Computability, Probability and Logic

Computability, Probability and Logic

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 10 juni 2015
om 14:30 uur precies

door

Rutger Kuyper

geboren op 10 februari 1989
te Haarlem

Promotoren:

Prof. dr. M. Gehrke	Université Paris Diderot – Paris 7, Frankrijk
Prof. dr. P. Stevenhagen	Universiteit Leiden

Copromotor:

Dr. S.A. Terwijn

Manuscriptcommissie:

Prof. dr. B.J.J. Moonen	
Prof. dr. R.G. Downey	Victoria University of Wellington, Nieuw-Zeeland
Prof. dr. A. Sorbi	Università degli Studi di Siena, Italië

The research in this thesis was supported by NWO/DIAMANT grant 613.009.011 and by John Templeton Foundation grant 15619: 'Mind, Mechanism and Mathematics: Turing Centenary Research Project'.

ISBN: 978-94-6259-611-5

Contents

Acknowledgements

During the past three years I have had the pleasure of meeting many people from all over the world. Even across cultural differences, they all still had a lot in common: they are all incredibly friendly, kind and cheerful people who are very welcoming of a newcomer to their field. I would like to use this opportunity to thank all of them, apologising in advance to those who are not explicitly mentioned.

Let me begin by thanking DIAMANT and NWO, who funded a large part of my Ph.D. project. I also want to personally thank Peter Stevenhagen, who appointed me as a Ph.D. candidate at the University of Leiden when things fell through in Nijmegen. I am also very grateful to the John Templeton Foundation and the judges of the Turing Centenary Research Scholarship competition, for awarding me an incredibly generous scholarship which allowed me to travel all around the world. Being able to talk with so many other computability theorists has been immensely inspiring.

I especially want to thank Rod Downey, Noam Greenberg and Adam Day for hosting me at the Victoria University of Wellington for two months. Wellington is rightfully called one of the centres of computability theory and I had a great time there. The hospitality you offered and the great wines you served still vividly live in my memories.

There are many people who have either hosted me for a short visit, or whom I had inspiring and enjoyable conversations with at the various conferences I attended. In particular I would like to thank Laurent Bienvenu, Mingzhong Cai, Santiago Figueira, Melvin Fitting, Alex Galicki, Jeroen Goudsmit, Kojiro Higuchi, Rupert Hölzl, Iskander Kalimullin, Takayuki Kihara, Steffen Lempp, Alexander Melnikov, Joe Miller, Russell Miller, Kenshi Miyabe, Benoit Monin, Antonio Montalbán, André Nies, Paul Shafer, Andrea Sorbi, Frank Stephan, Jeroen Van der Meeren and Greg Yang. I would also like to extend my gratitude to my coauthors Denis Hirschfeldt, Carl Jockusch, Jr. and Paul Schupp. I am also very grateful to my fellow Ph.D. candidates in Nijmegen and all my colleagues in the Turing Centenary Research Project.

Finally, let me thank my promotors Mai Gehrke and Peter Stevenhagen for their support, and most importantly of all, let me thank my thesis advisor Sebastiaan Terwijn. Bas, thank you for always making time for me and for being a great help through these three years.

CHAPTER 1

Introduction

In this thesis we discuss some topics that spring from the interactions between the fields of computability, probability and logic. Before we discuss these interactions, let us first put these fields in their respective historical contexts.

Logic in a broad sense goes back a very long time, all the way to the ancient Greeks; the first person to thoroughly study logic was Aristotle. However, modern mathematical logic, which is the kind of logic referred to in the title of this thesis, starts with Frege and Russell in the late nineteenth century. The end of the nineteenth century and the early twentieth century were characterised by the rapid development of the field. That logic had become an important part of mathematics is clear from Hilbert's influential list of twenty-three mathematical problems presented in 1900: three of the problems on this list are in mathematical logic, and quite strikingly, the first two problems on the list are.

In particular, we should mention Hilbert's second problem: to prove the consistency of arithmetic. A major result in mathematical logic is Gödel's celebrated incompleteness theorem from 1931, which shows that such a proof cannot be carried out within arithmetic itself. A closely related problem is the *Entscheidungsproblem* posed by Hilbert and Ackermann in 1928, asking if there is an algorithm which decides the validity of formulas in the mathematical framework of first-order logic. This question was answered in the negative as well, independently by Church and Turing, in 1936. To give a negative answer, they first had to develop a satisfactory formal definition of what an algorithm is. Although their definitions turned out to be equivalent, it is Turing's definition which more evidently captures the informal notion of an algorithm, as admitted by Church himself [19]. The study of computation born from these logical investigations eventually turned into the flourishing field now called computability theory (*née* recursion theory), which is the computability referred to in the title.

Likewise, early probability theory goes back a long time, to the seventeenth century, but its modern foundations were laid by Kolmogorov in 1933. Kolmogorov in fact worked in all of computability, probability and logic: not only did he work in probability theory, as just described, but he also worked on topics such as Kolmogorov complexity (which is related to computability theory and connects to probability through the Coding Theorem) and on semantics for intuitionistic logic (in a way that would eventually lead to the Medvedev and Muchnik lattices, which lie in the intersection of computability and logic).

In this thesis we investigate three topics which are a combination of computability, probability and logic. More precisely, part I concerns the combination of computability and logic, in the form of the Medvedev and Muchnik lattices; part II explores algorithmic randomness and genericity, two topics which arise from the combination of computability and probability; and part III concerns the combination of probability and logic, in the form of a probability logic called ε-logic.

This thesis contains a collection of eight papers the author has (co-)written during his time as a Ph.D. candidate. Two of these are co-authored with Terwijn (work from these papers appears in chapters 6, 8 and 9), and one of these papers was written together with Hirschfeldt, Jockusch and Schupp (the work from this paper appears in chapter 7). Except for some new cross-references between these eight papers, the content remains largely unaltered from their published versions. The original source is cited in the beginning of each chapter.

1.1. Computability and logic: the Medvedev and Muchnik lattices

There have been several attempts to connect computation and logic, which all start from the Brouwer–Heyting–Kolmogorov interpretation for intuitionistic logic. Informally, Brouwer's intuitionistic logic is two-valued classical logic, except that proofs by contradiction are not allowed; a rigorous formalisation of it was given by Heyting. The Brouwer–Heyting–Kolmogorov interpretation gives an informal relation between proofs in this logic on one hand and constructions on the other hand. Probably the most well-known attempt to formalise this is Kleene's realisability [53], which is one way to connect computation with intuitionistic logic. Unfortunately, this original concept of Kleene realisability captures a proper extension of intuitionistic propositional logic (IPC), but various modifications to this concept have been studied (see e.g. the recent book by van Oosten [89]). There are various other approaches to formalising the Brouwer–Heyting–Kolmogorov interpretation, which we can not all name here, but let us mention the Curry–Howard isomorphism [20, 43] and Artemov's Logic of Proofs [2].

In this thesis we study an approach which is based on Kolmogorov's contribution to the Brouwer–Heyting–Kolmogorov interpretation. His idea was that proving a statement in intuitionistic logic is like solving a *problem*, where the exact definition of a problem is left informal, but should be seen as the kind of problem a mathematician would normally study. For example, a problem could be "proof the Riemann hypothesis" and to solve it would be to write a proof. However, Kolmogorov's notion of a problem is more general; for example, he also considers "solve the equation f(x) = 0" and "draw a circle through three given points" to be problems. There are several natural operations on problems: for example, there is the operation "and", where to solve a problem A "and" B means to solve both A and B simultaneously, and there is an operation "implies", where to solve A "implies" B means to solve B given an arbitrary solution of A. Kolmogorov connected these operations to the logical connectives and argued that proving in intuitionistic logic corresponds to solving problems.

This informal idea was formalised by Medvedev [79], using concepts from computability theory. He interprets a problem as a set of functions from the natural numbers to itself; a solution to this problem is then a computable element of this set. There is also a natural notion of reducing problems to each other: a problem A should reduce to a problem B if the solutions of B *uniformly* compute a solution of A. Muchnik [84] introduced a variant of this, by dropping the uniformity requirement. These approaches, which came to be known as the *Medvedev* and *Muchnik lattices*, can be used to talk about intuitionistic logic because they are so-called *Brouwer algebras*, algebraic structures which have a naturally associated propositional theory that lies between IPC and classical logic. Unfortunately, the theory associated to the Medvedev and Muchnik lattices turned out not to be IPC, because they both satisfy the weak law of the excluded middle $\neg p \vee \neg\neg p$.

While this might make it seem like we cannot use the Medvedev and Muchnik lattices to capture IPC, it turns out that there is in fact a way to do so. As already suggested in Rogers [98], Skvortsova [106] studied principal factors of the Medvedev lattice and looked at their theory. Quite remarkably, she showed that there is a principal factor of the Medvedev lattice which does capture exactly IPC. Sorbi and Terwijn [113] later showed that the same holds for the Muchnik lattice. While her result is certainly extraordinary, one could possibly object to the fact that the principal factors constructed in these proofs are constructed in an ad hoc manner and do not naturally appear in mathematics. In particular they do not have a clear computational interpretation, hence they do not yield the precise connection between IPC and computability we were looking for. This leads to the question, posed in Terwijn [118]: are there any natural principal factors which have IPC as their theory?

In chapter 3 we present an answer in the case of the Muchnik lattice. That is, we present several natural factors of the Muchnik lattice which capture IPC, which can be defined using well-known notions from computability theory. More precisely, we define these factors using concepts from computability theory such as lowness, 1-genericity below $\emptyset'$, hyperimmune-freeness and computable traceability, but since our framework is general, our results could be adapted to suit other concepts as well. Thus, not only do we obtain a purely computational semantics for IPC, but these results also put the above-mentioned computability-theoretic concepts in a new light.

Next, in chapter 4 we study if there are any natural factors of the Medvedev lattice which capture IPC. We present progress towards a positive answer, by showing that there are principal factors of the Medvedev lattice capturing IPC which have a stronger connection to computability theory than the one given by Skvortsova. These factors arise from the computability-theoretic notion of a *computably independent set*; that is, a set A such that for every natural number i we have that the i^{th} column of A is not computed by the other columns. The existence of computably independent sets was first proven by Kleene and Post [55]. In fact, almost all sets are computably independent: both in the measure-theoretic sense, because every 1-random is computably independent by van Lambalgen's theorem [71], and also in the Baire category sense, because every 1-generic is computably independent by the genericity analogue of van Lambalgen's theorem (see e.g. [25, Theorem 8.20.1]).

In the same chapter, we also study a question posed by Sorbi and Terwijn in [112]. Sorbi [110] showed that the theory of the Medvedev lattice is equal to Jankov's logic Jan, the deductive closure of IPC plus the weak law of the excluded middle $\neg p \vee \neg\neg p$. Sorbi and Terwijn were able to show that for many principal factors of the Medvedev lattice its theory is contained in Jankov's logic, which lead them to ask if this is the case for all non-trivial principal factors of the Medvedev lattice. More evidence was given by Shafer [104], who widened the class of principal factors for which the theory is contained in Jankov's logic. We show that the answer is indeed positive.

Finally, in chapter 5 we study an extension of the Medvedev lattice to first-order logic. Just like Medvedev and Muchnik, we try to keep in spirit with Kolmogorov's ideas from [58]. Even though Medvedev and Muchnik did not discuss first-order logic, Kolmogorov did briefly discuss the universal quantifier in his paper:

> "Im allgemeinen bedeutet, wenn x eine Variable (von beliebiger Art) ist und $a(x)$ eine aufgabe bezeichnet, deren Sinn von dem Werte von x abhängt, $(x)a(x)$ die Aufgabe "eine allgemeine Methode für die Lösung von $a(x)$ bei jedem einzelnen Wert von x anzugeben". Man soll dies so verstehen: Die aufgabe $(x)a(x)$ zu lösen, bedeutet, imstande sein, für jeden gegebenen Einzelwert x_0 von x die Aufgabe $a(x_0)$ nach einer endlichen Reihe von im voraus (schon vor der Wahl von x_0) bekannten Schritten zu lösen."[1]

It is important to note that, when Kolmogorov says that the steps should be fixed before x_0 is set, he probably does not mean that we should have one solution that works for every x_0; instead, the solution is allowed to depend on x_0, but it should do so uniformly. This belief is supported by one of the informal examples of a problem he gives: "given one solution of $ax^2 + bx + c = 0$, give the other solution". Of course there is no procedure to transform one solution to the other one which does not depend on the parameters a, b and c; however, there is one which does so uniformly. More evidence can be found in Kolmogorov's discussion of the law of the excluded middle, where he says that a solution of the problem $\forall a(a \vee \neg a)$, where a quantifies over all problems, should be "a general method which for any problem a allows one either to find its solution or to derive a contradiction from the existence of such a solution" and that "unless the reader considers himself omniscient, he will perhaps agree that [this formula] cannot be in the list of problems that he has solved". In other words, a solution of $\forall a(a \vee \neg a)$ should be a solution of $a \vee \neg a$ for every problem a which is allowed to depend on a, and it should be uniform because we are not omniscient.

In chapter 5 we formalise this idea in the spirit of Medvedev, using the notion of a *first-order hyperdoctrine* from categorical logic, which naturally extends the notion of a Brouwer algebra. We study its theory and look at intervals to try and obtain analogous results to Skvortsova's result mentioned above, i.e. we try to see if we can obtain intuitionistic first-order logic (IQC) as the theory of a suitable

[1]In the English translation [59] this reads as follows: "In the general case, if x is a variable (of any kind) and $a(x)$ denotes a problem whose meaning depends on the values of x, then $(x)a(x)$ denotes the problem "find a general method for solving the problem $a(x)$ for each specific value of x". This should be understood as follows: the problem $(x)a(x)$ is solved if the problem $a(x_0)$ can be solved for each given specific value of x_0 of the variable x by means of a finite number of steps which are fixed in advance (before x_0 is set)."

factor or interval. We show that the answer is false, by showing that there is an analogue of Tennenbaum's theorem [116] that every computable model of Peano arithmetic is the standard model. On the other hand, we provide a partial positive characterisation of exactly which intuitionistic theories can be obtained.

1.2. Computability and probability: algorithmic randomness and genericity

In *algorithmic randomness*, we study notions of effective randomness. That is, we study what it means for an infinite binary sequence to be random. Roughly, we say that an infinite binary string x is random if it is not in any 'effective' set of measure 0. There are various ways to define 'effective', which yield different notions of randomness, of which Martin-Löf randomness is the most studied. Perhaps surprisingly, there are two alternative definitions of randomness which can be shown to be equivalent: instead of calling x random if it is not in any effective measure 0 set, we can also say that x is random if there is no effective compression method which compresses x by a non-trivial amount. For the other equivalent definition, imagine we are in a casino with some initial capital and are allowed to bet on $x(0), x(1), \ldots$ in succession, where at time t we are allowed to divide our money in bets on 0 and 1, and our new capital is twice the bet on $x(t)$. Then x is random if there is no effective betting strategy which earns arbitrarily much money while betting on x like this.

Recently, several researchers have studied connections between notions of randomness and the differentiability of functions. For example, Brattka, Miller and Nies [11], building on work by Demuth [22], have shown that a real number $x \in [0, 1]$ is Martin-Löf random if and only if every computable function of bounded variation is differentiable at x. Bounded variation is a notion from analysis; instead of giving the formal definition we mention that on a closed interval like $[0, 1]$ a function is of bounded variation if and only if it is the difference of two monotonically increasing functions. Other papers in which similar connections between randomness and differentiability are studied include Freer, Kjos-Hanssen, Nies and Stephan [30], Pathak, Rojas and Simpson [91] and Rute [100].

Instead of using measure to define randomness, one can also use Baire category and say that x is random if it is not in any 'effective' meagre set. In computability theory, such reals are usually called *generic*, with 1-genericity roughly being the equivalent of Martin-Löf randomness (which is also called 1-randomness). Given the Erdös-Sierpinski duality (which roughly states that in the non-effective case meagre and measure 0 have very similar properties), one would expect genericity and randomness to have similar properties. In the effective setting this turns out to often be the case, but not always. Mirroring the connections between randomness and differentiability mentioned above, we show that there is also a connection between differentiable functions and 1-genericity: namely, x is 1-generic if and only if every differentiable computable function has continuous derivative at x. We discuss this in chapter 6. In that chapter we also show that nothing changes if we replace the derivative by the nth derivative, or if we replace computable by polynomial time computable.

In chapter 7 we study the interplay of algorithmic randomness with *coarse reducibility*, which is one of the notions recently introduced by Jockusch and Schupp to study the concept of being 'almost computable', based on the notion of generic-case computability from complexity theory introduced by Kapovich, Myasnikov, Schupp and Shpilrain [48]. This reducibility induces a degree structure, the so-called coarse degrees, in which the Turing degrees embed in a natural way. We study how this embedding interacts with the coarse degrees of random sets, which turns out to be connected to the notion of K-triviality. We also show that the degrees of two sets which are weak 3-random relative to each other form a minimal pair in the coarse degrees.

1.3. Probability and logic: ε-logic

In *probability logic*, one combines logic with probability theory. There are several different motivations for doing this, which in turn lead to several different approaches. As discussed in the introduction of Kuyper and Terwijn [68], these approaches roughly fall into two categories: there are those that consider probability distributions over classes of models, assigning each model a certain probability (the "probabilities over models" approach), and there are those that consider models which each have their own probability distribution over the universe of the model (the "models with probabilities" approach).

In this thesis we study a logic from the second category called *ε-logic*, introduced by Terwijn [117], which is related to Valiant's celebrated notion of *pac-learning* [122]. This logic is *learnable* in a sense closely related to pac-learning: given some unknown structure and some oracle that allows us to randomly sample elements from the structure and that tells us all the basic properties of our sample, we want to decide if a given (first-order) expression holds in the structure or not after taking only a finite sample. In classical first-order logic we cannot do this, because to decide if a universal statement holds, we cannot get around checking all elements instead of only finitely many. In ε-logic, this problem is solved by interpreting universal quantifiers not as 'for all' but as 'with probability at least $1-\varepsilon$', while keeping the definition of the existential quantifier as it is classically. Terwijn [117] has shown that this logic is indeed learnable.

Of the many different approaches, let us name a few. In the "probabilities over models" approach we have for example the papers Carnap [16] and Scott and Krauss [103]. Another logic in this category is a logic introduced earlier by Valiant [123] which is also related to his pac-model. However, there are major differences between Valiant's logic and ε-logic: not only does his logic fall under the "probabilities over models" approach instead of under the "models with probabilities" approach, Valiant also only studies finite models (while we look at models of arbitrary cardinality), and his syntax captures only a fragment of first-order logic.

A logic which is actually more closely related to ε-logic is Keisler's logic $\mathcal{L}_{\omega P}$, surveyed in Keisler [51], and which falls under the "models with probabilities" approach. Instead of the classical quantifiers, this logic has quantifiers of the form $(Px \geq r)$ which should be read as "holds for at least measure r many x". While this logic does not attempt to model probabilistic induction, and does not contain the classical universal and existential quantifiers, it turns out we can adapt some of the ideas used to prove results about $\mathcal{L}_{\omega P}$ to obtain similar results for ε-logic. Another example of work in the second category is H. Friedman's quantifier Q (which means "for almost all" in the sense of measure theory), which is discussed in Steinhorn [115]. Finally, we wish to mention a logic recently introduced by Goldbring and Towsner [34]. While this logic has a completely different motivation than ε-logic, namely to introduce a "logical framework for formalising connections between finitary combinators and measure theory or ergodic theory", the logic appears to be similar to ε-logic.

First, we introduce ε-logic in chapter 8. In the same chapter we also discuss some of the choices we make, especially with regard to the class of models we study. After that, in chapter 9, we discuss various aspects of the model theory of ε-logic. This model theory is very different from classical model theory: both the results are different, as well as many of the proof techniques used, which lean heavily on techniques from descriptive set theory. For example, we show that there are satisfiable sentences which do not have any countable model (i.e. the straightforward analogue of the downward Löwenheim–Skolem theorem does not hold), but every satisfiable theory is satisfied in some model of cardinality 2^ω; in fact we can build it on the unit interval with the Lebesgue measure. We also study compactness, showing that ε-logic is in general not compact, but that a weaker version of compactness does hold.

In chapter 10 we study the computational complexity of validity in ε-logic. In Terwijn [119] it was already shown that the set of ε-tautologies is undecidable in general, and that it coincides with classical validity if $\varepsilon = 0$. Building on a result of Hoover [42] for $\mathcal{L}_{\omega P}$, we show that validity in ε-logic is, in general, Π^1_1-hard (i.e. at least as complex as first-order arithmetic with second-order universal quantifiers). This shows that ε-logic is computationally much harder than first-order logic and that we cannot hope to find an effective calculus for it.

Next, in chapter 11 we study the computational complexity of satisfiability. Note that the complexity of satisfiability is not necessarily complementary to that of validity, as it is in classical logic, because our logic is paraconsistent (i.e. both a formula φ and its negation $\neg\varphi$ can hold at the same time). We study the fragment of ε-logic not containing equality or function symbols, i.e. containing only relation and constant symbols. Terwijn [119, Conjecture 5.3] conjectured ε-satisfiability for rational $\varepsilon \in [0, 1)$ to be decidable, but we show that it is Σ^1_1-complete for rational $\varepsilon \in (0, 1)$, thus refuting his conjecture. On the other hand, we show that 0-satisfiability is indeed decidable. These results are summarised in Table 1 below. Note that 1-logic is fairly trivial: every formula in prenex normal form containing a universal quantifier is trivially true, so the only interesting fragment in this case is the existential fragment, which is just the classical fragment.

Recently, Yang [125] has studied, among other things, the complexity of validity and satisfiability over finite models. In case $\varepsilon \in (0, 1) \cap \mathbb{Q}$ we get the same complexities one level down (i.e. finite ε-satisfiability is Σ^0_1-complete while finite ε-validity is Π^0_1-complete), but on the other hand finite 0-validity is Π^0_1-complete while finite 0-satisfiability remains decidable. He also showed ε-satisfiability and ε-validity for finite monadic relational languages are decidable in both the finite and the unrestricted case.

	$\varepsilon \in (0,1) \cap \mathbb{Q}$	$\varepsilon = 0$
ε-satisfiability	Σ^1_1-complete	decidable
ε-validity	Π^1_1-hard	Σ^0_1-complete

TABLE 1. Complexity of validity and satisfiability in ε-logic.

1.4. Notations and conventions

Our notation is mostly standard. We let ω denote the natural numbers, we let 2^ω denote the Cantor space and we let ω^ω the Baire space of functions from ω to ω. We let $2^{<\omega}$ denote the set of finite binary strings. For any set $\mathcal{A} \subseteq \omega^\omega$ we denote by $\overline{\mathcal{A}}$ its complement in ω^ω. When we say that a set is countable, we include the possibility that it is finite.

For finite strings σ, τ we denote by $\sigma \subseteq \tau$ that σ is a substring of τ, by $\sigma \subset \tau$ that σ is a proper substring of τ and by $\sigma \mid \tau$ that σ and τ are incomparable. The concatenation of σ and τ is denoted by $\sigma^\frown\tau$; for $n \in \omega$ we denote by $\sigma^\frown n$ the concatenation of σ with the string $\langle n \rangle$ of length 1.

We let $\emptyset'$ denote the halting problem. By $\{e\}^A(n)[s]\downarrow$ we mean that the eth Turing machine with oracle A and input n terminates in at most s steps. For functions $f, g \in \omega^\omega$ we denote by $f \oplus g$ the join of the functions f and g, i.e. $(f \oplus g)(2n) = f(n)$ and $(f \oplus g)(2n+1) = g(n)$. For any set $X \subseteq \omega^\omega$ we denote by $C(X)$ the upper cone $\{f \in \omega^\omega \mid \exists g \in X(f \geq_T g)\}$. We let $\langle a_1, \ldots, a_n \rangle$ denote a fixed computable bijection between ω^n and ω.

We denote by λ the Lebesgue measure on the unit interval $[0,1]$. When not mentioned otherwise $\mathcal{D}$ denotes a probability measure, or synonymously a probability distribution. When we say that some property holds for $\mathcal{D}$-almost all x, we mean that the set of x for which the property holds has $\mathcal{D}$-measure 1; when $\mathcal{D}$ is clear from the context we will omit it. For any measure $\mathcal{D}$ and any $n \in \omega$, we let $\mathcal{D}^n$ denote the product measure of n copies of $\mathcal{D}$, and we let $\mathcal{D}^\omega$ denote the product measure of countably infinitely many copies of $\mathcal{D}$.

For a poset $(X, \leq)$ and elements $x, y \in X$, we denote by $[x,y]_X$ the set of elements $u \in X$ satisfying $x \leq u \leq y$. We denote the join operation in lattices by $\oplus$ and the meet operation by $\otimes$. By a relational formula we mean a formula which only contains relation symbols and does not contain equality, function or constant symbols. For all formulas φ and ψ, the formula $\varphi \leftrightarrow \psi$ is short for $(\varphi \to \psi) \wedge (\psi \to \varphi)$.

For unexplained notions from computability theory, we refer to Odifreddi [87] or Downey and Hirschfeldt [25]. For the Muchnik and Medvedev lattices, we refer to the surveys of Sorbi [111] and Hinman [36], for lattice theory, we refer to Balbes and Dwinger [5], and for unexplained notions about Kripke semantics we refer to Chagrov and Zakharyaschev [17] and Troelstra and van Dalen [121]. For background in descriptive set theory we refer to Kechris [50] or Moschovakis [83]. Further background on (classical) Baire category theory can also be found in Oxtoby [90]. For more information about measure theory we refer the reader to

Bogachev [10], and for model theory we refer the reader to Chang and Keisler [18] and Hodges [41].

Part I

The Medvedev and Muchnik Lattices

CHAPTER 2

The Medvedev and Muchnik Lattices

As discussed in the introduction to this thesis, Kolmogorov [58] introduced an informal calculus of problems in an attempt to give a semantics for intuitionistic propositional logic IPC. This was later formalised by Medvedev [79] and Muchnik [84], who introduced the *Medvedev* and *Muchnik* lattices, which are two computability-theoretic structures. In this part of the thesis we will study the logical content of these two structures.

In this first chapter we will introduce the Medvedev and Muchnik lattices and briefly discuss some of the properties and theorems which are central to the rest of this part of the thesis.

2.1. Prerequisites

First, let us recall the definitions of the Medvedev and Muchnik lattices.

Definition 2.1.1. (Medvedev [79]) Let $\mathcal{A}, \mathcal{B} \subseteq \omega^\omega$ (we will call subsets of ω^ω *mass problems*). We say that $\mathcal{A}$ *Medvedev reduces to* $\mathcal{B}$ (denoted by $\mathcal{A} \leq_{\mathcal{M}} \mathcal{B}$) if there exists a Turing functional Φ such that $\Phi(\mathcal{B}) \subseteq \mathcal{A}$. If both $\mathcal{A} \leq_{\mathcal{M}} \mathcal{B}$ and $\mathcal{B} \leq_{\mathcal{M}} \mathcal{A}$ we say that $\mathcal{A}$ and $\mathcal{B}$ are *Medvedev equivalent* (denoted by $\mathcal{A} \equiv_{\mathcal{M}} \mathcal{B}$). The equivalence classes of mass problems under Medvedev equivalence are called *Medvedev degrees*, and the class of all Medvedev degrees is denoted by $\mathcal{M}$. We call $\mathcal{M}$ the *Medvedev lattice*.

Definition 2.1.2. (Muchnik [84]) Let $\mathcal{A}, \mathcal{B} \subseteq \omega^\omega$. We say that $\mathcal{A}$ *Muchnik reduces to* $\mathcal{B}$ (notation: $\mathcal{A} \leq_w \mathcal{B}$) if for every $g \in \mathcal{B}$ there exists an $f \in \mathcal{A}$ such that $f \leq_T g$. If $\mathcal{A} \leq_w \mathcal{B}$ and $\mathcal{B} \leq_w \mathcal{A}$ we say that $\mathcal{A}$ and $\mathcal{B}$ are *Muchnik equivalent* (notation: $\mathcal{A} \equiv_w \mathcal{B}$). The equivalence classes under Muchnik equivalence are called *Muchnik degrees* and the set of Muchnik degrees is denoted by $\mathcal{M}_w$. We call $\mathcal{M}_w$ the *Muchnik lattice*.

To avoid confusion, we do not use $\vee$ for the join (least upper bound) or $\wedge$ for the meet (greatest lower bound) in lattices, because later on we will see that the join corresponds to the logical conjunction $\wedge$ and that the meet corresponds to the logical disjunction $\vee$. Instead, we use $\oplus$ for join and $\otimes$ for meet.

Definition 2.1.3. (McKinsey and Tarski [77]) A *Brouwer algebra* is a bounded distributive lattice together with a binary *implication operator* $\to$ satisfying:

$$a \oplus c \geq b \text{ if and only if } c \geq a \to b$$

i.e. $a \to b$ is the least element c satisfying $a \oplus c \geq b$.

First, we give a simple example of a Brouwer algebra.

DEFINITION 2.1.4. Let $(X, \leq)$ be a poset. We say that a subset $Y \subseteq X$ is *upwards closed* or is an *upset* if for all $y \in Y$ and all $x \in X$ with $x \geq y$ we have $x \in Y$. Similarly, we say that $Y \subseteq X$ is *downwards closed* or a *downset* if for all $y \in Y$ and all $x \in X$ with $x \leq y$ we have $x \in Y$.

We denote by $\mathcal{O}(X)$ the collection of all upwards closed subsets of X, ordered under reverse inclusion $\supseteq$.

PROPOSITION 2.1.5. *$\mathcal{O}(X)$ is a Brouwer algebra under the operations $U \oplus V = U \cap V$, $U \otimes V = U \cup V$ and*

$$U \to V = \{x \in X \mid \forall y \geq x (y \in U \Rightarrow y \in V)\}.$$

PROOF. The upwards closed sets of a poset form a topology (because they are closed under arbitrary unions and intersections). The result now follows from Balbes and Dwinger [5, IX.3, Example 4].[1] □

It turns out that the Medvedev and Muchnik lattices are both Brouwer algebras.

PROPOSITION 2.1.6. (Medvedev [79],Muchnik [84]) *The Medvedev and Muchnik lattices are Brouwer algebras under the operations induced by:*

$$\begin{aligned}
\mathcal{A} \oplus \mathcal{B} &= \{f \oplus g \mid f \in \mathcal{A} \text{ and } g \in \mathcal{B}\} \\
\mathcal{A} \otimes \mathcal{B} &= \mathcal{A} \cup \mathcal{B} \\
\mathcal{A} \to \mathcal{B} &= \{g \in \omega^\omega \mid \forall f \in \mathcal{A} \exists h \in \mathcal{B} (f \oplus g \geq_T h)\}
\end{aligned}$$

In fact, the Muchnik lattice is of the simple form described in Definition 2.1.4 above.

PROPOSITION 2.1.7. *The Muchnik lattice is isomorphic to the lattice of upsets of the Turing degrees.*

PROOF. We use a proof inspired by Muchnik's proof that the Muchnik degrees can be embedded in the Medvedev degrees (preserving 0, 1 and minimal upper bounds) from [84]. For every $\mathcal{A} \subseteq \omega^\omega$, we have that $\mathcal{A} \equiv_w C(\mathcal{A}) := \{f \in \omega^\omega \mid \exists g \in \mathcal{A}(g \leq_T f)\}$. Now it is directly verified that the mapping sending $\mathcal{A}$ to $C(\mathcal{A})$ induces an order isomorphism between $\mathcal{M}_w$ and $\mathcal{O}(\mathcal{D})$ (as defined in Definition 2.1.4). Finally, every order isomorphism between Brouwer algebras is automatically a Brouwer algebra isomorphism, see Balbes and Dwinger [5, IX.4, Exercise 3]. □

The main reason Brouwer algebras are interesting is because we can use them to give algebraic semantics for IPC, as witnessed by the next definition and the results following after it.

DEFINITION 2.1.8. ([78]) Let $\varphi(x_1, \ldots, x_n)$ be a propositional formula with free variables among $x_1, \ldots, x_n$, let $\mathscr{B}$ be a Brouwer algebra and let $b_1, \ldots, b_n \in \mathscr{B}$. Let ψ be the formula in the language of Brouwer algebras obtained from φ by replacing logical disjunction $\vee$ by $\otimes$, logical conjunction $\wedge$ by $\oplus$, logical implication

[1] Throughout most of the literature, including Balbes and Dwinger, results are proved for Heyting algebras, the order-dual of Brouwer algebras. However, all results we cite directly follow for Brouwer algebras in the same way.

$\to$ by Brouwer implication $\to$ and the false formula $\bot$ by 1 (we view negation $\neg\alpha$ as $\alpha \to \bot$). We say that $\varphi(b_1, \dots, b_n)$ *holds in* $\mathscr{B}$ if $\psi(b_1, \dots, b_n) = 0$. Furthermore, we define the *theory* of $\mathscr{B}$ (notation: $\mathrm{Th}(\mathscr{B})$) to be the set of those formulas which hold for every valuation, i.e.

$$\mathrm{Th}(\mathscr{B}) = \{\varphi(x_1, \dots, x_m) \mid \forall b_1, \dots, b_m \in \mathscr{B}(\varphi(b_1, \dots, b_m) \text{ holds in } \mathscr{B})\}.$$

The following soundness result is well-known and directly follows from the observation that all rules in some fixed deduction system for IPC preserve truth.

PROPOSITION 2.1.9. ([78, Theorem 4.1]) *For every Brouwer algebra $\mathscr{B}$:* IPC $\subseteq$ $\mathrm{Th}(\mathscr{B})$.

PROOF. See e.g. Chagrov and Zakharyaschev [17, Theorem 7.10]. $\square$

Conversely, the class of Brouwer algebras is complete for IPC.

THEOREM 2.1.10. *([78, Theorem 4.3])*

$$\bigcap\{\mathrm{Th}(\mathscr{B}) \mid \mathscr{B} \textit{ a Brouwer algebra}\} = \mathrm{IPC}$$

Thus, Brouwer algebras can be used to provide algebraic semantics for IPC. Therefore, it would be nice if the computationally motivated Medvedev and Muchnik lattices have IPC as their theory, so that they would provide computational semantics for IPC. Unfortunately the weak law of the excluded middle $\neg p \vee \neg\neg p$ holds in both of them, as can be easily verified. In fact, their theory is exactly Jankov's logic Jan consisting of the deductive closure of IPC plus the weak law of the excluded middle, as shown by Sorbi [110].

However, as mentioned in the introduction we can still recover IPC by looking at principal factors.

PROPOSITION 2.1.11. *Let $\mathscr{B}$ be a Brouwer algebra and let $x, y \in \mathscr{B}$. Then the interval $[x, y]_{\mathscr{B}} = \{z \in \mathscr{B} \mid x \leq z \leq y\}$ is a sublattice of $\mathscr{B}$. Furthermore, it is a Brouwer algebra under the implication*

$$u \to_{[x,y]_{\mathscr{B}}} v = (u \to_{\mathscr{B}} v) \oplus x.$$

PROPOSITION 2.1.12. *Let $\mathscr{B}$ be a Brouwer algebra and let $x \in \mathscr{B}$. Then $\mathscr{B}/\{z \in \mathscr{B} \mid z \geq x\}$, which we will denote by $\mathscr{B}/x$, is isomorphic as a bounded distributive lattice to $[0, x]_{\mathscr{B}}$. In particular, $\mathscr{B}/x$ is a Brouwer algebra.*

Taking such a factor essentially amounts to moving from the entire algebra to just the interval $[0, x]_{\mathcal{M}_w}$ of elements below x. Because the top element of $[0, x]_{\mathcal{M}_w}$ is smaller than the top element of $\mathcal{M}_w$ if $x \neq 1$, the interpretation of negation $\neg b$, which is defined as $b \to 1$, also differs between these two algebras. Thus, taking a factor roughly corresponds to changing the negation.

Quite remarkably, Skvortsova has shown that there is an $\mathcal{A} \in \mathcal{M}$ such that the theory of $\mathcal{M}/\mathcal{A} = \mathrm{IPC}$, and the same was later shown to hold for the Muchnik lattice.

THEOREM 2.1.13. (Skvortsova [106]) *There exists a mass problem $\mathcal{A} \subseteq \omega^\omega$ such that* $\mathrm{Th}(\mathcal{M}/\mathcal{A}) = \mathrm{IPC}$.

THEOREM 2.1.14. (Sorbi and Terwijn [113]) *There exists a mass problem* $\mathcal{A} \subseteq \omega^\omega$ *such that* $\mathrm{Th}(\mathcal{M}_w/\mathcal{A}) = \mathrm{IPC}$.

However, as explained in the introduction, while these results are extraordinary, one could object to the fact that these mass problems $\mathcal{A}$ are constructed in an ad hoc manner. In chapter 3 we show that there is a natural way to choose such an $\mathcal{A}$ for the Muchnik lattice, and in chapter 4 we present progress towards obtaining such a natural $\mathcal{A}$ for the Medvedev lattice.

Finally, let us mention one easy lemma which we will use in this thesis.

LEMMA 2.1.15. *Let* $\mathscr{B}, \mathscr{C}$ *be Brouwer algebras and let* $\alpha : \mathscr{B} \to \mathscr{C}$ *be a surjective homomorphism. Then* $\mathrm{Th}(\mathscr{B}) \subseteq \mathrm{Th}(\mathscr{C})$.

PROOF. Let $\varphi(x_1, \dots, x_n) \notin \mathrm{Th}(\mathscr{C})$. Fix $c_1, \dots, c_n \in \mathscr{C}$ such that we have $\varphi(c_1, \dots, c_n) \neq 0$. Fix $b_1, \dots, b_n \in \mathscr{B}$ such that $\gamma(b_i) = c_i$. Then

$$\alpha(\varphi(b_1, \dots, b_n)) = \varphi(\alpha(b_1), \dots, \alpha(b_n)) = \varphi(c_1, \dots, c_n) \neq 0$$

because α is a homomorphism. Thus $\varphi(b_1, \dots, b_n) \neq 0$ and therefore we see that $\varphi \notin \mathrm{Th}(\mathscr{B})$. $\square$

CHAPTER 3

Natural Factors of the Muchnik Lattice Capturing IPC

As discussed in chapter 2, the Muchnik lattice does not capture intuitionistic propositional logic IPC. However, as shown by Sorbi and Terwijn [113], there is a principal factor of it which does capture IPC. Unfortunately, this factor is defined in an ad hoc manner and does not have any clear interpretation stemming from concepts normally studied in mathematics. In this chapter we present several natural examples of principal factors of the Muchnik lattice capturing IPC, which can be defined using concepts from computability theory.

The structure of this chapter is as follows. In section 3.1 we will describe our framework of *splitting classes*. In section 3.2 we show that our framework is non-trivial by proving that the low functions and the functions of 1-generic degree below $\emptyset'$ fit in our framework. Next, in section 3.3 we prove that splitting classes naturally induce a factor of the Muchnik lattice which captures IPC. Finally, in section 3.4 we consider whether two other concepts from computability theory give us splitting classes: hyperimmune-freeness and computable traceability.

This chapter is based on Kuyper [65].

3.1. Splitting classes

As announced above, we will present our results in a general framework so that additional examples can easily be obtained. Our framework of *splitting classes* abstracts exactly what we need for our proof in section 3.3 to work. It roughly says that $\mathcal{A}$ is a splitting class if, given some function $f \in \mathcal{A}$, we can construct functions $h_0, h_1 \in \mathcal{A}$ above it whose join is not in $\mathcal{A}$ while 'avoiding' a given finite set of other functions in $\mathcal{A}$. This is made precise below.

Definition 3.1.1. Let $\mathcal{A} \subseteq \omega^\omega$ be a non-empty countable class which is downwards closed under Turing reducibility. We say that $\mathcal{A}$ is a *splitting class* if for every $f \in \mathcal{A}$ and every finite subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exist $h_0, h_1 \in \mathcal{A}$ such that $h_0, h_1 \geq_T f$, $h_0 \oplus h_1 \notin \mathcal{A}$ and for all $g \in \mathcal{B}$: $g \oplus h_0, g \oplus h_1 \notin \mathcal{A}$.

Note that, because every splitting class $\mathcal{A}$ is downwards closed under Turing reducibility, we in particular have that $\mathcal{A}$ is closed under Turing equivalence, i.e. if $f \in \mathcal{A}$ and $g \equiv_T f$ then also $g \in \mathcal{A}$.

We emphasise that we required a splitting class to be countable. There are also interesting examples which satisfy the requirements except for the countability: for example, in section 3.4 we will see that this is the case for the set of hyperimmune-free functions. In that section we will also discuss how to suitably generalise the concept to classes of higher cardinality.

It turns our that in order to show that something is a splitting class it will be easier to prove that one of the two alternative formulations given by the next proposition holds.

PROPOSITION 3.1.2. *Let $\mathcal{A} \subseteq \omega^\omega$ be a non-empty countable class which is downwards closed under Turing reducibility. Then the following are equivalent:*

(i) *$\mathcal{A}$ is a splitting class.*
(ii) *For every $f \in \mathcal{A}$ and every finite subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exists $h \in \mathcal{A}$ such that $h >_T f$ and for all $g \in \mathcal{B}$: $g \oplus h \notin \mathcal{A}$.*
(iii) *For every $f \in \mathcal{A}$ there exists $h \in \mathcal{A}$ such that $h \not\leq_T f$, and for every $f \in \mathcal{A}$, every finite subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ and every $h_0 \in \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exists $h_1 \in \mathcal{A}$ such that $h_1 >_T f$, $h_0 \oplus h_1 \notin \mathcal{A}$ and for all $g \in \mathcal{B}$: $h_1 \not\geq_T g$.*

PROOF. (i) $\to$ (ii): Let $h_0, h_1 \in \mathcal{A}$ be such that $h_0, h_1 \geq_T f$, $h_0 \oplus h_1 \notin \mathcal{A}$ and for all $g \in \mathcal{B}$: $g \oplus h_0, g \oplus h_1 \notin \mathcal{A}$. Let $h = h_0$. Because $h \equiv_T f$ would imply that $h_0 \oplus h_1 \equiv_T h_1 \in \mathcal{A}$ we see that $h >_T f$ and therefore we are done.

