$

follows by transfer: for any rational α greater than $\lim_{P \to \infty} \frac{1}{P^l} \sum_{\vec{v} \in [1,P]^l} g(T^{\vec{v}} x_0)$ and for large enough P, α is greater than the average at P, so for all nonstandard c, α is greater than the average. Similarly for α less than the limit. Putting these together, for any $g \in \mathcal{G}$,

$$\int g d\nu = \int \tilde{g} d\mu$$

Define

$$\tilde{T}_i(x_1, \ldots, x_i, \ldots, x_l) = (x_1, \ldots, (x_i + 1) \bmod c, \ldots, x_l)$$

It follows that $\tilde{T}_i \tilde{g} = \widetilde{T_i g}$, and by ordinary properties of limits, $\tilde{\ }$ commutes with sums and products. Therefore in particular,

$$\int \left[A_N(f_1, \ldots, f_l) - A_M(f_1, \ldots, f_l) \right]^2 d\nu = \int \left[A_N(\tilde{f}_1, \ldots, \tilde{f}_l) - A_M(\tilde{f}_1, \ldots, \tilde{f}_l) \right]^2 d\mu$$

At each point $x_{\overline{k}}$ in $(\prod_{i\neq k}[1,c], \mathcal{B}_{\overline{k}}, \mu \upharpoonright \mathcal{B}_{\overline{k}})$, the Loeb measure construction induces a measure $\mu_{k,x_{\overline{k}}}$ generated by setting

$$\mu_{k,x_{\overline{k}}}(B) := st\left(\frac{1}{c}\sum_{n\in[1,c]}\chi_B(x_{\overline{k}}, n)\right)$$

for internal B.

Finally, let $\mathbb{X}' := ([1,c], \mathcal{B}', \mu')$ be given by the Loeb measure construction on $[1,c]$, and let $\gamma : [1,c]^l \to [1,c]$ be $\gamma(x_1,\ldots,x_l) = \sum_i x_i \mod c$. The function γ is measurable (since it is internal), and measure-preserving (since it maps exactly c^{l-1} points of $[1,c]^l$ to each point of $[1,c]$). For each $n \in [1,c]$, we may define

$$\mu'_n(B) := st\left(\frac{1}{c^{l-1}}\sum_{\vec{v}\in[1,c]^l\mid\sum v_i = n \mod c}\chi_B(\vec{v})\right)$$

for internal B and extend this to a measure on $\mathcal{B}$ by the Loeb measure construction.

$\mathbb{X}$ is isomorphic to a product of $\mathbb{X}'$ with the Loeb measure on $[1,c]^{l-1}$. A theorem of Keisler [49, 50] states that when U and V are hyperfinite sets and f is a measurable function on $Loeb(U\times V)$, the functions $v \mapsto f(u,v)$ are measurable for each u and $\int\int f dv du = \int f d(u\times v)$. In particular, this means that for any measurable B, $\mu(B) = \int \mu_n(B)d\mu'(n)$, so a measure disintegration exists.

For any $k \le l$ and any $x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_l \in \prod_{i\neq k}[1,c]$, $\gamma_{\vec{x}}$ is a measure-preserving bijection from $[1,c]$ to itself mapping measurable sets to measurable sets, and therefore an isomorphism. □

Using the ergodic decomposition, we may reduce the main theorem to the case where $\mathbb{X}$ is ergodic, and then use Theorem 7.5 to reduce to the following case:

Theorem 7.6. *Let $\mathbb{X} = (\prod_{i\le l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ be a cleanly factored dynamical system such that each T_i has the form*

$$T_i(x_1,\ldots,x_i,\ldots,x_l) = (x_1,\ldots,T'_i x_i,\ldots x_l)$$

Then for any $f_1,\ldots,f_l$ in $L^\infty(\mathbb{X})$, $A_N(f_1,\ldots,f_l)$ converges in the L^2 norm.

In order to prove this theorem, we need a slightly stronger inductive hypothesis, which is what we will actually prove.

Lemma 7.7. *Let $\mathbb{Y}$ be an arbitrary measure space, and let $\mathbb{X} = (\prod_{i\le l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ be a cleanly factored dynamical system such that each T_i has the form*

$$T_i(x_1,\ldots,x_i,\ldots,x_l) = (x_1,\ldots,T'_i x_i,\ldots x_l)$$

Then for any $f_1,\ldots,f_l$ in $L^\infty(\mathbb{X}\times\mathbb{Y})$, $A_N(f_1,\ldots,f_l)$ converges in the L^2 norm.

For the remainder of the chapter, assume $\mathbb{X}$ has this form and that $\mathbb{X}'$ is the factor witnessing that $\mathbb{X}$ is cleanly factored, and let γ be the projection onto this factor. By restricting to the factor generated by the countably many translations of the functions f_i, we may assume $\mathbb{X}$ and $\mathbb{X}'$ are separable.

7.2 Diagonal Averages

Note that the projection γ we have constructed is consistent with the transformations T_i, in the sense that $\gamma(x) = \gamma(y)$ implies $\gamma(T_i x) = \gamma(T_i y)$. Furthermore, since γ is $T_i T_j^{-1}$-invariant, $\gamma(x) = \gamma(y)$ implies that $\gamma(T_i x) = \gamma(T_j y)$, even if $i \neq j$.

Definition 7.8. *Define T_{l+1} on $\mathcal{X}'$ such that for each $x' \in \mathcal{X}'$, if $\gamma(x) = x'$ then $\gamma(T_i x) = T_{l+1}x'$.*

With the particular construction we have given, this definition makes sense pointwise. For arbitrary $\mathcal{X}'$ cleanly factoring arbitrary $\mathcal{X}$, this is true only almost everywhere.

We wish to reduce Lemma 7.7 to the case where $\mathcal{X}$ is ergodic. In order to apply the usual theorem for the existence of an ergodic decomposition (see [40]), the measure space must be a standard Borel space. It will be easier to take advantage of the fact that we are working with the L^2 norm, and get a weaker ergodic decomposition that suffices for our purposes. Let $\mathcal{C}$ be the factor consisting of sets which are

T_i-invariant for each i and fix representations of $E(f \mid \mathcal{C})$ for each $f \in L^2(\mathcal{X})$. Let ν be the restriction of μ to $\mathcal{C}$. For each point $x \in \prod X_i$, we can define a measure μ_x by $\int f d\mu_x = E(f \mid \mathcal{C})(x)$ with the property that $\iint f d\mu_x d\nu(x) = \int f d\mu$. Furthermore, the map $x \mapsto \mu_x$ is T_i-invariant for each i, since $\mathcal{C}$ is, and μ_x is ergodic for almost every x. (This argument is the first step of the ordinary proof of the ergodic decomposition, as given in [40], Theorem 3.42.)