(ii) $\to$ (iii): First, for every $f \in \mathcal{A}$ we can find $h \in \mathcal{A}$ such that $h \not\leq_T f$ by applying (ii) with $\mathcal{B} = \emptyset$. Next, using (ii) determine $h_1 \in \mathcal{A}$ such that $h_1 >_T f$ and for all $g \in \mathcal{B} \cup \{h_0\}$: $g \oplus h_1 \notin \mathcal{A}$. Then the only thing we still need to show is that $h \not\geq_T g$ for all $g \in \mathcal{B}$. However, $h \geq_T g$ would imply $h \oplus g \equiv_T h \in \mathcal{A}$, a contradiction.

(iii) $\to$ (ii): Fix $g_1 \in \mathcal{A}$ such that $g_1 \not\leq_T f$. Let $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ be finite. Without loss of generality, we may assume that $g_1 \in \mathcal{B}$; in particular, we may assume that $\mathcal{B}$ is non-empty. So, let $\mathcal{B} = \{g_1, \dots, g_n\}$. We inductively define a sequence $h_{1,0} <_T h_{1,1} <_T \cdots <_T h_{1,n}$ of functions in $\mathcal{A}$. First, we let $h_{1,0} = f$. Next, to obtain $h_{1,i+1}$ from $h_{1,i}$, apply (iii) to find a function $h_{1,i+1} >_T h_{1,i}$ such that $h_{1,i+1} \oplus g_{i+1} \notin \mathcal{A}$ and for all $i + 2 \leq j \leq n$ we have $g_j \not\leq_T h_{1,i+1}$. Then $h := h_{1,n}$ is as desired.

(ii) $\to$ (i): Using (ii), we can find $h_0 \in \mathcal{A}$ such that $h_0 >_T f$ and $g \oplus h_0 \notin \mathcal{A}$ for all $g \in \mathcal{B}$. By applying (ii) a second time, we can now find $h_1 \in \mathcal{A}$ such that $h_1 >_T f$ and for all $g \in \mathcal{B} \cup \{h_0\}$: $g \oplus h_1 \notin \mathcal{A}$. Then h_0 and h_1 are as desired. □

3.2. Low and 1-generic below $\emptyset'$ are splitting classes

Before we show that splitting classes allow us to capture IPC as a factor of the Muchnik lattice, we want to demonstrate that our framework of splitting classes is non-trivial. To this end, we will show that the class of low functions, and that the class of functions of 1-generic degree below $\emptyset'$ together with the computable functions, are splitting classes. We will denote the first class by $\mathcal{A}_{\text{low}}$ and the second class by $\mathcal{A}_{\text{gen}\leq\emptyset'}$. We remark that the second class naturally occurs as the class of functions that are low for EX (as proved in Slaman and Solovay [108]).

Because these kinds of arguments are usually given as constructions on sets (or elements of Cantor space) rather than the functions (or elements of Baire space) which occur in the Muchnik lattice, we will work with sets instead of functions in this section. However, we do not use the compactness of Cantor space anywhere and therefore it is only a notational matter.

First, we recall some basic facts about 1-genericity over a set.

DEFINITION 3.2.1. (Jockusch [14, p. 125]) Let $A, B \subseteq \omega$. We say that B is *1-generic* if for every $e \in \omega$ there exists $\sigma \subseteq B$ such that either $\{e\}^\sigma(e)\downarrow$ or for all $\tau \supseteq \sigma$ we have $\{e\}^\tau(e)\uparrow$.

More generally, we say that B is *1-generic over A* if for every $e \in \omega$ there exists $\sigma \subseteq B$ such that either $\{e\}^{A\oplus\sigma}(e)\downarrow$ or for all $\tau \supseteq \sigma$ we have $\{e\}^{A\oplus\tau}(e)\uparrow$.

LEMMA 3.2.2. (Folklore) *Let B be 1-generic over A. Then:*

(i) *If A is 1-generic, then $A \oplus B$ is 1-generic.*
(ii) *If A is low and $B \leq_T \emptyset'$, then $A \oplus B$ is low.*

PROOF. (i): Assume A is 1-generic. Let $e \in \omega$. We need to find a $\sigma \subseteq A \oplus B$ such that either $\{e\}^\sigma(e)\downarrow$ or such that for all $\tau \supseteq \sigma$ we have $\{e\}^\tau(e)\uparrow$.

If $\{e\}^{A\oplus B}(e)\downarrow$, we can choose $\sigma \subseteq A \oplus B$ such that $\{e\}^\sigma(e)\downarrow$. Otherwise, since B is 1-generic over A, we can determine $\sigma_B \subseteq B$ such that for all $\tau_B \supseteq \sigma_B$ we have $\{e\}^{A\oplus\tau_B}(e)\uparrow$. Fix an index $\tilde{e}$ such that for all $C \subseteq \omega$ and all $x \in \omega$:

$$\{\tilde{e}\}^C(x)\downarrow \Leftrightarrow \exists \tau_B \supseteq \sigma_B \{e\}^{C\oplus\tau_B}(e)\downarrow .$$

We first note that $\{\tilde{e}\}^A(x)\uparrow$ by our choice of σ_B. Therefore, using the 1-genericity of A, determine $\sigma_A \subseteq A$ such that for all $\tau_A \supseteq \sigma_A$ we have $\{\tilde{e}\}^{\tau_A}(\tilde{e})\uparrow$. By choice of $\tilde{e}$ we then have for for all $\tau_A \supseteq \sigma_A$ that $\forall \tau_B \supseteq \sigma_B \{e\}^{\tau_A\oplus\tau_B}(e)\uparrow$, which is the same as saying that for all $\tau \supseteq \sigma_A \oplus \sigma_B$ we have $\{e\}^\tau(e)\uparrow$. This is exactly what we needed to show.

(ii) We show that both $(A \oplus B)'$ and its complement $\overline{(A \oplus B)'}$ are c.e. in $A' \oplus B \equiv_T \emptyset'$. To this end, we note that $e \in (A \oplus B)'$ if and only if

$$\exists \sigma_A \subseteq A \exists \sigma_B \subseteq B \left(\{e\}^{\sigma_A\oplus\sigma_B}(e)\downarrow\right)$$

which is c.e. in $A \oplus B \leq_T A' \oplus B$. Next, using the fact that B is 1-generic over A, we see that $e \notin (A \oplus B)'$ if and only if

$$\exists \sigma_B \subseteq B \forall \tau_B \supseteq \sigma_B \left(\{e\}^{A\oplus\tau_B}(e)\uparrow\right)$$

which is c.e. in $A' \oplus B$. The result now follows by the relativised Post's theorem. ☐

THEOREM 3.2.3. $\mathcal{A}_{\mathrm{low}}$ *and* $\mathcal{A}_{\mathrm{gen}\leq\emptyset'}$ *are splitting classes.*

PROOF. The first class is clearly downwards closed; for the second class this is proved in Haught [35] (but also follows from the fact mentioned above that $\mathcal{A}_{\mathrm{gen}\leq\emptyset'}$ consists of exactly those functions which are low for EX).

First, we consider the class of low functions. By Proposition 3.1.2, we can show that the low functions form a splitting class by proving that for every low A and every finite $\mathcal{B} \subseteq \{B \in \omega^\omega \mid B \text{ low and } B \not\leq_T A\}$ there exists a set $C \not\leq_T A$ such that $A \oplus C$ is low and such that for all $B \in \mathcal{B}$ we have that $B \oplus (A \oplus C) \equiv_T \emptyset'$. (Note that $C \not\leq_T A$ ensures that $A \oplus C >_T A$, while $B \oplus (A \oplus C) \equiv_T \emptyset'$ ensures that $B \oplus (A \oplus C)$ is neither 1-generic nor low.) Lemma 3.2.2 tells us that we can make $A \oplus C$ low by ensuring that $C \leq \emptyset'$ and that C is 1-generic over A. Thus, it is enough if we can show:

(1) *If A is low and $\mathcal{B} \subseteq \{B \in \omega^\omega \mid B \leq_T \emptyset'$ and $B \not\leq_T A\}$ is finite, then there exists a set $C \leq_T \emptyset'$ which is 1-generic over A such that $C \not\leq_T A$ and for all $B \in \mathcal{B}$: $B \oplus (A \oplus C) \equiv_T \emptyset'$.*

In fact, we then also immediately get the result for the class of functions of 1-generic degree below $\emptyset'$. Namely, let $A \leq_T \emptyset'$ be of 1-generic degree and let $\mathcal{B} \subseteq \{B \in \omega^\omega \mid B \leq_T \emptyset'$ and $B \not\leq_T A\}$ be finite. Just as above, it would be enough to have a set $C \leq_T \emptyset'$ such that $C \not\leq_T A$, $A \oplus C$ is of 1-generic degree and for all $B \in \mathcal{B}$: $B \oplus (A \oplus C) \equiv_T \emptyset'$. Note that this expression is invariant under replacing A with a Turing equivalent set, so because A is of 1-generic degree we may without loss of generality assume A to be 1-generic. Then, because $A \leq \emptyset'$ is 1-generic, it is also low. So we can find a set C as in (1). By Lemma 3.2.2 we then have that $A \oplus C$ is 1-generic, and therefore C is exactly as desired.

To prove (1) we modify the proof of the Posner and Robinson Cupping Theorem [95]. Let $\mathcal{B} = \{B_1, \dots, B_k\}$. For every $B_i \in \mathcal{B}$, since $B_i \leq \emptyset'$ we can approximate B_i by a computable sequence $B_i^0, B_i^1, \dots$ of finite sets. We now let α_i be the *computation function* defined by letting $\alpha_i(n)$ be the least $m \geq n$ such that $B_i^m \upharpoonright (n+1) = B_i \upharpoonright (n+1)$. Then $B_i \equiv_T \alpha_i$. Now let $\alpha = \min(\alpha_1, \dots, \alpha_k)$. Then, by Lemma 6 of [95], any function g which dominates α computes some B_i. Thus, we see that no function computable in A can dominate α.

We will now construct a set C as in (1) by a finite extension argument, i.e. as $C = \bigcup_{n \in \omega} \sigma_n$. Fix any computable sequence $\tau_0, \tau_1, \dots$ of mutually incomparable finite strings (for example, $\tau_n = \langle 0^n 1 \rangle$, the string consisting of n times a 0 followed by a 1). We start with $\sigma_0 = \emptyset$. To define σ_{e+1} given σ_e, let n be the least $m \in \omega$ such that either (where the quantifiers are over finite strings):

$$\forall \sigma \supseteq \sigma_e{}^\frown\tau_m \left(\{e\}^{A \oplus \sigma}(e) \uparrow \right) \tag{2}$$

or

$$\exists \sigma \supseteq \sigma_e{}^\frown\tau_m \left(|\sigma| \leq \alpha(m) \wedge \{e\}^{A \oplus \sigma}(e)[|\sigma|] \downarrow \right). \tag{3}$$

Such an m exists: otherwise, for every $l \in \omega$ we could let $\beta(l)$ be the least $s \in \omega$ such that

$$\exists \sigma \supseteq \sigma_e{}^\frown\tau_l \left(\{e\}^{A \oplus \sigma}(e)[|\sigma|] \downarrow \wedge |\sigma| = s \right).$$

For every l such an s exists because (2) does not hold for l, while such an s has to be strictly bigger than $\alpha(l)$ because (3) also does not hold. So, β would be a function computable in A which dominates α, of which we have shown above that it cannot exist.

Now, if case (2) holds for n, then we let $\sigma_{e+1} = \sigma_e{}^\frown\tau_n{}^\frown\emptyset'(e)$. Otherwise, we let $\sigma_{e+1} = \sigma^\frown\emptyset'(e)$, where σ is the least σ such that (3) is satisfied.

The construction is computable in $A' \oplus B_1 \oplus \cdots \oplus B_k \leq_T \emptyset'$: the set of $m \in \omega$ for which (2) holds is co-c.e. in A, while for (3) this is computable in $\alpha \leq_T B_1 \oplus \cdots \oplus B_k$ and A. Therefore, $C \leq_T \emptyset'$ holds.

Furthermore, per construction of σ_{e+1} we have either $\{e\}^{A \oplus \sigma_{e+1}}(e) \downarrow$, or for all $\tau \supseteq \sigma_{e+1}$ we have $\{e\}^{A \oplus \tau}(e) \uparrow$. So, C is 1-generic over A.

Next, for every $1 \leq i \leq k$ the construction is computable in $(A \oplus C) \oplus B_i$: to determine σ_{e+1} given σ_e, use C to find the unique $n \in \omega$ such that $C \supseteq \sigma_e{}^\frown\tau_n$. We can now compute in A and B_i if there exists some string $\sigma \supseteq \sigma_e{}^\frown\tau_n$ of length at most $\alpha_i(n)$ such that $\{e\}^{A\oplus\sigma}(e)[|\sigma|]\downarrow$: if so, let σ be the least such string and then $\sigma_{e+1} = B \upharpoonright |\sigma| + 1$. Otherwise, $\sigma_{e+1} = B \upharpoonright |\sigma_e| + 1$. Then we also see that $\emptyset'$ is computable in $(A \oplus C) \oplus B_i$, because $\emptyset'(e)$ is the last element of σ_{e+1}. Since also $A, B_i, C \leq_T \emptyset'$ we see that $(A \oplus C) \oplus B_i \equiv_T \emptyset'$.

Finally, because for every low A there exists some low $B_0 >_T A$ (because we can construct a 1-generic relative to A in $A' \equiv_T \emptyset'$ and apply Lemma 3.2.2) we may without loss of generality assume that such a B_0 is in $\mathcal{B}$. Then we have $B_0 \oplus (A \oplus C) \equiv_T \emptyset'$, as shown above. Now, if it were the case that $C \leq_T A$, then $\emptyset' \equiv_T B_0 \oplus (A \oplus C) \equiv_T B_0$, which contradicts B_0 being low. So $C \not\leq_T A$, which is the last thing we needed to show. ☐

3.3. The theory of a splitting class

We will now show that the theory of a splitting class equals IPC. We start by moving away from our algebraic viewpoint to Kripke semantics (we refer the reader to Chagrov and Zakharyaschev [17] for an introduction to Kripke frames). The following folklore result contains the crucial idea we need for this.

THEOREM 3.3.1. *For any poset $(X, \leq)$, the theory of $(X, \leq)$ as a Kripke frame is the same as theory of the lattice of upsets of X as a Brouwer algebra.*

PROOF. See e.g. Chagrov and Zakharyaschev [17, Theorem 7.20] for the order-dual result for Heyting algebras. ☐

PROPOSITION 3.3.2. *Let $\mathcal{A} \subseteq \omega^\omega$ be downwards closed under Turing reducibility. Then $\mathcal{M}_w/\overline{\mathcal{A}}$ (i.e. $\mathcal{M}_w$ modulo the principal filter generated by $\overline{\mathcal{A}}$) is isomorphic to the lattice of upsets $\mathcal{O}(\mathcal{A})$ of $\mathcal{A}$. In particular,* $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}}) = \mathrm{Th}(\mathcal{A})$ *(the first as Brouwer algebra, the second as Kripke frame).*

PROOF. By Proposition 2.1.7, $\mathcal{M}_w$ is isomorphic to the lattice of upsets $\mathcal{O}(\mathcal{D})$ of the Turing degrees $\mathcal{D}$, by sending each set $\mathcal{B} \subseteq \omega^\omega$ to $C(\mathcal{B})$. Since $\overline{\mathcal{A}}$ is upwards closed, we see that the isomorphism sends $\overline{\mathcal{A}}$ to itself. Therefore, $\mathcal{M}_w/\overline{\mathcal{A}}$, which is isomorphic to the initial segment $[\omega^\omega, \overline{\mathcal{A}}]_{\mathcal{M}_w}$ of $\mathcal{M}_w$, is isomorphic to the initial segment $[\omega^\omega, \overline{\mathcal{A}}]_{\mathcal{O}(\mathcal{D})}$. Finally, $[\omega^\omega, \overline{\mathcal{A}}]_{\mathcal{O}(\mathcal{D})}$ is easily seen to be isomorphic to $\mathcal{O}(\mathcal{A})$, by sending each set $\mathcal{B} \in \mathcal{O}(\mathcal{A})$ to $\mathcal{B} \cup \overline{\mathcal{A}}$. The result now follows from the previous theorem. ☐

Thus, if we take the factor of $\mathcal{M}_w$ given by the principal filter generated by $\overline{\mathcal{A}}$, we get exactly the theory of the Kripke frame $(\mathcal{A}, \leq_T)$. The rest of this section will be used to show that for splitting classes this theory is exactly IPC. To this end, we need the right kind of morphisms for Kripke frames, called *p-morphisms*.

DEFINITION 3.3.3. (de Jongh and Troelstra [21]) Let $(X_1, \leq_1)$, $(X_2, \leq_2)$ be Kripke frames. A surjective function $f : (X_1, \leq_1) \to (X_2, \leq_2)$ is called a *p-morphism* if

(1) f is an order homomorphism: $x \leq_1 y \to f(x) \leq_2 f(y)$,
(2) $\forall x \in X_1 \forall y \in X_2 (f(x) \leq_2 y \to \exists z \in X_1 (x \leq_1 z \wedge f(z) = y))$.

PROPOSITION 3.3.4. *If there exists a p-morphism from $(X_1, \leq_1)$ to $(X_2, \leq_2)$, then* $\mathrm{Th}(X_1, \leq_1) \subseteq \mathrm{Th}(X_2, \leq_2)$.

PROOF. See e.g. Chagrov and Zakharyaschev [17, Corollary 2.17]. □

THEOREM 3.3.5. (Smoryński [109]) $\mathrm{Th}(2^{<w}) = \mathrm{IPC}$.

PROOF. See e.g. Chagrov and Zakharyaschev [17, Corollary 2.33]. □

So, if we want to show that the theory of $\mathcal{M}_w/\overline{\mathcal{A}}$ equals IPC, it is enough to show that there exists a p-morphism from $\mathcal{A}$ to $2^{<\omega}$. We next show that this is indeed possible for splitting classes.

PROPOSITION 3.3.6. *Let $\mathcal{A}$ be a splitting class. Then there exists a p-morphism $\alpha : (\mathcal{A}, \leq_T) \to 2^{<\omega}$.*

PROOF. Instead of building a p-morphism from $\mathcal{A}$, we will build it from $\mathcal{A}/{\equiv_T}$ (which is equivalent to building one from $\mathcal{A}$, since any order homomorphism has to send T-equivalence classes to equal strings). For ease of notation we will write $\mathcal{A}$ for $\mathcal{A}/\equiv_T$ during the remainder of this proof.

Fix an enumeration $\mathbf{a}_0, \mathbf{a}_1, \dots$ of $\mathcal{A}$. We will build a sequence $\alpha_0 \subseteq \alpha_1 \subseteq \dots$ of finite, partial order homomorphisms from $\mathcal{A}$ to $2^{<\omega}$, which additionally satisfy that if $\mathbf{a}, \mathbf{b} \in \mathrm{dom}(\alpha_i)$ and $\alpha_i(\mathbf{a}) \mid \alpha_i(\mathbf{b})$, then $\mathbf{a} \oplus \mathbf{b} \notin \mathcal{A}$.

We satisfy the following requirements:

- R_0: $\alpha_0(\mathbf{0}) = \emptyset$ (where $\mathbf{0}$ is the least Turing degree)
- R_{2n+1}: $\mathbf{a}_n \in \mathrm{dom}(\alpha_{2n+1})$
- R_{2n+2}: there are $\mathbf{c}_0, \mathbf{c}_1 \in \mathrm{dom}(\alpha_{2n+2})$ with $\mathbf{c}_0, \mathbf{c}_1 \geq_T \mathbf{a}_n$ and $\alpha_{2n+2}(\mathbf{c}_0) = \alpha_{2n+1}(\mathbf{a}_n)^\frown 0$, $\alpha_{2n+2}(\mathbf{c}_1) = \alpha_{2n+1}(\mathbf{a}_n)^\frown 1$.

First, we show that for such a sequence the function $\alpha = \bigcup_{n \in \omega} \alpha_n$ is a p-morphism $\alpha : (\mathcal{A}, \leq_T) \to 2^{<\omega}$. First, the odd requirements ensure that α is total. Furthermore, α is an order homomorphism because the α_i are. To show that α is a p-morphism, let $\mathbf{a} \in \mathcal{A}$ and let $\alpha(\mathbf{a}) \subseteq y$; we need to find some $\mathbf{a} \leq_T \mathbf{b} \in \mathcal{A}$ such that $\alpha(\mathbf{b}) = y$. Because $\alpha(\mathbf{a}) \subseteq y$ we know that $y = \alpha(\mathbf{a})^\frown y'$ for some finite string y'. We may assume y' to be of length 1, the general result then follows by induction. Now, if we let $n \in \omega$ be such that $\mathbf{a} = \mathbf{a}_n$ then $\mathbf{a}_n \in \mathrm{dom}(\alpha_{2n+1})$, so requirement R_{2n+2} tells us that there are functions $\mathbf{c}_0, \mathbf{c}_1 \geq \mathbf{a}_n$ with $\alpha_{2n+2}(\mathbf{c}_0) = \alpha(\mathbf{a})^\frown 0$ and $\alpha_{2n+2}(\mathbf{c}_1) = \alpha(\mathbf{a})^\frown 1$. Now either $\alpha(\mathbf{c}_0) = y$ or $\alpha(\mathbf{c}_1) = y$, which is what we needed to show. That α is surjective directly follows from the fact that $\emptyset$ is in its range and that it satisfies property (2) of a p-morphism.

Now, we show how to actually construct the sequence. First, α_0 is already defined. Next assume α_{2n} has been constructed, we will construct α_{2n+1} extending α_{2n} such that $\mathbf{a}_n \in \mathrm{dom}(\alpha_{2n+1})$. The set

$$X := \{\alpha_{2n}(\mathbf{b}) \mid \mathbf{b} \in \mathrm{dom}(\alpha_{2n}) \text{ and } \mathbf{b} \leq_T \mathbf{a}_n\}$$

is totally ordered under $\subseteq$. Since, if $\mathbf{b}, \mathbf{c} \leq_T \mathbf{a}_n$ then $\mathbf{b} \oplus \mathbf{c} \leq_T \mathbf{a}_n$. Now, if $\alpha_{2n}(\mathbf{b})$ and $\alpha_{2n}(\mathbf{c})$ are incomparable then we assumed that $\mathbf{b} \oplus \mathbf{c} \notin \mathcal{A}$. This contradicts the assumption that $\mathcal{A}$ is downwards closed. So, we can define $\alpha_{2n+1}(\mathbf{a}_n)$ to be the largest element of X.

We show that α_{2n+1} is an order homomorphism; we then also automatically know that it is well-defined. Thus, let $\mathbf{b}_1, \mathbf{b}_2 \in \mathrm{dom}(\alpha_{2n+1})$ with $\mathbf{b}_1 \leq_T \mathbf{b}_2$. If they are both already in $\mathrm{dom}(\alpha_{2n})$, then the induction hypothesis on α_{2n} already tells us that $\alpha_{2n+1}(\mathbf{b}_1) \subseteq \alpha_{2n+1}(\mathbf{b}_2)$. If $\mathbf{b}_1 \in \mathrm{dom}(\alpha_{2n})$ and $\mathbf{b}_2 = \mathbf{a}_n$, then $\alpha_{2n}(\mathbf{b}_1) \in X$, so by definition of $\alpha_{2n+1}(\mathbf{a}_n)$ we directly see that $\alpha_{2n+1}(\mathbf{b}_1) \subseteq \alpha_{2n+1}(\mathbf{a}_n)$. Finally, we consider the case that $\mathbf{b}_1 = \mathbf{a}_n$ and $\mathbf{b}_2 \in \mathrm{dom}(\alpha_{2n})$. To show that $\alpha_{2n+1}(\mathbf{a}_n) \subseteq \alpha_{2n+1}(\mathbf{b}_2) = \alpha_{2n}(\mathbf{b}_2)$ it is enough to show that all elements of X are below $\alpha_{2n}(\mathbf{b}_2)$, because $\alpha_{2n+1}(\mathbf{a}_n)$ is the largest element of the set X. Therefore, let $\mathbf{b} \in \mathrm{dom}(\alpha_{2n})$ be such that $\mathbf{b} \leq_T \mathbf{a}_n$. Then we have that $\mathbf{b} \leq_T \mathbf{a}_n \leq_T \mathbf{b}_2$, and since α_{2n} is an order homomorphism this implies that $\alpha_{2n}(\mathbf{b}) \leq_T \alpha_{2n}(\mathbf{b}_2)$, as desired.

Finally, we need to show that if $\mathbf{c} \in \mathrm{dom}(\alpha_{2n})$ is such that $\alpha_{2n+1}(\mathbf{c})$ and $\alpha_{2n+1}(\mathbf{a}_n)$ are incomparable, then $\mathbf{c} \oplus \mathbf{a}_n \notin \mathcal{A}$. If $\alpha_{2n+1}(\mathbf{c})$ and $\alpha_{2n+1}(\mathbf{a}_n)$ are incomparable, there has to be some $\mathbf{b} \leq_T \mathbf{a}_n$ with $\mathbf{b} \in \mathrm{dom}(\alpha_{2n})$ such that $\alpha_{2n}(\mathbf{c})$ and $\alpha_{2n}(\mathbf{b})$ are incomparable (because $\alpha_{2n+1}(\mathbf{a}_n)$ is the largest element of X). However, then by induction hypothesis $\mathbf{b} \oplus \mathbf{c} \notin \mathcal{A}$ and because $\mathcal{A}$ is downwards closed this also implies that $\mathbf{c} \oplus \mathbf{a}_n \notin \mathcal{A}$.

We now assume that α_{2n+1} has been defined and consider requirement R_{2n+2}. Let $\mathcal{B} = \{\mathbf{b} \in \mathrm{dom}(\alpha_{2n+1}) \mid \mathbf{b} \not\leq_T \mathbf{a}_n\}$. Since $\mathcal{A}$ is a splitting class there exist $\mathbf{c}_0, \mathbf{c}_1 \in \mathcal{A}$ such that $\mathbf{c}_0, \mathbf{c}_1 \geq \mathbf{a}_n$, $\mathbf{c}_0 \oplus \mathbf{c}_1 \notin \mathcal{A}$ and for all $\mathbf{b} \in \mathcal{B}$ we have $\mathbf{b} \oplus \mathbf{c}_0, \mathbf{b} \oplus \mathbf{c}_1 \notin \mathcal{A}$. Now extend α_{2n+1} by letting $\alpha_{2n+2}(\mathbf{c}_0) = \alpha_{2n+1}(\mathbf{a}_n)^\frown 0$ and $\alpha_{2n+2}(\mathbf{c}_1) = \alpha_{2n+1}(\mathbf{a}_n)^\frown 1$.

First, we show that α_{2n+2} is an order homomorphism. Let $\mathbf{b}_1, \mathbf{b}_2 \in \mathrm{dom}(\alpha_{2n+2})$ and $\mathbf{b}_1 \leq_T \mathbf{b}_2$. We again distinguish several cases:

- $\mathbf{b}_1, \mathbf{b}_2 \in \mathrm{dom}(\alpha_{2n+1})$: this directly follows from the fact that α_{2n+1} is an order homomorphism by induction hypothesis.
- $\mathbf{b}_1, \mathbf{b}_2 \in \{\mathbf{c}_0, \mathbf{c}_1\}$: since $\mathbf{c}_0 \oplus \mathbf{c}_1 \notin \mathcal{A}$ and therefore differs from both $\mathbf{c}_0$ and $\mathbf{c}_1$, this can only happen if $\mathbf{b}_1 = \mathbf{b}_2$, so this case is trivial.
- $\mathbf{b}_1 \in \{\mathbf{c}_0, \mathbf{c}_1\}$, $\mathbf{b}_2 \in \mathrm{dom}(\alpha_{2n+1})$: note that $\mathbf{c}_0, \mathbf{c}_1 >_T \mathbf{a}_n$ (otherwise $\mathbf{c}_0 \oplus \mathbf{c}_1 \in \{\mathbf{c}_0, \mathbf{c}_1\} \subseteq \mathcal{A}$), so we see that $\mathbf{b}_2 >_T \mathbf{a}$, and then by construction of $\mathbf{c}_0$ and $\mathbf{c}_1$ we know that $\mathbf{b}_2 \oplus \mathbf{c}_0, \mathbf{b}_2 \oplus \mathbf{c}_1 \notin \mathcal{A}$. This contradicts $\mathbf{b}_1 \leq_T \mathbf{b}_2$, so this case is impossible.
- $\mathbf{b}_1 \in \mathrm{dom}(\alpha_{2n+1})$, $\mathbf{b}_2 \in \{\mathbf{c}_0, \mathbf{c}_1\}$: if $\mathbf{b}_1 \not\leq_T \mathbf{a}_n$, then again by construction of $\mathbf{c}_0$ and $\mathbf{c}_1$ we have that $\mathbf{b}_2 = \mathbf{b}_1 \oplus \mathbf{b}_2 \notin \mathcal{A}$ which is a contradiction. So $\mathbf{b}_1 \leq_T \mathbf{a}_n$ and therefore $\alpha_{2n+2}(\mathbf{b}_1) = \alpha_{2n+1}(\mathbf{b}_1) \subseteq \alpha_{2n+1}(\mathbf{a}_n) \subseteq \alpha_{2n+2}(\mathbf{b}_2)$.

Finally, we show that if $\mathbf{b} \in \mathrm{dom}(\alpha_{2n+2})$ is such that $\alpha_{2n+2}(\mathbf{b})$ and $\alpha_{2n+2}(\mathbf{c}_1)$ are incomparable, then $\mathbf{b} \oplus \mathbf{c}_1 \notin \mathcal{A}$ (the same then follows analogously for $\mathbf{c}_2$). If $\mathbf{b} = \mathbf{c}_2$ this is clear from the definition of α_{2n+2}. Otherwise, we have $\mathbf{b} \in \mathrm{dom}(\alpha_{2n+1})$. If it were the case that $\mathbf{b} \leq_T \mathbf{a}_n$, then $\alpha_{2n+2}(\mathbf{b}) = \alpha_{2n+1}(\mathbf{b}) \subseteq \alpha_{2n+1}(\mathbf{a}_n) \subseteq \alpha_{2n+2}(\mathbf{c}_1)$, a contradiction. Thus $\mathbf{b} \not\leq_T \mathbf{a}_n$, and therefore $\mathbf{b} \oplus \mathbf{c}_1 \notin \mathcal{A}$ by construction of $\mathbf{c}_1$. □

THEOREM 3.3.7. *For any splitting class $\mathcal{A}$:* $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}}) = \mathrm{IPC}$.

PROOF. From Proposition 3.3.2, Proposition 3.3.4, Theorem 3.3.5 and Proposition 3.3.6. □

Therefore, combining this with the results from section 3.2 we now see:

THEOREM 3.3.8. $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{low}}}) = \mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{gen}\leq\emptyset'}}) = \mathrm{IPC}$.

3.4. Further splitting classes

3.4.1. Hyperimmune-free functions. In this section, we will look at some other classes and consider if they are splitting classes. First, we look at the class of hyperimmune-free functions. Recall that a function f is *hyperimmune-free* if every $g \leq_T f$ is dominated by a computable function. We can see a problem right away: the class of hyperimmune-free functions is well-known to be uncountable, while we required splitting classes to be countable. We temporarily remedy this by only looking at the hyperimmune-free functions which are low$_2$ (where a function f is low$_2$ if $f'' \equiv_T \emptyset''$); after the proof, we will discuss how we might be able to look at the entire class.[1]

As in section 3.2 we will present our constructions as constructions on Cantor space rather than Baire space for the reasons discussed in that section.

THEOREM 3.4.1. *The class $\mathcal{A}_{\mathrm{HIF},\mathrm{low}_2}$ of hyperimmune-free functions which are low$_2$ is a splitting class. In particular,* $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{HIF},\mathrm{low}_2}}) = \mathrm{IPC}$.

PROOF. We prove that (iii) of Proposition 3.1.2 holds. That for every hyperimmune-free low$_2$ set A there exists a hyperimmune-free low$_2$ set B such that $B \not\leq_T A$ (or that there even exists one such that $B >_T A$) is well-known, see Miller and Martin [81, Theorem 2.1]. We prove the second part of (iii) from Proposition 3.1.2. Our construction uses the tree method of Miller and Martin [81].

Let $A \leq_T \emptyset''$ be hyperimmune-free and low$_2$, let

$$\mathcal{B} \subseteq \{B \subseteq \omega \mid B \not\leq_T A, B \leq_T \emptyset'' \text{ and } B \text{ HIF}\}$$

be a finite subset and let $C_0 \leq_T \emptyset''$ be a hyperimmune-free (low$_2$) set not below A. We need to construct a hyperimmune-free set $A <_T C_1 \leq_T \emptyset''$ such that $C_0 \oplus C_1$ is not of hyperimmune-free degree (i.e. of hyperimmune degree) and such that for all $B \in \mathcal{B}$ we have that $C_1 \not\geq_T B$.

First, we remark that we may assume that not only $C_0 \not\leq_T A$, but even that $C_0 \not\leq_T A'$. Indeed, assume $C_0 \leq_T A'$. If $C_0 \geq_T A$ then we see that $A < C_0 \leq A'$ so by Miller and Martin [81, Theorem 1.2] we see that C_0 is of hyperimmune degree, contrary to our assumption. So, $C_0 \mid_T A$. However, then $A <_T A \oplus C_0 \leq A'$ and as before we then see that $A \oplus C_0$ is already of hyperimmune degree, so we may take C_1 to be any hyperimmune-free set strictly above A which is low$_2$; as mentioned above such a set exists by [81, Theorem 2.1].

Without loss of generality we may even assume that C_0 is not c.e. in A': we may replace C_0 by $C_0 \oplus \overline{C_0}$, which is of the same Turing degree as C_0, and is not c.e. in A' because otherwise C_0 would be computable in A', a contradiction.

Let $\mathcal{B} = \{B_1, \dots, B_n\}$ and fix a computable enumeration α of $n \times \omega$. We will construct a sequence $T_0 \supseteq T_1 \supseteq \dots$ of A-computable binary trees in the sense of Shoenfield [105] (see e.g. Odifreddi [87, Definition V.5.1]) such that:

[1] There are different natural countable subsets of the hyperimmune-free degrees which form splitting classes; for example, instead of the low$_2$ hyperimmune-free functions we could also take the hyperimmune-free functions f for which there exists an $n \in \omega$ such that $f \leq_T \emptyset^{(n)}$. This follows from the proof of Theorem 3.4.1. However, since our main reason to look at these countable subclasses is to view them as a stepping stone towards the class of all hyperimmune-free functions, we will not pursue this topic further.

(i) T_0 is the full binary tree.
(ii) For all D on T_{4e+1}: $D \neq \{e\}^A$.
(iii) For T_{4e+2}, one of the following holds:
   (a) For all D on T_{4e+2}, $\{e\}^{A\oplus D}$ is not total.
   (b) For all D on T_{4e+2}, $\{e\}^{A\oplus D}$ is total and

$$\forall n \forall \sigma (|\sigma| = n \rightarrow \{e\}^{A\oplus T_{4e+2}(\sigma)}(n)[|T_{4e+2}(\sigma)|] \downarrow).$$

   Furthermore, this choice is computable in $\emptyset''$.
(iv) For all D on T_{4e+3}, $\{\alpha_2(e)\}^{A\oplus D} \neq B_{\alpha_1(e)}$.
(v) T_{4e+4} is the full subtree of T_{4e+3} above $T_{4e+3}(\langle\emptyset''(e)\rangle)$.
(vi) For every infinite branch D on all of the trees T_i, the sequence $T_0 \supseteq T_1 \supseteq \dots$ is computable in $C_0 \oplus (A' \oplus D)$.
(vii) The sequence $T_0 \supseteq T_1 \supseteq \dots$ is computable in $\emptyset''$.

For now, assume we can construct such a sequence. Let $D = \bigcup_{i\in\omega} T_i(\emptyset)$, then D is an infinite branch lying on all of the T_i. Let $C_1 = A \oplus D$. Then the requirements (ii) guarantee that $D \not\leq_T A$ and therefore $C_1 >_T A$. By (vii) we also have that $C_1 \leq_T \emptyset''$. Furthermore, the requirements (iii) enforce that C_1 is hyperimmune-free relative to A by Miller and Martin [81] (see Odifreddi [87, Proposition V.5.6]), and because A is itself hyperimmune-free it is directly seen that C_1 is hyperimmune-free. The requirements (iv) ensure that $C_1 \not\geq_T B_i$ for all $B_i \in \mathcal{B}$.

Next, we have that $(C_0 \oplus C_1)' \geq_T C_0 \oplus (A' \oplus D) \geq_T \emptyset''$: by requirement (vi) the sequence T_i is computable in $C_0 \oplus (A' \oplus D)$, while by requirement (v) we have that $T_{4e+4}(\emptyset) = T_{4e+3}(\emptyset)^\frown\emptyset''(e)$ which allows us to recover $\emptyset''(e)$. So, $C_0 \oplus C_1$ is not low$_2$. In fact, $C_0 \oplus C_1$ is not even hyperimmune-free: by a theorem of Martin [76] we know that $(C_0 \oplus C_1)' \geq_T \emptyset''$ implies that $C_0 \oplus C_1$ computes a function which dominates every total computable function, and therefore $C_0 \oplus C_1$ is not hyperimmune-free, as desired.

Finally, we show that C_1 is low$_2$. By requirement (iv) and requirement (vii) we have that $\emptyset'' \geq_T \{e \in \omega \mid \{e\}^{C_1} \text{ is total}\}$. Since the latter has the same Turing degree as C_1'', this shows that C_1 is indeed low$_2$.

We now show how to actually construct such a sequence of computable binary trees. Let T_0 be the full binary tree. Next, assume T_{4e} has already been defined. To fulfil requirement (ii), observe that $T_{4e}(0)$ and $T_{4e}(1)$ are incompatible, so at least one of them has to differ from $\{e\}^A$. If the first differs from $\{e\}^A$ we take T_{4e+1} to be the full subtree above $T_{4e}(0)$, and otherwise we take the full subtree above $T_{4e}(1)$.

Next, assume T_{4e+1} has been defined, we will construct T_{4e+2} fulfilling requirement (iii). Let n be the smallest $m \in \omega$ such that either

$$m \notin C_0 \wedge \exists \sigma \supseteq \langle 0^m 1\rangle \exists x \forall \tau \supseteq \sigma \left(\{e\}^{A\oplus T_{4e+1}(\tau)}(x) \uparrow\right) \tag{4}$$

or

$$m \in C_0 \wedge \forall \sigma \supseteq \langle 0^m 1\rangle \forall x \exists \tau \supseteq \sigma \left(\{e\}^{A\oplus T_{4e+1}(\tau)}(x) \downarrow\right), \tag{5}$$

where as before $\langle 0^m 1\rangle$ denotes the string consisting of m times a 0 followed by a 1.

Such an m exists: indeed, if such an m did not exist, then

$$C_0 = \left\{ m \in \omega \mid \exists \sigma \supseteq \langle 0^m 1 \rangle \exists x \forall \tau \supseteq \sigma \left(\{e\}^{A \oplus T_{4e+1}(\tau)}(x) \uparrow \right) \right\}$$

and therefore C_0 is c.e. in A', which contradicts our assumption above.

If (4) holds for n, let $\sigma \supseteq \langle 0^n 1 \rangle$ be the smallest such string and let T_{4e+2} be the full subtree above $T_{4e+1}(\sigma)$. Otherwise, we inductively define $T_{4e+2} \subseteq T_{4e+1}$. First, if we let τ be the least $\tilde{\tau} \supseteq \langle 0^n 1 \rangle$ such that $\{e\}^{A \oplus T_{4e+1}(\tilde{\tau})}(0)[|T_{4e+1}(\tilde{\tau})|] \downarrow$, then we let $T_{4e+2}(0) = T_{4e+1}(\tau)$. Inductively, given $T_{4e+2}(\sigma)$, let ρ be such that $T_{4e+2}(\sigma) = T_{4e+1}(\rho)$. Now, if we let τ be the least $\tilde{\tau} \supseteq \rho$ such that $\{e\}^{A \oplus T_{4e+1}(\tilde{\tau})}(|\sigma| + 1)[|T_{4e+1}(\tilde{\tau})|] \downarrow$, we let $T_{4e+2}(\sigma^\frown 0) = T_{4e+1}(\tau^\frown 0)$ and $T_{4e+2}(\sigma^\frown 1) = T_{4e+1}(\tau^\frown 1)$.

For the requirements (iv) we do something similar. Let $\tilde{e} = \alpha_2(e)$. First, we build a subtree $S \subseteq T_{4e+2}$ such that either there is no $\tilde{e}$-splitting relative to A on S (i.e. for all strings σ, τ on S and all $x \in \omega$, if $\{\tilde{e}\}^{A \oplus \sigma}(x) \downarrow$ and $\{\tilde{e}\}^{A \oplus \tau}(x) \downarrow$, then their values are equal), or $S(0)$ and $S(1)$ are an $\tilde{e}$-splitting relative to A (in fact, S will even be an $\tilde{e}$-splitting tree relative to A). Let n be the smallest $m \in \omega$ such that

$$\begin{aligned} & m \notin C_0 \wedge \exists \sigma \supseteq \langle 0^m 1 \rangle \forall \tau, \tau' \supseteq \sigma \forall x \Big(\{\tilde{e}\}^{A \oplus T_{4e+2}(\tau)}(x) \downarrow \wedge \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau')}(x) \downarrow \\ & \qquad \rightarrow \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau)}(x) = \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau')}(x) \Big) \end{aligned} \tag{6}$$

or

$$\begin{aligned} & m \in C_0 \wedge \forall \sigma \supseteq \langle 0^m 1 \rangle \exists \tau, \tau' \supseteq \sigma \exists x \Big(\{\tilde{e}\}^{A \oplus T_{4e+2}(\tau)}(x) \downarrow \wedge \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau')}(x) \downarrow \\ & \qquad \wedge \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau)}(x) \neq \{\tilde{e}\}^{A \oplus T_{4e+2}(\tau')}(x) \Big) \end{aligned} \tag{7}$$

That such an m exists can be shown in the same way as above. If (6) holds for n, let σ be the smallest such string and let S be the full subtree above $T_{4e+2}(\sigma)$. Then there are no $\tilde{e}$-splittings relative to A on S. Otherwise, we can inductively build S: let $S(\emptyset) = T_{4e+2}(\langle 0^n 1 \rangle)$ and if $S(\sigma)$ is already defined we can take $S(\sigma^\frown 0)$ and $S(\sigma^\frown 1)$ to be two $\tilde{e}$-splitting extensions relative to A of $S(\sigma)$ on T_{4e+2}.

If there are no $\tilde{e}$-splittings relative to A on S, then we can take $T_{4e+3} = S$. Since, assume $\{\tilde{e}\}^{A \oplus D} = B_i$ for some $B_i \in \mathcal{B}$. Then, by Spector's result [114] (see e.g. Odifreddi [87, Proposition V.5.9]) we have that $B_i \leq_T A$, contrary to assumption.

Otherwise we can find an $x \in \omega$ such that $\{\tilde{e}\}^{A \oplus S(\langle 0 \rangle)}(x)$ and $\{\tilde{e}\}^{A \oplus S(\langle 1 \rangle)}(x)$ both converge, but such that their value differs. Then either $\{\tilde{e}\}^{A \oplus S(\langle 0 \rangle)}(x) \neq B_{\alpha_1(e)}$ and we take T_{4e+3} to be the full subtree above $S(\langle 0 \rangle)$, or $\{\tilde{e}\}^{A \oplus S(\langle 1 \rangle)}(x) \neq B_{\alpha_1(e)}$ and we take T_{4e+3} to be the full subtree above $S(\langle 1 \rangle)$. Then T_{4e+3} satisfies requirement (iv).

Finally, how to define T_{4e+4} from T_{4e+3} is already completely specified by requirement (v). This completes the definitions of all the T_i. Note that all steps in the construction are computable in $A'' \equiv_T \emptyset''$.

So, the last thing we need to show is that requirement (vi) is satisfied, i.e. that for any infinite branch D on all T_i the construction is computable in $C_0 \oplus (A' \oplus D)$. This is clear for the construction of T_{4e+1} from T_{4e}. For the construction of T_{4e+2} from T_{4e+1} the only real problem is that we need to choose between (4) and (5).

However, because D is on T_{4e+2}, we can uniquely determine $n \in \omega$ such that $T_{4e+1}(\langle 0^n 1 \rangle)$ is an initial segment of D. Then (4) holds if and only if $n \notin C_0$ and (5) holds if and only if $n \in C_0$. So, we can decide which alternative was taken using C_0. Furthermore, if (4) holds then we can use A' to calculate the string σ used in the computation of T_{4e+2}.

For T_{4e+3} we can do something similar for the tree S used in the definition of T_{4e+3}, and using D we can determine if we took T_{4e+3} to be the subtree above $S(\langle 0 \rangle)$ or $S(\langle 1 \rangle)$. Finally, using D it is also easily decided which alternative we took for T_{4e+4}, because T_{4e+4} is the full subtree above $T_{4e+3}(\langle i \rangle)$ for the unique $i \in \{0, 1\}$ such that $T_{4e+3}(\langle i \rangle) \subseteq D$. Therefore we see that the construction is indeed computable in $C_0 \oplus (A' \oplus D)$, which completes our proof. $\square$

This result is slightly unsatisfactory because we restricted ourselves to the hyperimmune-free which are low$_2$. Because the entire class of hyperimmune-free functions $\mathcal{A}_{\mathrm{HIF}}$ is also downwards closed we directly see from the proof above that the only real problem is the uncountability, i.e. $\mathcal{A}_{\mathrm{HIF}}$ satisfies all properties of a splitting class except for the countability. Our next result shows that, if we assume the continuum hypothesis, we can still show that the theory of the factor given by $\mathcal{A}_{\mathrm{HIF}}$ is IPC.

DEFINITION 3.4.2. Let $\mathcal{A} \subseteq \omega^\omega$ be a non-empty class of cardinality $\aleph_1$ which is downwards closed under Turing reducibility. We say that $\mathcal{A}$ is an $\aleph_1$ *splitting class* if for every $f \in \mathcal{A}$ and every countable subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exist $h_0, h_1 \in \mathcal{A}$ such that $h_0, h_1 \geq_T f$, $h_0 \oplus h_1 \notin \mathcal{A}$ and for all $g \in \mathcal{B}$: $g \oplus h_0, g \oplus h_1 \notin \mathcal{A}$.

PROPOSITION 3.4.3. *Let $\mathcal{A} \subseteq \omega^\omega$ be a non-empty class of cardinality $\aleph_1$ which is downwards closed under Turing reducibility. Then the following are equivalent:*

(i) *$\mathcal{A}$ is an $\aleph_1$ splitting class.*
(ii) *For every $f \in \mathcal{A}$ and every countable subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exists $h \in \mathcal{A}$ such that $h >_T f$ and for all $g \in \mathcal{B}$: $g \oplus h \notin \mathcal{A}$.*

Furthermore, if every countable chain in $\mathcal{A}$ has an upper bound in $\mathcal{A}$, these two are also equivalent to:

(iii) *For every $f \in \mathcal{A}$ there exists $h \in \mathcal{A}$ such that $h \not\leq_T f$, and for every $f \in \mathcal{A}$, every countable subset $\mathcal{B} \subseteq \{g \in \mathcal{A} \mid g \not\leq_T f\}$ and every $h_0 \in \{g \in \mathcal{A} \mid g \not\leq_T f\}$ there exists $h_1 \in \mathcal{A}$ such that $h_1 >_T f$, $h_0 \oplus h_1 \notin \mathcal{A}$ and for all $g \in \mathcal{B}$: $h_1 \not\geq_T g$.*

PROOF. In almost exactly the same way as Proposition 3.1.2. For the implication (iii) $\to$ (ii) we define an infinite sequence $h_{1,0} <_T h_{1,1} <_T \ldots$ instead of a finite one, and then let h be an upper bound in $\mathcal{A}$ of this chain. $\square$

THEOREM 3.4.4. *For any $\aleph_1$ splitting class $\mathcal{A}$:* $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}}) = \mathrm{IPC}$.

PROOF. We can generalise the construction in Proposition 3.3.6 to a transfinite construction over $\aleph_1$. However, instead of building a p-morphism to $2^{<\omega}$ we show that we can build a p-morphism to every finite tree T. It is known from folklore that this is already enough to show that the theory is IPC (see e.g. Chagrov and Zakharyaschev [17, Corollary 2.33]).[2]

Fix an enumeration $(\mathbf{a}_\gamma)_{\gamma<\aleph_1}$ of $\mathcal{A}$. This time we will build a sequence $(\alpha_\gamma)_{\gamma<\aleph_1}$ of partial order homomorphisms from $\mathcal{A}$ to T with countable domain, which is increasing in the sense that $\alpha_\gamma \subseteq \alpha_{\tilde{\gamma}}$ if $\gamma \leq \tilde{\gamma}$. As before, it should additionally satisfy that if $\mathbf{a}, \mathbf{b} \in \mathrm{dom}(\alpha_\gamma)$ and $\alpha_\gamma(\mathbf{a}) \mid \alpha_\gamma(\mathbf{b})$, then $\mathbf{a} \oplus \mathbf{b} \notin \mathcal{A}$.

Fix some bijection $\zeta : \{0,1\} \times \aleph_1 \to \aleph_1 \setminus \{0\}$ satisfying that $\zeta(1,\gamma) > \zeta(0,\gamma)$ for every $\gamma < \aleph_1$. We satisfy the requirements:

- R_0: $\alpha_0(\mathbf{0}) = \emptyset$
- $R_{(0,\gamma)}$: $\mathbf{a}_\gamma \in \mathrm{dom}(\alpha_{\zeta(0,\gamma)})$
- $R_{(1,\gamma)}$: if $\sigma := \alpha_{\zeta(0,\gamma)}(\mathbf{a}_\gamma)$ is not maximal in T and has children $\tau_1, \dots, \tau_n$, then there are $\mathbf{c}_1, \dots, \mathbf{c}_n \in \mathrm{dom}(\alpha_{\zeta(1,\gamma)})$ with $\mathbf{c}_1, \dots, \mathbf{c}_n \geq_T \mathbf{a}_\gamma$ and $\alpha_{\zeta(1,\gamma)}(\mathbf{c}_i) = \tau_i$ for every $1 \leq i \leq n$.

That these requirements give us the required p-morphisms follows in the same way as in Proposition 3.3.6. The construction of the sequence $(\alpha_\gamma)_{\gamma<\aleph_1}$ also proceeds in almost the same way, apart from three minor details. First, if γ is a limit ordinal it does not have a clear predecessor, so we cannot say that α_γ should extend its predecessor. Instead, we construct α_γ as an extension of $\bigcup_{\tilde{\gamma}<\gamma} \alpha_{\tilde{\gamma}}$ (note that this union is countable because $\gamma < \aleph_1$, and hence $\bigcup_{\tilde{\gamma}<\gamma} \alpha_{\tilde{\gamma}}$ has countable domain).

Secondly, the domains of the α_γ are no longer finite but are now countable, which means that in the construction for requirement $R_{(1,\gamma)}$ we now need to consider countable sets $\mathcal{B}$ instead of just finite sets $\mathcal{B}$. However, this is exactly why we changed our definition of an $\aleph_1$ splitting class to allow countable sets $\mathcal{B}$ instead of just finite sets $\mathcal{B}$.

Finally, instead of dealing with only two children we may have to deal with up to n children at the same time; however, that we can also split into n points instead of just into two follows in the same way as the proofs of Propositions 3.1.2 and 3.4.3. $\square$

THEOREM 3.4.5. *Assume* CH. *Then* $\mathcal{A}_{\mathrm{HIF}}$ *is an* $\aleph_1$ *splitting class. In particular,* $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{HIF}}}) = \mathrm{IPC}$.

PROOF. First, $\mathcal{A}_{\mathrm{HIF}}$ has cardinality $\aleph_1$ by CH. Next, every countable chain in $\mathcal{A}_{\mathrm{HIF}}$ has an upper bound in $\mathcal{A}_{\mathrm{HIF}}$ (Miller and Martin [81, Theorem 2.2]), so we can use the equivalence of (i) and (iii) of Proposition 3.4.3. Thus, it is sufficient if we show that the construction in Theorem 3.4.1 not only applies to just finite sets $\mathcal{B}$, but also to countable sets $\mathcal{B}$. However, this is readily verified. $\square$

In particular, we see that it is consistent (relative to ZFC) to have that $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{HIF}}}) = \mathrm{IPC}$. Unfortunately, we currently do not know if this already follows from ZFC or if it is independent of ZFC.

QUESTION 3.4.6. Does $\mathrm{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\mathrm{HIF}}}) = \mathrm{IPC}$ follow from ZFC?

[2] In the published version, the author incorrectly claimed that it is enough only to consider finite binary trees. This is false. However, as shown below the original proof also works for all finite trees, so the result still holds without any major changes.

3.4.2. Computably traceable functions. A class that is closely related to the hyperimmune-free functions is the class $\mathcal{A}_{\text{trace}}$ of computably traceable functions. We first recall its definition.

DEFINITION 3.4.7. (Terwijn and Zambella [120]) A set $T \subseteq \omega \times \omega$ is called a *trace* if all sections $T^{[k]} = \{n \in \omega \mid (k, n) \in T\}$ are finite. A *computable trace* is a trace such that the function which maps k to the canonical index of $T^{[k]}$ is computable. A trace T *traces* a function g if $g(k) \in T^{[k]}$ for every $k \in \omega$. A *bound* is a function $h : \omega \to \omega$ that is non-decreasing and has infinite range. If $|T^{[k]}| \leq h(k)$ for all $k \in \omega$, we say that h is a *bound for* T.

Finally, a function f is called *computably traceable* if there exists a computable bound h such that all (total) functions $g \leq_T f$ are traced by a computable trace bounded by h.

Computable traceability can be seen as a uniform kind of hyperimmune-freeness. If f is computably traceable, then it is certainly hyperimmune-free: if $g \leq_T f$ is traced by some computable trace T, then for the computable function $\tilde{g}(k) = \max\left(T^{[k]}\right)$ we have $g \leq \tilde{g}$. Conversely, if f is hyperimmune-free and $g \leq_T f$, then g has a computable trace: fix some computable $\tilde{g} \geq g$ and let $T_g = \{(k, m) \mid m \leq \tilde{g}(k)\}$. However, these traces T_g need not be bounded by any uniform computable bound h. Computable traceability asserts that such a uniform bound does exist. It can be shown that there are hyperimmune-free functions which are not computably traceable, see Terwijn and Zambella [120].

The computably traceable functions naturally occur in algorithmic randomness. In [120] it is shown that the computably traceable functions are precisely those functions which are low for Schnorr null, and in Kjos-Hanssen, Nies and Stephan [52] it is shown that this class also coincides with the functions which are low for Schnorr randomness.

Terwijn and Zambella also showed that the usual Miller and Martin tree construction of hyperimmune-free degrees actually already yields a computably traceable degree. Combining their techniques with the next lemma, we can directly see that our constructions of hyperimmune-free degrees above can also be used to construct computably traceable degrees.

LEMMA 3.4.8. *Let A be computably traceable, and let B be computably traceable relative to A. Then B is computably traceable.*

PROOF. Let h_1 be a computable bound for the traces of functions computed by A and let $h_2 \leq_T A$ be a bound for the traces of functions computed by B. Because A is hyperimmune-free (as discussed above) h_2 is bounded by a computable function $\tilde{h}_2$.We claim: every function computed by B has a trace bounded by the computable function $h_1 \cdot \tilde{h}_2$.

To this end, let $g \leq_T B$. Fix a trace $T \leq_T A$ for g which is bounded by h_2 (and hence is also bounded by $\tilde{h}_2$). Then the function mapping k to the canonical index of $T^{[k]}$ is computable in A, so because A is computably traceable we can determine a computable trace S for this function which is bounded by h_1.

Finally, denote by $D_{e,n}$ the (at most) n smallest elements of the set D_e corresponding to the canonical index e; i.e. $D_{e,n}$ consists of the n smallest elements of D_e if $|D_e| \geq n$, and $D_{e,n} = D_e$ otherwise. Now let U be the computable trace such that $U^{[k]} = \bigcup_{e \in S^{[k]}} D_{e,\tilde{h}_2(k)}$. Then U is clearly bounded by $h_1 \cdot \tilde{h}_2$. It also traces g, because $g(k) \in T^{[k]}$ and for some $e \in S^{[k]}$ we have $T^{[k]} = D_{e,\tilde{h}_2(k)}$. $\square$

THEOREM 3.4.9. *The class $\mathcal{A}_{\text{trace,low}_2}$ of computably traceable functions which are low$_2$ is a splitting class. In particular,* $\text{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\text{trace,low}_2}}) = \text{IPC}$.

PROOF. As in Theorem 3.4.1. $\square$

THEOREM 3.4.10. *Assume* CH. *Then $\mathcal{A}_{\text{trace}}$ is an $\aleph_1$ splitting class. In particular,* $\text{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\text{trace}}}) = \text{IPC}$.

PROOF. As in Theorem 3.4.5. $\square$

QUESTION 3.4.11. Does $\text{Th}(\mathcal{M}_w/\overline{\mathcal{A}_{\text{trace}}}) = \text{IPC}$ follow from ZFC?

CHAPTER 4

Natural Factors of the Medvedev Lattice Capturing IPC

In the previous chapter we discussed natural principal factors of the Muchnik lattice capturing IPC. In this chapter we turn towards the uniform Medvedev lattice: we know there are principal factors of the Medvedev lattice capturing IPC, but are there any natural ones?

In this chapter we present progress towards an affirmative answer to the question formulated above, by showing that there are principal factors of the Medvedev lattice capturing IPC which are more natural than the one given by Skvortsova. These factors arise from the computability-theoretic notion of a *computably independent set*: that is, a set A such that for every $i \in \omega$ we have that $\bigoplus_{j \neq i} A^{[j]} \not\geq_T A^{[i]}$, where $A^{[i]}$ is the i^{th} column of A, i.e. $A^{[i]}(n) = A(\langle i, n \rangle)$. We can now state the main theorem of this chapter.

THEOREM 4.0.1. *Let A be a computably independent set. Then*

$$\mathrm{Th}\left(\mathcal{M}/\left\{i^\frown f \mid f \geq_T A^{[i]}\right\}\right) = \mathrm{IPC}.$$

We note that the factor from Theorem 4.0.1 is not nearly as natural as the factors for the Muchnik lattice described in the previous chapter. On the other hand, the factor from Theorem 4.0.1 is far more natural than the one given by Skvortsova: our factor is easily definable from just a computably independent set, which occurs naturally in computability theory. Furthermore, while Skvortsova used a deep result by Lachlan, we manage to work around this and therefore our proof is more elementary.

We also study a question posed by Sorbi and Terwijn in [112]. As mentioned above, the theory of the Medvedev lattice is equal to Jankov's logic Jan, the deductive closure of IPC plus the weak law of the excluded middle $\neg p \vee \neg\neg p$. Let $0'$ be the mass problem consisting of all non-computable functions. Recall that we say that a mass problem is *Muchnik* if it is upwards closed under Turing reducibility. In [112] it is shown that for all Muchnik $\mathcal{B} >_{\mathcal{M}} 0'$ the theory of the factor $\mathcal{M}/\mathcal{B}$ is contained in Jan. Therefore, Sorbi and Terwijn asked: is $\mathrm{Th}(\mathcal{M}/\mathcal{B})$ contained in Jan for all mass problems $\mathcal{B} >_{\mathcal{M}} 0'$? We give a positive answer to this question.

This chapter is based on Kuyper [67].

4.1. Upper implicative semilattice embeddings of $\mathcal{P}(I)$ into $\mathcal{M}$

As a first step, we will describe a method to embed Boolean algebras of the form $\mathcal{P}(I)$, ordered under reverse inclusion $\supseteq$, into the Medvedev lattice $\mathcal{M}$ as an upper implicative semilattice (i.e. preserving $\oplus$, $\rightarrow$, 0 and 1). It should be noted that we will only need this for finite I, and Skvortsova [107, Lemma 7] already showed that such embeddings exist. However, Skvortsova used Lachlan's result [70] that every countable distributive lattice can be order-theoretically embedded as an initial segment of the Turing degrees. Because we want natural factors of the Medvedev lattice, we want to avoid the use of this theorem. Our main result of this section will show that there are various natural embeddings of $\mathcal{P}(I)$ into $\mathcal{M}$. These embeddings are induced by so-called *strong upwards antichains*, where the notion of a strong upwards antichain is the order-dual of the notion of an antichain normally used in forcing.

Definition 4.1.1. Let $\mathcal{A} \subseteq \omega^\omega$ be downwards closed under Turing reducibility and let $(f_i)_{i \in I} \in \mathcal{A}^I$. Then we say that $(f_i)_{i \in I}$ is a *strong upwards antichain in $\mathcal{A}$* if for all $i \neq j$ we have that $f_i \oplus f_j \notin \mathcal{A}$.

Henceforth we will mean by *antichain* a *strong upwards antichain.*

Example 4.1.2. We give some examples of countably infinite antichains.

(i) Take $\mathcal{A}$ to be the computable functions together with the functions of minimal degree, and $f_0, f_1 \dots$ any sequence of functions of distinct minimal Turing degree.
(ii) Let $f_0, f_1, \dots$ be pairwise incomparable under Turing reducibility and take $\mathcal{A}$ to be the lower cone of $\{f_i \mid i \in \omega\}$.

The next theorem shows that each antichain induces an upper implicative semilattice embedding of $\mathcal{P}(I)$ in a natural way.

Theorem 4.1.3. *Let $\mathcal{A} \subseteq \omega^\omega$ be downwards closed under Turing reducibility, let $(f_i)_{i \in I}$ be an antichain in $\mathcal{A}$, and let $\mathcal{B} = \overline{\mathcal{A}} \cup C(\{f_i \mid i \in I\})$. Then the map α given by $\alpha(X) = \overline{\mathcal{A}} \cup C(\{f_i \mid i \in X\})$ is an upper implicative semilattice embedding of $(\mathcal{P}(I), \supseteq)$ into $\left[\mathcal{B}, \overline{\mathcal{A}}\right]_{\mathcal{M}}$.*

Proof. For ease of notation, if $X \subseteq I$ we will denote by $C(X)$ the set $C(\{f_i \mid i \in X\})$.

We have:

$$\alpha(X \cap Y) = \overline{\mathcal{A}} \cup C(X \cap Y).$$

On the other hand, because $\alpha(X)$ and $\alpha(Y)$ are upwards closed their join is just intersection (see Skvortsova [107, Lemma 5]), and therefore:

$$\alpha(X) \oplus \alpha(Y) \equiv_{\mathcal{M}} \overline{\mathcal{A}} \cup (C(X) \cap C(Y)).$$

Clearly, $\alpha(X \cap Y) \subseteq \overline{\mathcal{A}} \cup (C(X) \cap C(Y))$. Conversely, let $g \in \overline{\mathcal{A}} \cup (C(X) \cap C(Y))$. If $g \notin \mathcal{A}$ then clearly $g \in \alpha(X \cap Y)$. So, assume $g \in \mathcal{A}$. Let $i \in X, j \in Y$ be such that $g \geq_T f_i$ and $g \geq_T f_j$. Then $f_i \oplus f_j \leq_T g \in \mathcal{A}$ so $f_i \oplus f_j \in \mathcal{A}$. Since $(f_i)_{i \in I}$ is an antichain in $\mathcal{A}$ this can only be the case if $i = j$, so we see that $g \in \alpha(X \cap Y)$.

We also have, again by [107, Lemma 5]:

$$\begin{aligned}
&\alpha(X) \to_{\left[\mathcal{B},\overline{\mathcal{A}}\right]_{\mathcal{M}}} \alpha(Y) \\
&\quad \equiv_{\mathcal{M}} \mathcal{B} \oplus \{g \mid \forall h \in \alpha(X)(g \oplus h \in \alpha(Y))\} \\
&\quad \equiv_{\mathcal{M}} \{g \in \mathcal{B} \mid \forall i \in X \forall h \geq_T f_i \exists j \in Y(g \oplus h \in \mathcal{A} \to g \oplus h \geq_T f_j)\} \\
&\quad = \overline{\mathcal{A}} \cup \{g \in C\left(\{f_i \mid i \in I\}\right) \\
&\qquad\qquad\qquad \mid \forall i \in X \forall h \geq_T f_i \exists j \in Y(g \oplus h \in \mathcal{A} \to g \oplus h \geq_T f_j)\}.
\end{aligned}$$

Fix any $g \in \mathcal{A} \cap C\left(\{f_i \mid i \in I\}\right)$ such that

$$\forall i \in X \forall h \geq_T f_i \exists j \in Y(g \oplus h \in \mathcal{A} \to g \oplus h \geq_T f_j). \tag{8}$$

Then we know that there is some $k \in I$ such that $g \geq_T f_k$. We claim: $k \notin X$ or $k \in Y$.

Namely, assume $k \in X$ and $k \notin Y$. Then, by (8) (with $h = g$) there exists some $j \in Y$ such that $g \geq_T f_j$, and since $k \notin Y$ we know that $j \neq k$. But then $f_k \oplus f_j \leq_T g \in \mathcal{A}$ so $f_k \oplus f_j \in \mathcal{A}$, a contradiction with the fact that $(f_i)_{i \in I}$ is an antichain in $\mathcal{A}$.

Conversely, if $g \in \mathcal{A}$ is such that $g \geq_T f_k$ for some $k \notin X$ or some $k \in Y$, then (8) holds: namely, if $k \notin X$ then we have for all $i \in X$ that $g \oplus f_i \notin \mathcal{A}$ because $(f_i)_{i \in I}$ is an antichain in $\mathcal{A}$, while if $k \in Y$ we have that $g \oplus f_i \geq_T f_k$.

So, from this we see:

$$\begin{aligned}
\alpha(X) \to_{\left[\mathcal{B},\overline{\mathcal{A}}\right]_{\mathcal{M}}} \alpha(Y) &\equiv_{\mathcal{M}} \overline{\mathcal{A}} \cup C((I \setminus X) \cup Y) \\
&= \alpha((I \setminus X) \cup Y) \\
&= \alpha(X \to_{\mathcal{P}(I)} Y).
\end{aligned}$$

□

4.2. From embeddings of $\mathcal{P}(\omega)$ to factors capturing IPC

In this section we will show how to construct a more natural factor of the Medvedev lattice with IPC as its theory; that is, we will prove Theorem 4.0.1. For this proof we will use several ideas from Skvortsova's construction of a factor of the Medvedev lattice which has IPC as its theory, given in Skvortsova [107]. We combine these ideas with our own to get to the factor in Theorem 4.0.1. First, let us discuss canonical subsets of a Brouwer algebra.

Definition 4.2.1. ([107, p. 134]) Let $\mathscr{B}$ be a Brouwer algebra and let $\mathscr{C} \subseteq \mathscr{B}$. Then we call $\mathscr{C}$ *canonical* if:

(i) All elements in $\mathscr{C}$ are meet-irreducible,
(ii) $\mathscr{C}$ is closed under joins and implications (i.e. it is a sub-upper implicative semilattice),
(iii) For all $a \in \mathscr{C}$ and $b, c \in \mathcal{B}$ we have $a \to (b \otimes c) = (a \to b) \otimes (a \to c)$.

Proposition 4.2.2. ([107, Corollary to Lemma 6]) *The set of Muchnik degrees is a canonical subset of $\mathcal{M}$.*

Corollary 4.2.3. *The range of α from Theorem 4.1.3 is canonical in $[\alpha(I), \alpha(\emptyset)]_{\mathcal{M}}$.*

PROOF. The range of α consists of Muchnik degrees, so (i) holds by Proposition 4.2.2. Furthermore, α is an upper implicative semilattice embedding, and therefore (ii) also holds. Finally, if $\mathcal{C}_0, \mathcal{C}_1 \in [\alpha(I), \alpha(\emptyset)]_{\mathcal{M}}$ and $X \subseteq I$, then we see, using Proposition 4.2.2:

$$\begin{aligned}
\alpha(X) &\to_{[\alpha(I),\alpha(\emptyset)]_{\mathcal{M}}} (\mathcal{C}_0 \otimes \mathcal{C}_1) \\
&= (\alpha(X) \to_{\mathcal{M}} (\mathcal{C}_0 \otimes \mathcal{C}_1)) \oplus \alpha(I) \\
&\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \mathcal{C}_0) \otimes (\alpha(X) \to_{\mathcal{M}} \mathcal{C}_1)) \oplus \alpha(I) \\
&\equiv_{\mathcal{M}} (\alpha(X) \to_{[\alpha(I),\alpha(\emptyset)]_{\mathcal{M}}} \mathcal{C}_0) \otimes (\alpha(X) \to_{[\alpha(I),\alpha(\emptyset)]_{\mathcal{M}}} \mathcal{C}_1).
\end{aligned}$$ □

PROPOSITION 4.2.4. ([107, Lemma 2]) *If $\mathscr{C}$ is a canonical set in a Brouwer algebra $\mathscr{B}$, then the smallest sub-Brouwer algebra of $\mathscr{B}$ containing $\mathscr{C}$ is $\{a_1 \otimes \cdots \otimes a_n \mid a_i \in \mathscr{C}\}$, and it is isomorphic to the free Brouwer algebra over the upper implicative semilattice $\mathscr{C}$ through an isomorphism fixing $\mathscr{C}$.*

In particular, we see:

COROLLARY 4.2.5. *If we let α be the embedding of $(\mathcal{P}(I), \supseteq)$ from Theorem 4.1.3, then $\{\alpha(X_1) \otimes \cdots \otimes \alpha(X_n) \mid X_i \in \mathcal{P}(I)\}$ is a sub-Brouwer algebra of $[\alpha(I), \alpha(\emptyset)]_{\mathcal{M}}$ which is isomorphic to the free Brouwer algebra over the upper implicative semilattice $(\mathcal{P}(I), \supseteq)$.*

PROOF. From Corollary 4.2.3 and Proposition 4.2.4. □

Let $\mathscr{B}_n$ be the Brouwer algebra of the upwards closed subsets of $(\mathcal{P}(\{1,\dots,n\}) \setminus \{\emptyset\}, \supseteq)$ ordered under reverse inclusion $\supseteq$, i.e. the elements of $\mathscr{B}_n$ are those $A \subseteq \mathcal{P}(\{1,\dots,n\}) \setminus \emptyset$ such that if $X \in A$ and $Y \in \mathcal{P}(\{1,\dots,n\}) \setminus \{\emptyset\}$ is such that $X \supseteq Y$, then $Y \in A$. We can use $\mathscr{B}_n$ to capture IPC in the following way:

PROPOSITION 4.2.6. ([107, the remark following Lemma 3])

$$\bigcap_{n>0} \bigcap_{x \in \mathscr{B}_n} \mathrm{Th}(\mathscr{B}_n / x) = \mathrm{IPC}.$$

PROOF. Let $\mathrm{LM} = \bigcap_{n>0} \mathrm{Th}(\mathscr{B}_n)$, the *Medvedev logic of finite problems.* Given a set of formulas X, let X^+ denote the set of positive (i.e. negation-free) formulas in X. Then $\mathrm{LM}^+ = \mathrm{IPC}^+$, see Medvedev [80].

Now, let $\varphi(x_1,\dots,x_m)$ be any formula. Let $\varphi'(x_1,\dots,x_{m+1})$ be the formula where x_{m+1} is a fresh variable and where $\bot$ is replaced by $x_1 \wedge \cdots \wedge x_{m+1}$, so φ' is negation-free. Then, if $\varphi \notin \mathrm{IPC}$, we have $\varphi' \notin \mathrm{IPC}^+$ (see Jankov [47]), so there are $n \in \omega$ and $x_1,\dots,x_{m+1} \in \mathscr{B}_n$ such that $\varphi'(x_1,\dots,x_{m+1}) \neq 0$. Let $x = x_1 \oplus \cdots \oplus x_{m+1}$, then $\varphi \notin \mathrm{Th}(\mathscr{B}_n/x)$. □

Furthermore, it is easy to obtain these $\mathscr{B}_n$ as free distributive lattices over upper implicative semilattices, as expressed by the following proposition.

PROPOSITION 4.2.7. ([107, Lemma 3]) *The Brouwer algebra $\mathscr{B}_n$ is isomorphic to the free distributive lattice over the upper implicative semilattice $(\mathcal{P}(\{1,\dots,n\}), \supseteq)$.*

Corollary 4.2.8. *Let I be a set of size n. If we let α be the embedding of $(\mathcal{P}(I), \supseteq)$ from Theorem 4.1.3, then $\{\alpha(X_1) \otimes \cdots \otimes \alpha(X_m) \mid m \in \omega \wedge \forall i \leq m(X_i \in \mathcal{P}(I))\}$ is a sub-Brouwer algebra of $[\alpha(I), \alpha(\emptyset)]_{\mathcal{M}}$ isomorphic to $\mathscr{B}_n$.*

Proof. From Corollary 4.2.5 and Proposition 4.2.7. □

The following lemma allows us to compare the theories of different intervals.

Lemma 4.2.9. ([107, Lemma 4]) *In any Brouwer algebra $\mathscr{B}$: if $x, y, z \in \mathscr{B}$ are such that $x \oplus z = y$, then $\mathrm{Th}([0, z]_{\mathscr{B}}) \subseteq \mathrm{Th}([x, y]_{\mathscr{B}})$.*

Proof. Let $\gamma : [0, z]_{\mathscr{B}} \to [x, y]_{\mathscr{B}}$ be given by $\gamma(u) = x \oplus u$. This map is well-defined, since if $u \leq z$, then $x \oplus u \leq x \oplus z = y$. Clearly γ preserves $\oplus$ and $\otimes$, while for $\to$ we have:

$$\gamma(u \to_{[0,z]_{\mathscr{B}}} v) = (u \to_{\mathscr{B}} v) \oplus x = ((u \oplus x) \to_{\mathscr{B}} (v \oplus x)) \oplus x = \gamma(u) \to_{[x,y]_{\mathscr{B}}} \gamma(v).$$

Furthermore, γ is surjective, so the result now follows from Lemma 2.1.15. □

Before we get to the proof of Theorem 4.0.1 we need one theorem from computability theory.

Theorem 4.2.10. *Let $A, E \in 2^\omega$ be such that $E \geq_T A'$. Let $B_0, B_1, \dots \in 2^\omega$ be uniformly computable in E and such that $A \not\geq_T B_i$. Then there exists a set $D \geq_T A$ such that $D' \leq_T E$ and such that for all $i \in \omega$ we have $D \oplus B_i \geq_T E$.*

Proof. This follows from relativising Posner and Robinson [95, Theorem 3] to A. □

Finally, we need an easy lemma on extending computably independent sets. For ease of notation, let us assume that our pairing function is such that $(A \oplus B)^{[2i]} = A^{[i]}$ and $(A \oplus B)^{[2i+1]} = B^{[i]}$.

Lemma 4.2.11. *Let A be a computably independent set. Then there exists a set B such that $A \oplus B$ is computably independent.*

Proof. Our requirements are as follows:

$$R_{\langle e,2n\rangle} : A^{[n]} \neq \{e\}^{\bigoplus_{i\neq 2n}(A\oplus B)^{[i]}}$$
$$R_{\langle e,2n+1\rangle} : B^{[n]} \neq \{e\}^{\bigoplus_{i\neq 2n+1}(A\oplus B)^{[i]}}.$$

We build B by the finite extension method, i.e. we define strings $\sigma_0 \subseteq \sigma_1 \subseteq \dots$ and let $B = \bigcup_{s\in\omega} \sigma_s$. For ease of notation, define $\sigma_{-1} = \emptyset$. At stage s, we deal with requirement R_s. There are two cases:

- $s = \langle e, 2n\rangle$: if there is a string σ extending σ_{s-1} and an $m \in \omega$ such that $\{e\}^{\bigoplus_{i\neq 2n}(A\oplus\sigma)^{[i]}}(m)\downarrow \neq A^{[n]}(m)$, take σ_s to be the least such σ. Otherwise, let $\sigma_s = \sigma_{s-1}$.
- $s = \langle e, 2n + 1\rangle$: if there exists a string σ extending σ_{s-1} such that $\{e\}^{\bigoplus_{i\neq 2n+1}(A\oplus\sigma)^{[i]}}(|\sigma_{s-1}|+1)\downarrow$, take the least such σ and let σ_s be the least string extending σ_{s-1} which coincides with σ outside the n^{th} column and such that $\sigma_s^{[n]}(|\sigma_{s-1}|+1) = 1 - \{e\}^{\bigoplus_{i\neq 2n+1}(A\oplus\sigma)^{[i]}}(|\sigma_{s-1}|+1)$. Otherwise, let $\sigma_s = \sigma_{s-1}$.

We claim: B is as required. To this end, we verify the requirements:

- $R_{\langle e,2n\rangle}$: towards a contradiction, assume $A^{[n]} = \{e\}^{\bigoplus_{i\neq 2n}(A\oplus B)^{[i]}}$. Let $s = \langle e, 2n\rangle$. By construction we then know for every σ extending σ_{s-1} and every $m \in \omega$ that, if $\{e\}^{\bigoplus_{i\neq 2n}(A\oplus\sigma)^{[i]}}(m)\downarrow$, we have $\{e\}^{\bigoplus_{i\neq 2n}(A\oplus\sigma)^{[i]}}(m) = A^{[n]}(m)$. Furthermore, for every $m \in \omega$ there is a string σ extending σ_{s-1} such that $\{e\}^{\bigoplus_{i\neq 2n}(A\oplus\sigma)^{[i]}}(m)\downarrow$: just take a suitably long initial segment of B. However, this means that $\bigoplus_{i\neq n} A^{[i]} \geq_T A^{[n]}$, which contradicts A being computably independent.
- $R_{\langle e,2n+1\rangle}$: let $s = \langle e, 2n+1\rangle$. Then by our construction we know that, if $\{e\}^{\bigoplus_{i\neq 2n+1}(A\oplus B)^{[i]}}(|\sigma_{s-1}|+1)\downarrow$, then it differs from $B^{[n]}(|\sigma_{s-1}|+1)$. □

We can now prove Theorem 4.0.1.

THEOREM 4.0.1. *Let A be a computably independent set. Then*

$$\mathrm{Th}\left(\mathcal{M}/\left\{i^\frown f \mid f \geq_T A^{[i]}\right\}\right) = \mathrm{IPC}.$$

PROOF. Fix $n \in \omega$ and $x \in \mathscr{B}_n$. Let $I = \{1, \dots, n\}$. For now assume we have some downwards closed $\mathcal{A}$ and an antichain $D_1, \dots, D_n \in \mathcal{A}$. Then Corollary 4.2.8 tells us that

$$\{\alpha(Y_1) \otimes \cdots \otimes \alpha(Y_m) \mid m \in \omega \wedge \forall i \leq m(Y_i \in \mathcal{P}(I))\}$$

is a subalgebra of $\left[\overline{\mathcal{A}} \cup C(\{D_1, \dots, D_n\}), \overline{\mathcal{A}}\right]_{\mathcal{M}}$ isomorphic to $\mathscr{B}_n$. So, there are $X_1, \dots, X_k \subseteq I$ such that we can embed $\mathscr{B}_n/x$ as subalgebra of

$$\left[\overline{\mathcal{A}} \cup C(\{D_1, \dots, D_n\}), \alpha(X_1) \otimes \cdots \otimes \alpha(X_k)\right]_{\mathcal{M}}.$$

If we would additionally have that

$$\left(\overline{\mathcal{A}} \cup C(\{D_1, \dots, D_n\})\right) \oplus \left\{i^\frown f \mid f \geq_T A^{[i]}\right\} \equiv_{\mathcal{M}} \alpha(X_1) \otimes \cdots \otimes \alpha(X_k), \tag{9}$$

then Lemma 4.2.9 tells us that

$$\begin{aligned}
&\mathrm{Th}\left(\mathcal{M}/\left\{i^\frown f \mid f \geq_T A^{[i]}\right\}\right)\\
&\qquad\subseteq \mathrm{Th}\left(\left[\overline{\mathcal{A}} \cup C(\{D_1, \dots, D_n\}), \alpha(X_1) \otimes \cdots \otimes \alpha(X_k)\right]_{\mathcal{M}}\right)\\
&\qquad\subseteq \mathrm{Th}(\mathscr{B}_n/x).
\end{aligned}$$

Now, if we would be able to do this for arbitrary $n \in \omega$ and $x \in \mathscr{B}_n$, then Proposition 4.2.6 tells us that

$$\mathrm{Th}\left(\mathcal{M}/\left\{i^\frown f \mid f \geq_T A^{[i]}\right\}\right) = \mathrm{IPC},$$

so then we would be done.

Thus, it suffices to show that for all $n \in \omega$ and all $X_1, \dots, X_k \subseteq \{1, \dots, n\}$ there exists a downwards closed $\mathcal{A}$ and an antichain $D_1, \dots, D_n \in \mathcal{A}$ such that (9) holds. Fix a B for A as in Lemma 4.2.11. Let $\mathcal{A} = \omega^\omega \setminus C(\{(A \oplus B)'\})$. For every $1 \leq i \leq n$ fix a $D_i \geq_T \left(\bigoplus_{1\leq j\leq k, i\in X_j} A^{[j]}\right) \oplus B^{[i]}$ such that $D_i' \leq_T (A \oplus B)'$, such that $D_i \oplus A^{[j]} \geq_T (A \oplus B)'$ for every $j \in \{1 \leq j \leq k \mid i \notin X_j\} \cup \{k+1, k+2, \dots\}$ and such that $D_i \oplus B^{[j]} \geq_T (A \oplus B)'$ for every $j \neq i$, which exists by Theorem 4.2.10.

We claim: $\{D_1, \dots, D_n\}$ is an antichain in $\mathcal{A}$. Clearly, $D_1, \dots, D_n \in \mathcal{A}$. Next, let $1 \leq i < j \leq n$. Then:

$$D_i \oplus D_j \geq_T B^{[i]} \oplus D_j \geq_T (A \oplus B)',$$

so $D_i \oplus D_j \notin \mathcal{A}$.

Thus, we need to show that (9) holds. First, let $g \in \overline{\mathcal{A}} \cup C(\{D_1, \dots, D_n\})$ and let $f \geq_T A^{[j]}$. If $j > k$, then either $g \geq_T D_i$ for some $1 \leq i \leq n$ and $f \oplus g \geq_T A^{[j]} \oplus D_i \geq_T (A \oplus B)'$, or $g \geq_T (A \oplus B)'$ and then also $f \oplus g \geq_T (A \oplus B)'$. In both cases we see that $f \oplus g \in \overline{\mathcal{A}} \subseteq \alpha(X_1)$.