We may carry out the same construction on $\mathcal{X}'$ and observe that this preserves the clean factoring property, so it suffices to prove Lemma 7.7 in the case where μ is ergodic.

We wish to extend $\mathcal{X} \times \mathcal{X}'$ to ensure that for each L^2 function f on $\mathcal{X}$, the functions $x_{\overline{k}}, x' \mapsto f(x_{\overline{k}}, \gamma^{-1}_{x_{\overline{k}}}(x'))$ are measurable with integral $\int f d\mu$. If $\mathcal{X}$ were simply a product $\prod_{i \leq l} \mathcal{X}'$, this would occur automatically. Since this is not necessarily the case, however, we must copy over all the additional sets by swapping coordinates.

Formally[1], take the measure algebra $\mathcal{B}^*$ to make measurable all functions $g(x, x')$ such that $x \mapsto \int g(x, x') d\mu'(x')$ is L^∞, and define

$$\int g(x, x') d\nu = \int g(x, x') d\mu'(x') d\mu$$

Define $\mathcal{X}^* := (\prod_{i \leq l} X_i \times X', \mathcal{B}^*, \nu)$.

Importantly, this retains a measure disintegration onto each coordinate:

$$\int f(x, x', y) d\nu = \iint f(x, x', y) d\nu_{k,(x,x')_{\overline{k}}} d\nu_k$$

where $\nu_{k,(x,x')_{\overline{k}}}$ is the pushforward of $\mu_{k,x_{\overline{k}}}$ under γ when $k < l+1$ and $\nu_{l+1,x}(\cdot)$ is $\int \cdot \, \delta_x d\mu$.

Definition 7.9. *By abuse of notation, we take T_i, $i \leq l+1$, to be transformations on $\mathcal{X}^* \times \mathcal{Y}$ where $T_i(x, x', y)$ is given by $(T_i x, x', y)$ if $i \leq l$ and $T_{l+1}(x, x', y) := (x, T_{l+1} x', y)$.*

Since we will only refer to the product measures with $\mathcal{Y}$, and to limit the proliferation of measures, henceforth we let ν denote the measure on $\mathcal{X}^* \times \mathcal{Y}$ and μ denote the measure on $\mathcal{X} \times \mathcal{Y}$. We write μ_k and ν_k for the restriction of μ and ν to the σ-algebra of T_k-invariant sets.

To briefly summarize the construction up to this point, given a measure space $\mathcal{Z}$ and L^∞ functions $f_1, \ldots, f_l$, we have constructed a space $\mathcal{X}^* = (\prod_{i \leq l} X_i \times X', \mathcal{B}^*, \nu, T_1, \ldots, T_{l+1})$ with functions $\tilde{f}_1, \ldots, \tilde{f}_l$ such that:

- Convergence of $A_N(\tilde{f}_1, \ldots, \tilde{f}_l)$ implies convergence of $A_N(f_1, \ldots, f_n)$
- The transformations T_i each act only on the i-th coordinate
- The space $\mathcal{X}^*$ has a measure disintegration onto each coordinate, and onto the space of T_i-invariant functions for each i
- There is a function $\gamma : \prod_{i \leq l} X_i \to X'$ cleanly factoring the space $(\prod_{i \leq l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$

Definition 7.10. *Let $e \subseteq [1, l+1]$. We say $f \in L^2(\mathcal{X}^* \times \mathcal{Y})$ is e-measurable if it is T_i-invariant for each $i \notin e$. We define $\mathcal{I}_d := \{e \subseteq [1, l+1]] \mid |e| = d\}$. We say f has complexity d if it is a finite sum of functions of the form $\prod_{e \in \mathcal{I}_d} g_e$ where each g_e is e-measurable.*

Lemma 7.11. *If $f \in L^2(\mathcal{X}^* \times \mathcal{Y})$ is e-measurable for some e with $|e| < l+1$ then $f(x, \gamma(x), y)$ is an L^2 function and $||f(x, \gamma(x), y)||_{L^2(\mathcal{X} \times \mathcal{Y})} = ||f||_{L^2(\mathcal{X}^* \times \mathcal{Y})}$.*

Proof. For any $i \notin e$,

$$\begin{aligned}\int [f(x, \gamma(x), y)]^2 d\mu &= \iint [f(x_{\overline{i}}, x_i, \gamma_{x_{\overline{i}}}(x_i), y)]^2 d\mu_{i, x_{\overline{i}}} d\mu_i \\ &= \iint [f(x_{\overline{i}}, \gamma^{-1}_{x_{\overline{i}}}(x'), x', y)]^2 d\mu' d\mu_i\end{aligned}$$

[1] John Griesmer suggested this simplified definition of $\mathcal{B}^*$

since γ_{x_i} is an isomorphism. Since f is T_i-invariant and x_i is ergodic with respect to T_i, this is equal to

$$\iint [f(x_{\bar{i}}, x_i, x', y)]^2 d\mu' d\mu$$

Recall that $x_{\bar{i}}, x_i$ is identical to the vector x. But the measure ν was constructed so this is precisely

$$\int [f(x, x', y)]^2 d\nu$$

□

In particular, this means that $f(x, \gamma(x), y)$ is an L^2 function when $f \in L^2(\mathcal{X}^* \times \mathcal{Y})$ has complexity d for some $d < l+1$.