Thus, we may assume that $j \leq k$. We claim: $f \oplus g \in \alpha(X_j)$. Indeed, if $g \geq_T D_i$ for some $i \in X_j$, then $f \oplus g \geq_T D_i$ and $C(D_i) \subseteq \alpha(X_j)$, while if $g \geq_T D_i$ for some $i \notin X_j$, then $f \oplus g \geq_T A^{[j]} \oplus D_i \geq_T (A \oplus B)'$, and finally, if $g \geq_T (A \oplus B)'$ then clearly $f \oplus g \geq_T (A \oplus B)'$. Thus, we see that $f \oplus g$ computes an element of $\alpha(X_1) \otimes \cdots \otimes \alpha(X_k)$, and that this computation is in fact uniform in $(j^\frown f) \oplus g$.

For the other direction, note that for fixed $1 \leq i \leq k$ we have that

$$C(X_i) \subseteq C(\{D_1, \dots, D_n\})$$

and also that

$$C(X_i) \subseteq \left\{f \mid f \geq_T A^{[i]}\right\}$$

because for every $j \in X_i$ we have that $D_j \geq_T A^{[i]}$. □

4.3. Relativising the construction

We will next show that Skvortsova's construction can be performed below every mass problem $\mathcal{B} >_{\mathcal{M}} 0'$. This also implies that for every $\mathcal{B} >_{\mathcal{M}} 0'$ we have that $\mathrm{Th}(\mathcal{M}/\mathcal{B}) \subseteq \mathrm{Jan}$, answering a question by Sorbi and Terwijn; see Corollary 4.3.3 below.

First, note that for every $\mathcal{B} > 0'$ we can find a countable mass problem $\mathcal{E} \subseteq 0'$ such that $\mathcal{E} \not\geq_{\mathcal{M}} \mathcal{B}$ (e.g. by taking one function for every $n \in \omega$ witnessing that $\Phi_n(0') \not\subseteq \mathcal{B}$). Then the set $\{A \mid \forall f \in \mathcal{E}(A \not\geq_T f)\}$ has measure 1 (by Sacks's result [101] that upper cones in the Turing degrees have measure 0, see e.g. Downey and Hirschfeldt [25, Corollary 8.12.2]), so it contains a 1-random set; in particular it contains a computably independent set A. In this section we will show that we can use such sets to obtain factors with theory IPC below $\mathcal{B}$, by relativising Theorem 4.0.1.

However, we first show that we can relativise Theorem 4.1.3 below $\mathcal{B}$.

THEOREM 4.3.1. *Let $\mathcal{B}$ be a mass problem, let $\mathcal{E}$ be a mass problem such that $\mathcal{E} \not\geq_{\mathcal{M}} \mathcal{B}$ and let $\mathcal{D} = \mathcal{E} \to_{\mathcal{M}} \mathcal{B}$. Let $\mathcal{A} \subseteq \omega^\omega$ be a mass problem which is downwards closed under Turing reducibility such that $\mathcal{E} \subseteq \overline{\mathcal{A}}$. Let $(f_i)_{i \in I}$ be an antichain in $\mathcal{A}$. Then the map β given by $\beta(X) = (\overline{\mathcal{A}} \cup \{g \mid \exists i \in X(g \geq_T f_i)\}) \otimes \mathcal{D}$ is an upper implicative semilattice embedding of $(\mathcal{P}(I), \supseteq)$ into $[\beta(I), \beta(\emptyset)]_{\mathcal{M}}$ with range canonical in $[\beta(I), \beta(\emptyset)]_{\mathcal{M}}$.*

PROOF. First, note that $\mathcal{E} \not\geq_{\mathcal{M}} \mathcal{D}$, since if $\mathcal{E} \geq_{\mathcal{M}} \mathcal{D}$ then

$$\mathcal{E} \equiv_{\mathcal{M}} \mathcal{E} \oplus \mathcal{D} = \mathcal{E} \oplus (\mathcal{E} \to \mathcal{B}) \geq_{\mathcal{M}} \mathcal{B},$$

a contradiction.

As in the proof of Theorem 4.1.3, if $X \subseteq I$ we will denote by $C(X)$ the set $C(\{f_i \mid i \in X\})$. By Theorem 4.1.3, the function $\alpha : \mathcal{P}(I) \to \mathcal{M}/\overline{\mathcal{A}}$ given by $\alpha(X) = \overline{\mathcal{A}} \cup C(X)$ is an upper implicative semilattice embedding of $(\mathcal{P}(I), \supseteq)$ into $\left[\overline{\mathcal{A}} \cup C(I), \overline{\mathcal{A}}\right]_{\mathcal{M}}$. Note that $\mathcal{E} \subseteq \overline{\mathcal{A}}$ and therefore $\mathcal{E} \subseteq \alpha(X)$ for every $X \subseteq I$.

Now let $\beta : \mathcal{P}(I) \to \mathcal{M}/\overline{\mathcal{A}}$ be the function given by $\beta(X) = \alpha(X) \otimes \mathcal{D}$. Then the range of β is certainly contained in $[\beta(I), \beta(\emptyset)]_{\mathcal{M}}$. We prove that β is in fact an upper implicative semilattice embedding into $[\beta(I), \beta(\emptyset)]_{\mathcal{M}}$ with canonical range.

- β is injective: assume $\beta(X) \leq_{\mathcal{M}} \beta(Y)$. Thus, we have $\alpha(X) \otimes \mathcal{D} \leq_{\mathcal{M}} \alpha(Y) \otimes \mathcal{D}$. In particular we have that $\alpha(X) \otimes \mathcal{D} \leq_{\mathcal{M}} \alpha(Y)$, say via Φ_n. We claim: $\Phi_n(\alpha(Y)) \subseteq 0^\frown\alpha(X)$.

  Namely, assume towards a contradiction that $\Phi_n(f) \in 1^\frown\mathcal{D}$ for some $f \in \alpha(Y)$. Determine $\sigma \subseteq f$ such that $\Phi_n(\sigma)(0) = 1$. As noted above we have that $\mathcal{E} \subseteq \alpha(Y)$, and since $\alpha(Y)$ is Muchnik we therefore see that $\sigma^\frown\mathcal{E} \subseteq \alpha(Y)$. However, then we can reduce $\mathcal{E}$ to $1^\frown\mathcal{D}$ by sending $g \in \mathcal{E}$ to $\Phi_n(\sigma^\frown g)$, and therefore $\mathcal{E} \geq_{\mathcal{M}} \mathcal{D}$, a contradiction.

  Thus, $\alpha(X) \leq_{\mathcal{M}} \alpha(Y)$, and since α is an upper implicative semilattice embedding this tells us that $X \supseteq Y$.
- β preserves joins: we have

$$\begin{aligned}\beta(X \oplus Y) = \alpha(X \oplus Y) \otimes \mathcal{D} &\equiv_{\mathcal{M}} (\alpha(X) \oplus \alpha(Y)) \otimes \mathcal{D} \\ &\equiv_{\mathcal{M}} (\alpha(X) \otimes \mathcal{D}) \oplus (\alpha(Y) \otimes \mathcal{D}) = \beta(X) \oplus \beta(Y).\end{aligned}$$

- β preserves implications: we have

$$\begin{aligned}&\beta(X) \to_{[\beta(I),\beta(\emptyset)]_{\mathcal{M}}} \beta(Y) \\ &\quad = ((\alpha(X) \otimes \mathcal{D}) \to_{\mathcal{M}} (\alpha(Y) \otimes \mathcal{D})) \oplus \beta(I) \\ &\quad \equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} (\alpha(Y) \otimes \mathcal{D})) \oplus (\mathcal{D} \to_{\mathcal{M}} (\alpha(Y) \otimes \mathcal{D}))) \oplus \beta(I) \\ &\quad \equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} (\alpha(Y) \otimes \mathcal{D})) \oplus \omega^\omega) \oplus \beta(I) \\ &\quad \equiv_{\mathcal{M}} (\alpha(X) \to_{\mathcal{M}} (\alpha(Y) \otimes \mathcal{D})) \oplus \beta(I).\end{aligned}$$

  Next, using Proposition 4.2.2 we see:

$$\begin{aligned}&\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \otimes (\alpha(X) \to_{\mathcal{M}} \mathcal{D})) \oplus \beta(I) \\ &= ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \otimes (\alpha(X) \to_{\mathcal{M}} (\mathcal{E} \to_{\mathcal{M}} \mathcal{B}))) \oplus \beta(I) \\ &\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \otimes ((\alpha(X) \oplus \mathcal{E}) \to_{\mathcal{M}} \mathcal{B})) \oplus \beta(I).\end{aligned}$$

As noted above, we have $\mathcal{E} \subseteq \alpha(X)$, and therefore:

$$\begin{aligned}
&\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \otimes (\mathcal{E} \to_{\mathcal{M}} \mathcal{B})) \oplus \beta(I)\\
&\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \otimes \mathcal{D}) \oplus (\alpha(I) \otimes \mathcal{D}).\\
&\equiv_{\mathcal{M}} ((\alpha(X) \to_{\mathcal{M}} \alpha(Y)) \oplus \alpha(I)) \otimes \mathcal{D}\\
&= \big(\alpha(X) \to_{[\alpha(I),\alpha(\emptyset)]_{\mathcal{M}}} \alpha(Y)\big) \otimes \mathcal{D}\\
&= \alpha(X \to_{\mathcal{P}(I)} Y) \otimes \mathcal{D}\\
&= \beta(X \to_{\mathcal{P}(I)} Y).
\end{aligned}$$

- β has canonical range:
  (i) Let $X \subseteq I$, we show that that $\beta(X)$ is meet-irreducible in $[\beta(I), \beta(\emptyset)]_{\mathcal{M}}$. Indeed, let $\mathcal{C}_0, \mathcal{C}_1 \leq_{\mathcal{M}} \beta(\emptyset)$ be such that $\mathcal{C}_0 \otimes \mathcal{C}_1 \leq_{\mathcal{M}} \alpha(X) \otimes \mathcal{D}$. Then $\mathcal{C}_0 \otimes \mathcal{C}_1 \leq_{\mathcal{M}} \alpha(X)$, and since $\alpha(X)$ is Muchnik, we see from Proposition 4.2.2 that $\mathcal{C}_0 \leq_{\mathcal{M}} \alpha(X)$ or $\mathcal{C}_1 \leq_{\mathcal{M}} \alpha(X)$. Since $\mathcal{C}_0, \mathcal{C}_1 \leq_{\mathcal{M}} \beta(\emptyset) \leq_{\mathcal{M}} \mathcal{D}$ this shows that in fact $\mathcal{C}_0 \leq_{\mathcal{M}} \beta(X)$ or $\mathcal{C}_1 \leq_{M} \beta(X)$.
  (ii) The range of β is clearly closed under implication and joins.
  (iii) Let $X \subseteq \omega$ and let $\mathcal{C}_0, \mathcal{C}_1 \in [\beta(I), \beta(\emptyset)]_{\mathcal{M}}$. Then we have:

$$\begin{aligned}
&\beta(X) \to_{[\beta(I),\beta(\emptyset)]_{\mathcal{M}}} (\mathcal{C}_0 \otimes \mathcal{C}_1)\\
&\qquad= (\alpha(X) \otimes \mathcal{D}) \to_{[\beta(I),\beta(\emptyset)]_{\mathcal{M}}} (\mathcal{C}_0 \otimes \mathcal{C}_1)\\
&\qquad= ((\alpha(X) \otimes \mathcal{D}) \to_{\mathcal{M}} (\mathcal{C}_0 \otimes \mathcal{C}_1)) \oplus \beta(I)\\
&\qquad\equiv_{\mathcal{M}} (\alpha(X) \to_{\mathcal{M}} (\mathcal{C}_0 \otimes \mathcal{C}_1)) \oplus \beta(I),
\end{aligned}$$

  because $\mathcal{C}_0$ and $\mathcal{C}_1$ are below $\beta(\emptyset)$ and hence below $\mathcal{D}$. Since $\alpha(X)$ is Muchnik, we now see from Proposition 4.2.2:

$$\begin{aligned}
&= ((\alpha(X) \to_{\mathcal{M}} \mathcal{C}_0) \otimes (\alpha(X) \to_{\mathcal{M}} \mathcal{C}_1)) \oplus \beta(I)\\
&\equiv_{\mathcal{M}} (\beta(X) \to_{[\beta(I),\beta(\emptyset)]_{\mathcal{M}}} \mathcal{C}_0) \otimes (\beta(X) \to_{[\beta(I),\beta(\emptyset)]_{\mathcal{M}}} \mathcal{C}_1).
\end{aligned}$$ □

We can now prove there is a principal factor of the Medvedev lattice with theory IPC below a given $\mathcal{B} > 0'$.

Theorem 4.3.2. *Let $\mathcal{B}$ be a mass problem, let $\mathcal{E}$ be a countable mass problem such that $\mathcal{E} \not\geq_{\mathcal{M}} \mathcal{B}$ and let $\mathcal{D} = \mathcal{E} \to \mathcal{B}$ (so, $\mathcal{D} \leq_{\mathcal{M}} \mathcal{B}$). Let A be a computably independent set such that for all $f \in \mathcal{E}$ we have $A \not\geq_T f$. Then*

$$\mathrm{Th}\left(\mathcal{M}/\left(\left\{i^\frown g \mid g \geq_T A^{[i]} \text{ or } g \in C(\mathcal{E})\right\} \otimes \mathcal{D}\right)\right) = \mathrm{IPC}.$$

Proof. The proof largely mirrors that of Theorem 4.0.1. Let $\mathcal{E} = \{f_0, f_1, \dots\}$, let E_i be the graph of f_i and let U be such that $U^{[0]} = A$ and $U^{[i+1]} = E_i$. Then $A, E_0, E_1, \dots$ is uniformly computable in U.

We need to make a slight modification to Lemma 4.2.11: we not only want $A \oplus B$ to be computably independent, but we also need to make sure that $A \oplus B \not\geq_T f$ for every $f \in \mathcal{E}$. This modification is straightforward and we omit the details. The requirements on D_i are slightly different: we now want for every $1 \leq i \leq k$ that $D_i \geq_T \bigoplus_{1 \leq j \leq k, i \in X_j} A^{[j]} \oplus B^{[i]}$, that $D_i' \leq_T (U \oplus B)'$, that $D_i \oplus A^{[j]} \geq_T (U \oplus B)'$ for every $j \in \{1 \leq j \leq k \mid i \notin X_j\} \cup \{k+1, k+2, \dots\}$, that $D_i \oplus B^{[j]} \geq_T (U \oplus B)'$

for every $j \neq i$ and that $D_i \oplus E_j \geq_T (U \oplus B)'$ for all $j \in \omega$; this is still possible by Theorem 4.2.10. We change the definition of $\mathcal{A}$ into $\mathcal{A} = \omega^\omega \setminus C(\{(U \oplus B)'\} \cup \mathcal{E})$. Then we still have $D_i \in \mathcal{A}$, because $D_i \geq_T f_j$ would imply that $D_i \geq_T D_i \oplus E_j \geq_T (U \oplus B)'$, a contradiction. Finally, replace α with the β of Theorem 4.3.1 and change (9) into

$$\left(\left(\overline{\mathcal{A}} \cup C(\{D_1, \ldots, D_n\})\right) \otimes \mathcal{D}\right) \oplus \left(\left\{i^\frown g \mid g \geq_T A^{[i]} \text{ or } g \in C(\mathcal{E})\right\} \otimes \mathcal{D}\right)$$
$$\equiv_{\mathcal{M}} \beta(X_1) \otimes \cdots \otimes \beta(X_k).$$

Then the whole proof of Theorem 4.0.1 goes through. □

In particular, this allows us to give a positive answer to the question mentioned at the beginning of this section.

Corollary 4.3.3. *Let* $\mathcal{B} >_{\mathcal{M}} 0'$. *Then* $\mathrm{Th}(\mathcal{M}/\mathcal{B}) \subseteq \mathrm{Jan}$.

Proof. Since an intermediate logic is contained in Jan if and only if its positive fragment coincides with IPC (see Jankov [46]), we need to show that, denoting the positive fragment by $^+$, we have that $\mathrm{Th}^+(\mathcal{M}/\mathcal{B}) \subseteq \mathrm{IPC}^+$. By Theorem 4.3.2 there exists a $\mathcal{C} \leq_{\mathcal{M}} \mathcal{B}$ such that $\mathrm{Th}(\mathcal{M}/\mathcal{C}) = \mathrm{IPC}$. Then $\mathcal{M}/\mathcal{C}$ is a subalgebra of $\mathcal{M}/\mathcal{B}$, except for the fact that the top element is not necessarily preserved. However, it can be directly verified that for any two Brouwer algebras $\mathscr{C}$ and $\mathscr{B}$ for which $\mathscr{C}$ is a $(\oplus, \otimes, \rightarrow, 0)$-subalgebra of $\mathscr{B}$ we have for all positive formulas $\varphi(x_1, \ldots, x_n)$ and all elements $b_1, \ldots, b_n \in \mathscr{B}$ that the interpretation of φ at $b_1, \ldots, b_n$ is the same in both $\mathscr{C}$ and $\mathscr{B}$. Since we can refute every positive formula φ which is not in IPC^+ in $\mathcal{M}/\mathcal{C}$, we can therefore refute it in $\mathcal{M}/\mathcal{B}$ using the same valuation. In other words, $\mathrm{Th}^+(\mathcal{M}/\mathcal{B}) \subseteq \mathrm{Th}^+(\mathcal{M}/\mathcal{C}) = \mathrm{IPC}^+$, as desired. □

CHAPTER 5

First-Order Logic in the Medvedev Lattice

In this chapter we will discuss an extension of the Medvedev lattice to first-order logic, using the notion of a *first-order hyperdoctrine* from categorical logic to define what we will call the *hyperdoctrine of mass problems*.

We will give a short overview of the necessary definitions and properties in section 5.1. After that, in section 5.2 we will introduce the *degrees of ω-mass problems*, which combine the idea of Medvedev that 'solving' should be interpreted as 'computing' with the idea of Kolmogorov that 'solving' should be uniform in the variables. Using these degrees of ω-mass problems, we will introduce the hyperdoctrine of mass problems in section 5.3.

Next, in section 5.4 we study the intermediate logic which this hyperdoctrine of mass problems gives us, and we start looking at subintervals of it to try and obtain analogous results to Skvortsova's [106] remarkable result that intuitionistic propositional logic can be obtained from a factor of the Medvedev lattice. In section 5.5 we show that even in these intervals we cannot get every intuitionistic theory, by showing that there is an analogue of Tennenbaum's theorem [116] that every computable model of Peano arithmetic is the standard model. Finally, in section 5.6 we prove a partial positive result on which theories can be obtained in subintervals of the hyperdoctrine of mass problems, through a characterisation using Kripke models.

This chapter is based on Kuyper [63].

5.1. Categorical semantics for IQC

In this section we will discuss the notion of *first-order hyperdoctrine*, as formulated by Pitts [92], based on the important notion of hyperdoctrine introduced by Lawvere [72]. These first-order hyperdoctrines can be used to give sound and complete categorical semantics for IQC. Our notion of first-order logic in the Medvedev lattice will be based on this, so we will discuss the basic definitions and the basic properties before we proceed with our construction. We use the formulation from Pitts [94] (but we use Brouwer algebras instead of Heyting algebras, because the Medvedev lattice is normally presented as a Brouwer algebra).

Let us first give the definition of a first-order hyperdoctrine. After that we will discuss an easy example and discuss how first-order hyperdoctrines interpret first-order intuitionistic logic. We will not discuss all details and the full motivation behind this definition, instead referring the reader to the works by Pitts [92, 94]. However, we will discuss some of the motivation behind this definition in Remark 5.1.9 below.

DEFINITION 5.1.1. ([94, Definition 2.1]) Let $\mathbf{C}$ be a category such that for every object $X \in \mathbf{C}$ and every $n \in \omega$, the n-fold product X^n of X exists. A *first-order hyperdoctrine* $\mathcal{P}$ over $\mathbf{C}$ is a contravariant functor $\mathcal{P} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Poset}$ from $\mathbf{C}$ into the category $\mathbf{Poset}$ of partially ordered sets and order homomorphisms, satisfying:

(i) For each object $X \in \mathbf{C}$, the partially ordered set $\mathcal{P}(X)$ is a Brouwer algebra;
(ii) For each morphism $f : X \to Y$ in $\mathbf{C}$, the order homomorphism $\mathcal{P}(f) : \mathcal{P}(Y) \to \mathcal{P}(X)$ is a homomorphism of Brouwer algebras;
(iii) For each diagonal morphism $\Delta_X : X \to X \times X$ in $\mathbf{C}$ (i.e. a morphism such that $\pi_1 \circ \Delta_X = \pi_2 \circ \Delta_X = 1_X$), the upper adjoint to $\mathcal{P}(\Delta_X)$ at the bottom element $0 \in \mathcal{P}(X)$ exists. In other words, there is an element $=_X \in \mathcal{P}(X \times X)$ such that for all $A \in \mathcal{P}(X \times X)$ we have
$$\mathcal{P}(\Delta_X)(A) \leq 0 \text{ if and only if } A \leq =_X.$$
(iv) For each product projection $\pi : \Gamma \times X \to \Gamma$ in $\mathbf{C}$, the order homomorphism $\mathcal{P}(\pi) : \mathcal{P}(\Gamma) \to \mathcal{P}(\Gamma \times X)$ has both an upper adjoint $(\exists x)_\Gamma$ and a lower adjoint $(\forall x)_\Gamma$, i.e.:
$$\mathcal{P}(\pi)(B) \leq A \text{ if and only if } B \leq (\exists x)_\Gamma(A)$$
$$A \leq \mathcal{P}(\pi)(B) \text{ if and only if } (\forall x)_\Gamma(A) \leq B.$$
Moreover, these adjoints are natural in Γ, i.e. given $s : \Gamma \to \Gamma'$ in $\mathbf{C}$ we have
$$\begin{tikzcd} \mathcal{P}(\Gamma' \times X) \arrow[r, "\mathcal{P}(s \times 1_X)"'] \arrow[d, "(\exists x)_{\Gamma'}"'] & \mathcal{P}(\Gamma \times X) \arrow[d, "(\exists x)_\Gamma"'] \\ \mathcal{P}(\Gamma') \arrow[r, "\mathcal{P}(s)"'] & \mathcal{P}(\Gamma) \end{tikzcd} \qquad \begin{tikzcd} \mathcal{P}(\Gamma' \times X) \arrow[r, "\mathcal{P}(s \times 1_X)"'] \arrow[d, "(\forall x)_{\Gamma'}"'] & \mathcal{P}(\Gamma \times X) \arrow[d, "(\forall x)_\Gamma"'] \\ \mathcal{P}(\Gamma') \arrow[r, "\mathcal{P}(s)"'] & \mathcal{P}(\Gamma). \end{tikzcd}$$
This condition is called the *Beck-Chevalley condition.*

We will also denote $P(f)$ by f^*.

REMARK 5.1.2. We emphasise that the adjoints $(\exists x)_\Gamma$ and $(\forall x)_\Gamma$ only need to be order homomorphisms, and that they do no need to preserve the lattice structure. This should not come as a surprise: after all, the universal quantifier does not distribute over logical disjunction, and neither does the existential quantifier distribute over conjunction.

EXAMPLE 5.1.3. ([94, Example 2.2]) Let $\mathscr{B}$ be a complete Brouwer algebra. Then $\mathscr{B}$ induces a first-order hyperdoctrine $\mathcal{P}$ over the category $\mathbf{Set}$ of sets and functions as follows. We let $\mathcal{P}(X)$ be $\mathscr{B}^X$, which is again a Brouwer algebra under coordinate-wise operations. Furthermore, for each function $f : X \to Y$ we let $\mathcal{P}(f)$ be the function which sends $(B_y)_{y \in Y}$ to the set given by $A_x = B_{f(x)}$. The equality predicates $=_X$ are given by
$$=_X(x, z) = \begin{cases} 0 & \text{if } x = z \\ 1 & \text{otherwise.} \end{cases}$$

For the adjoints we use the fact that $\mathscr{B}$ is complete: given $B \in \mathcal{P}(\Gamma \times X)$ we let

$$((\forall x)_\Gamma(B))_\gamma = \bigoplus_{x \in X} B_{(\gamma,x)}$$

and

$$((\exists x)_\Gamma(B))_\gamma = \bigotimes_{x \in X} B_{(\gamma,x)}.$$

Then $\mathcal{P}$ is directly verified to be a first-order hyperdoctrine.

Remark 5.1.4. We obtain a special case of Example 5.1.3 when we take $\mathscr{B}$ to be the Muchnik lattice. In that case we obtain a fragment of the first-order part of the structure recently studied by Basu and Simpson [7], who independently studied an interpretation of higher-order intuitionistic logic based on the Muchnik lattice.

Thus, if we have a sequence of problems $\mathcal{B}_0, \mathcal{B}_1, \ldots$, we have

$$(\forall x)_1((\mathcal{B}_i)_{i \in \omega}) = \bigoplus_{i \in \omega} \mathcal{B}_i = \{f \in \omega^\omega \mid \forall i \in \omega \exists g \in \mathcal{B}_i (f \geq_T g)\},$$

in other words a solution of the problem $\forall x(\mathcal{B}(x))$ computes a solution of every $\mathcal{B}_i$ but does so non-uniformly.

If, as in [7], we take each $\mathcal{B}_i$ to be the canonical representative of its Muchnik degree, i.e. we take $\mathcal{B}_i$ to be upwards closed under Turing reducibility, then we have that

$$(\forall x)_1((\mathcal{B}_i)_{i \in \omega}) = \bigoplus_{i \in \omega} \mathcal{B}_i = \bigcap_{i \in \omega} \mathcal{B}_i,$$

i.e. a solution of the problem $\forall x(\mathcal{B}(x))$ is a single solution that solves every $\mathcal{B}_i$. Thus, depending on the view one has on the Muchnik lattice, either the solution is allowed to depend on x but non-uniformly, or it is not allowed to depend on x at all. So, Basu and Simpson's approach does not follow Kolmogorov's philosophy that the interpretation of the universal quantifier should depend uniformly on the variable. On the other hand, our approach will follow this philosophy.

Of course, an important advantage of Basu and Simpson's approach is that it is suitable for higher-order logic, while we can only deal with first-order logic. Another important difference between our work and theirs is that we start from the Medvedev lattice, while they take the Muchnik lattice as their starting point.

Next, let us discuss how first-order intuitionistic logic can be interpreted in first-order hyperdoctrines. Most of the literature on this subject deals with multi-sorted first-order logic; however, to keep the notation easy and because we do not intend to discuss multi-sorted logic in the Medvedev case, we will give the definition only for single-sorted first-order logic.

Definition 5.1.5. (Pitts [92, p. B2]) Let $\mathcal{P}$ be a first-order hyperdoctrine over $\mathbf{C}$ and let Σ be a first-order language. Then a *structure* $\mathfrak{M}$ for Σ in $\mathcal{P}$ consists of:

(i) an object $M \in \mathbf{C}$ (the universe),
(ii) a morphism $[\![f]\!]_\mathfrak{M} : M^n \to M$ in $\mathbf{C}$ for every n-ary function symbol f in Σ,
(iii) an element $[\![R]\!]_\mathfrak{M} \in \mathcal{P}(M^n)$ for every n-ary relation in Σ.

Case (iii) is probably the most interesting part of this definition, since it says that elements of $\mathcal{P}(M^n)$ should be seen as generalised n-ary predicates on M.

DEFINITION 5.1.6. ([92, Table 6.4]) Let t be a first-order term in a language Σ and let $\mathfrak{M}$ be a structure in a first-order hyperdoctrine $\mathcal{P}$. Let $\vec{x} = (x_1, \dots, x_n)$ be a context (i.e. an ordered list of distinct variables) containing all free variables in t. Then we define the interpretation $[\![t(\vec{x})]\!]_{\mathfrak{M}} \in M^n \to M$ inductively as follows:

(i) If t is a variable x_i, then $[\![t(\vec{x})]\!]_{\mathfrak{M}}$ is the projection of M^n to the ith coordinate.
(ii) If t is $f(s_1, \dots, s_m)$ for f in Σ, then $[\![t(\vec{x})]\!]_{\mathfrak{M}}$ is $[\![f]\!]_{\mathfrak{M}} \circ ([\![s_1(\vec{x})]\!]_{\mathfrak{M}}, \dots, [\![s_m(\vec{x})]\!]_{\mathfrak{M}})$.

Thus, we identify a term with the function mapping a valuation of the variables occurring in the term to the value of the term when evaluated at that valuation.

DEFINITION 5.1.7. ([92, Table 8.2]) Let φ be a first-order formula in a language Σ and let $\mathfrak{M}$ be a structure in a first-order hyperdoctrine $\mathcal{P}$. Let $\vec{x} = (x_1, \dots, x_n)$ be a context (i.e. an ordered list of distinct variables) containing all free variables in φ. Then we define the interpretation $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}} \in \mathcal{P}(M^n)$ in the context Γ inductively as follows:

(i) If φ is $R(t_1, \dots, t_m)$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is $([\![t_1(\vec{x})]\!]_{\mathfrak{M}}, \dots, [\![t_m(\vec{x})]\!]_{\mathfrak{M}})^*([\![R]\!]_{\mathfrak{M}})$.
(ii) If φ is $t_1 = t_2$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $([\![t_1(\vec{x})]\!]_{\mathfrak{M}}, [\![t_2(\vec{x})]\!]_{\mathfrak{M}})^*(=_M)$.
(iii) If φ is $\top$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $0 \in \mathcal{P}(M^n)$; i.e. the smallest element of $\mathcal{P}(M^n)$.
(iv) If φ is $\bot$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $1 \in \mathcal{P}(M^n)$; i.e. the largest element of $\mathcal{P}(M^n)$.
(v) If φ is $\psi \vee \theta$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $[\![\psi(\vec{x})]\!]_{\mathfrak{M}} \otimes [\![\theta(\vec{x})]\!]_{\mathfrak{M}}$.
(vi) If φ is $\psi \wedge \theta$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $[\![\psi(\vec{x})]\!]_{\mathfrak{M}} \oplus [\![\theta(\vec{x})]\!]_{\mathfrak{M}}$.
(vii) If φ is $\psi \to \theta$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $[\![\psi(\vec{x})]\!]_{\mathfrak{M}} \to [\![\theta(\vec{x})]\!]_{\mathfrak{M}}$.
(viii) If φ is $\exists y.\psi$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $(\exists y)_{M^n}([\![\psi(\vec{x}, y)]\!]_{\mathfrak{M}})$.
(ix) If φ is $\forall y.\psi$, then $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}}$ is defined as $(\forall y)_{M^n}([\![\psi(\vec{x}, y)]\!]_{\mathfrak{M}})$.

DEFINITION 5.1.8. ([92, Definition 8.4]) Let φ be a formula in a language Σ and a context $\vec{x} = (x_1, \dots, x_n)$, and let $\mathfrak{M}$ be a structure in a first-order hyperdoctrine $\mathcal{P}$. Then we say that $\varphi(\vec{x})$ is *satisfied* if $[\![\varphi(\vec{x})]\!]_{\mathfrak{M}} = 0$ in $\mathcal{P}(M^n)$. We let the *theory* of $\mathfrak{M}$ be the set of sentences which are satisfied (in the empty context), and we denote this by $\mathrm{Th}(\mathfrak{M})$. Given a language Σ, we let the *theory* of $\mathcal{P}$ be the intersection of the theories of all structures $\mathfrak{M}$ for Σ in $\mathcal{P}$, and we denote this theory by $\mathrm{Th}(\mathcal{P})$.

REMARK 5.1.9. Let us make some remarks on the definitions given above.

- As mentioned above, we identify terms $t(\vec{x})$ with functions $[\![t(\vec{x})]\!]_{\mathfrak{M}}$, and m-ary predicates $R(y_1, \dots, y_m)$ are elements of $\mathcal{P}(M^n)$. Since we required our category $\mathbf{C}$ to contain n-fold products, if we have terms $t_1, \dots, t_m$, then $([\![t_1(\vec{x})]\!]_{\mathfrak{M}}, \dots, [\![t_m(\vec{x})]\!]_{\mathfrak{M}}) : M^n \to M^m$, so
$$([\![t_1(\vec{x})]\!]_{\mathfrak{M}}, \dots, [\![t_m(\vec{x})]\!]_{\mathfrak{M}})^* : \mathcal{P}(M^m) \to \mathcal{P}(M^n).$$
This should be seen as the *substitution of* $t_1(\vec{x}), \dots, t_m(\vec{x})$ *for* $y_1, \dots, y_m$, which explains case (i) and (ii).
- Quantifiers are interpreted as adjoints, which is an idea due to Lawvere. For example, for the universal quantifier this says that
$$[\![\psi]\!]_{\mathfrak{M}} \geq [\![\forall x \varphi(x)]\!]_{\mathfrak{M}} \Leftrightarrow [\![\psi(x)]\!]_{\mathfrak{M}} \geq [\![\varphi(x)]\!]_{\mathfrak{M}},$$

where we assume x does not occur freely in ψ. Reading $\geq$ as $\vdash$, the two implications are essentially the introduction and elimination rules for the universal quantifier.

- The Beck-Chevalley condition is necessary to ensure that substitutions commute with the quantifiers (modulo restrictions on bound variables).

Let us introduce a notational convention: *when the structure is clear from the context, we will omit the subscript* $\mathfrak{M}$ *in* $[\![-]\!]_{\mathfrak{M}}$. Having finished giving the definition of first-order hyperdoctrines, let us just mention that they are sound and complete for intuitionistic first-order logic IQC.

PROPOSITION 5.1.10. ([92, Proposition 8.8]) *Structures in first-order hyperdoctrines are* sound *for IQC, i.e. the deductive closure of* Th($\mathfrak{M}$) *in IQC is equal to* Th($\mathfrak{M}$).

THEOREM 5.1.11. (Pitts [93, Corollary 5.31]) *The class of first-order hyperdoctrines is* complete *for IQC.*

5.2. The degrees of ω-mass problems

In this section, we will introduce an extension of the Medvedev lattice, which we will need to define our first-order hyperdoctrine based on the Medvedev lattice. As mentioned in the introduction, Kolmogorov mentioned in his paper that solving the problem $\forall x \varphi(x)$ is the same as solving the problem $\varphi(x)$ for all x, *uniformly in* x. We formalise this in the spirit of Medvedev and Muchnik in the following way.

DEFINITION 5.2.1. An *ω-mass problem* is an element $(\mathcal{A}_i)_{i\in\omega} \in (\mathcal{P}(\omega^\omega))^\omega$. Given two ω-mass problems $(\mathcal{A}_i)_{i\in\omega}, (\mathcal{B}_i)_{i\in\omega}$, we say that $(\mathcal{A}_i)_{i\in\omega}$ *reduces* to $(\mathcal{B}_i)_{i\in\omega}$ (notation: $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$) if there exists a partial Turing functional Φ such that for every $n \in \omega$ we have $\Phi(n^\frown \mathcal{B}_n) \subseteq \mathcal{A}_n$. If both $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$ and $(\mathcal{B}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} (\mathcal{A}_i)_{i\in\omega}$ we say that $(\mathcal{A}_i)_{i\in\omega}$ and $(\mathcal{B}_i)_{i\in\omega}$ are *equivalent* (notation: $(\mathcal{A}_i)_{i\in\omega} \equiv_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$). We call the equivalence classes of this equivalence the *degrees of ω-mass problems* and denote the set of the degrees of ω-mass problems by $\mathcal{M}_\omega$.

DEFINITION 5.2.2. Let $(\mathcal{A}_i)_{i\in\omega}, (\mathcal{B}_i)_{i\in\omega}$ be ω-mass problems. We say that $(\mathcal{A}_i)_{i\in\omega}$ *weakly reduces* to $(\mathcal{B}_i)_{i\in\omega}$ (notation: $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_{w\omega}} (\mathcal{B}_i)_{i\in\omega}$) if for every sequence $(g_i)_{i\in\omega}$ with $g_i \in \mathcal{B}_i$ there exists a partial Turing functional Φ such that for every $n \in \omega$ we have $\Phi(n^\frown g_n) \in \mathcal{A}_n$. If both $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_{w\omega}} (\mathcal{B}_i)_{i\in\omega}$ and $(\mathcal{B}_i)_{i\in\omega} \leq_{\mathcal{M}_{w\omega}} (\mathcal{A}_i)_{i\in\omega}$ we say that $(\mathcal{A}_i)_{i\in\omega}$ and $(\mathcal{B}_i)_{i\in\omega}$ are *weakly equivalent* (notation: $(\mathcal{A}_i)_{i\in\omega} \equiv_{\mathcal{M}_{w\omega}} (\mathcal{B}_i)_{i\in\omega}$). We call the equivalence classes of weak equivalence the *weak degrees of ω-mass problems* and denote the set of the weak degrees of ω-mass problems by $\mathcal{M}_{w\omega}$.

The next proposition tells us that $\mathcal{M}_\omega$ is a Brouwer algebra, like the Medvedev lattice.

PROPOSITION 5.2.3. *The degrees of ω-mass problems form a Brouwer algebra.*

Proof. We claim that $\mathcal{M}_\omega$ is a Brouwer algebra under the component-wise operations on $\mathcal{M}$, i.e. the operations induced by:

$$\begin{aligned}((\mathcal{A}_i)_{i\in\omega} \oplus (\mathcal{B}_i)_{i\in\omega})_n &= \{f \oplus g \mid f \in \mathcal{A}_n, g \in \mathcal{B}_n\}\\ ((\mathcal{A}_i)_{i\in\omega} \otimes (\mathcal{B}_i)_{i\in\omega})_n &= 0^\frown\mathcal{A}_n \cup 1^\frown\mathcal{B}_n\\ ((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega})_n &= \{e^\frown f \mid \forall g \in \mathcal{A}_n(\Phi_e(g \oplus f) \in \mathcal{B}_n).\end{aligned}$$

The proof of this is mostly analogous to the proof for the Medvedev lattice, so we will only give the proof for the implication. Let us first show that $(\mathcal{A}_i)_{i\in\omega} \oplus ((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}) \geq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$. Define a Turing functional Φ by

$$\Phi(n^\frown(g \oplus (e^\frown f))) = \Phi_e(g \oplus f).$$

Then Φ witnesses that $(\mathcal{A}_i)_{i\in\omega} \oplus ((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}) \geq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$.

Conversely, let $(\mathcal{C}_i)_{i\in\omega}$ be such that $(\mathcal{A}_i)_{i\in\omega} \oplus (\mathcal{C}_i)_{i\in\omega} \geq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$. Let $e \in \omega$ be such that Φ_e witnesses this fact. Let φ be a computable function sending n to an index for the functional mapping h to $\Phi_e(n^\frown h)$. Let Ψ be the functional sending $n^\frown f$ to $\varphi(n)^\frown f$. Then $(\mathcal{C}_i)_{i\in\omega} \geq_{\mathcal{M}_\omega} (\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}$ through Ψ. $\square$

However, it turns out that this fails for $\mathcal{M}_{w\omega}$: it is still a distributive lattice, but it is not a Brouwer algebra.

Proposition 5.2.4. *The weak degrees of ω-mass problems form a distributive lattice, but not a Brouwer algebra. In particular, they do not form a complete lattice.*

Proof. It is easy to see that $\mathcal{M}_{w\omega}$ is a distributive lattice under the same operations as $\mathcal{M}_\omega$. Towards a contradiction, assume $\mathcal{M}_{w\omega}$ is a Brouwer algebra, under some implication $\to$. Let $f, g \in \omega^\omega$ be two functions of incomparable Turing degree. Let $(\mathcal{A}_i)_{i\in\omega}$ be given by $\mathcal{A}_i = \{h \mid h \equiv_T f\}$ and let $(\mathcal{B}_i)_{i\in\omega}$ be given by $\mathcal{B}_i = \{f \oplus g\}$. For every $j \in \omega$, let $(\mathcal{C}_i^j)_{i\in\omega}$ be given by $\mathcal{C}_i^j = \{g\}$ for $i = j$, and $\mathcal{C}_i^j = \{f \oplus g\}$ otherwise.

Then, for every $j \in \omega$ we have $(\mathcal{A}_i)_{i\in\omega} \oplus (\mathcal{C}_i^j)_{i\in\omega} \geq_{\mathcal{M}_{w\omega}} (\mathcal{B}_i)_{i\in\omega}$: given a sequence $(h_i)_{i\in\omega}$ with $h_i \in \mathcal{A}_i$, let e be such that $\Phi_e(h_j) = f$. Now let $\Phi(n^\frown(s \oplus t))$ be t for $n \neq j$ and $\Phi_e(s) \oplus t$ otherwise. This Φ is the required witness.

So, since we assumed $\to$ makes $\mathcal{M}_{w\omega}$ into a Brouwer algebra, we know that every $((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}) \leq_{\mathcal{M}_{w\omega}} (\mathcal{C}_i^j)_{i\in\omega}$ for every $j \in \omega$. Thus, for every $j \in \omega$ there is some $g_j \leq_T g$ in $((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega})_j$. For every $j \in \omega$, fix a $\sigma_j \in \omega^{<\omega}$ such that there exists an $n \in \omega$ with $\Phi_j(j^\frown(\sigma_j \oplus g_j))(n)\downarrow \neq (f \oplus g)(n)$, which exists because g, and therefore $g_j \leq_T g$, does not compute f. Now let $f_j = {\sigma_j}^\frown f$. Then we have $(f_i)_{i\in\omega} \in (\mathcal{A}_i)_{i\in\omega}$ and $(g_i)_{i\in\omega} \in (\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}$, but for every $j \in \omega$ we have that $\Phi_j(j^\frown(f_j \oplus g_j)) \notin \mathcal{B}_j$. Thus $(\mathcal{A}_i)_{i\in\omega} \oplus ((\mathcal{A}_i)_{i\in\omega} \to (\mathcal{B}_i)_{i\in\omega}) \not\geq_{\mathcal{M}_{w\omega}} (\mathcal{B}_i)_{i\in\omega}$, a contradiction. $\square$

Finally, let us show that $\mathcal{M}_\omega$ and $\mathcal{M}_{w\omega}$ are extensions of the Medvedev and Muchnik lattices, in the sense that the latter embed into the first. Furthermore, we show that the countable products of $\mathcal{M}$ and $\mathcal{M}_w$ are quotients of $\mathcal{M}_\omega$ and $\mathcal{M}_{w\omega}$.