Definition 7.12. *If $f \in L^\infty(\mathcal{X}^* \times \mathcal{Y})$ has complexity d, define $\Delta_N f \in L^\infty(\mathcal{X} \times \mathcal{Y})$ by*

$$\Delta_N f := \frac{1}{N} \sum_{n=1}^{N} f(x, T_{l+1}^n \gamma(x), y)$$

We can reduce the question of the convergence of A_N to the convergence of Δ_N:

Definition 7.13. *If $f \in L^2(\mathcal{X} \times \mathcal{Y})$, define $f^i(x, x', y) := f(x_{\bar{i}}, \gamma_{x_{\bar{i}}}^{-1}(x'), y)$.*

Note that $f^i(x, T_{l+1}^n x', y) = f(x_{\bar{i}}, \gamma_{x_{\bar{i}}}^{-1}(T_{l+1}^n x'), y) = f(x_{\bar{i}}, T_i^n \gamma_{x_{\bar{i}}}^{-1}(x'), y)$.

Lemma 7.14. *Let $f_1, \ldots, f_l$ be given. $A_N(f_1, \ldots, f_l)$ converges in the L^2 norm iff $\Delta_N \prod_{i \in \{1,\ldots,l\}} f_i^i$ converges in the L^2 norm.*

Proof.

$$\begin{aligned}
\Delta_N \prod f_i^i(x, y) =& \frac{1}{N} \sum_{n=1}^{N} \prod_i f_i^i(x, T_{l+1}^n \gamma(x), y) \\
=& \frac{1}{N} \sum_{n=1}^{N} \prod_i f_i(x_{\bar{i}}, \gamma_{x_{\bar{i}}}^{-1}(T_{l+1}^n \gamma(x)), y) \\
=& \frac{1}{N} \sum_{n=1}^{N} \prod_i f_i(x_{\bar{i}}, T_i^n \gamma_{x_{\bar{i}}}^{-1}(\gamma(x)), y) \\
=& \frac{1}{N} \sum_{n=1}^{N} \prod_i f_i(x_{\bar{i}}, T_i^n x_i, y) \\
=& A_N(f_1, \ldots, f_l)(x, y)
\end{aligned}$$

□

Each f_i^i is $[1, l+1] \setminus \{i\}$-measurable, so to prove the main theorem, it suffices to prove convergence of $\Delta_N g$ for functions of complexity $d < l+1$.

While $\Delta_N f$ was defined as a function in $L^\infty(\mathcal{X} \times \mathcal{Y})$, we will sometimes view it as the function in $L^\infty(\mathcal{X}^* \times \mathcal{Y})$ where x' is a dummy variable.

Lemma 7.15. *If g and f are $L^\infty(\mathcal{X}^* \times \mathcal{Y})$ functions with complexity $d < l+1$ and g is T_{l+1}-invariant then $\Delta_N g f = g \Delta_N f$.*

Proof. Immediate from the definition. □

Lemma 7.16. *Suppose g has complexity 1. Then $\Delta_N g$ converges in the L^2 norm.*

Proof. If for almost every $y \in Y$, we have convergence for $x \mapsto g(x,y)$ then we may apply the dominated convergence theorem to obtain convergence over $\mathcal{X}^* \times \mathcal{Y}$. Since Δ_N distributes over sums, we may further assume that g has the form $\prod_i g_i$ where each g_i is $\{i\}$-measurable. Then $\Delta_N g = \prod_{i \neq l+1} g_i \Delta_N g_{l+1}$, and by the previous lemma, it suffices to show that $\Delta_N g_{l+1}$ converges. But this follows immediately from the mean ergodic theorem. □

Because the inductive step generalizes the proof of the ordinary mean ergodic theorem, it is instructive to consider the form of that proof. The key step is proving that the function g_{l+1} can be partitioned into two components; these components are usually described as an invariant component $g_\perp$ and a component $g_\top$ in the limit of functions of the form $u - T_{l+1}u$. Unfortunately, this characterization of the second set does not generalize. There is an alternative characterization, namely that $g_\top$ has the property that $||\Delta_N g_\top||$ converges to 0. This turns out to be harder to work with (and, in particular, this characterization does not seem to give a pointwise version of the theorem), but it can be extended to a higher complexity versions.

We will argue as follows: take a function of complexity d in the form $\prod g_e$ with each g_e e-measurable, and argue that each g_e can be written in the form $g_{e,\perp} + g_{e,\top}$, where $g_{e,\top}$ is suitably random, so that $||\Delta_N g_{e,\top} \prod h_{e'}|| \to 0$, while $g_{e,\perp}$ is essentially of complexity $d-1$. If we observe that constant functions have complexity 0, the usual proof of the mean ergodic theorem has the same form.

7.3 The Inductive Step

We now return to the proof of Theorem 7.7. Let $\mathcal{X} = (\prod_{i \leq l} X_i, \mathcal{B}, \mu, T_1, \ldots, T_l)$ cleanly factored by $\mathcal{X}'$ be given, and let $\mathcal{Y}$ be an arbitrary measure space. Recall that $\mathcal{I}_n$ is the set of subsets of $[1, l+1]$ with cardinality n. If e is a subset of $[1, l+1]$, we write $\overline{e}$ for the complement of e, that is, $[1, l+1] \setminus e$. We continue to be concerned with functions belonging to $L^\infty(\mathcal{X}^*)$.

Definition 7.17. *Let $e_0 \subseteq [1, l+1]$ contain $l+1$. Z_{e_0} is the subspace of the e_0-measurable functions g such that for every sequence $\langle g_e \rangle_{e \in \mathcal{I}_{|e_0|} \setminus \{e_0\}}$ with each g_e e-measurable,*

$$||\Delta_N g \prod_e g_e|| \to 0$$

as N goes to ∞.

D_{e_0} is the set of e_0-measurable functions generated by projections onto the e_0-measurable sets of weak limit points of sequences of the form

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i \in e_0} b_{i,N}(x_{\overline{k}}, T_k^n \gamma_{x_{\overline{k}}}^{-1}(x'), x', y)$$

as N goes to infinity, for some $k \notin e_0$, where each b_i is $[1, l+1] \setminus \{i\}$-measurable.

Lemma 7.18. *If g is e_0-measurable where $l+1 \in e_0$, $|e_0| < d+1$, and $g \notin Z_{e_0}$ then there is an $h \in D_{e_0}$ such that $\int ghd\mu > 0$.*

Proof. Let an e_0-measurable $g \notin Z_{e_0}$ be given. Then there is a sequence $\langle g_e \rangle_{e \in \mathcal{I}_{|e_0|} \setminus \{e_0\}}$ where each g_e is e-measurable and some $\epsilon > 0$ such that $||\Delta_N(g \prod_{e \in \mathcal{I}_{|e_0|}, e \neq e_0} g_e)|| > \epsilon$ for infinitely many N. Set $f_N := \Delta_N(g \prod_{e \in \mathcal{I}_{|e_0|}, e \neq e_0} g_e)$. For each such N, we have

$$\int f_N \Delta_N (g \prod_{e \in \mathcal{I}_{|e_0|}, e \neq e_0} g_e) d\mu > \epsilon^2$$

This means

$$\int \frac{1}{N} \sum_{n=1}^{N} f_N(x,y) g(x, T_{l+1}^n \gamma(x), y) \prod_{e \in \mathcal{I}_{|e_0|}, e \neq e_0} g_e(x, T_{l+1}^n \gamma(x), y) d\mu > \epsilon^2$$

For each $e \neq e_0$, there is some $i \in e_0 \setminus e$, so we may assign to each g_e some i such that g_e is independent of x_i and collect the g_e into terms $b_{i,N}$ (independent on N, in fact), each a product of some of the g_e, such that b_i is independent of x_i. Since f_N is $[1,l]$-measurable, we may also fold f_N into $b_{l+1,N}$, and we have therefore shown that there exist functions $b_{i,N}$ which are $[1,l+1] \setminus \{i\}$-measurable such that

$$\int \frac{1}{N}\sum_{n=1}^{N} g(x, T_{l+1}^n \gamma(x), y) \prod_{i\in e_0} b_{i,N}(x, T_{l+1}^n\gamma(x), y) d\mu > \epsilon^2$$

Choosing some $k \notin e_0$, and letting $g'(x_{\overline{k}}, x', y) := g(x, x', y)$ for almost any x_k, this becomes

$$\int g'(x_{\overline{k}}, x', y)\frac{1}{N}\sum_{n=1}^{N}\prod_{i\in e_0} b_{i,N}(x_{\overline{k}}, T_k^n \gamma_{x_{\overline{k}}}^{-1}(x'), x', y) d\nu_k > \epsilon^2$$

for infinitely many N. Choosing a subsequence S of these N such that

$$h' := \lim_{N\in S}\frac{1}{N}\sum_{n=1}^{N}\prod_{i} b_{i,N}(x_{\overline{k}}, T_k^n \gamma_{x_{\overline{k}}}^{-1}(x'), x', y)$$

converges, the projection h of h' onto the e_0-measurable sets witnesses the lemma. (In particular, since g is e_0-measurable, $\int g h d\mu = \int g h' d\mu > 0$.)

□

Lemma 7.19. *Every e_0-measurable function g may be written in the form $g_\perp + g_\top$ where $g_\perp \in D_{e_0}$ and $g_\top \in Z_{e_0}$.*

Proof. Consider the projection of g onto D_{e_0}. By the previous lemma, if $g - E(g \mid D_{e_0})$ is not in Z_{e_0} then there is an $h \in D_{e_0}$ such that $\int h(g - E(g \mid D_{e_0}))d\mu > 0$; this is a contradiction, so $g - E(g \mid D_{e_0})$ belongs to Z_{e_0}. □

We could proceed to show that this decomposition is unique, but this is not necessary for the proof.

The following is a corollary of the nonstandard proof of Theorem 6.5 above:

Corollary 7.20. *Let $\mathcal{X} = (X, \mathcal{B}, \mu, T_1, \ldots, T_l)$ be a separable measure space and let b be a real number. For $s \leq k$, let $\mathcal{X}_s$ be a factor of $\mathcal{X}$ and $\{b_{m,m',s}\}_{m\leq m'\in\mathbb{N}}$ be a sequence of $L^\infty(\mathcal{X}_s)$ functions bounded (in the L^∞ norm) by b. Let $\{m_t\}_{t\in\mathbb{N}}$ be a sequence such that*

$$\frac{1}{m_t}\sum_{i=1}^{m_t}\prod_{s\leq k} b_{i,m_t,s}$$

converges weakly to f. Then there is a space $\mathcal{Y} = (Y, \mathcal{D}, \sigma)$ and functions $\tilde{b}_s \in L^\infty(\mathcal{X}_s \times \mathcal{Y})$ such that $f(x) = \int \prod \tilde{b}_s(x, y)d\sigma$ for almost every x and, if for some transformations $S_1, \ldots, S_k$ generated from the $T_1, \ldots, T_l$ and integers $n_1, \ldots, n_k$

$$\frac{1}{m_t}\sum_{i=1}^{m_t}\prod_{s\leq k} b_{i,m_t,s} \circ T_s^{n_s}$$

converges weakly to f' then $f' = \int \prod \tilde{b}_s(x, y) \circ T_s^{n_s} d\sigma$

Proof. All but the last condition is an immediate application of Theorem 6.5, using the technique of Corollary 6.9 to show that $\tilde{b}_s$ actually belongs to the product. The last condition must be checked separately, but follows immediately by the same transfer argument. □

The following technical lemma is another variation on this theme, combining the method of the previous corollary with the method of Theorem 7.5:

Lemma 7.21. *Let $\mathbb{Y}$ be a separable measure spaces and let b be real number. Let k, k' be given. Suppose that for every $w \in \mathbb{Z}^l$, $w' \in \mathbb{Z}$, there are functions $\{b_{w,w',s,s'}\}_{s\leq k, s'\leq k'}$.*

Then there is a measure space $\mathbb{W} = (\prod_{i\leq l} W_i, \mathcal{B}, \nu, T_1, \ldots, T_l)$, and space $\mathbb{W}'$ cleanly factoring $\mathbb{W}$ (witnessed by $\delta : \mathbb{W} \to \mathbb{W}'$), and an extension $\mathbb{W}^*$ of $\mathbb{W} \times \mathbb{W}'$, together with functions $\tilde{b}_{s,s'} \in \mathbb{W}^* \times \mathbb{Y}$ such that for any N, N',

$$||\Delta_N \prod_{s\leq k} \int \prod_{s'\leq k'} \tilde{b}_{s,s'} dy - \Delta_{N'} \prod_{s\leq k} \int \prod_{s'\leq k'} \tilde{b}_{s,s'} dy||_{L^2(\mathbb{W})} =$$

$$\lim_{M\to\infty} \frac{1}{M^l} \sum_{w\in M^l} \left[\frac{1}{N} \sum_{n=1}^{N} \prod_{s\leq k} \int \prod_{s'\leq k'} b_{w,\sum_i w_i+n,s,s'} dy - \frac{1}{N'} \sum_{n=1}^{N'} \prod_{s\leq k} \int \prod_{s'\leq k'} b_{w,\sum_i w_i+n,s,s'} dy \right]^2$$

and if, for some s, s', $b_{w,w',s,s'}$ is not dependent on w_i then $\tilde{b}_{s,s'}$ does not depend on W_i.