PROPOSITION 5.2.5. *There is a Brouwer algebra embedding of $\mathcal{M}$ into $\mathcal{M}_\omega$ and a lattice embedding of $\mathcal{M}_w$ into $\mathcal{M}_{w\omega}$, both given by*

$$\alpha(\mathcal{A})_n = \mathcal{A}.$$

PROOF. Direct, using the fact that the diagonal of $\mathcal{M}_\omega$, i.e. $\{(\mathcal{A}_i)_{i\in\omega} \in \mathcal{M}_\omega \mid \forall n, m(\mathcal{A}_n = \mathcal{A}_m)\}$, is isomorphic to the diagonal of $\mathcal{M}^\omega$, which is directly seen to be isomorphic to $\mathcal{M}$. The same holds for $\mathcal{M}_{w\omega}$ and $\mathcal{M}_w$. □

PROPOSITION 5.2.6. *There is a Brouwer algebra homomorphism of $\mathcal{M}_\omega$ onto $\mathcal{M}^\omega$ and a lattice homomorphism of $\mathcal{M}_{w\omega}$ onto ${\mathcal{M}_w}^\omega$.*

PROOF. Follows directly from the fact that all operations on $\mathcal{M}_\omega$ and $\mathcal{M}_{w\omega}$ are component-wise, and the fact that the reducibilities on $\mathcal{M}_\omega$ and $\mathcal{M}_{w\omega}$ are stronger than those on $\mathcal{M}^\omega$ respectively $\mathcal{M}_w^\omega$. □

5.3. The hyperdoctrine of mass problems

In this section, we will introduce our first-order hyperdoctrine based on $\mathcal{M}$ and $\mathcal{M}_\omega$, which we will call the *hyperdoctrine of mass problems* $\mathcal{P}_\mathcal{M}$. We will take the category **C** to be the category with objects $\{1\}, \{1,2\}, \dots$ and ω, and with functions the computable functions between them. We will define $\mathcal{P}_\mathcal{M}(\omega)$ to be $\mathcal{M}_\omega$. Now, let us look at how to define $\mathcal{P}_\mathcal{M}(\alpha) = \alpha^*$ for functions $\alpha : \omega \to \omega$.

DEFINITION 5.3.1. Let $\alpha : \omega \to \omega$. Then $\alpha^* : \mathcal{P}(\omega^\omega)^\omega \to \mathcal{P}(\omega^\omega)^\omega$ is the function given by

$$(\alpha^*((\mathcal{A}_i)_{i\in\omega}))_n = \mathcal{A}_{\alpha(n)}.$$

PROPOSITION 5.3.2. *Let $\alpha : \omega \to \omega$ be a computable function. Then α^* induces a well-defined function on $\mathcal{M}_\omega$ by sending $(\mathcal{A}_i)_{i\in\omega}$ to $\alpha^*((\mathcal{A}_i)_{i\in\omega})$, which is in fact a Brouwer algebra homomorphism.*

PROOF. We need to show that if $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$, then $\alpha^*((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} \alpha^*((\mathcal{B}_i)_{i\in\omega})$. Let Φ witness that $\mathcal{A} \leq_{\mathcal{M}_\omega} \mathcal{B}$. Let Ψ be the partial Turing functional sending $n^\frown f$ to $\Phi(\alpha(n)^\frown f)$. Then Ψ witnesses that $\alpha^*((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} \alpha^*((\mathcal{B}_i)_{i\in\omega})$. That α^* is a Brouwer algebra homomorphism follows easily from the fact that the operations on $\mathcal{M}_\omega$ are component-wise. □

Next, we will show that for every computable α we have that α^* has both upper and lower adjoints, which will certainly suffice to satisfy condition (iv) of Definition 5.1.1.

PROPOSITION 5.3.3. *Let $\alpha : \omega \to \omega$ be a computable function. Then $\alpha^* : \mathcal{M}_\omega \to \mathcal{M}_\omega$ has an upper adjoint $\exists_\alpha$ and a lower adjoint $\forall_\alpha$.*

PROOF. Let us first consider the upper adjoint. We define:

$$(\exists_\alpha((\mathcal{A}_i)_{i\in\omega}))_m = \{n^\frown f \mid f \in \mathcal{A}_n \wedge \alpha(n) = m\}.$$

Then $\exists_\alpha$ is a well-defined function on $\mathcal{M}_\omega$. Namely, assume $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$, say through Φ. Let Ψ be the partial functional sending $m^\frown n^\frown h$ to $n^\frown \Phi(n^\frown h)$, then Ψ witnesses that $\exists_\alpha((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} \exists_\alpha((\mathcal{B}_i)_{i\in\omega})$.

We claim: $\exists_\alpha$ is an upper adjoint for α^*, i.e. $\alpha^*((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$ if and only if $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} \exists_\alpha((\mathcal{B}_i)_{i\in\omega})$. First, let us assume that $\alpha^*((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega}$

$(\mathcal{B}_i)_{i\in\omega}$; say through Φ. Let Ψ be the functional sending $j^\frown i^\frown h$ to $\Phi(i^\frown h)$. We claim: for every $m \in \omega$, $\Psi(m^\frown(\exists_\alpha((\mathcal{B}_i)_{i\in\omega}))_m) \subseteq \mathcal{A}_m$. Indeed, let $n^\frown f \in (\exists_\alpha((\mathcal{B}_i)_{i\in\omega}))_m$. Then $\alpha(n) = m$ and $f \in \mathcal{B}_n$. Thus, per choice of Φ we know that

$$\Psi(m^\frown n^\frown f) = \Phi(n^\frown f) \in \alpha^*((\mathcal{A}_i)_{i\in\omega})_n = \mathcal{A}_{\alpha(n)} = \mathcal{A}_m.$$

Conversely, assume $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} \exists_\alpha((\mathcal{B}_i)_{i\in\omega})$; say through Ψ. Let Φ be the functional sending $i^\frown h$ to $\Psi(\alpha(i)^\frown i^\frown h)$. Let $n \in \omega$. We claim:

$$\Phi(n^\frown \mathcal{B}_n) \subseteq (\alpha^*((\mathcal{A}_i)_{i\in\omega}))_n = \mathcal{A}_{\alpha(n)}.$$

Indeed, let $f \in \mathcal{B}_n$. Then $n^\frown f \in (\exists_\alpha((\mathcal{B}_i)_{i\in\omega}))_{\alpha(n)}$. Thus:

$$\Phi(n^\frown f) = \Psi(\alpha(n)^\frown n^\frown f) \in \mathcal{A}_{\alpha(n)}.$$

Next, we consider the lower adjoint. We define:

$$(\forall_\alpha((\mathcal{A}_i)_{i\in\omega}))_m = \left\{ \bigoplus_{n\in\omega} f_n \mid \forall n((\alpha(n) = m \wedge f_n \in \mathcal{A}_n) \vee (\alpha(n) \neq m \wedge f_n = 0)) \right\}.$$

Then $\forall_\alpha$ is a well-defined function on $\mathcal{M}_\omega$, as can be proven in a similar way as for $\exists_\alpha$. We claim that it is a lower adjoint for α^*, i.e. $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} \alpha^*((\mathcal{B}_i)_{i\in\omega})$ if and only if $\forall_\alpha((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$.

First, assume $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} \alpha^*((\mathcal{B}_i)_{i\in\omega})$, say through Φ. Let $m \in \omega$ and let $g \in \mathcal{B}_m$. Now let

$$f = \bigoplus_{n\in\omega} f_n$$

where $f_n = \Phi(n^\frown g)$ if $\alpha(n) = m$, and $f_n = 0$ otherwise. Note that, if $\alpha(n) = m$, then $g \in \mathcal{B}_m = (\alpha^*((\mathcal{B}_i)_{i\in\omega}))_n$, so $\Phi(n^\frown g) \in \mathcal{A}_n$. Thus, $f \in (\forall_\alpha(h_a))_m$. Note that this reduction is uniform in g and m, so $\forall_\alpha((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$.

Conversely, assume $\forall_\alpha((\mathcal{A}_i)_{i\in\omega}) \leq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i\in\omega}$, say through Ψ. Let $n \in \omega$ and let $g \in \alpha^*((\mathcal{B}_i)_{i\in\omega})_n = \mathcal{B}_{\alpha(n)}$. Then $\Psi(\alpha(n)^\frown g) \in (\forall_\alpha((\mathcal{A}_i)_{i\in\omega}))_{\alpha(n)}$. Since clearly $\alpha(n) = \alpha(n)$, it follows that $\Psi(\alpha(n)^\frown g)^{[n]} \in \mathcal{A}_n$. Again this reduction is uniform in n and g, so $(\mathcal{A}_i)_{i\in\omega} \leq_{\mathcal{M}_\omega} \alpha^*((\mathcal{B}_i)_{i\in\omega})$. □

Remark 5.3.4. Note that, if $\alpha : \omega \to \omega$ is is the projection to the first coordinate (i.e. the function mapping $\langle n, m\rangle$ to n), then

$$\forall_\alpha((\mathcal{A}_i)_{i\in\omega}) \equiv_{\mathcal{M}_\omega} \left(\left\{ \bigoplus_{m\in\omega} f_m \mid f_m \in \mathcal{A}_{\langle i,m\rangle} \right\}\right)_{i\in\omega}.$$

We will tacitly identify these two. Similarly,

$$\exists_\alpha((\mathcal{A}_i)_{i\in\omega}) \equiv_{\mathcal{M}_\omega} \left(\{m^\frown f_m \mid f_m \in \mathcal{A}_{\langle i,m\rangle}\}\right)_{i\in\omega}.$$

We now generalise this notion to include all the functions in our category $\mathbf{C}$. We will define $\mathcal{P}_{\mathcal{M}}(\{1,\dots,n\})$ to be the n-fold product $\mathcal{M}^n$.

Definition 5.3.5. Let $X, Y \in \{\{1\},\{1,2\},\dots\} \cup \{\omega\}$. Let $\alpha : X \to Y$ be computable. Then $\alpha^* : \mathcal{P}_{\mathcal{M}}(Y) \to \mathcal{P}_{\mathcal{M}}(X)$ is the function given by

$$\alpha^*((\mathcal{A}))_i = \mathcal{A}_{\alpha(i)}.$$

PROPOSITION 5.3.6. *The functions from Definition 5.3.5 are well-defined Brouwer algebra homomorphisms.*

PROOF. As in Proposition 5.3.2. ☐

PROPOSITION 5.3.7. *Let* $X, Y \in \{\{1\}, \{1,2\}, \dots\} \cup \{\omega\}$ *and let* $\alpha : X \to Y$ *be computable. Then* α^* *has both upper and lower adjoints.*

PROOF. As in Proposition 5.3.3. ☐

Thus, everything we have done above leads us to the following definition.

DEFINITION 5.3.8. Let $\mathbf{C}$ be the category with objects $\{1\}, \{1,2\}, \dots$ and ω and functions the computable functions between them. Let $\mathcal{P}_{\mathcal{M}}$ be the functor sending a finite set $\{1, \dots, n\}$ to $\mathcal{M}^n$, ω to $\mathcal{M}_\omega$ and α to α^*. We call this the *hyperdoctrine of mass problems.*

We now verify that the remaining conditions of Definition 5.1.1 hold for $\mathcal{P}_{\mathcal{M}}$.

THEOREM 5.3.9. *The functor* $\mathcal{P}_{\mathcal{M}}$ *from Definition 5.3.8 is a first-order hyperdoctrine.*

PROOF. First note that $\mathbf{C}$ is closed under all n-fold products, because ω^n is isomorphic to ω through some fixed computable function $\langle a_1, \dots, a_n \rangle$, and similarly $\{1, \dots, m\}^n$ is isomorphic to $\{1, \dots, mn\}$.

We now verify the conditions from Definition 5.1.1. Condition (i) follows from Proposition 5.2.3. Condition (ii) follows from Proposition 5.3.6. For condition (iii), use the fact that diagonal morphisms are computable together with Proposition 5.3.7. From the same theorem we know that the projections have lower and upper adjoints. Thus, we only need to verify that the Beck-Chevalley condition holds for them to verify condition (iv). Consider the diagram

$$\begin{tikzcd}
\mathcal{P}_{\mathcal{M}}(\Gamma' \times X) \arrow[r, "(s \times 1_X)^*"'] \arrow[d, "(\exists x)_{\Gamma'}"'] & \mathcal{P}_{\mathcal{M}}(\Gamma \times X) \arrow[d, "(\exists x)_{\Gamma}"'] \\
\mathcal{P}_{\mathcal{M}}(\Gamma') \arrow[r, "s^*"'] & \mathcal{P}_{\mathcal{M}}(\Gamma),
\end{tikzcd}$$

we need to show that it commutes.

We have:

$$((\exists x)_\Gamma((s \times 1_X)^*((\mathcal{A}_i)_{i \in \Gamma' \times X})))_n = \{m^\frown \langle n, m \rangle^\frown f \mid f \in \mathcal{A}_{\langle s(n), m \rangle}\}$$

and

$$(s^*((\exists x)_{\Gamma'}((\mathcal{A}_i)_{i \in \Gamma' \times X})))_n = \{\langle s(n), m \rangle^\frown m^\frown f \mid f \in \mathcal{A}_{\langle s(n), m \rangle}\}$$

by Remark 5.3.4. Then

$$s^*((\exists x)_{\Gamma'}((\mathcal{A}_i)_{i \in \Gamma \times X})) \leq_{\mathcal{M}_\omega} ((\exists x)_\Gamma((s \times 1_X)^*((\mathcal{A}_i)_{i \in \Gamma \times X})))$$

through the functional sending $i^\frown k^\frown \langle n, m \rangle^\frown f$ to $\langle s(n), m \rangle^\frown m^\frown f$, and the opposite inequality holds through the functional sending $n^\frown \langle l, m \rangle^\frown k^\frown f$ to $m^\frown \langle n, m \rangle^\frown f$.

Next, consider

$$\begin{tikzcd}
\mathcal{P}_{\mathcal{M}}(\Gamma' \times X) \arrow[r, "(s\times 1_X)^*"'] \arrow[d, "(\forall x)_{\Gamma'}"'] & \mathcal{P}_{\mathcal{M}}(\Gamma \times X) \arrow[d, "(\forall x)_{\Gamma}"'] \\
\mathcal{P}_{\mathcal{M}}(\Gamma') \arrow[r, "s^*"'] & \mathcal{P}_{\mathcal{M}}(\Gamma),
\end{tikzcd}$$

we need to show that this also commutes.

Again by Remark 5.3.4 we have:

$$\begin{aligned}
((\forall x)_{\Gamma}((s \times 1_X)^*((\mathcal{A}_i)_{i \in \Gamma' \times X})))_n &= \left\{ \bigoplus_{m \in \omega} f_m \mid f_m \in \mathcal{A}_{\langle s(n), m \rangle} \right\} \\
&= (s^*((\forall x)_{\Gamma'}((\mathcal{A}_i)_{i \in \Gamma' \times X})))_n,
\end{aligned}$$

as desired. □

For future reference, we state the following lemma which directly follows from the formula for the upper adjoint given in the proof of Proposition 5.3.3.

LEMMA 5.3.10. *For any X, the equality $=_X$ in $\mathcal{P}_{\mathcal{M}}$ is given by:*

$$(=_X)_{\langle n,m \rangle} = \begin{cases} \omega^\omega & \text{if } n = m \\ \emptyset & \text{otherwise.} \end{cases}$$

PROOF. From the formula given for the upper adjoint in the proof of Proposition 5.3.3, and the definition of $=_X$ in a first-order hyperdoctrine in Definition 5.1.1. □

Finally, let us give an easy example of a structure in $\mathcal{P}_{\mathcal{M}}$. More examples will follow in the next sections.

EXAMPLE 5.3.11. Consider the language consisting of a constant 0, a unary function S and binary functions $+$ and $\cdot$. We define a structure $\mathfrak{M}$ in $\mathcal{P}_{\mathcal{M}}$. Let the universe M be ω. Take the interpretation to be the standard model, i.e. $[\![0]\!] = 0$, $[\![S]\!](n) = S(n)$, $[\![+]\!](n,m) = n + m$ and $[\![\cdot]\!](n,m) = n \cdot m$. Then the sentences which hold in $\mathfrak{M}$ are exactly those which have a computable realiser in Kleene's second realisability model, see Kleene and Vesley [56, p. 96].

Thus, the hyperdoctrine of mass problems can be seen as an extension of Kleene's second realisability model with computable realisers. There is also a topos which can be seen as an extension of this model, namely the Kleene–Vesley topos, see e.g. van Oosten [89]. However, this topos does not follow Kolmogorov's philosophy that the interpretation of the universal quantifier should be uniform in the variable. On the other hand, a topos can interpret much more than just first-order logic.

Note that our category **C** only contains countable sets. On one hand this could be seen as a restriction, but on the other hand this should not come as a surprise since we are dealing with computability. That it is not that much of a restriction is illustrated by the rich literature on computable model theory dealing with computable, countable models.

5.4. Theory of the hyperdoctrine of mass problems

Given a first-order language Σ, we wonder what the theory of $\mathcal{P}_{\mathcal{M}}$ is. In particular, we want to know: is the theory of $\mathcal{P}_{\mathcal{M}}$ equal to first-order intuitionistic logic IQC? To this, the answer is 'no' in general: it is well-known that the weak law of the excluded middle $\neg\varphi \vee \neg\neg\varphi$ holds in the Medvedev lattice; therefore $\neg\varphi \vee \neg\neg\varphi$ holds for sentences in $\mathcal{P}_{\mathcal{M}}$. However, recall from Theorem 2.1.13 that for the Medvedev lattice Skvortsova has shown that there is an $\mathcal{A} \in \mathcal{M}$ such that $\mathrm{Th}(\mathcal{M}/\mathcal{A}) = \mathrm{IPC}$. Thus, Skortsova's result tells us that there is a principal factor of the Medvedev lattice which captures exactly intuitionistic propositional logic.

There is a natural way to extend principal factors to the hyperdoctrine of mass problems: given $\mathcal{A}$ in $\mathcal{M}$, let $\mathcal{P}_{\mathcal{M}/\mathcal{A}}$ be as in Definition 5.3.8, but with $\mathcal{M}$ replaced by $\mathcal{M}/\mathcal{A}$, and $\mathcal{M}_\omega$ replaced by $\mathcal{M}_\omega/(\mathcal{A}, \mathcal{A}, \dots)$. It is directly verified that $\mathcal{P}_{\mathcal{M}/\mathcal{A}}$ is also a first-order hyperdoctrine. Thus, there is a first-order analogue to the problem studied by Skortsova in the propositional case: is there an $\mathcal{A} \in \mathcal{M}$ such that the sentences that hold in $\mathcal{P}_{\mathcal{M}/\mathcal{A}}$ are exactly those that are deducible in IQC?

First, note that equality is always decidable (i.e. $\forall x, y(x = y \vee \neg x = y)$ holds) by the analogue of Lemma 5.3.10 (with ω^ω replaced by $\mathcal{A}$). So, can we get the theory to equal IQC plus decidable equality? Surprisingly, the answer turns out to be 'no' in general. Recall that for a poset X and $x, y \in X$ with $x \leq y$ we have that the interval $[x, y]_X$ denotes the set of elements $z \in X$ with $x \leq z \leq y$. If $\mathscr{B}$ is a Brouwer algebra then so is $[x, y]_{\mathscr{B}}$, with lattice operations as in $\mathscr{B}$ and implication given by

$$u \to_{[x,y]_{\mathscr{B}}} v = (u \to_{\mathscr{B}} v) \oplus x.$$

If $x = 0$, this gives us exactly the factor $\mathscr{B}/y$.

We can use this to introduce a specific kind of intervals in the hyperdoctrine of mass problems.

Definition 5.4.1. Let $\mathcal{A}, \mathcal{B} \in \mathcal{M}$. Then the *interval* $[\mathcal{B}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ is the first-order hyperdoctrine defined as in Definition 5.3.8, but with $\mathcal{M}$ replaced by $[\mathcal{B}, \mathcal{A}]_{\mathcal{M}}$, and $\mathcal{M}_\omega$ replaced by $[(\mathcal{B}, \mathcal{B}, \dots), (\mathcal{A}, \mathcal{A}, \dots)]_{\mathcal{M}_\omega}$.

It can be directly verified that this is a first-order hyperdoctrine; if one is not convinced this also follows from the more general Theorem 5.4.5 below.

The axiom schema CD, consisting of all formulas of the form $\forall z(\varphi(z) \vee \psi) \to \forall z(\varphi(z)) \vee \psi$, has been studied because it characterises the Kripke frames with constant domain. Our first counterexample is based on the fact that a specific instance of this schema holds in every structure in an interval of $\mathcal{M}$ with finite universe.

Proposition 5.4.2. *Consider the language consisting of one nullary relation R, one unary relation S and equality. Then in every structure in every interval $[\mathcal{B}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ the formula*

$$\forall x, y, z(x = y \vee x = z \vee y = z) \wedge \forall z(S(z) \vee R) \to \forall z(S(z)) \vee R$$

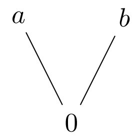
holds. However, this formula is not in IQC.

Proof. Let $\mathfrak{M}$ be a structure in $[\mathcal{B}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$. Note that by the analogue of Lemma 5.3.10 we know that if $\forall x, y, z(x = y \vee x = z \vee y = z)$ does not hold, then it gets interpreted as $\mathcal{A}$ and then the formula certainly holds. However, $\forall x, y, z(x = y \vee x = z \vee y = z)$ can only hold if $\mathcal{M}$ has at most two elements. Let us first assume $\mathfrak{M}$ has two elements. Let f be an element of $[\![\forall z(S(z) \vee R)]\!]$. Then $f = f_1 \oplus f_2$, with $f_1 \in [\![S(z) \vee R]\!]_1$ and $f_2 \in [\![S(z) \vee R]\!]_2$.[1] There are two cases: either both f_1 and f_2 start with a 0 and we can compute an element of $[\![\forall z S(z)]\!]$, or one of them starts with a 1 in which case we can compute an element of $[\![R]\!]$. Since the reduction is uniform in f, we see that

$$[\![\forall z(S(z) \vee R)]\!] \geq_{\mathcal{M}} [\![\forall z(S(z)) \vee R]\!],$$

and thus the formula given in the statement of the proposition holds. If $\mathfrak{M}$ has only one element, a similar proof yields the same result.

To show that the formula is not in IQC, consider the following Kripke frame:



.

Let $\mathfrak{K}_0$ have universe $\{1\}$ and let $\mathfrak{K}_a, \mathfrak{K}_b$ have universe $\{1, 2\}$. Let $S(1)$ be true everywhere, let $S(2)$ be true only at a and let R be true only at b. Then $\mathfrak{K}$ is a Kripke model refuting the formula in the statement of the proposition. □

Note that the schema CD can be refuted in $\mathcal{P}_{\mathcal{M}}$, as long as we allow models over infinite structures: namely, let $\varphi(z) = S(z)$ and $\psi(z) = R$. We build a structure $\mathfrak{M}$ with ω as universe. Let A be computably independent. Let $[\![S]\!]_n = A^{[n+1]}$ and let $[\![R]\!] = A^{[0]}$. Towards a contradiction, assume CD holds in this structure and let Φ witness $[\![\forall z(S(z) \vee R)]\!] \geq_{\mathcal{M}} [\![\forall z(S(z)) \vee R]\!]$. Now the function f given by $f^{[n]} = 0^\frown A^{[n+1]}$ is in $[\![\forall z(S(z) \vee R)]\!]$, so $\Phi(f) \in [\![\forall z(S(z)) \vee R]\!]$. Because A is computably independent f cannot compute $A^{[0]}$, so $\Phi(f)(0) = 0$. Let u be the use of this computation and let g be the function such that $g^{[n]} = f^{[n]}$ for $n \leq u$ and $g^{[n]} = A^{[0]}$ for $n > u$. Then $\Phi(g)(0) = 0$ so g computes $A^{[u+1]}$, contradicting A being computably independent.

Thus, one might object to our counterexample for being too unnatural by restricting the universe to be finite. However, the next example shows that even without this restriction we can find a counterexample.

PROPOSITION 5.4.3. *Consider the language consisting of a unary relations R. Then in every structure in every interval $[\mathcal{B}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ the formula*

$$(\forall x(S(x) \vee \neg S(x)) \wedge \neg \forall x(\neg S(x))) \rightarrow \exists x(\neg\neg S(x)).$$

holds. However, this formula is not in IQC.

PROOF. Towards a contradiction, assume $\mathfrak{M}$ is some structure satisfying the formula. Let $f \in [\![\forall x(S(x) \vee \neg S(x))]\!]$ and let $g \in [\![\neg \forall x(\neg S(x))]\!]$. If for every $n \in \mathfrak{M}$ we have $f^{[n]}(0) = 1$ then f computes an element of $[\![\forall x \neg S(x)]\!]$, which together with g computes an element of the top element $\mathcal{A}$ so then we are done. Otherwise

[1]Note that $[\![S(z) \vee R]\!] \in \mathcal{M}^2$, so $[\![S(z) \vee R]\!]_1$ and $[\![S(z) \vee R]\!]_2$ denote the first respectively second component.

we can compute from f some $n \in \mathfrak{M}$ with $f^{[n]}(0) = 0$. Let $\tilde{f}$ be $f^{[n]}$ without the first bit. Let e be an index for the functional sending $(k^\frown h_1) \oplus h_2$ to $\Phi_k(h_2 \oplus h_1)$. Then if $k^\frown h_1 \in [\![\neg S(x)]\!]_n$ we have

$$\Phi_e((k^\frown h_1) \oplus \tilde{f}) = \Phi_k(\tilde{f} \oplus h_1) \in \mathcal{A},$$

so $e^\frown \tilde{f} \in [\![\neg\neg S(x)]\!]_n$. Therefore $n^\frown e^\frown \tilde{f} \in [\![\exists x(\neg\neg S(x))]\!]$. So

$$[\![\forall x(S(x) \vee \neg S(x))]\!] \oplus [\![\neg\forall x(\neg S(x))]\!] \geq_{\mathcal{M}_\omega} [\![\exists x(\neg\neg S(x))]\!].$$

To show that the formula is not in IQC, consider the following Kripke frame:

$$\begin{tikzcd} a \arrow[d, dash] \\ 0 \end{tikzcd}.$$

Let $\mathfrak{K}_0$ have universe $\{1\}$ and let $\mathfrak{K}_a$ have universe $\{1, 2\}$. Let $S(1)$ be false everywhere and let $S(2)$ be true only at a. Then $\mathfrak{K}$ is a Kripke model refuting the formula in the statement of the proposition. $\square$

What the last theorem really says is not that our approach is hopeless, but that instead of looking at intervals $[\mathcal{B}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$, we should look at more general intervals. Right now we are taking the bottom element $\mathcal{B}$ to be the same for each $i \in \omega$. Compare this with what happens if in a Kripke model we take the domain at each point to be the same: then CD holds in the Kripke model. Proposition 5.4.2 should therefore not come as a surprise (although it is surprising that the full schema can be refuted). Instead, we should allow $\mathcal{B}_i$ to vary (subject to some constraints); roughly speaking $\mathcal{B}_i$ then expresses the problem of 'showing that i exists' or 'constructing i'. This motivates the next definition.

Definition 5.4.4. Let $\mathcal{A} \in \mathcal{M}$ and $(\mathcal{B}_i)_{i \geq -1} \in \mathcal{M}_\omega$ be such that

$$(\mathcal{A}, \mathcal{A}, \dots) \geq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i \in \omega} \geq_{\mathcal{M}_\omega} (\mathcal{B}_{-1}, \mathcal{B}_{-1}, \dots)$$

and such that $\mathcal{B}_i \not\geq_\mathcal{M} \mathcal{A}$ for all $i \geq -1$. We define the *interval* $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ as follows. Let $\mathbf{C}$ be the category with as objects $\{\{1, \dots, m\}^n \mid n, m \in \omega\} \cup \{\omega, \omega^2, \dots\}$. Let the functions in $\mathbf{C}$ be the computable functions α which additionally satisfy that $\mathcal{B}_y \geq_\mathcal{M} \mathcal{B}_{\alpha(y)}$ for all $y \in \mathrm{dom}(\alpha)$ uniformly in y, where we define $\mathcal{B}_{(y_1, \dots, y_n)}$ to be $\mathcal{B}_{y_1} \oplus \dots \oplus \mathcal{B}_{y_n}$. Then $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ consists of the functor $\mathcal{P}$ sending $\{1, \dots, m\}^n$ to the Brouwer algebra $\left[(\mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n})_{(a_1, \dots, a_n) \in \{1, \dots, m\}^n}, (\mathcal{A}, \mathcal{A}, \dots)\right]_{\mathcal{M}^{mn}}$, sending ω^n to the Brouwer algebra $[(\mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n})_{\langle a_1, \dots, a_n \rangle \in \omega}, (\mathcal{A}, \mathcal{A}, \dots, \mathcal{A})]_{\mathcal{M}_\omega}$ and sending every function $\alpha : Y \to Z$ to $\mathcal{P}_\mathcal{M}(\alpha) \oplus (\mathcal{B}_i)_{i \in Y}$, where we implicitly identify ω^n with ω and $\{1, \dots, m\}^n$ with $\{1, \dots, mn\}$ through some fixed computable bijection $\langle a_1, \dots, a_n \rangle$.

Theorem 5.4.5. *The interval $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ is a first-order hyperdoctrine.*

Proof. First, note that the base category $\mathbf{C}$ is closed under n-fold products: indeed, the n-fold product of Y is just Y^n, and the projections are computable functions satisfying the extra requirement. Furthermore, if $\alpha_1, \dots, \alpha_n : Y \to Z$ are in $\mathbf{C}$, then $(\alpha_1, \dots, \alpha_n) : Y^n \to Z$ in in $\mathbf{C}$ because for all $y_1, \dots, y_n \in Y$ we have

$$\mathcal{B}_{(y_1,\dots,y_n)} = \mathcal{B}_{y_1} \oplus \cdots \oplus \mathcal{B}_{y_n} \geq_{\mathcal{M}} \mathcal{B}_{\alpha(y_1)} \cdots \oplus \mathcal{B}_{\alpha(y_n)} = \mathcal{B}_{(\alpha_1,\dots,\alpha_n)(y_1,\dots,y_n)},$$

with reductions uniform in $y_1, \dots, y_n$. Finally, for each α in $\mathbf{C}$ we have that $\mathcal{P}_{\mathcal{M}}(\alpha) \oplus 0_{\mathcal{P}(Y)}$ is a Brouwer algebra homomorphism: that joins and meets are preserved follows by distributivity, that the top element is preserved follows directly from $(\mathcal{A}, \mathcal{A}, \dots) \geq_{\mathcal{M}_\omega} (\mathcal{B}_i)_{i \in \omega} \geq_{\mathcal{M}} (\mathcal{B}_{-1}, \mathcal{B}_{-1}, \dots)$ and that the bottom element is preserved follows from the assumption that $\mathcal{B}_y \geq_{\mathcal{M}} \mathcal{B}_{\alpha(y)}$ for all $y \in \mathrm{dom}(\alpha)$ uniformly in y. That implication is preserved is more work: let $\alpha : X \to Y$. Throughout the remainder of the proof we will implicitly identify ω^n with ω and $\{1, \dots, m\}^n$ with $\{1, \dots, mn\}$ through some fixed bijection $\langle a_1, \dots, a_n \rangle$. Now:

$$\begin{array}{ll} & ((\mathcal{P}_{\mathcal{M}}(\alpha)((\mathcal{C}_i)_{i \in Y}))_j \oplus \mathcal{B}_j) \to_{[\mathcal{B}_j, \mathcal{A}]_{\mathcal{M}}} ((\mathcal{P}_{\mathcal{M}}(\alpha)((\mathcal{D}_i)_{i \in Y}))_j \oplus \mathcal{B}_j) \\ = & ((\mathcal{C}_{\alpha(j)} \oplus \mathcal{B}_j) \to (\mathcal{D}_{\alpha(j)} \oplus \mathcal{B}_j)) \oplus \mathcal{B}_j \\ \equiv_{\mathcal{M}} & (\mathcal{C}_{\alpha(j)} \to \mathcal{D}_{\alpha(j)}) \oplus \mathcal{B}_j \\ = & (\mathcal{P}_{\mathcal{M}}(\alpha)((\mathcal{C}_i)_{i \in Y} \to (\mathcal{D}_i)_{i \in Y}))_j \oplus \mathcal{B}_j, \end{array}$$

with uniform reductions.

Thus, we need to verify that the product projections have adjoints; in fact, we will show that every morphism α in the base category $\mathbf{C}$ has adjoints. Let $\alpha : X \to Y$. We claim: $\mathcal{P}_{\mathcal{M}}(\alpha) \oplus (\mathcal{B}_i)_{i \in X}$ has as an upper adjoint $\exists_\alpha$ and as a lower adjoint the map sending $(\mathcal{C}_i)_{i \in X}$ to $\forall_\alpha((\mathcal{B}_i \to_{\mathcal{M}} \mathcal{C}_i)_{i \in X}) \oplus (\mathcal{B}_i)_{i \in Y}$, where $\exists_\alpha$ and $\forall_\alpha$ are as in Proposition 5.3.3. Indeed, we have:

$$(\mathcal{D}_i)_{i \in Y} \leq_{\mathcal{M}_\omega} \exists_\alpha((\mathcal{C}_i)_{i \in X}) \Leftrightarrow (\mathcal{D}_{\alpha(i)})_{i \in X} \leq_{\mathcal{M}_\omega} (\mathcal{C}_i)_{i \in X}$$

and because $(\mathcal{C}_i)_{i \in X} \in [(\mathcal{B}_i)_{i \in X}, (\mathcal{A}, \mathcal{A}, \dots)]_{\mathcal{M}_\omega}$:

$$\Leftrightarrow (\mathcal{B}_i)_{i \in X} \oplus (\mathcal{D}_{\alpha(i)})_{i \in X} \leq_{\mathcal{M}_\omega} (\mathcal{C}_i)_{i \in X} \Leftrightarrow \mathcal{P}_{\mathcal{M}}(\alpha)((\mathcal{D}_i)_{i \in Y}) \oplus (\mathcal{B}_i)_{i \in X} \leq_{\mathcal{M}_\omega} (\mathcal{C}_i)_{i \in X}.$$

Similarly, for $\forall$ we have:

$$\begin{aligned} & \forall_\alpha((\mathcal{B}_i \to_{\mathcal{M}} \mathcal{C}_i)_{i \in X}) \oplus (\mathcal{B}_i)_{i \in Y} \leq_{\mathcal{M}_\omega} (\mathcal{D}_i)_{i \in Y} \\ \Leftrightarrow & \forall_\alpha((\mathcal{B}_i \to_{\mathcal{M}} \mathcal{C}_i)_{i \in X}) \leq_{\mathcal{M}_\omega} (\mathcal{D}_i)_{i \in Y} \\ \Leftrightarrow & (\mathcal{B}_i \to_{\mathcal{M}} \mathcal{C}_i)_{i \in X} \leq_{\mathcal{M}_\omega} (\mathcal{D}_{\alpha(i)})_{i \in X} \\ \Leftrightarrow & (\mathcal{C}_i)_{i \in X} \leq_{\mathcal{M}_\omega} (\mathcal{B}_i \oplus \mathcal{D}_{\alpha(i)})_{i \in X} \\ \Leftrightarrow & (\mathcal{C}_i)_{i \in X} \leq_{\mathcal{M}_\omega} \mathcal{P}_{\mathcal{M}}(\alpha)((\mathcal{D}_i)_{i \in Y}) \oplus (\mathcal{B}_i)_{i \in X}. \end{aligned}$$

Finally, we need to verify that $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ satisfies the Beck-Chevalley condition. We have (writing α^* for the image of the morphism α under the functor for $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$):

$$((\exists x)_\Gamma((s \times 1_X)^*((\mathcal{C}_i)_{i \in \Gamma' \times X})))_n = \{m^\frown \langle n, m \rangle^\frown f \mid f \in \mathcal{C}_{(s(n),m)} \oplus \mathcal{B}_n \oplus \mathcal{B}_m\}$$

and

$$(s^*((\exists x)_{\Gamma'}((\mathcal{C}_i)_{i \in \Gamma' \times X})))_n = \{\langle s(n), m \rangle^\frown m^\frown f \mid f \in \mathcal{C}_{(s(n),m)} \oplus \mathcal{B}_n \oplus \mathcal{B}_{s(n)}\}.$$

As in the proof of Theorem 5.3.9 we have

$$s^*((\exists x)_{\Gamma'}((\mathcal{C}_i)_{i \in \Gamma' \times X})) \leq_{\mathcal{M}_\omega} ((\exists x)_\Gamma((s \times 1_X)^*((\mathcal{C}_i)_{i \in \Gamma' \times X}))).$$

The opposite inequality is also almost the same as in the proof of Theorem 5.3.9, except that we now need to use that $\mathcal{C}_{(s(n),m)}$ uniformly computes an element of $\mathcal{B}_{(s(n),m)}$ and hence of $\mathcal{B}_m$.

For the other part of the Beck-Chevalley condition we have:

$$\begin{aligned}
& (((\forall x)_\Gamma((\mathcal{B}_i)_{i\in\Gamma'\times X} \to (s\times 1_X)^*((\mathcal{C}_i)_{i\in\Gamma'\times X}))) \oplus (\mathcal{B}_i)_{i\in\Gamma})_n \\
= \quad & \left\{\bigoplus_{m\in X} f_m \mid f_m \in \mathcal{B}_n \oplus \mathcal{B}_m \to \left(\mathcal{B}_n \oplus \mathcal{B}_m \oplus \mathcal{C}_{(s(n),m)}\right)\right\} \oplus \mathcal{B}_n \\
\equiv_{\mathcal{M}} \quad & \left\{\bigoplus_{m\in X} f_m \mid f_m \in \mathcal{B}_m \to \mathcal{C}_{(s(n),m)}\right\} \oplus \mathcal{B}_n.
\end{aligned}$$

Now, using the fact that $\mathcal{B}_{s(n)}$ uniformly reduces to $\mathcal{B}_n$:

$$\begin{aligned}
\equiv_{\mathcal{M}} \quad & \left\{\bigoplus_{m\in X} f_m \mid f_m \in (\mathcal{B}_{s(n)} \oplus \mathcal{B}_m) \to \mathcal{C}_{(s(n),m)}\right\} \oplus \mathcal{B}_{s(n)} \oplus \mathcal{B}_n \\
= \quad & (s^*((\forall x)_{\Gamma'}((\mathcal{B}_i)_{i\in\Gamma'\times X} \to (\mathcal{C}_i)_{i\in\Gamma'\times X}) \oplus (\mathcal{B}_i)_{i\in\Gamma'}))_n,
\end{aligned}$$

as desired. □

In Propositions 5.6.8 and 5.6.9 below we will show that we can refute the formulas from Propositions 5.4.2 and 5.4.3 in these more general intervals. Next, let us rephrase Lemma 5.3.10 for our intervals.

LEMMA 5.4.6. *Given any X in the base category $\mathbf{C}$ of $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$, let 0_X and 1_X be the bottom respectively top elements of the Brouwer algebra corresponding to X. Then the equality $=_X$ in $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ is given by:*

$$(=_X)_{\langle n,m\rangle} = \begin{cases} 0_X & \text{if } n = m \\ 1_X & \text{otherwise.} \end{cases}$$

PROOF. From the formula given for the upper adjoint in the proof of Theorem 5.4.5, and the definition of $=_X$ in a first-order hyperdoctrine in Definition 5.1.1. □

As a final remark, note that we cannot vary $\mathcal{A}$ (i.e. make intervals of the form $[(\mathcal{B}_i)_{i\geq -1}, (\mathcal{A}_i)_{i\geq -1}]_{\mathcal{P}_\mathcal{M}}$): if we did, then to make α^* into a homomorphism we would need to meet with $\mathcal{A}_i$. While joining with $\mathcal{B}_i$ was not a problem, if we meet with $\mathcal{A}_i$ the implication will in general not be preserved.

5.5. Heyting arithmetic in the hyperdoctrine of mass problems

In the previous section we introduced the general intervals $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$. However, it turns out that even these intervals cannot capture every theory in IQC, which we will show by looking at models of Heyting arithmetic. Our approach is based on the following classical result about computable classical models of Peano arithmetic.

THEOREM 5.5.1. (Tennenbaum [116]) *There is no computable non-standard model of Peano arithmetic.*

PROOF. (Sketch) Let A, B be two c.e. sets which are computably inseparable and for which PA proves that they are disjoint (e.g. take $A = \{e \in \omega \mid \{e\}(e)\downarrow = 0\}$ and $B = \{e \in \omega \mid \{e\}(e)\downarrow = 1\}$). Let $\varphi(e) = \exists s \varphi'(e, s)$ define A and let $\psi(e) = \exists s \psi'(e, s)$ define B, where φ, ψ are Δ^0_0-formulas which are monotone in s. Now consider the following formulas:

$$\alpha_1 = \forall e, s \forall s' \geq s((\varphi'(e, s) \to \varphi'(e, s')) \wedge (\psi'(e, s) \to \psi'(e, s')))$$
$$\alpha_2 = \forall e, s(\neg(\varphi'(e, s) \wedge \psi'(e, s)))$$
$$\alpha_3 = \forall n, p \exists! a, b(b < p \wedge ap + b = n)$$
$$\alpha_4 = \forall n \exists m \forall e < n(\varphi'(e, n) \leftrightarrow \exists a < n(ap_e = m)),$$

where p_e denotes the eth prime.

These are all provable in PA. The first formula tells us that φ' and ψ' are monotone in s. The second formula expresses that A and B are disjoint. The third formula says that the Euclidean algorithm holds. The last formula tells us that for every n, we can code the elements of $A[n] \cap [0, n)$ as a single number. We can prove this inductively, by letting m be the product of those p_e such that $e \in A[n] \cap [0, n)$.

Thus, every non-standard model of Peano arithmetic also satisfies these formulas. Towards a contradiction, let $\mathfrak{M}$ be a computable non-standard model of PA. Let $n \in M$ be a non-standard element, i.e. $n > k$ for every standard k. Let $m \in M$ be such that

$$\mathfrak{M} \models \forall e < n(\varphi'(e, n) \leftrightarrow \exists a < n.ap_e = m).$$

If $e \in A$, then $\varphi'(e, s)$ holds in the standard model for large enough standard s, and since $\mathfrak{M}$ is a model of Q and φ' is Δ^0_0 we see that also $\mathfrak{M} \models \varphi'(e, s)$ for large enough standard s. By monotonicity, we therefore have $\mathfrak{M} \models \varphi'(e, n)$. Thus, $\mathfrak{M} \models \exists a < n.ap_e = m$.

Conversely, if $e \in B$, then $\mathfrak{M} \models \psi'(e, s)$ for large enough standard s, so by monotonicity we see that $\mathfrak{M} \models \psi'(e, n)$. Therefore, $\mathfrak{M} \models \neg\varphi'(e, n)$ by α_2. Thus, $\mathfrak{M} \models \neg(\exists a < n.ap_e = m)$. So, the set $C = \{e \in \omega \mid \mathfrak{M} \models \exists a, b < n.ap_{S^e(0)} = m\}$ separates A and B.

However, C is also computable: because the Euclidean algorithm holds in $\mathfrak{M}$, we know that there exist unique a, b with $b < p_{S^e(0)}$ such that $ap_{S^e(0)} + b = m$. Since $\mathfrak{M}$ is computable we can find those a and b computably. Now e is in C if and only if $b = 0$. This contradicts A and B being computably separable. □

When looking at models of arithmetic, we often use that fairly basic systems (like Robinson's Q) already represent the computable functions (a fact which we used in the proof of Tennenbaum's theorem above). In other words, this tells us that there is not much leeway to change the truth of Δ^0_1-statements. The next two lemmas show that in a language without any relations except equality (like arithmetic), as long as our formulas are Δ^0_1, their truth value in the hyperdoctrine of mass problems is essentially classical; in other words, there is also no leeway to make their truth non-classical.

LEMMA 5.5.2. *Let Σ be a language without relations (except possibly equality). Let $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ be an interval and let $\mathfrak{M}$ be a structure for Σ in*

$[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$. *Let* $\varphi(x_1, \dots, x_n)$ *be a* Δ^0_0*-formula and let* $a_1, \dots, a_n \in M$. *Then we have either*

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus B_{a_n}$$

or

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A},$$

with the first holding if and only if $\varphi(a_1, \dots, a_n)$ *holds classically in the classical model induced by* $\mathfrak{M}$ *(i.e. the classical model with universe* M *and functions as in* $\mathfrak{M}$*).*

Furthermore, it is decidable which of the two cases holds, and the reduction is uniform in $a_1, \dots, a_n$.

PROOF. We prove this by induction on the structure of φ.

- φ is of the form $t(x_1, \dots, x_n) = s(x_1, \dots, x_n)$: by Lemma 5.4.6 we know that $[\![t(x_1, \dots, x_n) = s(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle}$ is either $\mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus B_{a_n}$ or $\mathcal{A}$, with the first holding if and only if $t(a_1, \dots, a_n) = s(a_1, \dots, a_n)$ holds classically. Since all functions are computable and equality is true equality, it is decidable which of the two cases holds.
- φ is of the form $\psi(x_1, \dots, x_n) \wedge \chi(x_1, \dots, x_n)$: there are three cases:
  - If both $[\![\psi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle}$ and $[\![\chi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle}$ are equivalent to $\mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus B_{a_n}$, then $[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus B_{a_n}$,
  - If $[\![\psi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A}$, then $[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A}$ by sending $f \oplus g$ to f,
  - If $[\![\chi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A}$, then $[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A}$ by sending $f \oplus g$ to g.

  This case distinction is decidable because the induction hypothesis tells us that the truth of ψ and χ is decidable.
- φ is of the form $\psi(x_1, \dots, x_n) \to \chi(x_1, \dots, x_n)$: this follows directly from the fact that, in any Brouwer algebra with top element 1 and bottom element 0, we have $0 \to 1 = 1$ and $0 \to 0 = 1 \to 1 = 1 \to 0 = 0$. The case distinction is again decidable by the induction hypothesis.

The other cases are similar. $\square$

LEMMA 5.5.3. *Let* $\Sigma, \mathcal{A}, (\mathcal{B}_i)_{i\geq -1}$ *and* $\mathfrak{M}$ *be as in Lemma 5.5.2. Let* T *be some theory satisfied which is satisfied by* $\mathfrak{M}$, *i.e.* $[\![\psi]\!]_\mathfrak{M} = \mathcal{B}_{-1}$ *for every* $\psi \in T$. *Let* $\varphi(x_1, \dots, x_n)$ *be a formula which is* Δ^0_1 *over* T *and let* $a_1, \dots, a_n \in M$. *Then either*

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \mathcal{B}_{a_n}$$

or

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n\rangle} \equiv_\mathcal{M} \mathcal{A},$$

with the first holding if and only if $\varphi(a_1, \dots, a_n)$ *holds classically in* $\mathfrak{M}$.

Furthermore, it is decidable which of the two cases holds, and the reduction is uniform in $a_1, \dots, a_n$.

PROOF. Let

$$\varphi \Leftrightarrow \forall y_1,\dots,y_m \psi(x_1,\dots,x_n,y_1,\dots,y_m) \Leftrightarrow \exists y_1,\dots,y_m \chi(x_1,\dots,x_n,y_1,\dots,y_m),$$

where ψ and χ are Δ^0_0-formulas. Then by soundness (see Proposition 5.1.10) we know that

(10)
$$[\![\forall y_1,\dots,y_m \psi(x_1,\dots,x_n,y_1,\dots,y_m)]\!] \equiv_{\mathcal{M}_\omega} [\![\exists y_1,\dots,y_m \chi(x_1,\dots,x_n,y_1,\dots,y_m)]\!].$$

Let $a_1,\dots,a_n \in M$. We claim: there are some $b_1,\dots,b_m$ such that either

$$(11) \qquad [\![\psi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n,b_1,\dots,b_m\rangle} \equiv_{\mathcal{M}} \mathcal{A}$$

or

(12)
$$[\![\chi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n,b_1,\dots,b_m\rangle} \equiv_{\mathcal{M}} \mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n} \oplus \mathcal{B}_{b_1} \oplus \dots \oplus \mathcal{B}_{b_m}.$$

Indeed, otherwise we see from Lemma 5.5.2 and some easy calculations that

$$[\![\forall y_1,\dots,y_m \psi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n}$$

and

$$[\![\exists y_1,\dots,y_m \chi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{A},$$

which contradicts (10).

Thus, again by Lemma 5.5.2, we can find $b_1,\dots,b_m$ computably such that either (11) or (12) holds. First, if (11) holds, then it can be directly verified that

$$[\![\forall y_1,\dots,y_m \psi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{A},$$

while if (12) holds, then it can be directly verified that

$$[\![\exists y_1,\dots,y_m \chi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n},$$

with all the reductions uniform in $a_1,\dots,a_n$. □

Next, we slightly extend this to Π^0_1-formulas and Σ^0_1-formulas, although at the cost of dropping the uniformity.

LEMMA 5.5.4. *Let $\Sigma, \mathcal{A}, (\mathcal{B}_i)_{i\geq -1}$ and $\mathfrak{M}$ be as in Lemma 5.5.2 above. Let $\varphi(x_1,\dots,x_n)$ be a Π^0_1-formula or a Σ^0_1-formula and let $a_1,\dots,a_n \in M$. Then we have either*

$$[\![\varphi(x_1,\dots,x_n)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{B}_{-1} \oplus \mathcal{B}_{a_1} \oplus \dots \mathcal{B}_{a_n}$$

or

$$[\![\varphi(x_1,\dots,x_n)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \mathcal{A},$$

with the first holding if and only if $\varphi(a_1,\dots,a_n)$ holds classically in $\mathfrak{M}$.[2]

PROOF. Let $\varphi(x_1,\dots,x_n) = \forall y_1,\dots,y_m \psi(x_1,\dots,x_n,y_1,\dots,y_n)$ with ψ a Δ^0_0-formula. First, let us assume $\varphi(a_1,\dots,a_n)$ holds classically. Thus, for all $b_1,\dots,b_m \in M$ we know that $\psi(a_1,\dots,a_n,b_1,\dots,b_m)$ holds classically. By Lemma

[2] However, unlike the previous two lemmas, the reductions need not be uniform in $a_1,\dots,a_n$.

5.5.2 we then know that $\psi(a_1,\dots,a_n,b_1,\dots,b_m)$ gets interpreted as $\mathcal{B}_{a_1}\oplus\cdots\oplus\mathcal{B}_{a_n}\oplus\mathcal{B}_{b_1}\oplus\cdots\oplus\mathcal{B}_{b_m}$ (by a reduction uniform in $b_1,\dots,b_m$). Now note that

$$\begin{aligned}
&[\![\varphi]\!]_{\langle a_1,\dots,a_n\rangle}\\
\equiv_{\mathcal{M}}\quad&\bigoplus_{\langle b_1,\dots,b_m\rangle\in\omega}\Big((\mathcal{B}_{a_1}\oplus\cdots\oplus\mathcal{B}_{a_n}\oplus\mathcal{B}_{b_1}\oplus\cdots\oplus\mathcal{B}_{b_m})\\
&\to_{\mathcal{M}}[\![\psi(x_1,\dots,x_n,y_1,\dots,y_m)]\!]_{\langle a_1,\dots,a_n,b_1,\dots,b_m\rangle}\Big)\\
&\oplus(\mathcal{B}_{-1}\oplus\mathcal{B}_{a_1}\oplus\cdots\oplus\mathcal{B}_{a_n})\\
\equiv_{\mathcal{M}}\quad&\mathcal{B}_{-1}\oplus\mathcal{B}_{a_1}\oplus\cdots\oplus\mathcal{B}_{a_n}.
\end{aligned}$$

Now, let us assume $\varphi(a_1,\dots,a_n)$ does not hold classically. Let $b_1,\dots,b_m\in M$ be such that $\psi(a_1,\dots,a_n,b_1,\dots,b_m)$ does not hold classically. By Lemma 5.5.2 we know that $\psi(a_1,\dots,a_n,b_1\dots,b_m)$ gets interpreted as $\mathcal{A}$. Then it is directly checked that in fact

$$[\![\varphi(x_1,\dots,x_n)]\!]_{\langle a_1,\dots,a_n\rangle}\geq_{\mathcal{M}}\mathcal{A},$$

as desired.

The proof for Σ^0_1-formulas φ is similar. □

Now, we will prove an analogue of Theorem 5.5.1 for the hyperdoctrine of mass problems.

THEOREM 5.5.5. *Let Σ be the language of arithmetic (i.e. the language consisting of a function symbol for every primitive recursive function, and equality). There is a finite theory $T\supset Q$ derivable in HA such that for every interval $[(\mathcal{B}_i)_{i\geq-1},\mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ and every classically true Π^0_1-sentence or Σ^0_1-sentence χ we have that every structure $\mathfrak{M}$ in $[(\mathcal{B}_i)_{i\geq-1},\mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ satisfies $\bigwedge T\to\chi$. In particular this holds for $\chi=\mathrm{Con(PA)}$ and so for the language of arithmetic we have* $\mathrm{Th}\left([(\mathcal{B}_i)_{i\geq-1},\mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}\right)\neq\mathrm{IQC}$.

PROOF. Our proof is inspired by the proof of Theorem 5.5.1 given above. Let A, B, φ' and ψ' as in that proof. We first define a theory T' which consists of Q together with the formulas

$$\begin{aligned}
&\forall e,s\forall s'\geq s((\varphi'(e,s)\to\varphi'(e,s'))\wedge(\psi'(e,s)\to\psi'(e,s')))\\
&\forall n,s(\neg\varphi'(n,s)\wedge\psi'(n,s))\\
&\forall n,p\exists!a,b(b<p\wedge ap+b=n)\\
&\forall n\exists m\forall k,s<n(\varphi'(k,s)\leftrightarrow\exists a,b<n.ap_k=m).
\end{aligned}$$

Then T' is deducible in Peano arithmetic; in particular it holds in the standard model. Note that T' is equivalent to a Π^0_2-formula. Furthermore, note that there are computable Skolem functions (for example, take the function mapping n to the least witness). Thus, we can get rid of the existential quantifiers; for example, we can replace

$$\forall n,p\exists a,b(b<p\wedge ap+b=n)$$

by

$$\forall n,p(g(n,p)<p\wedge f(n,p)p+g(n,p)=n)$$

where f is the symbol representing the primitive recursive function sending (n,p) to n divided by p, and g is the symbol representing the primitive recursive function sending (n, p) to the remainder of the division of n by p. We can also turn Q into a Π^0_1-theory using the predecessor function.

So, let T consist of a Π^0_1-formula which is equivalent to T', together with Π^0_1 defining axioms for the finitely many computable functions we used. Then T is certainly deducible in PA, but it is also deducible in HA because every Π^0_2-sentence which is in PA in also in HA, see e.g. Troelstra and van Dalen [121, Proposition 3.5].

Now, if $\llbracket \bigwedge T \rrbracket \equiv_{\mathcal{M}} \mathcal{A}$, we are done. We may therefore assume this is not the case. Then, by Lemma 5.5.4 we see that T holds classically in $\mathfrak{M}$. Therefore T' also holds classically in $\mathfrak{M}$, and by the proof of Theorem 5.5.1 we see that $\mathfrak{M}$ is classically the standard model. Therefore χ holds classically in $\mathfrak{M}$ so we see by Lemma 5.5.4 that $\llbracket \chi \rrbracket \equiv_{\mathcal{M}} \mathcal{B}_{-1}$. □

5.6. Decidable frames

In the last section we saw that even in our intervals $[(\mathcal{B}_i)_{i \geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ we cannot generally obtain IQC. However, note that Heyting arithmetic, like Peano arithmetic is undecidable. We therefore wonder: what happens if we look at decidable theories? In the classical case, we know that every decidable theory has a decidable model. The intuitionistic case was studied by Gabbay [32] and Ishihara, Khoussainov and Nerode [44, 45], culminating in the following result.

Definition 5.6.1. A Kripke model is *decidable* if the underlying Kripke frame is computable, the universe at every node is computable and the forcing relation

$$w \Vdash \varphi(a_1, \dots, a_n)$$

is computable.

Definition 5.6.2. A theory is *decidable* if its deductive closure is computable and equality is decidable, i.e.

$$\forall x, y(x = y \vee \neg x = y)$$

holds.

Theorem 5.6.3. ([44, Theorem 5.1]) *Every decidable theory T has a decidable Kripke model, i.e. a decidable Kripke model whose theory is exactly the set of sentences deducible from T.*[3]

Our next result shows how to encode such decidable Kripke models in intervals of the hyperdoctrine of mass problems. Unfortunately we do not know how to deal with arbitrary decidable Kripke frames; instead we have to restrict to those without infinite ascending chains. As we will see later in this section, this nonetheless still proves to be useful.

---

[3]In [44] this result is stated for first-order languages without equality and function symbols. However, we can apply the original result to the language with an additional binary predicate R representing equality and to the theory T' consisting of T extended with the equality axioms. Using this equality we can now also represent functions by relations in the usual way.

THEOREM 5.6.4. *Let $\mathfrak{K}$ be a decidable Kripke model which is based on a Kripke frame without infinite ascending chains. Then there is an interval $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ and a structure $\mathfrak{M}$ in $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ such that the theory of $\mathfrak{M}$ is exactly the theory of $\mathfrak{K}$.*

Furthermore, if we allow infinite ascending chains, then this still holds for the fragments of the theories without universal quantifiers.

PROOF. Let $T = \{t_0, t_1, \dots\}$ be a computable representation of the poset T on which $\mathfrak{K}$ is based. Let $f_0, f_1, \dots$ be an antichain in the Turing degrees and let $\mathcal{D} = \{g \mid \exists i (g \leq_T f_i)\}$. Consider the collection $\mathcal{V} = \{C(\{f_i \mid i \in I\} \cup \overline{\mathcal{D}}) \mid I \subseteq \omega\}$. By Theorem 4.1.3, this is a sub-implicative semilattice of $[C(\{f_i \mid i \in \omega\}) \cup \overline{\mathcal{D}}, \overline{\mathcal{D}}]_\mathcal{M}$. We will use the mass problems $C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$ to represent the points t_j of the Kripke frame T. If T were finite, we would only have to consider a finite sub-upper semilattice of $\mathcal{V}$, and by Skvortsova [106, Lemma 2] the meet-closure of this would be exactly the Brouwer algebra of upwards closed subsets of T. However, since in our case T might be infinite, we need to suitably generalise this to arbitrary 'meets'.

Let us now describe how to do this. First, we define $\mathcal{A}$:

$$\mathcal{A} = \{k_1{}^\frown k_2{}^\frown \left(C(\{f_i \mid i \notin \{k_1, k_2\}\}) \cup \overline{\mathcal{D}}\right) \mid t_{k_1} \text{ and } t_{k_2} \text{ are incomparable})\}$$

if T is not a chain, and $\mathcal{A} = \overline{\mathcal{D}}$ otherwise. The idea behind $\mathcal{A}$ is that if t_{k_1} and t_{k_2} are incomparable in T, then there should be no mass problem representing a point above their representations.

Now, let $\mathcal{U}$ be the collection of upwards closed subsets of T. We then define the map $\alpha : \mathcal{U} \to \mathcal{M}$ by:

$$\alpha(Y) = \bigcup \{j^\frown \left(\left(C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}\right) \otimes \mathcal{A}\right) \mid t_j \in Y\},$$

and $\alpha(\emptyset) = \mathcal{A}$. Now let $\mathcal{B}_{-1} = \alpha(T)$ and let $\mathcal{B}_i = \alpha(Z_i)$, where Z_i is the set of nodes where i is in the domain of $\mathfrak{K}$. Then $\alpha : \mathcal{U} \to [\mathcal{B}_{-1}, \mathcal{A}]$ as a function; we are not yet claiming that it preserves the Brouwer algebra structure. We will prove a stronger result for a suitable sub-collection of $\mathcal{U}$ below.

First, let us show that α is injective. Indeed, assume $\alpha(Y) \leq_\mathcal{M} \alpha(Z)$. We will show that $Y \supseteq Z$. By applying Lemma 5.6.5 below twice we then have that for every j with $t_j \in Z$ there exists a k with $t_k \in Y$ such that either $C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}} \leq_\mathcal{M} C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$ or $\mathcal{A} \leq_\mathcal{M} C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$. In the first case, towards a contradiction let us assume that $k \neq j$. Then f_k computes an element of $C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$ and therefore $f_k \in C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}}$ since the latter is upwards closed. However, this contradicts the fact that the f_i form an antichain in the Turing degrees. Thus, $k = j$ and therefore $t_j \in Y$.

In the latter case, we have that $C(\{f_i \mid i \notin \{k_1, k_2\}) \cup \overline{\mathcal{D}} \leq_\mathcal{M} C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$ for some $k_1, k_2 \in \omega$ for which t_{k_1} and t_{k_2} are incomparable. Without loss of generality, let us assume that $k_1 \neq j$. Then, reasoning as above, we see that $f_{k_1} \in C(\{f_i \mid i \notin \{k_1, k_2\}) \cup \overline{\mathcal{D}}$, a contradiction.

For ease of notation, let us assume the union of the universes of $\mathfrak{K}$ is ω; the general case follows in the same way. Let $\mathfrak{M}$ be the structure with functions as in $\mathfrak{K}$, and let the interpretation of a relation $[\![R(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n \rangle}$ be $\alpha(Y)$, where Y is exactly the set of nodes where $R(a_1, \dots, a_n)$ holds in $\mathfrak{K}$.

We show that $\mathfrak{M}$ is as desired. To this end, we claim: for every formula $\varphi(x_1, \ldots, x_n)$ and every sequence $a_1, \ldots, a_n$,

$$[\![\varphi(x_1, \ldots, x_n)]\!]_{\langle a_1, \ldots, a_n \rangle} \equiv_{\mathcal{M}} \alpha(Y),$$

where Y is exactly the set of nodes where $a_1, \ldots, a_n$ are all in the domain and $\varphi(a_1, \ldots, a_n)$ holds in the Kripke model $\mathfrak{K}$. Furthermore, we claim that this reduction is uniform in $a_1, \ldots, a_n$ and in φ. We prove this by induction on the structure of φ. First, if φ is atomic, this follows directly from the choice of the valuations, from the fact that $\mathfrak{K}$ is decidable and from Lemma 5.4.6.

Next, let us consider $\varphi(x_1, \ldots, x_n) = \psi(x_1, \ldots, x_n) \vee \chi(x_1, \ldots, x_n)$. Let U be the set of nodes where $\psi(a_1, \ldots, a_n)$ holds in $\mathfrak{K}$ and similarly let V be the set of nodes where $\chi(a_1, \ldots, a_n)$ holds. By induction hypothesis and by the definition of the interpretation of $\vee$ we have

$$\begin{aligned}
& [\![\psi(x_1, \ldots, x_n) \vee \chi(x_1, \ldots, x_n)]\!]_{\langle a_1, \ldots, a_n \rangle} \\
\equiv_{\mathcal{M}} \quad & \alpha(U) \otimes \alpha(V) \\
= \quad & \bigcup \{ j^\frown \big((C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}) \otimes \mathcal{A} \big) \mid t_j \in U \} \\
& \otimes \bigcup \{ j^\frown \big((C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}) \otimes \mathcal{A} \big) \mid t_j \in V \}.
\end{aligned}$$

We need to show that this is equivalent to

$$\alpha(Y) = \bigcup \{ j^\frown \big((C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}) \otimes \mathcal{A} \big) \mid t_j \in Y \},$$

where Y is the set of nodes where $\varphi(a_1, \ldots, a_n)$ holds. First, let $j^\frown f \in \alpha(Y)$. Then $\varphi(a_1, \ldots, a_n)$ holds at t_j. Thus, by the definition of truth in Kripke frames, we know that at least one of $\psi(a_1, \ldots, a_n)$ and $\chi(a_1, \ldots, a_n)$ holds in t_j, and because our frame is decidable we can compute which of them holds. So, send $j^\frown f$ to $0^\frown j^\frown f$ if $\psi(a_1, \ldots, a_n)$ holds, and to $1^\frown j^\frown f$ otherwise. Thus, $\alpha(U) \otimes \alpha(V) \leq_{\mathcal{M}} \alpha(Y)$. Conversely, if either $\psi(a_1, \ldots, a_n)$ or $\chi(a_1, \ldots, a_n)$ holds then $\varphi(a_1, \ldots, a_n)$ holds, so the functional sending $i^\frown j^\frown f$ to $j^\frown f$ witnesses that $\alpha(Y) \leq_{\mathcal{M}} \alpha(U) \otimes \alpha(V)$.

The proof for conjunction is similar. Next, let us consider implication. So, let $\varphi(x_1, \ldots, x_n) = \psi(x_1, \ldots, x_n) \to \chi(x_1, \ldots, x_n)$. Let U be the set of nodes where $\psi(a_1, \ldots, a_n)$ holds in $\mathfrak{K}$, let V be the set of nodes where $\chi(a_1, \ldots, a_n)$ holds and let Y be the set of nodes where $\varphi(a_1, \ldots, a_n)$ holds. By induction hypothesis, we know that

$$\begin{aligned}
& [\![\varphi(x_1, \ldots, x_n)]\!]_{\langle a_1, \ldots, a_n \rangle} \\
\equiv_{\mathcal{M}} \; & \alpha(U) \to_{[\mathcal{B}_{(a_1, \ldots, a_n)}, \mathcal{A}]} \alpha(V).
\end{aligned}$$

First, note that $\alpha(Y) \geq_{\mathcal{M}} \alpha(U) \to_{[\mathcal{B}_{(a_1, \ldots, a_n)}, \mathcal{A}]} \alpha(V)$ is equivalent to $\alpha(Y) \oplus \alpha(U) \geq_{\mathcal{M}} \alpha(V)$. So, let $k^\frown h \in \alpha(Y)$ and $j^\frown g \in \alpha(U)$. Then $t_k \in Y$, $h \in \big(C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}}\big) \otimes \mathcal{A}$, $t_j \in U$ and $g \in \big(C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}\big) \otimes \mathcal{A}$. We need to uniformly compute from this some $m \in \omega$ with $t_m \in Y$ and an element of $\big(C(\{f_i \mid i \notin p_m\}) \cup \overline{\mathcal{D}}\big) \otimes \mathcal{A}$. First, if either the first bit of h or g is 1, then h respectively g computes an element of $\mathcal{A}$. So, we may assume this is not the case. Then there are $i_1 \neq j$ and $i_2 \neq k$ such that $g \geq_T f_{i_1}$ and $h \geq_T f_{i_2}$. If $i_1 \neq i_2$ then

$h \oplus g \in \overline{\mathcal{D}}$, and if $i_1 = i_2$ then $h \oplus g \in C(\{f_i \mid i \notin \{k, j\}\})$. So, we have

$$h \oplus g \in C(\{f_i \mid i \notin \{k, j\}\}) \cup \overline{\mathcal{D}}.$$

There are now two cases: if t_k and t_j are incomparable then $k^\frown j^\frown(h \oplus g) \in \mathcal{A}$. Otherwise, compute $m \in \{k, j\}$ such that $t_m = \max(t_k, t_j)$. Then, because $t_k \in Y$ and $t_j \in U$, we know that $t_m \in V$ and that $h \oplus g \in C(\{f_i \mid i \neq m\}) \cup \overline{\mathcal{D}}$, which is exactly what we needed. Since this is all uniform we therefore see

$$\alpha(Y) \geq_{\mathcal{M}} [\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n \rangle}.$$

Conversely, take any element

$$(e^\frown g) \oplus h \in (\alpha(U) \to_{\mathcal{M}} \alpha(V)) \oplus \mathcal{B}_{(a_1, \dots, a_n)} = \alpha(U) \to_{[\mathcal{B}_{\langle a_1, \dots, a_n \rangle}, \mathcal{A}]} \alpha(V).$$

We need to compute an element of $\alpha(Y)$. Let Z be the collection of nodes where $a_1, \dots, a_n$ are all in the domain. Then h computes some element $\tilde{h} \in \alpha(Z)$, as follows from the definition of $\mathcal{B}_{(a_1, \dots, a_n)}$ and the fact that we have already proven the claim for conjunctions applied to $[\![x_1 = x_1 \wedge \dots \wedge x_n = x_n]\!]_{\langle a_1, \dots, a_n \rangle}$. If the second bit of $\tilde{h}$ is 1, then $\tilde{h}$ computes an element of $\mathcal{A}$ and therefore also computes an element of $\alpha(Y)$. So, we may assume it is 0. Let $k = \tilde{h}(0)$. First compute if $\varphi(a_1, \dots, a_n)$ holds in $\mathfrak{K}$ at the node t_k; if so, we know that $\tilde{h} \in \alpha(Y)$ so we are done. Otherwise, there must be a node $t_{\tilde{k}}$ (above t_k) such that $t_{\tilde{k}} \in U$ but $t_{\tilde{k}} \notin V$.

Let σ be the least string such that $\Phi(e)\left(g \oplus \left(\tilde{k}^\frown 0^\frown \sigma\right)\right)(0)\downarrow$ and such that $\Phi(e)\left(g \oplus \left(\tilde{k}^\frown 0^\frown \sigma\right)\right)(1)\downarrow$ and let $m = \Phi_e\left(g \oplus \left(\tilde{k}^\frown 0^\frown \sigma\right)\right)(0)$ (such a σ much exist, since there is some initial segment of $\tilde{k}^\frown 0^\frown f_{\tilde{k}+1} \in \alpha(U)$ for which this must halt by choice of g and e). Then we see, by choice of g and e that $t_m \in V$ and that

$$\begin{aligned} \{g\} \oplus C\left(\left\{f_i \mid i \neq \tilde{k}\right\}\right) &\geq_{\mathcal{M}} \{g\} \oplus \left(\sigma^\frown C\left(\left\{f_i \mid i \neq \tilde{k}\right\}\right)\right) \\ &\geq_{\mathcal{M}} \left(C(\{f_i \mid i \neq m\}) \cup \overline{\mathcal{D}}\right) \otimes \mathcal{A}. \end{aligned}$$

In fact, since the value at 1 has also already been decided by choice of σ, we even get that either

$$\{g\} \oplus C\left(\left\{f_i \mid i \neq \tilde{k}\right\}\right) \geq_{\mathcal{M}} \mathcal{A}$$

or

$$\{g\} \oplus C\left(\left\{f_i \mid i \neq \tilde{k}\right\}\right) \geq_{\mathcal{M}} C(\{f_i \mid i \neq m\}) \cup \overline{\mathcal{D}}.$$

In the first case, we are clearly done. Otherwise, we claim: $g \oplus \tilde{h} \in C(\{f_i \mid i \neq m) \cup \overline{\mathcal{D}}$. We distinguish several cases:

- If $\tilde{h} \in \overline{\mathcal{D}}$, then $g \oplus \tilde{h} \geq_T \tilde{h} \in \overline{\mathcal{D}}$ and $\overline{\mathcal{D}}$ is upwards closed.
- Otherwise, $\tilde{h} \geq_T f_i$ for some $i \neq k$. If $i \neq \tilde{k}$, then we have just seen that $g \oplus \tilde{h}$ computes an element of $C(\{f_i \mid i \neq m\}) \cup \overline{\mathcal{D}}$. Since the latter is upwards closed, we see that $g \oplus \tilde{h} \in C(\{f_i \mid i \neq m\}) \cup \overline{\mathcal{D}}$.
- If $\tilde{h} \geq_T f_{\tilde{k}}$, then $g \oplus \tilde{h} \geq_T \tilde{h} \in C(\{f_i \mid i \neq m)$: after all, $t_m \in V$ while $t_{\tilde{k}} \notin V$, so $\tilde{k} \neq m$.

Thus, $g \oplus \tilde{h}$ uniformly computes an element of $\alpha(Y)$, which is what we needed to show.

Now, let us consider the quantifiers. So, let $\varphi(x_1, \dots, x_n) = \forall y \psi(x_1, \dots, x_n, y)$. For every $b \in \omega$, let U_b be the set of nodes where $\psi(a_1, \dots, a_n, b)$ holds in $\mathfrak{K}$, and likewise let Y be the set of nodes where $\varphi(a_1, \dots, a_n)$ holds. We need to show that

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n \rangle} \equiv_{\mathcal{M}} \alpha(Y).$$

By definition of the interpretation of the universal quantifier and the induction hypothesis, we know that

$$\begin{aligned} [\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1, \dots, a_n \rangle} &\equiv_{\mathcal{M}} \left(\bigoplus_{b \in \omega} \mathcal{B}_{(a_1, \dots, a_n, b)} \to_{\mathcal{M}} \alpha(U_b) \right) \oplus \mathcal{B}_{a_1} \oplus \dots \oplus \mathcal{B}_{a_n} \\ &= \bigoplus_{b \in \omega} \mathcal{B}_{(a_1, \dots, a_n, b)} \to_{[\mathcal{B}_{(a_1, \dots, a_n)}, \mathcal{A}]_{\mathcal{M}}} \alpha(U_b). \end{aligned}$$

Let Z_b be the set of nodes where $a_1, \dots, a_n$ and b are in the domain, and let Z be the set of nodes where $a_1, \dots, a_n$ are in the domain. Then we get in the same way as above:

$$\equiv_{\mathcal{M}} \bigoplus_{b \in \omega} \alpha(Z_b) \to_{[\mathcal{B}_{(a_1, \dots, a_n)}, \mathcal{A}]_{\mathcal{M}}} \alpha(U_b).$$

Finally, let us introduce new predicates $R_b(x_1, \dots, x_n)$, which are defined to hold in $\mathfrak{K}$ if $\varphi(x_1, \dots, x_n, b)$ holds in $\mathfrak{K}$, and let us introduce new nullary predicates S_b which are defined to hold when all of $a_1, \dots, a_n$ and b are in the domain. Then, applying the fact that we have already proven the claim for implications to $[\![S_b \to R_b]\!]_{\langle a_1, \dots, a_n \rangle}$, we get

$$\equiv_{\mathcal{M}} \bigoplus_{b \in \omega} \alpha((Z_b \to U_b) \cap Z).$$

We now claim that this is equivalent to $\alpha(Y)$. We have $Y \subseteq (Z_b \to U_b) \cap Z$ by the definition of truth in Kripke frames, which suffices to prove that

$$\bigoplus_{b \in \omega} \alpha((Z_b \to U_b) \cap Z) \leq_{\mathcal{M}} \alpha(Y).$$

Conversely, let

$$\bigoplus_{b \in \omega} g_b \in \bigoplus_{b \in \omega} \alpha((Z_b \to U_b) \cap Z).$$

We show how to compute an element of $\alpha(Y)$ from this. If the second bit of g_0 is 1, then h computes an element of $\mathcal{A}$; thus, assume it is 0. Let $m_0 = g_0(0)$. First compute if $\varphi(a_1, \dots, a_n)$ holds in $\mathfrak{K}$ at the node $\gamma(t_{m_0})$; if so, we know that $g_0 \in \alpha(Y)$ so we are done. Therefore, we may assume this is not the case. So, we can compute a $b_1 \in \omega$ such that $t_{m_0} \notin Z_{b_1} \to U_{b_1}$ by the definition of truth in Kripke frames. Now consider g_{b_1}. If the second bit of g_{b_1} is 1, then g_{b_1} computes an element of $\mathcal{A}$ so we are done. Otherwise, let $m_1 = g_{b_1}(0)$. Then $t_{m_1} \in Z_{b_1} \to U_{b_1}$ and $g_{b_1} \in C(f_i \mid i \neq m_1) \cup \overline{\mathcal{D}}$. Then $m_1 \not\leq m_0$ because $t_{m_0} \notin Z_{b_1} \to U_{b_1}$. If m_1 is incomparable with m_0, then $m_0{}^\frown m_1{}^\frown(g_{b_1} \oplus h) \in \mathcal{A}$ so we are done. Thus, the only remaining case is when $m_1 > m_0$.

Iterating this argument, if it does not terminate after finitely many steps, we obtain a sequence $m_0 < m_1 < m_2 < \dots$. However, we assumed that our Kripke frame does not contain any infinite ascending chains, so the algorithm has to terminate after finitely many steps. Thus,

$$\bigoplus_{b\in\omega} \alpha((Z_b \to U_b) \cap Z) \geq_{\mathcal{M}} \alpha(Y).$$

We note that this is the only place in the proof where we use the assumption about infinite ascending chains.

Finally, we consider the existential quantifier. To this end, let $\varphi(x_1, \dots, x_n) = \exists y \psi(x_1, \dots, x_n, y)$. Let U_b and Z be as for the universal quantifier. Then the induction hypothesis tells us that

$$[\![\varphi(x_1, \dots, x_n)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \bigcup \{b^\frown \alpha(Y_b) \mid b \in \omega\}.$$

First, since $Y_b \subseteq Z$, we certainly have that $\alpha(Z) \leq_{\mathcal{M}} \bigcup\{b^\frown\alpha(Y_b) \mid b \in \omega\}$. Conversely, let $j^\frown f \in \alpha(Z)$. Then $f \in \big(C(f_i \mid i \neq j) \cup \overline{\mathcal{D}}\big) \otimes \mathcal{A}$ and $t_j \in Z$. Thus, there is some $b \in \omega$ such that $\psi(a_1, \dots, a_n, b)$ holds, and therefore by induction hypothesis $j^\frown f \in \alpha(Y_b)$. Furthermore, since $\mathfrak{K}$ is decidable we can compute such a b. Thus, $\alpha(Z) \geq_{\mathcal{M}} \bigcup\{b^\frown\alpha(Y_b) \mid b \in \omega\}$, which completes the proof of the claim.

Thus, by the claim we have that, for any sentence φ, that $[\![\varphi]\!] = \alpha(Y)$, where Y is the set of nodes where φ holds in the Kripke model $\mathfrak{K}$. Furthermore, α is injective so $\alpha(Y) = \mathcal{B}_{-1}$ if and only if $Y = T$. So, φ holds in $\mathfrak{M}$ if and only if $Y = T$ if and only if φ holds in $\mathfrak{K}$, which is what we needed to show.

For the second part of the theorem, note that we only used the assumption about infinite ascending chains in the part of the proof dealing with the universal quantifier. □

Lemma 5.6.5. *Let $\mathcal{C} \subseteq \omega^\omega$ be non-empty and upwards closed under Turing reducibility, let $\mathcal{E}_i \subseteq \omega^\omega$ and let $\bigcup\{i^\frown\mathcal{E}_i\} \leq_{\mathcal{M}} \mathcal{C}$. Then there is an $i \in \omega$ such that $\mathcal{E}_i \leq_{\mathcal{M}} \mathcal{C}$.*

Proof. Let $\Phi_e(\mathcal{C}) \subseteq \bigcup\{i^\frown\mathcal{E}_i\}$. Let σ be the least string such that $\Phi_e(\sigma)(0)\downarrow$. Such a string must exist, because $\mathcal{C}$ is non-empty. Let $i = \Phi_e(\sigma)(0)$. Then:

$$\mathcal{C} \geq_{\mathcal{M}} \sigma^\frown\mathcal{C} \geq_{\mathcal{M}} \mathcal{E}_i,$$

as desired. □

Our proof relativises if our language does not contain function symbols, which gives us the following result.

Theorem 5.6.6. *Let $\mathfrak{K}$ be a Kripke model for a language without function symbols which is based on a Kripke frame without infinite ascending chains. Then there is an interval $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ and a structure $\mathfrak{M}$ in $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ such that the theory of $\mathfrak{M}$ is exactly the theory of $\mathfrak{K}$.*

Furthermore, if we allow infinite ascending chains, then this still holds for the fragment of the theories without universal quantifiers.

PROOF. Let h be such that $\mathfrak{K}$ is h-decidable. We relativise the construction in the proof of Theorem 5.6.4 to h. We let all definitions be as in that proof, except where mentioned otherwise. This time we let f_i be an antichain over h, i.e. for all $i \neq j$ we have $f_i \oplus h \not\geq_T f_j$. We change the definition of $\mathcal{D}$ into $\{g \mid \exists i(g \leq_T f_i \oplus h)\}$ We let

$$\mathcal{A} = \{\big(k_1{}^\frown k_2{}^\frown \big(C(\{f_i \mid i \notin \{k_1,k_2\}\}) \cup \overline{\mathcal{D}}\big)\big) \oplus h \mid t_{k_1} \text{ and } t_{k_2} \text{ are incomparable}\}$$

if T is not a chain, and let $\mathcal{A} = \overline{\mathcal{D}} \oplus h$ otherwise. We let $\beta(Y) = \alpha(Y) \oplus \{h\}$ for all $Y \in \mathcal{U}$. Then β is still injective. Indeed, let us assume $\beta(Y) \leq_{\mathcal{M}} \beta(Z)$; we will show that $Y \supseteq Z$. By applying Lemma 5.6.7 below we see that for every j with $t_j \in Z$ there exists a k with $t_k \in Y$ such that either $\big(C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}}\big) \oplus \{h\} \leq_{\mathcal{M}} \big(C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}\big) \oplus \{h\}$ or $\mathcal{A} \leq_{\mathcal{M}} \big(C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}\big) \oplus \{h\}$. If the first holds, let us assume that $k \neq j$; we will derive a contradiction from this. Then $f_k \in C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$ and therefore $f_k \oplus h \in C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}}$ since this set is upwards closed. However, we know that the f_i form an antichain over h in the Turing degrees, which is a contradiction. So, $k = j$ and therefore $t_j \in Y$.

In the second case, we have that

$$C(\{f_i \mid i \notin \{k_1,k_2\}\}) \cup \overline{\mathcal{D}} \leq_{\mathcal{M}} \big(C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}\big) \oplus \{h\}$$

for some $k_1, k_2 \in \omega$ for which t_{k_1} and t_{k_2} are incomparable. Without loss of generality, we may assume that $k_1 \neq j$. Then, in the same way as above, we see that $f_{k_1} \oplus h \in C(\{f_i \mid i \notin \{k_1,k_2\}\}) \cup \overline{\mathcal{D}}$ which is again a contradiction.

We let $\mathcal{B}_{-1} = \beta(T)$ and we let $\mathcal{B}_i = \beta(Z_i)$, where Z_i is the set of nodes where i is in the domain of $\mathfrak{K}$. We claim: for every formula $\varphi(x_1,\dots,x_n)$ and every sequence $a_1,\dots,a_n$,

$$[\![\varphi(x_1,\dots,x_n)]\!]_{\langle a_1,\dots,a_n\rangle} \equiv_{\mathcal{M}} \beta(Y),$$

where Y is exactly the set of nodes where $a_1,\dots,a_n$ are all in the domain and $\varphi(a_1,\dots,a_n)$ holds in the Kripke model $\mathfrak{K}$. The proof is the same as before, except that this time we use that all mass problems we deal with are above $\mathcal{B}_{-1} = \alpha(T) \oplus \{h\}$ and hence uniformly compute h. Thus, we can still decide all the properties about $\mathfrak{K}$ which we need during the proof. □

LEMMA 5.6.7. *Let $\mathcal{C} \subseteq \omega^\omega$ be non-empty and upwards closed under Turing reducibility, let $\mathcal{E}_i \subseteq \omega^\omega$, let $h \in \omega^\omega$ and let $\bigcup\{i^\frown \mathcal{E}_i\} \leq_{\mathcal{M}} \mathcal{C} \oplus \{h\}$. Then there is an $i \in \omega$ such that $\mathcal{E}_i \leq_{\mathcal{M}} \mathcal{C}$.*

PROOF. Let $\Phi_e(\mathcal{C}) \subseteq \bigcup\{i^\frown \mathcal{E}_i\}$. Let σ be the least string such that $\Phi_e(\sigma \oplus h)(0)\downarrow$. Such a string must exist, because $\mathcal{C}$ is non-empty. Let $i = \Phi_e(\sigma \oplus h)(0)$. Then:

$$\mathcal{C} \oplus h \geq_{\mathcal{M}} (\sigma^\frown \mathcal{C}) \oplus h \geq_{\mathcal{M}} \mathcal{E}_i,$$

as desired. □

We will now use Theorem 5.6.6 to show that we can refute the formulas discussed in section 5.4.