Proof. The proof is essentially the combination of the method of Theorem 7.5 with the previous corollary. We take a nonstandard extension of a universe containing $\mathbb{Y}$ and $\{b_{w,w',s,'}\}_{w,w',s,s'}$ and pick some nonstandard a. Then $\mathbb{W}^*$ is the Loeb measure on $[1, a]^{l+1}$, taking $\delta : \mathbb{W} \to \mathbb{W}'$ to be the map $w \mapsto \sum_i w_i$.

Then set $\tilde{b}_{s,s'}(w, w', y) := st \circ b^*_{w,w',s,s'}(y)$. Measurability with respect to $\mathbb{W}^* \times \mathbb{Y}$ follows from the argument given in Corllary 6.9. The particular equality required (and a much wider class of similar equalities) follow directly from transfer.

For the final condition, if for every $w_{\bar{i}}, w'$ and every w_i, w'_i, $b_{w_{\bar{i}},w_i,w',s,s'} = b_{w_{\bar{i}},w'_i,w',s,s'}$ then the same holds for b^*, so $\tilde{b}_{s,s'}$ is independent of the i-th coordinate. □

Lemma 7.22. *If $g = \prod_{e\in\mathcal{I}_{d+1}} g_e$ and each $g_e \in D_e$ then $\Delta_N g$ converges in the L^2 norm.*

Proof. For convenience, assume g is in the stricter form $\prod_{e\in\mathcal{I}_{d+1},l+1\in e} g_e$. This is without loss of generality, since if $h = \prod_{e\in\mathcal{I}_{d+1},l+1\notin e} g_e$ then we have

$$\Delta_N h \prod_{e\in\mathcal{I}_{d+1},l+1\in e} g_e = h\Delta_N \prod_{e\in\mathcal{I}_{d+1},l+1\in e} g_e$$

First, assume each g_e is a basic element of D_e; that is, there is a function g'_e such that g_e is the projection of g'_e onto $\mathcal{B}_{\overline{e_0}}$ and g'_e is a projection weak limit of an average of the form

$$\frac{1}{N} \sum_{n=1}^{N} \prod_i b^e_{i,N}(x_{\overline{k}}, T^n_k \gamma^{-1}_{x_{\overline{k}}}(x'), x', y)$$

Applying the corollary above, there exists a measure space $\mathbb{Z}_e$ and functions $\tilde{b}^e_i$ such that

$$g'_e(x_e, x', y) = \iint \prod_i \tilde{b}^e_i(x_{\overline{k}}, z, x', y) d\sigma_e d\mu_{\overline{e}}$$

Furthermore, for any i',

$$g'_e(T_{i'} x_{\overline{k}}, x', y) = \iint \prod_i \tilde{b}^e_i(T_{i'} x_{\overline{k}}, z, x', y) d\sigma_e d\mu_{\overline{e}}$$

but also $g'_e \circ T_{i'}$ is a projection of a weak limit of the average

$$\frac{1}{N} \sum_{n=1}^{N} \prod_i b^e_{i,N}(T_{i'} x_{\overline{k}}, T^{n+1}_k \gamma^{-1}_{x_{\overline{k}}}(x'), x', y)$$

But the contribution of the shift to the second parameter is negligible, and when $i = i'$, $b^e_{i,N}$ is $T_{i'}$-invariant, so $g'_e \circ T_{i'}$ is a projection of a weak limit of the average

$$\frac{1}{N} \sum_{n=1}^{N} b^e_{i,N}(x_{\overline{k}}, T^n_k \gamma^{-1}_{x_{\overline{k}}}(x'), y) \prod_{i\neq i} b^e_{i,N}(T_{i'} x_{\overline{k}}, T^n_k \gamma^{-1}_{x_{\overline{k}}}(x'), x', y)$$

But this means that

$$g'_e(T_{i'}x_e, x', y) = \int \tilde{b}^e_i(x_{\overline{k}}, z, x', y) \prod_{i \neq i'} \tilde{b}^e_i(T_{i'}x_{\overline{k}}, z, x', y) d\sigma_e d\mu_{\overline{e}}$$

Define

$$g_{w,w'}(x, x', y) := g(T^w x, T^{w'}_{l+1} x', y)$$

Then also

$$g_{w,w'}(x, x', y) = \prod_e \iint \prod_i \tilde{b}^e_i(T^w x_{\overline{k}}, x_{\overline{k}}, z, T^{w'}_{l+1} x', y) d\sigma_e dx_{\overline{e}}$$

But by the argument above, each b^e_i depends only on those w_j where j is in $e \setminus \{i\}$.

Applying the pointwise ergodic theorem, we may choose x generic for rational polynomial combinations of shifts of $g(x, \gamma(x), y)$. Then it suffices to show that for each ϵ, large enough N, and $N' > N$,

$$\lim_{M\to\infty} \frac{1}{M^l} \sum_{w \in M^l} \left[\frac{1}{N} \sum_{n=1}^{N} g_{w,w'+n}(x, \gamma(x), y) - \frac{1}{N} \sum_{n=1}^{N} g_{w,w'+n}(x, \gamma(x), y) \right]^2 < \epsilon$$

Applying Lemma 7.21, this becomes the claim that

$$||\Delta_N \tilde{g} - \Delta_{N'} \tilde{g}|| < \epsilon$$

where $\tilde{g}$ is a function on some $\mathbb{W}^* \times \mathbb{Y}$. (Indeed, Lemma 7.21 gives functions on a bigger space $\mathbb{W}^* \times \mathbb{Y} \times \prod_e \mathbb{Z}_e$, where $\tilde{g}$ is a product of an integral of a product of these functions.) It therefore suffices to show convergence for $\tilde{g}$.