PROPOSITION 5.6.8. *There is an interval* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *and a structure* $\mathfrak{M}$ *in* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *such that* $\mathfrak{M}$ *refutes the formula*

$$\forall x, y, z(x = y \vee x = z \vee y = z) \wedge \forall z(S(z) \vee R) \to \forall z(S(z)) \vee R$$

from Proposition 5.4.2.

PROOF. As shown in the proof of Proposition 5.4.2 there is a finite Kripke frame refuting the formula. Now apply Theorem 5.6.6. $\square$

PROPOSITION 5.6.9. *There is an interval* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *and a structure* $\mathfrak{M}$ *in* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *such that* $\mathfrak{M}$ *refutes the formula*

$$(\forall x(S(x) \vee \neg S(x)) \wedge \neg\forall x(\neg S(x))) \to \exists x(\neg\neg S(x)).$$

from Proposition 5.4.3.

PROOF. In the proof of Proposition 5.4.3 we showed that there is a finite Kripke frame refuting the given formula. So, the claim follows from Theorem 5.6.6. $\square$

Thus, moving to the more general intervals $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ did allow us to refute more formulas. Let us next note that Theorem 5.5.5 really depends on the fact that we chose the language of arithmetic to contain function symbols.

PROPOSITION 5.6.10. *Let* Σ *be the language of arithmetic, but formulated with relations instead of with function symbols. Let* T *be derivable in* PA *and let* χ *be a* Π^0_1*-sentence or* Σ^0_1*-sentence which is not derivable in* PA. *Then there is an interval* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *and a structure* $\mathfrak{M}$ *in* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *refuting* $\bigwedge T \to \chi$.

PROOF. Let $\mathfrak{K}$ be a classical model refuting $\bigwedge T \to \chi$, which can be seen as a Kripke model on a frame consisting of one point. Now apply Theorem 5.6.6. $\square$

Finally, let us consider the schema $\forall x\neg\neg\varphi(x) \to \neg\neg\forall x\varphi(x)$, called *Double Negation Shift (DNS)*. It is known that this schema characterises exactly the Kripke frames for which every node is below a maximal node (see Gabbay [33]), so in particular it holds in every Kripke frame without infinite chains. We will show that we can refute it in an interval of the hyperdoctrine of mass problems, even though Theorem 5.6.6 does not apply.

PROPOSITION 5.6.11. *Let* Σ *be the language containing one unary relation* R. *There is an interval* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *and a structure* $\mathfrak{M}$ *in* $[(\mathcal{B}_i)_{i\geq -1},\mathcal{A}]_{\mathcal{P}_\mathcal{M}}$ *such that* $\mathfrak{M}$ *refutes* $\forall x\neg\neg R(x) \to \neg\neg\forall x R(x)$.

PROOF. We let $\mathfrak{K}$ be the Kripke model based on the Kripke frame $(\omega, <)$, where n is in the domain at m if and only if $m \geq n$, and $R(n)$ holds at m if and only if $m > n$. Let everything be as in the proof of Theorem 5.6.4, except we change the definition of $\mathcal{A}$ into:

$$\bigcup \left\{ \left(C\left(\{f_i \mid i \notin X\}\right) \cup \overline{\mathcal{D}}\right) \oplus X \mid X \in 2^\omega \text{ is infinite} \right\},$$

where by X being infinite we mean that the subset $X \subseteq \omega$ represented by X is infinite. We claim: α is still injective under this modified definition of $\mathcal{A}$. Indeed, assume that

$$\mathcal{A} \leq_\mathcal{M} C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}},$$

say through Φ_e; we need to show that this still yields a contradiction. Let σ be the least string such that the right half of $\Phi_e(\sigma)$ has a 1 at a position different from j, say at position k; such a σ must exist since $\Phi_e(f_{j+1}) \in \mathcal{A}$. Then $\Phi_e(\sigma^\frown f_k) \in C(\{f_i \mid i \neq k\}) \cup \overline{\mathcal{D}}$, which is a contradiction.

All the other parts of the proof of Theorem 5.6.4 now go through as long as we look at formulas not containing existential quantifiers. Since $\forall x \neg\neg R(x)$ is intuitionistically equivalent to $\neg\exists x \neg R(x)$, we therefore see that

$$[\![\forall x \neg\neg R(x)]\!] \equiv_{\mathcal{M}} \mathcal{B}_{-1}.$$

We claim: $[\![\neg\forall x(R(x))]\!] \equiv_{\mathcal{M}} \mathcal{B}_{-1}$, which is enough to prove the proposition. Note that $[\![\forall x(R(x))]\!] \equiv_{\mathcal{M}} \mathcal{B}_{-1} \oplus \bigoplus_{m\in\omega}(\mathcal{B}_m \rightarrow_{\mathcal{M}} \mathcal{B}_{m+1})$. By introducing new predicates S_m which hold if and only if m is in the domain and looking at $[\![S_m \to S_{m+1}]\!]$, we therefore get that $[\![\forall x(R(x))]\!] \equiv_{\mathcal{M}} \bigoplus_{m\in\omega} \mathcal{B}_{m+1}$.

We claim that from every element $g \in \bigoplus_{m\in\omega} \mathcal{B}_{m+1}$ we can uniformly compute an element of $\mathcal{A}$. In fact, we show how to uniformly compute from g a sequence $k_0 < k_1 < \dots$ such that $g \in C(\{f_i \mid i \neq k_j\}) \cup \overline{\mathcal{D}}$ for every $j \in \omega$; then if we let $X = \{k_j \mid j \in \omega\}$ we have $g \oplus X \in \big(C(\{f_i \mid i \notin X) \cup \overline{\mathcal{D}}\big) \oplus X \subseteq \mathcal{A}$. For ease of notation let $k_{-1} = 0$. We show how to compute k_{i+1} if k_i is given. There are two possibilities:

- The second bit of $g^{[k_i]}$ is 0: take k_{i+1} to be the first bit of $g^{[k_i]}$; then $k_{i+1} > k_i$ by the definition of $\mathcal{B}_{k_i+1}$.
- The second bit of $g^{[k_i]}$ is 1: then $g^{[k_i]}$ computes an element of $\mathcal{A}$ and therefore computes infinitely many j such that $g^{[k_i]} \in C(\{f_i \mid i \neq j\}) \cup \overline{\mathcal{D}}$, so take k_{i+1} to be such a j which is greater than k_i. □

We do not know how to combine the proof of the last Proposition with the proofs of Theorems 5.6.4 and 5.6.6, because it makes essential use of the fact that the formula is refuted in a model on a frame which is a chain, and of the fact that the subformulas containing universal quantifiers hold either everywhere or nowhere in this model. So, we solved part of the following question, but the definitive answer is still open.

QUESTION 5.6.12. For which theories T is there an interval $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ and a structure $\mathfrak{M}$ in $[(\mathcal{B}_i)_{i\geq -1}, \mathcal{A}]_{\mathcal{P}_{\mathcal{M}}}$ such that the theory of $\mathfrak{M}$ is exactly T?

Part II

Algorithmic Randomness and Genericity

CHAPTER 6

Effective Genericity and Differentiability

This chapter is based on Kuyper and Terwijn [69].

6.1. Introduction

The notion of 1-genericity is an effective notion of genericity from computability theory that has been studied extensively, see e.g. Jockusch [14], or the textbooks Odifreddi [88] and Downey and Hirschfeldt [25]. 1-Genericity, or Σ^0_1-genericity in full, can be defined using computably enumerable (c.e.) sets of strings as forcing conditions. This notion captures a certain type of effective finite extension constructions that is common in computability theory. In this chapter we give a characterisation of 1-genericity in terms of familiar notions from computable analysis. This complements recent results by Brattka, Miller, and Nies [11] that characterise various notions of algorithmic randomness in terms of computable analysis. For example, in [11] it was proven (building on earlier work by Demuth [22]) that an element $x \in [0,1]$ is Martin-Löf random if and only if every computable function of bounded variation is differentiable at x. Note that the notion of Martin-Löf randomness, which one could also call Σ^0_1-randomness, is the measure-theoretic counterpart of the topological notion of 1-genericity.

The main result of this paper is as follows.

THEOREM 6.1.1. *A real $x \in [0,1]$ is 1-generic if and only if every differentiable computable function $f : [0,1] \to \mathbb{R}$ has continuous derivative at x.*

The two implications of this theorem will be proven in Theorems 6.4.3 and 6.5.2. Note that by "differentiable computable function" we mean a computable function that is classically differentiable, so that in particular the derivative need not be continuous.

Our result can be seen an effectivisation of a result by Bruckner and Leonard.

THEOREM 6.1.2. (Bruckner and Leonard [13, p. 27]) *A set $A \subseteq \mathbb{R}$ is the set of discontinuities of a derivative if and only if A is a meager $\mathbf{\Sigma}^0_2$ set.*

One might expect that, in analogy to Theorem 6.1.1, n times differentiable computable functions would characterise n-genericity. However, in section 6.7 we show that 1-genericity is also equivalent to the nth derivative of any n times differentiable computable function being continuous at x. In section 6.8 we consider differentiable polynomial time computable functions and show that again these characterise 1-genericity.

First, let us recall some relevant definitions. For a string $\sigma \in 2^{<\omega}$, we have

$$[\sigma] = \{x \in 2^\omega : \sigma \subseteq x\}.$$

The product topology on 2^ω, sometimes called the tree topology, or the finite information topology, has all sets of the form $[\sigma]$ as basic open sets. For a set $A \subseteq 2^{<\omega}$, we let

$$[A] = \bigcup_{\sigma \in A} [\sigma].$$

Thus every set A of finite strings defines an open subset of 2^ω. A subset of 2^ω is a Σ^0_1 *class*, or *effectively open*, if it is of the form $[A]$, with $A \subseteq 2^{<\omega}$ computably enumerable (c.e.). A set is a Π^0_1 *class*, or *effectively closed*, if it is the complement of a Σ^0_1 class. Thus Σ^0_1 and Π^0_1 classes form the first levels of the effective Borel hierarchy. As usual, the levels of the classical Borel hierarchy are denoted by boldface symbols $\boldsymbol{\Sigma}^0_n$ and $\boldsymbol{\Pi}^0_n$. These notions are defined in the same way for $[0,1]$, using rational intervals as basic opens. We denote the interior of a set $V \subseteq 2^\omega$ by $\mathrm{Int}(V)$.

6.2. 1-Genericity

First, let us recall what it means for an element $x \in 2^\omega$ to be 1-generic. We will then discuss 1-genericity for elements of $[0,1]$. A discussion of the properties of arithmetically generic and 1-generic sets can be found in Jockusch [14]. The "forcing-free" formulation of genericity we use here is due to Posner, see [14, p115].

Given a sequence $x \in 2^\omega$ and a set $A \subseteq 2^{<\omega}$, we say that x *meets* A if there exists $\sigma \sqsubseteq x$ such that $\sigma \in A$; equivalently, if $x \in [A]$. The set A is *dense along* x if for every $\sigma \sqsubseteq x$ there is an extension $\tau \sqsupseteq \sigma$ such that $[\tau] \subseteq [A]$; equivalently, if x is in the closure of the open set $[A]$. Thus, rephrasing Definition 3.2.1, an element $x \in 2^\omega$ is 1-generic if x meets every c.e. set $A \subseteq 2^{<\omega}$ that is dense along x.

We now reformulate the definition of 1-genericity into a form that will be convenient in what follows. This formulation is also better suited for the discussion of generic real numbers (as opposed to infinite strings).

LEMMA 6.2.1. *Let $A \subseteq 2^{<\omega}$ and let $V = 2^\omega \setminus [A]$. Then A is dense along x if and only if x is not in the interior of V. Therefore, A is dense along x and x does not meet A if and only if $x \in V \setminus \mathrm{Int}(V)$.*

PROOF. Note that A is dense along x if and only if every open set containing x has non-empty intersection with $[A]$. Thus, A is dense along x if and only if every open set disjoint from $[A]$ does not contain x. However, the open sets disjoint from $[A]$ are exactly the open sets contained in V, of which $\mathrm{Int}(V)$ is the largest. Thus A is dense along x if and only if $\mathrm{Int}(V)$ does not contain x. □

COROLLARY 6.2.2. *For any $x \in 2^\omega$ we have that x is 1-generic if and only if for every Π^0_1 class $V \subseteq 2^\omega$ we have $x \notin V \setminus \mathrm{Int}(V)$.*

One of the reasons this is interesting to mention explicitly is because a typical example of a nowhere dense set is a closed set with its interior removed, and the Π^0_1 sets are the simplest type of closed sets. Thus, the Corollary 6.2.2 says that x is 1-generic if it is not in any of the simple, typical nowhere dense sets. This way of looking at 1-generic sets complements the usual motivation of 1-genericity by forcing, and it also allows one to easily compare 1-genericity with weak 1-genericity (since x is weakly 1-generic if it is not in any Π^0_1-class with empty interior, see [25]).

With this equivalence in mind, we can now also define what it means for an element of $[0,1]$ to be 1-generic.

DEFINITION 6.2.3. Let $x \in [0,1]$. We say that x is 1-*generic* if for every Π^0_1 class $V \subseteq [0,1]$ we have $x \notin V \setminus \mathrm{Int}(V)$.

There is a natural 'almost-homeomorphism' between 2^ω and $[0,1]$: given an infinite sequence $x \in 2^\omega$ we have $0.x \in [0,1]$ (interpreting the sequence as a decimal expansion in binary), and conversely given $y \in [0,1]$ we can take the binary expansion of y containing infinitely many 0s, which gives us an element of 2^ω. Note that the problem of non-unique expansions only occurs for rationals, which are not 1-generic anyway. It is thus natural to ask if the notions of 1-genericity in these two spaces correspond via this mapping. The next proposition says this is indeed the case.

PROPOSITION 6.2.4. *For any irrational $x \in [0,1]$ we have that x is 1-generic if and only if its (unique) binary expansion is* 1*-generic in* 2^ω*.*

PROOF. Let 2^ω_- be the set of infinite binary sequences which contain infinitely many 0s and infinitely many 1s. Then the 'almost-homeomorphism' given above in fact restricts to a homeomorphism to 2^ω_- and $[0,1]_-$, where $[0,1]_-$ is $[0,1]$ without the dyadic rationals. Therefore, 1-genericity on 2^ω_- and $[0,1]_-$ (which are defined as in Definition 6.2.3) coincide.

Note that 2^ω_- is dense in 2^ω and that $[0,1]_-$ is dense in $[0,1]$. Thus, it is enough if we can show that if $Y \subseteq X$ is such that Y is dense in X, then 1-genericity on X and Y coincide for elements $y \in Y$. Given a Π^0_1 class $V \subseteq X$, let $W = V \cap Y$. Then W is a Π^0_1 class in Y. Conversely, every Π^0_1 class $W \subseteq Y$ is of the form $W = V \cap Y$ by definition of the induced topology.

We claim that $\mathrm{Int}_X(V) \cap Y = \mathrm{Int}_Y(W)$. Clearly, $\mathrm{Int}_X(V) \cap Y \subseteq \mathrm{Int}_Y(W)$. Conversely, if we let $\mathrm{Int}_Y(W) = U \cap Y$ for some open $U \subseteq X$, then $U \subseteq \mathrm{Int}_X(V \cup (X \setminus Y))$. Towards a contradiction, assume that $U \cap (X \setminus V) \neq \emptyset$, then this is a non-empty open set. However, we also have $U \cap (X \setminus V) \subseteq X \setminus Y$, which contradicts the fact that Y is dense in X. Thus, we see that $U \subseteq V$, and therefore $U \subseteq \mathrm{Int}_X(V)$. So, $\mathrm{Int}_X(V) \cap Y = \mathrm{Int}_Y(W)$.

So, we see that $y \notin V \setminus \mathrm{Int}_X(V)$ if and only if $y \notin W \setminus \mathrm{Int}_Y(W)$. This completes the proof. $\square$

6.3. Effective Baire class 1 functions

In this section we will discuss what it means for a function to be of effective Baire class 1, and discuss some of the basic properties of these functions. First, let us recall what it means for a function on the reals to be computable. Our definitions follow Moschovakis [83].

DEFINITION 6.3.1. Let $f : [0,1] \to \mathbb{R}$. We say that f is *computable* if for every basic open set U we have that $f^{-1}(U)$ is Σ^0_1 uniformly in U, i.e. if there exists a computable function $\alpha : \mathbb{Q} \times \mathbb{Q} \to \omega$ such that for all $q, r \in \mathbb{Q}$ we have that $f^{-1}((q,r))$ is equal to the Σ^0_1 class given by the index $\alpha(q,r)$.

Definition 6.3.1 is equivalent to the formulation with computable functionals, see e.g. the discussion in Pour-El and Richards [96].

Functions of effective Baire class 1 are obtained by weakening the above definition as follows.

Definition 6.3.2. A function $f : [0,1] \to \mathbb{R}$ is of *effective Baire class 1* if for every basic open set U we have that $f^{-1}(U)$ is Σ^0_2 uniformly in U.

Replacing Σ^0_2 by $\boldsymbol{\Sigma}^0_2$ in the above definition, we obtain what is known as a function of *(non-effective) Baire class 1*. Before we give an important example of an effective Baire class 1 function, let us first consider the following proposition, which gives an equivalent condition for a function to be of effective Baire class 1. This proposition mirrors the classical proposition by Lebesgue, Hausdorff and Banach which says that a function is of Baire class 1 if and only if it is a pointwise limit of continuous functions, see e.g. Kechris [50, p. 192]. (This does not hold for all Polish spaces; it holds for $f : X \to Y$ if either X is zero-dimensional or $Y = \mathbb{R}$.)

Proposition 6.3.3. *Let $f : [0,1] \to \mathbb{R}$. The following are equivalent:*

(i) *f is of effective Baire class 1,*
(ii) *f is the pointwise limit of a uniform sequence of computable functions, i.e. there exists a sequence $f_0, f_1, \dots$ of functions from $[0,1]$ to $\mathbb{R}$ converging pointwise to f and a computable function $\alpha : \omega \times \mathbb{Q} \times \mathbb{Q} \to \omega$ such that for all $q, r \in \mathbb{Q}$ and all $n \in \omega$ we have that $f_n^{-1}((q,r))$ is equal to the Σ^0_1 class given by the index $\alpha(n,q,r)$.*

Proof. (ii) $\to$ (i): Let $f_0, f_1, \dots$ be a sequence of uniformly computable functions converging to f and let U be any basic open set. Then $U = \bigcup_{i \in \omega, V_i \subseteq U} V_i$, where $V_0, V_1, \dots$ is a computable enumeration of the closed intervals with rational endpoints. We claim:

$$f^{-1}(U) = \bigcup_{V_i \subseteq U} \bigcup_{n \in \omega} \bigcap_{m \geq n} f_m^{-1}(V_i),$$

which is clearly Σ^0_2 uniformly in U.

To prove the claim, let $x \in f^{-1}(U)$. Then $f(x) \in U$, so there exists $V_i \subseteq U$ such that $f(x) \in \mathrm{Int}(V_i)$, say $(f(x) - \varepsilon, f(x) + \varepsilon) \subseteq V_i$. Let $n \in \omega$ be such that for every $m \geq n$ we have that $|f_m(x) - f(x)| < \varepsilon$. Then for every $m \geq n$ we have that $x \in f_m^{-1}(V_i)$, which proves the first inclusion.

Conversely, let $n \in \omega$, $V_i \subseteq U$ and $x \in \bigcap_{m \geq n} f_m^{-1}(V_i)$. Then for every $m \geq n$ we have $f_m(x) \in V_i$, and since V_i is closed we then also have $f(x) = \lim_{m \to \infty} f_m(x) \in V_i \subseteq U$, which completes the proof of the claim.

(i) $\to$ (ii): This follows by effectivising Kechris [50, Theorem 24.10]; this result is also mentioned (without proof) in Moschovakis [83, Exercise 3.E.14]. Since this implication is not used anywhere in this thesis, we will not go into further detail. □

Using this proposition, we can now give an important example of effective Baire class 1 functions: derivatives of computable functions. This also explains our interest in them.

COROLLARY 6.3.4. *Let $f : [0,1] \to \mathbb{R}$ be a differentiable computable function. Then f' is of effective Baire class* 1.

PROOF. Let $f_n(x) = 2^n(f(x+2^{-n}) - f(x))$. To account for the problem that for $x+2^{-n} > 1$ the value $f(x+2^{-n})$ is not defined, we let $f(y) = -f(2-y)+2f(1)$ for $y > 1$ (i.e. we flip and mirror f on $[1,2]$). Then the sequence $f_0, f_1, \dots$ is uniformly computable and converges pointwise to f', so f' is of effective Baire class 1 by Proposition 6.3.3. □

6.4. Continuity of Baire class 1 functions

At the basis of this section lies the following important classical result.

THEOREM 6.4.1. (Baire) *Let $f : [0,1] \to \mathbb{R}$ be of (non-effective) Baire class 1. Then the points of discontinuity of f form a meagre $\mathbf{\Sigma}^0_2$ set.*

PROOF. See Kechris [50, Theorem 24.14] or Oxtoby [90, Theorem 7.3]. □

We will now effectivise this result.

THEOREM 6.4.2. *Let $f : [0,1] \to \mathbb{R}$ be of effective Baire class* 1*. Then f is continuous at every* 1*-generic point.*

PROOF. We effectivise the proof from Kechris [50, Theorem 24.14]. Let $U_0, U_1, \dots$ be an effective enumeration of the basic open sets. Now f is continuous at x if and only if the inverse image of every neighbourhood of $f(x)$ is a neighbourhood of x. Thus, f is discontinuous at x if and only if there exists an open set U containing $f(x)$ such that every open set contained in $f^{-1}(U)$ does not contain x. Hence

$$\{x \in [0,1] \mid f \text{ is discontinous at } x\} = \bigcup_{n\in\omega} f^{-1}(U_n) \setminus \mathrm{Int}(f^{-1}(U_n)).$$

Now, let x be such that f is discontinuous at x and let n be such that $x \in f^{-1}(U_n) \setminus \mathrm{Int}(f^{-1}(U_n))$. Because f is of effective Baire class 1, we know that $f^{-1}(U_n)$ is Σ^0_2. So, let $f^{-1}(U_n) = \bigcup_{i\in\omega} V_i$, where each V_i is Π^0_1. Then it is directly verified that

$$f^{-1}(U_n) \setminus \mathrm{Int}(f^{-1}(U_n)) \subseteq \bigcup_{i\in\omega} (V_i \setminus \mathrm{Int}(V_i)).$$

Let i be such that $x \in V_i \setminus \mathrm{Int}(V_i)$. Then x is not 1-generic by Definition 6.2.3. □

Combining this result with the fact that derivatives of computable functions are of effective Baire class 1, we get the first implication of Theorem 6.1.1 as a consequence.

THEOREM 6.4.3. *If $f : [0,1] \to \mathbb{R}$ is a computable function, then f' is continuous at every 1-generic real.*

PROOF. From Corollary 6.3.4 and Theorem 6.4.2. □

6.5. Functions discontinuous at non-1-generics

In this section we will prove the second implication of Theorem 6.1.1. To this end, we will build, for each Π^0_1 class V, a Volterra-style differentiable computable function whose derivative will fail to be continuous at the points whose non-1-genericity is witnessed by V. We have to be careful in order to make this function computable.

THEOREM 6.5.1. *Let V be a Π^0_1 class. Then there exists a differentiable computable function $f : [0,1] \to \mathbb{R}$ such that f' is discontinuous at every $x \in V \setminus \operatorname{Int}(V)$.*

PROOF. In the construction of f below, we first define auxiliary functions g and h.

Construction. We define an auxiliary function g, with the property that g is differentiable and computable, and g' is continuous on $(0,1)$ and discontinuous at 0 and 1.

Define the function h on $[0,1]$ by $h(0) = 0$ and

$$h(x) = x^2 \sin(\frac{1}{x^2})$$

for $x > 0$. Then h is computable and differentiable, with derivative $h'(0) = 0$ and

$$h'(x) = 2x \sin(\frac{1}{x^2}) - 2\frac{1}{x} \cos(\frac{1}{x^2})$$

when $x > 0$. Note that h' is discontinuous at $x = 0$. Fix a computable $x_0 \in (0, \frac{1}{2}]$ such that $h'(x_0) = 0$. Such an x_0 exists, because h' has isolated roots, and isolated roots of computable functions are computable. Now define g on $[0,1]$ by

$$g(x) = \begin{cases} 0 & \text{if } x = 0 \\ h(x) & \text{if } x \in (0, x_0] \\ h(x_0) & \text{if } x \in [x_0, 1 - x_0] \\ h(1-x) & \text{if } x \in [1 - x_0, 1) \\ 0 & \text{if } x = 1. \end{cases}$$

Then g is a differentiable computable function, with derivative

$$g'(x) = \begin{cases} 0 & \text{if } x = 0 \\ h'(x) & \text{if } x \in (0, x_0] \\ 0 & \text{if } x \in [x_0, 1 - x_0] \\ -h'(1-x) & \text{if } x \in [1 - x_0, 1) \\ 0 & \text{if } x = 1. \end{cases}$$

In particular, we see that g' is continuous exactly on $(0,1)$. We will use g to construct f.

For the given Π^0_1 class V, let $U = [0,1] \setminus V$, and fix computable enumerations $q_0, q_1, \ldots$ and $r_0, r_1, \ldots$ of rational numbers in $[0,1]$ such that $U = \bigcup_{n\in\omega} [q_n, r_n]$

and such that the (q_n, r_n) are pairwise disjoint. We will construct f as a sum of a sequence $f_0, f_1, \dots$ of uniformly computable functions. We define f_n by:

$$f_n(x) = \begin{cases} 0 & \text{if } x \in [0, q_n] \\ \dfrac{r_n - q_n}{2^n}\, g\left(\dfrac{x - q_n}{r_n - q_n}\right) & \text{if } x \in [q_n, r_n] \\ 0 & \text{if } x \in [r_n, 1]. \end{cases} \tag{13}$$

Finally, we let $f = \sum_{n=0}^{\infty} f_n$.

Verification. We first show that f is computable. To this end, first observe that each f_n is supported on (q_n, r_n), and therefore the supports of the different f_n are disjoint. Furthermore, each f_n is bounded by 2^{-n}.

Let (a, b) be a basic open subset of $\mathbb{R}$. We distinguish two cases. First, assume $0 \notin (a, b)$. We assume $a > 0$, the case $b < 0$ is proven in a similar way. Let $n \in \omega$ be such that $2^{-n} < a$. Then, since the supports of the f_m are disjoint, and each f_m is bounded by 2^{-m}, we have

$$f^{-1}((a, b)) = (f_0 + \dots + f_n)^{-1}((a, b)),$$

which is Σ_1^0 because a finite sum of computable functions is computable.

In the second case, we have $0 \in (a, b)$. Let n be such that $|a|, |b| \geq 2^{-n}$. Then, again because the supports of the f_m are disjoint, we see that if x is not in the support of any f_m for $m \leq n$ then certainly $f(x) \in (a, b)$. Therefore we have

$$f^{-1}((a, b)) = (f_0 + \dots + f_n)^{-1}((a, b)) \cup \bigcap_{m \leq n} ([0, 1] \setminus [q_m, r_m]),$$

which is also Σ_1^0. It is clear that the case distinction is uniformly computable, so it follows that f is computable.

Next, we check that f is differentiable. We first note that every f_n is differentiable, because g is differentiable. Let $x \in [0, 1]$. We distinguish two cases. First, if x is in some (q_n, r_n) then it is immediate that f is differentiable at x with derivative $f_n'(x)$, because the intervals (q_n, r_n) are disjoint. Next, we consider the case where x is not in any interval (q_n, r_n). Note that in this case we have $f(x) = 0$. Fix $m \in \omega$. Then we have:

$$\lim_{y \to x} \left| \frac{f(y)}{y - x} \right| \leq \lim_{y \to x} \left| \frac{(f_0 + \dots + f_m)(y)}{y - x} \right| + \lim_{y \to x} \left| \frac{(f_{m+1} + f_{m+2} + \dots)(y)}{y - x} \right|.$$

Because $f_0 + \dots + f_m$ is differentiable at x with derivative 0, this is equal to:

$$\lim_{y \to x} \left| \frac{(f_{m+1} + f_{m+2} + \dots)(y)}{y - x} \right|. \tag{14}$$

To show that this limit is 0, we will prove that it is bounded by $\frac{1}{2^m(1 - x_0)}$ for every m. Let $y \in [0, 1]$ be distinct from x. Let us assume that $x < y$; the other case is proven in the same way. If y is not in any (q_n, r_n) for $n \geq m + 1$ then $(f_{m+1} + f_{m+2} + \dots)(y) = 0$. Otherwise, there is exactly one such n. Then:

$$\left| \frac{(f_{m+1} + f_{m+2} + \dots)(y)}{y - x} \right| = \left| \frac{f_n(y)}{y - x} \right| \leq \left| \frac{f_n(y)}{y - q_n} \right|,$$

where the last inequality follows from the fact that x does not lie in (q_n, r_n). We distinguish three cases. First, if $z = \frac{y-q_n}{r_n-q_n} \in (0, x_0]$, then

$$\left|\frac{f_n(y)}{y-q_n}\right| = \left|\frac{2^{-n}(r_n-q_n)g(z)}{y-q_n}\right| = \left|2^{-n} z \sin(z^{-2})\right| \leq 2^{-n} \leq \frac{1}{2^m(1-x_0)}.$$

Next, if $z \in [x_0, 1-x_0]$ (which is non-empty because $x_0 \leq \frac{1}{2}$), then

$$\left|\frac{f_n(y)}{y-q_n}\right| \leq \frac{2^{-n}(r_n-q_n)x_0^2}{y-q_n} = \frac{2^{-n}x_0^2}{z} \leq x_0 2^{-n} \leq \frac{1}{2^m(1-x_0)}$$

where we use the fact that $z \geq x_0$. Finally, if $z \in [1-x_0, 1]$, then

$$\left|\frac{f_n(y)}{y-q_n}\right| = \left|\frac{2^{-n}(r_n-q_n)h(1-z)}{y-q_n}\right| \leq \frac{1}{2^n z} \leq \frac{1}{2^n(1-x_0)} \leq \frac{1}{2^m(1-x_0)}.$$

Combining this with (14) we see that $\lim_{y\to x}\left|\frac{f(y)}{y-x}\right| \leq \frac{1}{2^m(1-x_0)}$. Since m was arbitrary this shows that f is differentiable at x, with derivative $f'(x) = 0$.

Finally, we need to verify that f' is discontinuous at x for all $x \in V \setminus \mathrm{Int}(V)$. Therefore, let $x \in V \setminus \mathrm{Int}(V)$. Then every open set W containing x has non-empty intersection $W \cap U$ (recall that $U = [0,1] \setminus V$), but this intersection does not contain x. We have shown above that $f'(x) = 0$. We will show that for every open interval I containing x there is a point $y \in I$ such that $f'(y) \leq -1$, which clearly shows that f' cannot be continuous at x. Fix an open interval I containing x. Then $I \cap U \neq \emptyset$, so there is an $n \in \omega$ such that $I \cap [q_n, r_n]$ is non-empty. Note that I contains x and therefore I cannot be a subinterval of $[q_n, r_n]$. Therefore there exists a $q_n < s < r_i$ such that either $[q_n, s) \subseteq I$ or $(s, r_n] \subseteq I$. We will assume the first case; the second case is proven in a similar way.

Note that on $[q_n, s)$ the function f' is equal to f_n'. For $y \in (q_n, s)$ we thus have:

$$f'(y) = 2^{-n} g'((y-q_n)/(r_n-q_n)).$$

So, we need to show that there is a $y \in (q_n, s)$ such that $g'((y-q_n)/(r_n-q_n)) \leq -2^n$, or equivalently, that there is a $z \in (0, (s-q_n)/(r_n-q_n))$ such that $g'(z) \leq -2^n$. Without loss of generality, $(s-q_n)/(r_n-q_n) < x_0$. Let $k \geq n$ be such that $2^{-k} \leq \frac{s-q_n}{r_n-q_n}$. Then:

$$\begin{aligned} g'\left(1/\left(2^k\sqrt{\pi}\right)\right) &= \frac{1}{2^{k-1}\sqrt{\pi}}\sin(2^{2k}\pi) - 2^{k+1}\sqrt{\pi}\cos(2^{2k}\pi) \\ &= -2^{k+1}\sqrt{\pi} \leq -2^k \leq -2^n. \end{aligned}$$

This completes the verification. □

THEOREM 6.5.2. *If $x \in [0,1]$ is such that every differentiable computable function $f : [0,1] \to \mathbb{R}$ has continuous derivative at x, then x is 1-generic.*

PROOF. If x is not 1-generic, then there is a Π^0_1 class V such that $x \in V \setminus \mathrm{Int}(V)$. Applying Theorem 6.5.1 to V gives a differentiable computable function f for which f' is discontinuous at x. □

6.6. n-Genericity

The notion of 1-genericity corresponds to the first level of the arithmetical hierarchy. Higher genericity notions can be defined using forcing conditions from higher levels of the arithmetical hierarchy. As for 1-genericity, an equivalent formulation can be given as follows, see Jockusch [14]:

Definition 6.6.1. An element $x \in 2^\omega$ is *n-generic* if x meets every Σ^0_n set of strings $A \subseteq 2^{<\omega}$ that is dense along x.

As usual, let $\emptyset'$ denote the halting set, and let $\emptyset^{(n)}$ denote the n-th jump. Since a Σ^0_n set of strings is the same as a Σ^0_1 set of strings relative to $\emptyset^{(n-1)}$, a set is n-generic if and only if it is 1-generic relative to $\emptyset^{(n-1)}$.

Corollary 6.2.2 relativises to:

Proposition 6.6.2. *For any $x \in 2^\omega$ we have that x is n-generic if and only if for every $\Pi^{0,\emptyset^{(n-1)}}_1$ class $V \subseteq 2^\omega$ we have $x \notin V \setminus \mathrm{Int}(V)$.*

Note that in general a $\Pi^{0,\emptyset^{(n-1)}}_1$ class in 2^ω is not the same as a Π^0_n class, since the latter need not even be closed. (And even if one *assumes* that the class is closed the notions are not the same, see [25, p76].)

Given this equivalence, we can now generalise Definition 6.2.3 to:

Definition 6.6.3. Let $x \in [0,1]$. We say that x is *n-generic* if for every $\Pi^{0,\emptyset^{(n-1)}}_1$ class $V \subseteq [0,1]$ we have $x \notin V \setminus \mathrm{Int}(V)$.

Further justification for this definition comes from the fact that Proposition 6.2.4 relativises: An irrational $x \in [0,1]$ is n-generic according to Definition 6.6.3 if and only if its binary expansion is n-generic in 2^ω.

It is straightforward to check that the results of all the previous sections relativise to an arbitrary oracle A. This gives the following relativised version of Theorem 6.1.1:

Theorem 6.6.4. *A real $x \in [0,1]$ is 1-generic relative to A if and only if for every differentiable A-computable function $f : [0,1] \to \mathbb{R}$, f' is continuous at x.*

Taking $A = \emptyset^{(n-1)}$, this immediately gives the following characterisation of n-genericity:

Corollary 6.6.5. *A real $x \in [0,1]$ is n-generic if and only if for every differentiable $\emptyset^{(n-1)}$-computable function $f : [0,1] \to \mathbb{R}$, f' is continuous at x.*

Also, taking all n together, we see that a real x is arithmetically generic if and only if every differentiable arithmetical function has continuous derivative at x.

6.7. Multiply differentiable functions

We have characterised 1-genericity using the continuity of the derivatives of (once) differentiable computable functions. One might wonder: what kind of effective genericity for x corresponds to every twice differentiable, computable function having continuous second derivative at x? Or, more generally, what corresponds to every n times differentiable, computable function having continuous

nth derivative at x? It turns out that the answer is always 1-genericity. To show this we will need the following proposition, which essentially tells us that the case for $n > 2$ collapses to the case $n = 2$.

PROPOSITION 6.7.1. *Let $f : [0,1] \to \mathbb{R}$ be computable and twice continuously differentiable. Then f' is computable.*

PROOF. See e.g. Pour-El and Richards [96, Theorem 1.2]. □

If the second derivative of a computable function exists, it is easy to see that it is of effective Baire class 2 (i.e. a pointwise limit of a computable sequence of functions of effective Baire class 1), by similar arguments as in the proof of Corollary 6.3.4 However, using the following proposition we can easily see that the second derivative of a computable function is in fact of effective Baire class 1.

PROPOSITION 6.7.2. *Let $f : [0,1] \to \mathbb{R}$ be twice differentiable. Then*

$$f''(x) = \lim_{h \to 0} \frac{f(x+h) + f(x-h) - 2f(x)}{h^2}.$$

PROOF. See e.g. Rudin [99, p. 115]. □

THEOREM 6.7.3. *Fix $n \geq 1$. Then a real $x \in [0,1]$ is 1-generic if and only if every n times differentiable, computable function $f : [0,1] \to \mathbb{R}$ has continuous nth derivative at x.*

PROOF. For $n = 1$ this is exactly Theorem 6.1.1. So, we may assume $n \geq 2$. First, if $x \in [0,1]$ is not 1-generic, then by Theorem 6.1.1 there is a differentiable, computable function $g : [0,1] \to \mathbb{R}$ such that g' is not continuous at x. Now let $h_1 = g$ and let h_i be a computable antiderivative of h_{i-1} for $2 \leq i \leq n$ (which exists by Ko [57, Theorem 5.29]). Then, if we let $f = h_n$, we see that f is an n times differentiable, computable function such that $f^{(n)}$ is discontinuous at x.

Conversely, if f is an n times differentiable, computable function, then $f^{(n-2)}$ is computable by Proposition 6.7.1. So, $f^{(n)}$ is of effective Baire class 1 by Proposition 6.7.2. Thus, if $f^{(n)}$ is discontinuous at x, then x is not 1-generic by Theorem 6.4.2. □

6.8. Complexity-theoretic considerations

In this section we discuss polynomial time computable real functions. The theory of these functions is developed in Ko [57], to which we refer the reader for the basic results and definitions. Briefly, a function $f : [0,1] \to \mathbb{R}$ is polynomial time computable if for any $x \in [0,1]$ we can compute an approximate value of $f(x)$ to within an error of 2^{-n} in time n^k for some constant k.

Most of the common functions from analysis, such as rational functions and the trigonometric functions, as well as their inverses, are all polynomial time computable, see e.g. Brent [12] and Weihrauch [124]. Also, the polynomial time computable functions are closed under composition. With this knowledge, it is not difficult to see that the construction of the function f in section 6.5 can be modified to yield a polynomial time computable function, rather than just a computable one. For this it is also needed that the complement of the Π^0_1 class V from Theorem 6.5.1

can be represented by a polynomial time computable set of strings. This is similar to the fact that every non-empty computably enumerable set is the range of a polynomial time computable function, simply by sufficiently slowing down the enumeration. Since the enumeration of $U = [0,1] \setminus V$ in the proof of Theorem 6.5.1 is now slower, the definition of f_n in (13) has to be adapted by replacing 2^n by $2^{t(n)}$, where $t(n)$ is the stage at which the interval (q_n, r_n) is enumerated into U. This modification ensures that the functions f_n are uniformly polynomial time computable, so that also the function $f = \sum_{n=0}^{\infty} f_n$, is polynomial time computable. Thus we obtain the following strengthening of Theorem 6.5.1:

THEOREM 6.8.1. *Let V be a Π^0_1 class. Then there exists a differentiable polynomial time computable function $f : [0,1] \to \mathbb{R}$ such that f' is discontinuous at every $x \in V \setminus \mathrm{Int}(V)$.*

We now have the following variant of Theorem 6.1.1:

THEOREM 6.8.2. *A real $x \in [0,1]$ is 1-generic if and only if for every differentiable polynomial time computable function $f : [0,1] \to \mathbb{R}$, f' is continuous at x.*

PROOF. The "only if" direction is immediate from Theorem 6.1.1. For the "if" direction; if x is not 1-generic, then there is a Π^0_1 class V such that $x \in V \setminus \mathrm{Int}(V)$. Theorem 6.8.1 then gives a differentiable polynomial time computable function f for which f' is discontinuous at x. □

CHAPTER 7

Coarse Reducibility and Algorithmic Randomness

This chapter is based on Hirschfeldt, Jockusch, Kuyper and Schupp [38].

7.1. Introduction

There are many natural problems with high worst-case complexity that are nevertheless easy to solve in most instances. The notion of "generic-case complexity" was introduced by Kapovich, Myasnikov, Schupp, and Shpilrain [48] as a notion that is more tractable than average-case complexity but still allows a somewhat nuanced analysis of such problems. That paper also introduced the idea of generic computability, which captures the idea of having a partial algorithm that correctly computes $A(n)$ for "almost all" n, while never giving an incorrect answer. Jockusch and Schupp [15] began the general computability-theoretic investigation of generic computability and also defined the idea of coarse computability, which captures the idea of having a total algorithm that always answers and may make mistakes, but correctly computes $A(n)$ for "almost all" n. We are here concerned with this latter concept. We first need a good notion of "almost all" natural numbers.