But $\tilde{g}$ has the form

$$\prod_e \iint \prod_i \hat{b}^e_i(w, x_e, x_{\overline{e}}, z, \delta(w'), \gamma(x_e, x_{\overline{e}}), y) d\sigma_e d\mu_{\overline{e}}$$

where $\hat{b}^e_i$ depends only on $w_{e\setminus\{i\}}$. It follows that $\tilde{g}$ has complexity $d-1$.

If the g_e are sums of basic elements of D_e, the result follows immediately. If g_e is a limit of such elements, each g_e can be written $g^0_e + g^1_e$ where g^e_0 is a finite sum of basic elements of D_e and the norm of g^1_e is small. Then $\prod g_e = \sum_{E \subseteq \mathcal{I}_d} \prod_{e\in E} g^e_0 \prod_{e \notin E} g^e_1$. When $E = \mathcal{I}_d$, the result follows from the result for finite sums. When $E \neq \mathcal{I}_d$, the product contains some g^1_e, and sine g^1_e is e-measurable, it follows that $||\Delta_N g_e|| \leq ||g_e||$. Since the $g_{e'}$ are bounded in the L^∞ norm, $||\Delta_N \prod_e g_e|| \leq b \prod_e ||g_e||$ for some constant b, so $\prod_{e\in E} g^e_0 \prod_{e\notin E} g^e_1$ has small norm if $E \neq \mathcal{I}_d$. □

Using this, it is possible to prove Theorem 7.7. If $g = \prod_{e \in \mathcal{I}_{d+1}} g_e(x, x', y)$ where each g_e is e-measurable then it suffices to show convergence at each y, since then the dominated convergence theorem implies convergence over the whole space. When $l+1 \notin e$, we have $\Delta_N g_e f = g_e \Delta_N f$, so it suffices to show that $\Delta_N g$ converges where g has the form

$$\prod_{e \in \mathcal{I}_{d+1}, l+1 \in e} g_e$$

Then write each g_e as $g_{e,\perp} + g_{e,\top}$. Expanding the product gives

$$\sum_{E \subseteq \{e \in \mathcal{I}_{d+1} | l+1 \in e\}} \prod_{e \notin E} g_{e,\perp} \prod_{e\in E} g_{e,\top}$$

where each $g_{e,\top}$ is in Z_e and each $g_{e,\perp}$ is in D_e. Since Δ_N distributes over sums, it suffices to show that each summand converges. When E is non-empty, $\Delta_N \prod_{e\notin E} g_{e,\perp} \prod_{e\in E} g_{e,\top}$ converges to the 0 function by the definition of Z_e. When E is empty, Lemma 7.22 applies.

Bibliography

[1] Peter Aczel. The type theoretic interpretation of constructive set theory: inductive definitions. In *Logic, methodology and philosophy of science, VII (Salzburg, 1983)*, volume 114 of *Stud. Logic Found. Math.*, pages 17–49. North-Holland, Amsterdam, 1986.

[2] Klaug Aehlig. *On fragments of analysis with strengths of finitely iterated inductive definitions*. PhD thesis, University of Munich, 2003.

[3] Toshiyasu Arai. Proof theory for theories of ordinals. Preprints, 1996-1997.

[4] Toshiyasu Arai. Proof theory for theories of ordinals. I. Recursively Mahlo ordinals. *Ann. Pure Appl. Logic*, 122(1-3):1–85, 2003.

[5] Toshiyasu Arai. Proof theory for theories of ordinals. II. Π_3-reflection. *Ann. Pure Appl. Logic*, 129 (1-3):39–92, 2004.

[6] Jeremy Avigad. A realizability interpretation for classical arithmetic. In Buss, Hájek, and Pudlák, editors, *Logic Colloquium '98*, number 13 in Lecture Notes in Logic, pages 57–90. AK Peters, 2000.

[7] Jeremy Avigad. Interpreting classical theories in constructive ones. *J. Symbolic Logic*, 65(4):1785–1812, 2000.

[8] Jeremy Avigad. Predicative functionals and an interpretation of $\widehat{\mathrm{ID}}_{<\omega}$. *Ann. Pure Appl. Logic*, 92(1): 1–34, 1998.

[9] Jeremy Avigad. A variant of the double-negation translation. CMU-PHIL 179, Carnegie Mellon University, 2006.

[10] Jeremy Avigad and Solomon Feferman. Gödel's functional ("Dialectica") interpretation. In *Handbook of proof theory*, volume 137 of *Stud. Logic Found. Math.*, pages 337–405. North-Holland, Amsterdam, 1998.

[11] Jeremy Avigad and Ksenija Simic. Fundamental notions of analysis in subsystems of second-order arithmetic. *Ann. Pure Appl. Logic*, 139(1-3):138–184, 2006.

[12] Jeremy Avigad and Henry Towsner. Functional interpretation and inductive definitions. http://arxiv.org/abs/0802.1938.

[13] Jeremy Avigad, Philipp Gerhardy, and Henry Towsner. Local stability of ergodic averages. http://www.arXiv.org/abs/0706.1512, 2007.

[14] Vitaly Bergelson. Ergodic theory and Diophantine problems. In *Topics in symbolic dynamics and applications (Temuco, 1997)*, volume 279 of *London Math. Soc. Lecture Note Ser.*, pages 167–205. Cambridge Univ. Press, Cambridge, 2000.

[15] Errett Bishop. An upcrossing inequality with applications. *Michigan Math. J.*, 13:1–13, 1966. ISSN 0026-2285.

[16] Errett Bishop. *Foundations of constructive analysis*. McGraw-Hill Book Co., New York, 1967.

[17] Errett Bishop. A constructive ergodic theorem. *J. Math. Mech.*, 17:631–639, 1967/1968.

[18] Errett Bishop and Douglas Bridges. *Constructive Mathematics*. Springer, Berlin, 1985.

[19] Wilfried Buchholz. Relating ordinals to proofs in a perspicuous way. In *Reflections on the foundations*

of mathematics (Stanford, CA, 1998), volume 15 of *Lect. Notes Log.*, pages 37–59. Assoc. Symbol. Logic, Urbana, IL, 2002.

[20] Wilfried Buchholz. Eine Erweiterung der Schnitteliminationsmethode. Habilitationsschrift, 1977.

[21] Wilfried Buchholz. The $\Omega_{\mu+1}$-rule. In *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies* Buchholz et al. [23], pages v+383.

[22] Wilfried Buchholz. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, chapter Ordinal analysis of ID_ν. Springer-Verlag, 1981.

[23] Wilfried Buchholz, Solomon Feferman, Wolfram Pohlers, and Wilfried Sieg. *Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies*, volume 897 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1981.

[24] Wolfgang Burr. A Diller-Nahm-style functional interpretation of KPω. *Arch. Math. Logic*, 39(8): 599–604, 2000.

[25] Samuel R. Buss, editor. *Handbook of proof theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1998.

[26] Jean-Pierre Conze and Emmanuel Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France*, 112(2):143–175, 1984.

[27] Justus Diller and Werner Nahm. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik endlicher Typen. *Arch. Math. Logik Grundlagenforsch.*, 16:49–66, 1974. ISSN 0003-9268.

[28] A.G. Dragalin. New forms of realizability and Markov's rule (Russian). *Doklady*, 251:534–537, 1980. Translation *SM* 21, pp. 461–464.

[29] Gabor Elek and Balazs Szegedy. Limits of hypergraphs, removal and regularity lemmas. a non-standard approach, 2007. URL `http://arxiv.org/abs/0705.2179`.

[30] Solomon Feferman. Ordinals associated with theories for one inductively defined set. unpublished, 1968.

[31] Fernando Ferreira and Paulo Oliva. Bounded functional interpretation. *Ann. Pure Appl. Logic*, 135 (1-3):73–112, 2005.

[32] Nikos Frantzikinakis and Bryna Kra. Convergence of multiple ergodic averages for some commuting transformations. *Ergodic Theory Dynam. Systems*, 25(3):799–809, 2005.

[33] Harvey M. Friedman. *Higher Set Theory*, volume 669 of *Lecture Notes in Mathematics*, chapter Classically and intuitionistically provable functions, pages 21–27. Springer, Berlin, 1978.

[34] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Analyse Math.*, 31:204–256, 1977.

[35] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. Princeton University Press, Princeton, N.J., 1981. ISBN 0-691-08269-3. M. B. Porter Lectures.

[36] H. Furstenberg, Y. Katznelson, and D. Ornstein. The ergodic theoretical proof of Szemerédi's theorem. *Bull. Amer. Math. Soc. (N.S.)*, 7(3):527–552, 1982.

[37] Gerhard Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Math. Ann.*, 112(1):493–565, 1936.

[38] Philipp Gerhardy and Ulrich Kohlenbach. General logical metatheorems for functional analysis. To appear in *Trans. Am. Math. Soc.*, 2005.

[39] Jean-Yves Girard. Une extension de l'interprétation de Gödel à l'analyse, et son application à l'élimination des coupures dans l'analyse et la théorie des types. In *Proceedings of the Second Scandinavian Logic Symposium (Univ. Oslo, Oslo, 1970)*, pages 63–92. Studies in Logic and the Foundations of Mathematics, Vol. 63, Amsterdam, 1971. North-Holland.

[40] Eli Glasner. *Ergodic theory via joinings*, volume 101 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003. ISBN 0-8218-3372-3.

[41] Kurt Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.

[42] Robert Goldblatt. *Lectures on the hyperreals*, volume 188 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998. ISBN 0-387-98464-X. An introduction to nonstandard analysis.

[43] Bernard Host and Bryna Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)*, 161(1):397–488, 2005.

[44] W. A. Howard. Functional interpretation of bar induction by bar recursion. *Compositio Math.*, 20: 107–124 (1968), 1968.

[45] W. A. Howard. A system of abstract constructive ordinals. *J. Symbolic Logic*, 37:355–374, 1972. ISSN 0022-4812.

[46] Gerhard Jäger. *Theories for admissible sets: a unifying approach to proof theory*, volume 2 of *Studies in Proof Theory. Lecture Notes*. Bibliopolis, Naples, 1986.

[47] A. G. Kachurovskiĭ. Rates of convergence in ergodic theorems. *Uspekhi Mat. Nauk*, 51(4(310)):73–124, 1996. ISSN 0042-1316.

[48] Yitzhak Katznelson. *An introduction to harmonic analysis*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2004. ISBN 0-521-83829-0; 0-521-54359-2.

[49] H. Jerome Keisler. Hyperfinite model theory. In *Logic Colloquium 76 (Oxford, 1976)*, pages 5–110. Studies in Logic and Found. Math., Vol. 87. North-Holland, Amsterdam, 1977.

[50] H. Jerome Keisler. An infinitesimal approach to stochastic analysis. *Mem. Amer. Math. Soc.*, 48 (297):x+184, 1984.

[51] S. C. Kleene. Arithmetical predicates and function quantifiers. *Trans. Amer. Math. Soc.*, 79:312–340, 1955.

[52] S. C. Kleene. Hierarchies of number-theoretic predicates. *Bull. Amer. Math. Soc.*, 61:193–213, 1955.

[53] Ulirch Kohlenbach and Paulo Oliva. Proof mining: a systematic way of analyzing proofs in mathematics. *Tr. Mat. Inst. Steklova*, 242(Mat. Logika i Algebra):147–175, 2003.

[54] Ulrich Kohlenbach. Some logical metatheorems with applications in functional analysis. *Trans. Amer. Math. Soc.*, 357(1):89–128 (electronic), 2005. ISSN 0002-9947.

[55] Ulrich Kohlenbach. Elimination of Skolem functions for monotone formulas in analysis. *Arch. Math. Logic*, 37(5-6):363–390, 1998. ISSN 0933-5846. Logic Colloquium '95 (Haifa).

[56] Ulrich Kohlenbach. Applied proof theory: Proof interpretations and their use in mathematics. to appear, .

[57] Ulrich Kohlenbach. Applied proof theory: Proof interpretations and their use in mathematics. Springer Monographs in Mathematics, to appear, .

[58] Ulrich Kohlenbach. Effective bounds from proofs in abstract functional analysis. To appear in B. Cooper et al., eds., *CiE 2005 New Computational Paradigms: Changing Conceptions of What is Computable.*, Springer, Berlin, .

[59] Georg Kreisel. On weak completeness of intuitionistic predicate logic. *J. Symbolic Logic*, 27:139–158, 1962.

[60] Georg Kreisel. On the interpretation of non-finitist proofs. I. *J. Symbolic Logic*, 16:241–267, 1951. ISSN 0022-4812.

[61] Georg Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In *Constructivity in mathematics: Proceedings of the colloquium held at Amsterdam, 1957 (edited by A. Heyting)*, Studies in Logic and the Foundations of Mathematics, pages 101–128, Amsterdam, 1959. North-Holland Publishing Co.

[62] Georg Kreisel. Generalized inductive definitions. Technical report, Stanford University, 1963. Stanford Report on the Foundations of Analysis (mimeographed).

[63] Ulrich Krengel. On the speed of convergence in the ergodic theorem. *Monatsh. Math.*, 86(1):3–6, 1978/79. ISSN 0026-9255.

[64] Ulrich Krengel. *Ergodic theorems*, volume 6 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1985. ISBN 3-11-008478-3. With a supplement by Antoine Brunel.

[65] E. Lesigne. Équations fonctionnelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France*, 121(3):315–351, 1993.

[66] Elon Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[67] László Lovász and Balázs Szegedy. Limits of dense graph sequences. *J. Combin. Theory Ser. B*, 96 (6):933–957, 2006. ISSN 0095-8956.

[68] Michael Möllerfeld. *Generalized Inductive Definitions*. PhD thesis, Westfälische Wilhelms-Universität Münster, 2003.

[69] M. G. Nadkarni. *Spectral theory of dynamical systems*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Verlag, Basel, 1998. ISBN 3-7643-5817-3.

[70] J. A. Nuber. A constructive ergodic theorem. *Trans. Amer. Math. Soc.*, 164:115–137, 1972. ISSN 0002-9947.

[71] Wolfram Pohlers. Cut elimination for impredicative infinitary systems. I. Ordinal analysis for ID_1. *Arch. Math. Logik Grundlag.*, 21(3-4):113–129, 1981.

[72] Wolfram Pohlers. An upper bound for the provability of transfinite induction in systems with N-times iterated inductive definitions. In ⊨ *ISILC Proof Theory Symposion (Proc. Internat. Summer Inst. and Logic Colloq., Kiel, 1974)*, pages 271–289. Lecture Notes in Math., Vol. 500. Springer, Berlin, 1975.

[73] Wolfram Pohlers. Proof theoretical analysis of ID_ν by the method of local predicativity. In *Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies* Buchholz et al. [23], pages v+383.

[74] Marian B. Pour-El and J. Ian Richards. *Computability in analysis and physics*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1989. ISBN 3-540-50035-9.

[75] Michael Rathjen. Recent advances in ordinal analysis: Π^1_2-CA and related systems. *Bull. Symbolic Logic*, 1(4):468–485, 1995.

[76] Kurt Schütte. Beweistheoretische Erfassung der unerdlichen Induktion in der Zahlentheorie. *Math. Ann.*, 122:369–389, 1951.

[77] Helmut Schwichtenberg. *Handbook of Mathematical Logic*, chapter Proof theory: Some aspects of cut-elimination, pages 867–895. North-Holland Publishing Co., Amsterdam, 1977.

[78] Anton Setzer. Well-ordering proofs for Martin-Löf type theory with W-type and one universe. *Annals of Pure and applied Logic*, 92:113–159, 1998.

[79] Joseph R. Shoenfield. *Mathematical logic*. Association for Symbolic Logic, Urbana, IL, 2001. Reprint of the 1973 second printing.

[80] Wilfried Sieg. *Trees in metamathematics: theories of inductive definitions and subsystems of analysis*. PhD thesis, Stanford University, 1977.

[81] Wilfried Sieg. *Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies*, chapter Inductive definitions, constructive ordinals, and normal derivations, pages 143–187. Springer-Verlag, 1981.

[82] Ksenija Simic. The pointwise ergodic theorem in subsystems of second-order arithmetic. *J. Symbolic Logic*, 72(1):45–66, 2007. ISSN 0022-4812.

[83] Stephen G. Simpson. *Subsystems of second order arithmetic*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1999. ISBN 3-540-64882-8.

[84] Clifford Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an

extension of principles formulated in current intuitionistic mathematics. In *Proc. Sympos. Pure Math., Vol. V*, pages 1–27. American Mathematical Society, Providence, R.I., 1962.

[85] Bas Spitters. A constructive view on ergodic theorems. *J. Symbolic Logic*, 71(2):611–623, 2006. ISSN 0022-4812.

[86] Thomas Streicher and Ulrich Kohlenbach. Shoenfield is Gödel after Krivine. *MLQ Math. Log. Q.*, 53 (2):176–179, 2007.

[87] W. W. Tait. Applications of the cut elimination theorem to some subsystems of classical analysis. In *Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968)*, pages 475–488. North-Holland, Amsterdam, 1970.

[88] Terence Tao. A correspondence principle between (hyper)graph theory and probability theory, and the (hyper)graph removal lemma. http://arxiv.org/abs/math/0602037, 2006.

[89] Terence Tao. Norm convergence of multiple ergodic averages for commuting transformations, 2007. URL `http://arxiv.org/abs/0707.1117v2`.

[90] Henry Towsner. A realizability interpretation for classical analysis. *Arch. Math. Logic*, 43(7):891–900, 2004.

[91] A. S. Troelstra. *Collected works. Vol. II*, chapter Introductory note to 1958 and 1972. The Clarendon Press Oxford University Press, New York, 1990.

[92] A.S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, 1973. Springer. With contributions by A.S. Troelstra, C.A. Smoryński, J.I. Zucker and W.A. Howard.

[93] Klaus Weihrauch. *Computable analysis*. Texts in Theoretical Computer Science. An EATCS Series. Springer-Verlag, Berlin, 2000. ISBN 3-540-66817-9. An introduction.

[94] Qing Zhang. On convergence of the averages $(1/N)\sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$. *Monatsh. Math.*, 122(3):275–300, 1996.

[95] Tamar Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20 (1):53–97 (electronic), 2007. ISSN 0894-0347.

[96] J. I. Zucker. Iterated inductive definitions, trees, and ordinals. In *Metamathematical investigation of intuitionistic arithmetic and analysis*, pages 392–453. Lecture Notes in Mathematics, Vol. 344. Springer, Berlin, 1973.