DEFINITION 7.1.1. Let $A \subseteq \omega$. The *density of A below n*, denoted by $\rho_n(A)$, is $\frac{|A\upharpoonright n|}{n}$. The *upper density* $\overline{\rho}(A)$ of A is $\limsup_n \rho_n(A)$. The *lower density* $\underline{\rho}(A)$ of A is $\liminf_n \rho_n(A)$. If $\overline{\rho}(A) = \underline{\rho}(A)$ then we call this quantity the *density* of A, and denote it by $\rho(A)$.

We say that D is a *coarse description* of X if $\rho(D \triangle X) = 0$, where $\triangle$ denotes symmetric difference. A set X is *coarsely computable* if it has a computable coarse description.

This idea leads to natural notions of reducibility.

DEFINITION 7.1.2. We say that Y is *uniformly coarsely reducible* to X, and write $Y \leq_{\mathrm{uc}} X$, if there is a Turing functional Φ such that if D is a coarse description of X, then Φ^D is a coarse description of Y. This reducibility induces an equivalence relation $\equiv_{\mathrm{uc}}$ on 2^ω. We call the equivalence class of X under this relation the *uniform coarse degree* of X.

Uniform coarse reducibility, generic reducibility (defined in [15]), and several related reducibilities have been termed *notions of robust information coding* by Dzhafarov and Igusa [27]. Work on such notions has mainly focused on their uniform versions. (One exception is a result on non-uniform ii-reducibility in Hirschfeldt and Jockusch [37].) However, their non-uniform versions also seem to be of interest. In particular, we will work with the following non-uniform version of coarse reducibility.

DEFINITION 7.1.3. We say that Y is *non-uniformly coarsely reducible* to X, and write $Y \leq_{\mathrm{nc}} X$, if every coarse description of X computes a coarse description of Y. This reducibility induces an equivalence relation $\equiv_{\mathrm{nc}}$ on 2^ω. We call the equivalence class of X under this relation the *non-uniform coarse degree* of X.

Note that the coarsely computable sets form the least degree in both the uniform and non-uniform coarse degrees. Uniform coarse reducibility clearly implies non-uniform coarse reducibility. We will show in the next section that, as one might expect, the converse fails. The development of the theory of notions of robust information coding and related concepts have led to interactions with computability theory (as in Jockusch and Schupp [15]; Downey, Jockusch and Schupp [24]; and Downey, Jockusch, McNicholl and Schupp [23]), reverse mathematics (as in Dzhafarov and Igusa [27] and Hirschfeldt and Jockusch [37]), and algorithmic randomness (as in Astor [3]).

In this chapter, we investigate connections between coarse reducibility and algorithmic randomness. In section 7.2, we describe natural embeddings of the Turing degrees into the uniform and non-uniform coarse degrees, and discuss some of their basic properties. In section 7.3, we show that no weakly 2-random set can be in the images of these embeddings by showing that if X is weakly 2-random and A is non-computable, then there is some coarse description of X that does not compute A. More generally, we show that if X is 1-random and A is computable from every coarse description of X, then A is K-trivial. Our main tool is a kind of compactness theorem for cone-avoiding descriptions. We also show that there do exist non-computable sets computable from every coarse description of some 1-random set, but that not all K-trivial sets have this property. In section 7.4, we give further examples of classes of sets that cannot be in the images of our embeddings. In section 7.5, we show that if two sets are relatively weakly 3-random then their coarse degrees form a minimal pair, in both the uniform and non-uniform cases, but that, at least for the non-uniform coarse degrees, the same is not true of every pair of relatively 2-random sets. These results are analogous to the fact that, for the Turing degrees, two relatively weakly 2-random sets always form a minimal pair, but two relatively 1-random sets may not. In section 7.6, we conclude with a few open questions.

For $S \subseteq 2^{<\omega}$, we write $[\![S]\!]$ for the open subset of 2^ω generated by S, that is, $[\![S]\!] = \{X : \exists n\,(X \upharpoonright n \in S)\}$. We denote the uniform measure on 2^ω by μ.

7.2. Coarsenings and embeddings of the Turing degrees

We can embed the Turing degrees into both the uniform and non-uniform coarse degrees, and our first connection between coarse computability and algorithmic randomness comes from considering such embeddings. While there may be several ways to define such embeddings, a natural way to proceed is to define a map $\mathcal{C} : 2^\omega \to 2^\omega$ such that $\mathcal{C}(A)$ contains the same information as A, but coded in a "coarsely robust" way. That is, we would like $\mathcal{C}(A)$ to be computable from A, and A to be computable from any coarse description of $\mathcal{C}(A)$.

In the case of the uniform coarse degrees, one might think that the latter reduction should be uniform, but that condition would be too strong: if $\Gamma^D = A$

for every coarse description D of $\mathcal{C}(A)$ then $\Gamma^\sigma(n)\downarrow \Rightarrow \Gamma^\sigma(n) = A(n)$ (since every string can be extended to a coarse description of $\mathcal{C}(A)$), which, together with the fact that for each n there is a σ such that $\Gamma^\sigma(n)\downarrow$, implies that A is computable. Thus we relax the uniformity condition slightly in the following definition.

Definition 7.2.1. A map $\mathcal{C} : 2^\omega \to 2^\omega$ is a *coarsening* if for each A we have $\mathcal{C}(A) \leq_{\mathrm{T}} A$, and for each coarse description D of $\mathcal{C}(A)$, we have $A \leq_{\mathrm{T}} D$. A coarsening $\mathcal{C}$ is *uniform* if there is a binary Turing functional Γ with the following properties for every coarse description D of $\mathcal{C}(A)$:

1. Γ^D is total.
2. Let $A_s(n) = \Gamma^D(n, s)$. Then $A_s = A$ for cofinitely many s.

Proposition 7.2.2. *Let $\mathcal{C}$ and $\mathcal{F}$ be coarsenings and A and B be sets. Then*

1. *$B \leq_{\mathrm{T}} A$ if and only if $\mathcal{C}(B) \leq_{\mathrm{nc}} \mathcal{C}(A)$,*
2. *If $\mathcal{C}$ is uniform then $B \leq_{\mathrm{T}} A$ if and only if $\mathcal{C}(B) \leq_{\mathrm{uc}} \mathcal{C}(A)$,*
3. *$\mathcal{C}(A) \equiv_{\mathrm{nc}} \mathcal{F}(A)$, and*
4. *If $\mathcal{C}$ and $\mathcal{F}$ are both uniform then $\mathcal{C}(A) \equiv_{\mathrm{uc}} \mathcal{F}(A)$.*

Proof. 1. Suppose that $\mathcal{C}(B) \leq_{\mathrm{nc}} \mathcal{C}(A)$. Then $\mathcal{C}(A)$ computes a coarse description D_1 of $\mathcal{C}(B)$. Thus $B \leq_{\mathrm{T}} D_1 \leq_{\mathrm{T}} \mathcal{C}(A) \leq_{\mathrm{T}} A$.

Now suppose that $B \leq_{\mathrm{T}} A$ and let D_2 be a coarse description of $\mathcal{C}(A)$. Then $\mathcal{C}(B) \leq_{\mathrm{T}} B \leq_{\mathrm{T}} A \leq_{\mathrm{T}} D_2$. Thus $\mathcal{C}(B) \leq_{\mathrm{nc}} \mathcal{C}(A)$.

2. Suppose that $\mathcal{C}$ is uniform and that $B \leq_{\mathrm{T}} A$. Let D_2 be a coarse description of $\mathcal{C}(A)$. Let A_s be as in Definition 7.2.1, with $D = D_2$. Then $\mathcal{C}(B) \leq_{\mathrm{T}} B \leq_{\mathrm{T}} A$, so let Φ be such that $\Phi^A = \mathcal{C}(B)$. Let $X \leq_{\mathrm{T}} D_2$ be defined as follows. Given n, search for an $s > n$ such that $\Phi^{A_s}(n)\downarrow$ and let $X(n) = \Phi^{A_s}(n)$. (Note that such an s must exist.) Then $X(n) = \Phi^A(n) = \mathcal{C}(B)(n)$ for almost all n, so X is a coarse description of $\mathcal{C}(B)$. Since X is obtained uniformly from D_2, we have $\mathcal{C}(B) \leq_{\mathrm{uc}} \mathcal{C}(A)$. The converse follows immediately from 1.

3. Let D_3 be a coarse description of $\mathcal{F}(A)$. Then $\mathcal{C}(A) \leq_{\mathrm{T}} A \leq_{\mathrm{T}} D_3$. Thus $\mathcal{C}(A) \leq_{\mathrm{nc}} \mathcal{F}(A)$. By symmetry, $\mathcal{C}(A) \equiv_{\mathrm{nc}} \mathcal{F}(A)$.

4. If $\mathcal{F}$ is uniform then the same argument as in the proof of 2 shows that we can obtain a coarse description of $\mathcal{C}(A)$ uniformly from D_3, whence $\mathcal{C}(A) \leq_{\mathrm{uc}} \mathcal{F}(A)$. If $\mathcal{C}$ is also uniform then $\mathcal{C}(A) \equiv_{\mathrm{uc}} \mathcal{F}(A)$ by symmetry. □

Thus uniform coarsenings all induce the same natural embeddings. It remains to show that uniform coarsenings exist. One example is given by Dzhafarov and Igusa [27]. We give a similar example. Let $I_n = [n!, (n+1)!)$ and let $\mathcal{I}(A) = \bigcup_{n \in A} I_n$; this map first appeared in Jockusch and Schupp [15]. Clearly $\mathcal{I}(A) \leq_{\mathrm{T}} A$, and it is easy to check that if D is a coarse description of $\mathcal{I}(A)$ then D computes A. Thus $\mathcal{I}$ is a coarsening.

To construct a uniform coarsening, let $\mathcal{H}(A) = \{\langle n, i\rangle : n \in A \ \wedge \ i \in \omega\}$ and define $\mathcal{E}(A) = \mathcal{I}(\mathcal{H}(A))$. The notation $\mathcal{E}$ denotes this particular coarsening throughout this chapter.

Proposition 7.2.3. *The map $\mathcal{E}$ is a uniform coarsening.*

Proof. Clearly $\mathcal{E}(A) \leq_{\mathrm{T}} A$. Now let D be a coarse description of $\mathcal{E}(A)$. Let $G = \{m : |D \cap I_m| > \frac{|I_m|}{2}\}$ and let $A_s = \{n : \langle n, s\rangle \in G\}$. Then $G =^* \mathcal{H}(A)$, so $A_s = A$ for all but finitely many s, and the A_s are obtained uniformly from D. $\square$

A first natural question is whether uniform coarse reducibility and non-uniform coarse reducibility are indeed different. We give a positive answer by showing that, unlike in the non-uniform case, the mappings $\mathcal{E}$ and $\mathcal{I}$ are not equivalent up to uniform coarse reducibility. Recall that a set X is *autoreducible* if there exists a Turing functional Φ such that for every $n \in \omega$ we have $\Phi^{X \setminus \{n\}}(n) = X(n)$. Equivalently, we could require that Φ not ask whether its input belongs to its oracle. We now introduce a Δ^0_2-version of this notion.

Definition 7.2.4. A set X is *jump-autoreducible* if there exists a Turing functional Φ such that for every $n \in \omega$ we have $\Phi^{(X\setminus\{n\})'}(n) = X(n)$.

Proposition 7.2.5. *Let X be such that $\mathcal{E}(X) \leq_{\mathrm{uc}} \mathcal{I}(X)$. Then X is jump-autoreducible.*

Proof. We must give a procedure for computing $X(n)$ from $(X \setminus \{n\})'$ that is uniform in X. Given an oracle for $X \setminus \{n\}$, we can uniformly compute $\mathcal{I}(X \setminus \{n\})$. Now $\mathcal{I}(X \setminus \{n\}) =^* \mathcal{I}(X)$, so $\mathcal{I}(X \setminus \{n\})$ is a coarse description of $\mathcal{I}(X)$. Since $\mathcal{E}(X) \leq_{\mathrm{uc}} \mathcal{I}(X)$ by assumption, from $\mathcal{I}(X \setminus \{n\})$ we can uniformly compute a coarse description D of $\mathcal{E}(X)$. Since $\mathcal{E}$ is a uniform coarsening by Proposition 7.2.3, from D we can uniformly obtain sets $A_0, A_1, \ldots$ with $A_s = X$ for all sufficiently large s. Composing these various reductions, from $X \setminus \{n\}$ we can uniformly compute sets $A_0, A_1, \ldots$ with $A_s = X$ for all sufficiently large s. Then from $(X \setminus \{n\})'$ we can uniformly compute $\lim_s A_s(n) = X(n)$, as needed. $\square$

We will now show that 2-generic sets are not jump-autoreducible, which will give us a first example separating uniform coarse reducibility and non-uniform coarse reducibility. For this we first show that no 1-generic set is autoreducible, which is an easy exercise.

Proposition 7.2.6. *If X is 1-generic, then X is not autoreducible.*

Proof. Suppose for the sake of a contradiction that X is 1-generic and is autoreducible via Φ. For a string σ, let $\sigma^{-1}(i)$ be the set of n such that $\sigma(n) = i$. If τ is a binary string, let $\tau \setminus \{n\}$ be the unique binary string μ of the same length such that $\mu^{-1}(1) = \tau^{-1}(1) \setminus \{n\}$. Let S be the set of strings τ such that $\Phi^{\tau\setminus\{n\}}(n)\downarrow \neq \tau(n)\downarrow$ for some n. Then S is a c.e. set of strings and X does not meet S. Since X is 1-generic, there is a string $\sigma \prec X$ that has no extension in S. Let $n = |\sigma|$, and let $\tau \succ \sigma$ be a string such that $\Phi^{\tau\setminus\{n\}}(n)\downarrow$. Such a string τ exists because $\sigma \prec X$ and Φ witnesses that X is autoreducible. Furthermore, we may assume that $\tau(n) \neq \Phi^{\tau\setminus\{n\}}$, since changing the value of $\tau(n)$ does not affect any of the conditions in the choice of τ. Hence τ is an extension of σ and $\tau \in S$, which is the desired contradiction. $\square$

Proposition 7.2.7. *If X is 2-generic, then X is not jump-autoreducible.*

Proof. Since X is 2-generic, X is 1-generic relative to $\emptyset'$. Hence, by relativising the proof of the previous proposition to $\emptyset'$, we see that X is not autoreducible relative to $\emptyset'$. However, the class of 1-generic sets is uniformly GL_1, i.e., there exists a single Turing functional Ψ such that for every 1-generic X we have $\Psi^{X\oplus\emptyset'} = X'$, as can be verified by looking at the usual proof that every 1-generic is GL_1 (see [14, Lemma 2.6]). Of course, if X is 1-generic, then $X \setminus \{n\}$ is also 1-generic for every n. Thus from an oracle for $(X \setminus \{n\}) \oplus \emptyset'$ we can uniformly compute $(X \setminus \{n\})'$. Now, if X is jump-autoreducible, from $(X \setminus \{n\})'$ we can uniformly compute $X(n)$. Composing these reductions shows that $X(n)$ is uniformly computable from $(X \setminus \{n\}) \oplus \emptyset'$, which contradicts our previous remark that X is not autoreducible relative to $\emptyset'$. □

Corollary 7.2.8. *If X is 2-generic, then $\mathcal{E}(X) \leq_{\mathrm{nc}} \mathcal{I}(X)$ but $\mathcal{E}(X) \not\leq_{\mathrm{uc}} \mathcal{I}(X)$.*

Proof. We know that $\mathcal{E}(X) \leq_{\mathrm{nc}} \mathcal{I}(X)$ from Proposition 7.2.2. The fact that $\mathcal{E}(X) \not\leq_{\mathrm{uc}} \mathcal{I}(X)$ follows from Propositions 7.2.5 and 7.2.7. □

It is natural to ask whether the same result holds for 2-random sets. In the proof above we used the fact that the 2-generic sets are uniformly GL_1. For 2-random sets this fact is almost true, as expressed by the following lemma. The proof is adapted from Monin [82], where a generalisation for higher levels of randomness is proved. Let $U_0, U_1, \ldots$ be a fixed universal Martin-Löf test relative to $\emptyset'$. The 2-*randomness deficiency* of a 2-random X is the least c such that $X \notin U_c$.

Lemma 7.2.9. *There is a Turing functional Θ such that, for a 2-random X and an upper bound b on the 2-randomness deficiency of X, we have $\Theta^{X\oplus\emptyset',b} = X'$.*

Proof. Let $\mathscr{V}_e = \{Z : e \in Z'\}$. The $\mathscr{V}_e$ are uniformly Σ^0_1 classes, so we can define a function $f \leq_{\mathrm{T}} \emptyset'$ such that $\mu(\mathscr{V}_e \setminus \mathscr{V}_e[f(e,i)]) < 2^{-i}$ for all e and i. Then each sequence $\mathscr{V}_e \setminus \mathscr{V}_e[f(e,0)], \mathscr{V}_e \setminus \mathscr{V}_e[f(e,1)], \ldots$ is an $\emptyset'$-Martin Löf test, and from b we can compute a number m such that if X is 2-random and b bounds the 2-randomness deficiency of X, then $X \notin \mathscr{V}_e \setminus \mathscr{V}_e[f(e,m)]$. Then $X \in \mathscr{V}_e$ if and only if $X \in \mathscr{V}_e[f(e,m)]$, which we can verify $(X \oplus \emptyset')$-computably. □

Proposition 7.2.10. *If X is 2-random, then X is not jump-autoreducible.*

Proof. Because X is 2-random, it is not autoreducible relative to $\emptyset'$, as can be seen by relativising the proof of Figueira, Miller and Nies [29] that no 1-random set is autoreducible. To obtain a contradiction, assume that X is jump-autoreducible through some functional Φ. It can be directly verified that there is a computable function f such that $f(n)$ bounds the randomness deficiency of $X \setminus \{n\}$. Now let $\Psi^{Y\oplus\emptyset'}(n) = \Phi^{\Theta^{Y\oplus\emptyset',f(n)}}(n)$. Then X is autoreducible relative to $\emptyset'$ through Ψ, a contradiction. □

Corollary 7.2.11. *If X is 2-random, then $\mathcal{E}(X) \leq_{\mathrm{nc}} \mathcal{I}(X)$ but $\mathcal{E}(X) \not\leq_{\mathrm{uc}} \mathcal{I}(X)$.*

Although we will not discuss generic reducibility after this section, it is worth noting that our maps $\mathcal{E}$ and $\mathcal{I}$ also allow us to distinguish generic reducibility from its non-uniform analogue. Let us briefly review the relevant definitions from [15]. A *generic description* of a set A is a partial function that agrees with A where defined, and whose domain has density 1. A set A is *generically reducible* to a set B, written $A \leq_{\mathrm{g}} B$, if there is an enumeration operator W such that if Φ is a generic description of B, then $W^{\mathrm{graph}(\Phi)}$ is the graph of a generic description of A. We can define the notion of *non-uniform generic reducibility* in a similar way: $A \leq_{\mathrm{ng}} B$ if for every generic description Φ of B, there is a generic description Ψ of A such that $\mathrm{graph}(\Psi)$ is enumeration reducible to $\mathrm{graph}(\Phi)$.

It is easy to see that $\mathcal{E}(X) \leq_{\mathrm{ng}} \mathcal{I}(X)$ for all X. On the other hand, we have the following fact.

PROPOSITION 7.2.12. *If $\mathcal{E}(X) \leq_{\mathrm{g}} \mathcal{I}(X)$ then X is autoreducible.*

PROOF. Let I_n be as in the definition of $\mathcal{I}$. Suppose that W witnesses that $\mathcal{E}(X) \leq_{\mathrm{g}} \mathcal{I}(X)$. We can assume that W^Z is the graph of a partial function for every oracle Z. Define a Turing functional Θ as follows. Given an oracle Y and an input n, let $\Phi(k) = Y(m)$ if $k \in I_m$ and $m \neq n$, and let $\Phi(k)\uparrow$ if $k \in I_n$. Let Ψ be the partial function with graph $W^{\mathrm{graph}(\Phi)}$. Search for an i and a $k \in I_{\langle n,i \rangle}$ such that $\Psi(k)\downarrow$. If such numbers are found then let $\Theta^Y(n) = \Psi(k)$. If $Y = X \setminus \{n\}$ then Φ is a generic description of $\mathcal{I}(X)$, so Ψ is a generic description of $\mathcal{E}(X)$, and hence $\Theta^Y(n)\downarrow = X(n)$. Thus X is autoreducible. $\square$

We finish this section by showing that, for both the uniform and the non-uniform coarse degrees, coarsenings of the appropriate type preserve joins but do not always preserve existing meets.

PROPOSITION 7.2.13. *Let $\mathcal{C}$ be a coarsening. Then $\mathcal{C}(A \oplus B)$ is the least upper bound of $\mathcal{C}(A)$ and $\mathcal{C}(B)$ in the non-uniform coarse degrees. The same holds for the uniform coarse degrees if $\mathcal{C}$ is a uniform coarsening.*

PROOF. By Proposition 7.2.2 we know that $\mathcal{C}(A \oplus B)$ is an upper bound for $\mathcal{C}(A)$ and $\mathcal{C}(B)$ in both the uniform and non-uniform coarse degrees. Let us show that it is the least upper bound. If $\mathcal{C}(A), \mathcal{C}(B) \leq_{\mathrm{nc}} G$ then every coarse description D of G computes both A and B, so $D \geq_{\mathrm{T}} A \oplus B \geq_{\mathrm{T}} \mathcal{C}(A \oplus B)$. Thus $G \geq_{\mathrm{nc}} \mathcal{C}(A \oplus B)$.

Finally, assume that $\mathcal{C}$ is a uniform coarsening and let $\mathcal{C}(A), \mathcal{C}(B) \leq_{\mathrm{uc}} G$. Let Φ be a Turing functional such that $\Phi^{A \oplus B} = \mathcal{C}(A \oplus B)$. Every coarse description H of G uniformly computes coarse descriptions D_1 of $\mathcal{C}(A)$ and D_2 of $\mathcal{C}(B)$. Since $\mathcal{C}$ is uniform, there are Turing functionals Γ and Δ such that, letting $A_s(n) = \Gamma^{D_1}(n, s)$ and $B_s(n) = \Gamma^{D_2}(n, s)$, we have that $A \oplus B = A_s \oplus B_s$ for all sufficiently large s. Let E be defined as follows. Given n, search for an $s \geq n$ such that $\Phi^{A_s \oplus B_s}(n)\downarrow$, and let $E(n) = \Phi^{A_s \oplus B_s}(n)$. If n is sufficiently large, then $E(n) = \Phi^{A \oplus B}(n) = \mathcal{C}(A \oplus B)(n)$, so E is a coarse description of $\mathcal{C}(A \oplus B)$. Since E is obtained uniformly from H, we have that $\mathcal{C}(A \oplus B) \leq_{\mathrm{uc}} G$. $\square$

LEMMA 7.2.14. *Let $\mathcal{C}$ be a uniform coarsening and let $Y \leq_{\mathrm{T}} X$. Then $Y \leq_{\mathrm{uc}} \mathcal{C}(X)$.*

PROOF. Let Φ be a Turing functional such that $\Phi^X = Y$. Let D be a coarse description of $\mathcal{C}(X)$ and let A_s be as in Definition 7.2.1. Now define $G(n)$ to be the value of $\Phi^{A_s}(n)$ for the least pair $\langle s, t \rangle$ such that $s \geq n$ and $\Phi^{A_s}(n)[t]\downarrow$. Then $G =^* Y$, so G is a coarse description of Y. $\square$

PROPOSITION 7.2.15. *Let $\mathcal{C}$ be a coarsening. Then $\mathcal{C}$ does not always preserve existing meets in the non-uniform coarse degrees. The same holds for the uniform coarse degrees if $\mathcal{C}$ is a uniform coarsening.*

PROOF. Let X, Y be relatively 2-random and Δ^0_3. Then X and Y form a minimal pair in the Turing degrees, while X and Y do not form a minimal pair in the non-uniform coarse degrees by Theorem 7.5.6 below. Since every coarse description of $\mathcal{C}(X)$ computes X we see that $\mathcal{C}(X) \geq_{\mathrm{nc}} X$ and $\mathcal{C}(Y) \geq_{\mathrm{nc}} Y$. Therefore $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ also do not form a minimal pair in the non-uniform coarse degrees.

Next, let $\mathcal{C}$ be a uniform coarsening. We have seen above that there exists some $A \leq_{\mathrm{nc}} \mathcal{C}(X), \mathcal{C}(Y)$ that is not coarsely computable. Then $A \leq_{\mathrm{T}} X, Y$, so $A \leq_{\mathrm{uc}} \mathcal{C}(X), \mathcal{C}(Y)$ by the previous lemma. Thus, $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ do not form a minimal pair in the uniform coarse degrees. $\square$

7.3. Randomness, K-triviality and robust information coding

It is reasonable to expect that the embeddings induced by $\mathcal{E}$ (or equivalently, by any uniform coarsening) are not surjective. Indeed, if $\mathcal{E}(A) \leq_{\mathrm{uc}} X$ then the information represented by A is coded into X in a fairly redundant way. If A is non-computable, it should follow that X cannot be random. As we will see, we can make this intuition precise.

DEFINITION 7.3.1. Let $X^{\mathfrak{c}}$ be the set of all A such that A is computable from every coarse description of X.

We will show that if X is weakly 2-random then $X^{\mathfrak{c}} = \mathbf{0}$, and hence $\mathcal{E}(A) \not\leq_{\mathrm{nc}} X$ for all non-computable A (since every coarse description of $\mathcal{E}(A)$ computes A). Since no 1-random set can be coarsely computable, it will follow that $X \not\equiv_{\mathrm{nc}} \mathcal{E}(B)$ and $X \not\equiv_{\mathrm{uc}} \mathcal{E}(B)$ for all B. We will first prove the following theorem. Let $\mathcal{K}$ be the class of K-trivial sets. (See [25] or [86] for more on K-triviality.)

THEOREM 7.3.2. *If X is 1-random then $X^{\mathfrak{c}} \subseteq \mathcal{K}$.*

By Downey, Nies, Weber and Yu [26], if X is weakly 2-random then it cannot compute any non-computable Δ^0_2 sets. Since $\mathcal{K} \subset \Delta^0_2$, our desired result follows from Theorem 7.3.2.

COROLLARY 7.3.3. *If X is weakly 2-random then $X^{\mathfrak{c}} = \mathbf{0}$, and hence $\mathcal{E}(A) \not\leq_{\mathrm{nc}} X$ for all noncomputable A. In particular, in both the uniform and non-uniform coarse degrees, the degree of X is not in the image of the embedding induced by $\mathcal{E}$.*

To prove Theorem 7.3.2, we use the fact, established by Hirschfeldt, Nies and Stephan [40], that A is K-trivial if and only if A is a base for 1-randomness, that is, A is computable in a set that is 1-random relative to A. The basic idea is to show that if X is 1-random and $A \in X^{\mathfrak{c}}$, then for each $k > 1$ there is a way to partition X into k many "slices" $X_0, \dots, X_{k-1}$ such that for each $i < k$, we have $A \leq_{\mathrm{T}} X_0 \oplus \cdots \oplus X_{i-1} \oplus X_{i+1} \oplus \cdots \oplus X_{k-1}$ (where the right-hand side of this inequality denotes $X_1 \oplus \cdots \oplus X_{k-1}$ when $i = 0$ and $X_0 \oplus \cdots \oplus X_{k-2}$ when $i = k-1$). It will then follow by van Lambalgen's Theorem (which will be discussed below) that each X_i is 1-random relative to $X_0 \oplus \cdots \oplus X_{i-1} \oplus X_{i+1} \oplus \cdots \oplus X_{k-1} \oplus A$, and hence, again by van Lambalgen's Theorem, that X is 1-random relative to A. Since $A \in X^{\mathfrak{c}}$ implies that $A \leq_{\mathrm{T}} X$, we will conclude that A is a base for 1-randomness, and hence is K-trivial. We begin with some notation for certain partitions of X.

Definition 7.3.4. Let $X \subseteq \omega$. For an infinite subset $Z = \{z_0 < z_1 < \cdots\}$ of ω, let $X \upharpoonright Z = \{n : z_n \in X\}$. For $k > 1$ and $i < k$, define

$$X_i^k = X \upharpoonright \{n : n \equiv i \bmod k\} \quad \text{and} \quad X_{\neq i}^k = X \upharpoonright \{n : n \not\equiv i \bmod k\}.$$

Note that $X_{\neq i}^k \equiv_{\mathrm{T}} X \setminus \{n : n \equiv i \bmod k\}$ and $\overline{\rho}(X \triangle (X \setminus \{n : n \equiv i \bmod k\})) \leq \frac{1}{k}$.

Van Lambalgen's Theorem [71] states that $Y \oplus Z$ is 1-random if and only if Y and Z are relatively 1-random. The proof of this theorem shows, more generally, that if Z is computable, infinite, and coinfinite, then X is 1-random if and only if $X \upharpoonright Z$ and $X \upharpoonright \overline{Z}$ are relatively 1-random. Relativising this fact and applying induction, we get the following version of van Lambalgen's Theorem.

Theorem 7.3.5. (van Lambalgen [71]) *The following are equivalent for all sets X and A, and all $k > 1$.*

1. *X is 1-random relative to A.*
2. *For each $i < k$, the set X_i^k is 1-random relative to $X_{\neq i}^k \oplus A$.*

The last ingredient we need for the proof of Theorem 7.3.2 is a kind of compactness principle, which will also be used to yield further results in the next section, and is of independent interest given its connection with the following concept defined in [39].

Definition 7.3.6. Let $r \in [0,1]$. A set X is *coarsely computable at density r* if there is a computable set C such that $\overline{\rho}(X \triangle C) \leq 1 - r$. The *coarse computability bound* of X is

$$\gamma(X) = \sup\{r : X \text{ is coarsely computable at density } r\}.$$

As noted in [39], there are sets X such that $\gamma(X) = 1$ but X is not coarsely computable. In other words, there is no principle of "compactness of computable coarse descriptions" principle. (Although Miller (see [39, Theorem 5.8]) showed that one can in fact recover such a principle by adding a further effectivity condition to the requirement that $\gamma(X) = 1$.) The following theorem shows that if we replace "computable" by "cone-avoiding", the situation is different.

Theorem 7.3.7. *Let A and X be arbitrary sets. Suppose that for each $\varepsilon > 0$ there is a set D_ε such that $\overline{\rho}(X \triangle D_\varepsilon) \leq \varepsilon$ and $A \not\leq_{\mathrm{T}} D_\varepsilon$. Then there is a coarse description D of X such that $A \not\leq_{\mathrm{T}} D$.*

Proof. The basic idea is that, given a Turing functional Φ and a string σ that is "close to" X, we can extend σ to a string τ that is "close to" X such that $\Phi^D \neq A$ for all D extending τ that are "close to" X. We can take τ to be any string "close to" X such that, for some n, either $\Phi^\tau(n)\downarrow \neq A(n)$ or $\Phi^\gamma(n)\uparrow$ for all γ extending τ that are "close to" X. If no such τ exists, we can obtain a contradiction by arguing that $A \leq_{\mathrm{T}} D_\varepsilon$ for sufficiently small ε, since with an oracle for D_ε we have access to many strings that are "close to" D_ε and hence to X, by the triangle inequality for Hamming distance. In the above discussion the meaning of "close to" is different in different contexts, but the precise version will be given below. Further, as the construction proceeds, the meaning of "close to" becomes so stringent that we guarantee that $\rho(X\triangle D) = 0$. We now specify the formal details.

We obtain D as $\bigcup_e \sigma_e$, where $\sigma_e \in 2^{<\omega}$ and $\sigma_0 \subsetneq \sigma_1 \subsetneq \cdots$. In order to ensure that $\rho(X\triangle D) = 0$, we require that for all e and all m in the interval $[|\sigma_e|, |\sigma_{e+1}|]$, either D and X agree on the interval $[|\sigma_e|, m)$ or $\rho_m(X\triangle D) \leq 2^{-|\sigma_e|}$, with the latter true for $m = |\sigma_{e+1}|$. This condition implies that $\rho_m(X\triangle D) \leq 2^{-|\sigma_e|}$ for all $m \in [|\sigma_{e+1}|, |\sigma_{e+2}|]$, and hence that $\rho(X\triangle D) = 0$.

Let σ and τ be strings and let ε be a positive real number. Call τ an ε-*good* extension of σ if τ properly extends σ and for all $m \in [|\sigma|, |\tau|]$, either X and τ agree on $[|\sigma|, m)$ or $\rho_m(\tau\triangle X) \leq \varepsilon$, with the latter true for $m = |\tau|$. In line with the previous paragraph, we require that σ_{e+1} be a $2^{-|\sigma_e|}$-good extension of σ_e for all e.

At stage 0, let σ_0 be the empty string. At stage $e+1$, we are given σ_e and choose σ_{e+1} as follows so as to force that $A \neq \Phi_e^D$. Let $\varepsilon = 2^{-|\sigma_e|}$.

Case 1. There is a number n and a string τ that is an ε-good extension of σ_e such that $\Phi_e^\tau(n)\downarrow \neq A(n)$. Let σ_{e+1} be such a τ.

Case 2. Case 1 does not hold and there is a number n and a string β that is an ε-good extension of σ_e such that $|\beta| \geq |\sigma_e| + 2$ and $\Phi_e^\tau(n)\uparrow$ for all $\frac{\varepsilon}{4}$-good extensions τ of β. Let σ_{e+1} be such a β.

We claim that either Case 1 or Case 2 applies. Suppose not. Let $D_{\frac{\varepsilon}{5}}$ be as in the hypothesis of the lemma, so that $\overline{\rho}(X\triangle D_{\frac{\varepsilon}{5}}) \leq \frac{\varepsilon}{5}$ and $A \not\leq_{\mathrm{T}} D_{\frac{\varepsilon}{5}}$. Let $c \geq |\sigma_e|+2$ be sufficiently large so that $\rho_m(X\triangle D_{\frac{\varepsilon}{5}}) \leq \frac{\varepsilon}{4}$ for all $m \geq c$ and σ_e has an $\frac{\varepsilon}{4}$-good extension β of length c. Note that the string obtained from σ_e by appending a sufficiently long segment of X starting with $X(|\sigma_e|)$ is an $\frac{\varepsilon}{4}$-good extension of σ_e, so such a β exists, and we assume it is obtained in this manner.

We now obtain a contradiction by showing that $A \leq_{\mathrm{T}} D_{\frac{\varepsilon}{5}}$. To calculate $A(n)$ search for a string γ extending β such that $\Phi_e^\gamma(n)\downarrow$, say with use u, and $\rho_m(D_{\frac{\varepsilon}{5}}\triangle\gamma) \leq \frac{\varepsilon}{2}$ for all $m \in [c, u)$. We first check that such a string γ exists. Since Case 2 does not hold, there is a string τ that is an $\frac{\varepsilon}{4}$-good extension of β such that $\Phi_e^\tau(n)\downarrow$. We claim that τ meets the criteria to serve as γ. We need only check that $\rho_m(D_{\frac{\varepsilon}{5}}\triangle\tau) \leq \frac{\varepsilon}{2}$ for all $m \in [c, u)$. Fix $m \in [c, u)$. Then

$$\rho_m(D_{\frac{\varepsilon}{5}}\triangle\tau) \leq \rho_m(D_{\frac{\varepsilon}{5}}\triangle X) + \rho_m(X\triangle\tau) \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{\varepsilon}{2}.$$

Next we claim that γ is an ε-good extension of σ_e. The string γ extends σ_e since it extends β, and β extends σ_e. Let $m \in [|\sigma_e, |\gamma|]$ be given. If $m < c$, then γ and X agree on the interval $[|\sigma_e|, m)$ because β and X agree on this interval and γ extends β. Now suppose that $m \geq c$. Then

$$\rho_m(\gamma\triangle X) \leq \rho_m(\gamma\triangle D_{\frac{\varepsilon}{5}}) + \rho_m(D_{\frac{\varepsilon}{5}}\triangle X) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4} < \varepsilon.$$

Since γ is an ε-good extension of σ_e for which $\Phi_e^\gamma(n)\downarrow$, and Case 1 does not hold, we conclude that $\Phi_e^\gamma(n) = A(n)$. The search for γ can be carried out computably in $D_{\frac{\varepsilon}{5}}$, so we conclude that $A \leq_{\mathrm{T}} D_{\frac{\varepsilon}{5}}$, contradicting our choice of $D_{\frac{\varepsilon}{5}}$. (Although β cannot be computed from $D_{\frac{\varepsilon}{5}}$, we may use it in our computation of $A(n)$ since it is a fixed string which does not depend on n.) This contradiction shows that Case 1 or Case 2 must apply.

Let $D = \bigcup_n \sigma_n$. Then $\rho(D \triangle X) = 0$, and $A \nleq_{\mathrm{T}} D$ since Case 1 or Case 2 applies at every stage. $\square$

Proof of Theorem 7.3.2. Let $A \in X^{\mathfrak{c}}$. By Theorem 7.3.7, there is an $\varepsilon > 0$ such that $A \leq_{\mathrm{T}} D_\varepsilon$ whenever $\overline{\rho}(X \triangle D_\varepsilon) \leq \varepsilon$. Let k be an integer such that $k > \frac{1}{\varepsilon}$. As noted in Definition 7.3.4, $X^k_{\neq i}$ is Turing equivalent to such a D_ε for each $i < k$, so we have $A \leq_{\mathrm{T}} X^k_{\neq i}$ for all $i < k$. By the unrelativised form of Theorem 7.3.5, each X^k_i is 1-random relative to $X^k_{\neq i}$, and hence relative to $X^k_{\neq i} \oplus A \equiv_{\mathrm{T}} X^k_{\neq i}$. Again by Theorem 7.3.5, X is 1-random relative to A. But $A \leq_{\mathrm{T}} X$, so A is a base for 1-randomness, and hence is K-trivial. $\square$

Weak 2-randomness is exactly the level of randomness necessary to obtain Corollary 7.3.3 directly from Theorem 7.3.2, because, as shown in [26], if a 1-random set is not weakly 2-random, then it computes a non-computable c.e. set. The corollary itself does hold of some 1-random sets that are not weakly 2-random, because if it holds of X then it also holds of any Y such that $\rho(Y \triangle X) = 0$. (For example, let X be 2-random and let Y be obtained from X by letting $Y(2^n) = \Omega(n)$ (where Ω is Chaitin's halting probability) for all n and letting $Y(k) = X(k)$ for all other k. By van Lambalgen's Theorem, Y is 1-random, but it computes Ω, and hence is not weakly 2-random.)

Nevertheless, Corollary 7.3.3 does not hold of all 1-random sets, as we now show.

Definition 7.3.8. Let $W_0, W_1, \ldots$ be an effective listing of the c.e. sets. A set A is *promptly simple* if it is c.e. and coinfinite, and there exist a computable function f and a computable enumeration $A[0], A[1], \ldots$ of A such that for each e, if W_e is infinite then there are n and s for which $n \in W_e[s] \setminus W_e[s-1]$ and $n \in A[f(s)]$. Note that every promptly simple set is non-computable.

We will show that if $X \leq_{\mathrm{T}} \emptyset'$ is 1-random then $X^{\mathfrak{c}}$ contains a promptly simple set, and there is a promptly simple set A such that $\mathcal{E}(A) \leq_{\mathrm{nc}} X$. (We do not know whether we can improve the last statement to $\mathcal{E}(A) \leq_{\mathrm{uc}} X$.) In fact, we will obtain a considerably stronger result by first proving a generalisation of the fact, due to Hirschfeldt and Miller (see [25, Theorem 7.2.11]), that if $\mathcal{T}$ is a Σ^0_3 class of measure 0, then there is a non-computable c.e. set that is computable from each 1-random element of $\mathcal{T}$.

For a binary relation $P(Y, Z)$ between elements of 2^ω, let $P(Y) = \{Z : P(Y, Z)\}$.

Theorem 7.3.9. *Let $\mathcal{S}_0, \mathcal{S}_1, \ldots$ be uniformly Π^0_2 classes of measure 0, and let $P_0(Y, Z), P_1(Y, Z), \ldots$ be uniformly Π^0_1 relations. Let $\mathcal{D}$ be the class of all Y for which there are numbers k, m and a 1-random set Z such that $Z \in P_k(Y) \subseteq \mathcal{S}_m$. Then there is a promptly simple set A such that $A \leq_{\mathrm{T}} Y$ for every $Y \in \mathcal{D}$.*

Proof. Let $(\mathcal{V}^m_n)_{m,n\in\omega}$ be uniformly Σ^0_1 classes such that $\mathcal{S}_m = \bigcap_n \mathcal{V}^m_n$. We may assume that $\mathcal{V}^m_0 \supseteq \mathcal{V}^m_1 \supseteq \cdots$ for all m. For each m, we have $\mu(\bigcap_n \mathcal{V}^n_m) =$

$\mu(\mathcal{S}_m) = 0$, so $\lim_n \mu(\mathcal{V}_n^m) = 0$ for each m. Let Θ be a computable relation such that $P_k(Y, Z) \equiv \forall l\, \Theta(k, Y \upharpoonright l, Z \upharpoonright l)$.

Define A as follows. At each stage s, if there is an $e < s$ such that no numbers have entered A for the sake of e yet, and an $n > 2e$ such that $n \in W_e[s] \setminus W_e[s-1]$ and $\mu(\mathcal{V}_n^m[s]) \leq 2^{-e}$ for all $m < e$, then for the least such e, put the least corresponding n into A. We say that n enters A for the sake of e.

Clearly, A is c.e. and coinfinite, since at most e many numbers less than $2e$ ever enter A. Suppose that W_e is infinite. Let $t > e$ be a stage such that all numbers that will ever enter A for the sake of any $i < e$ are in $A[t]$. There must be an $s \geq t$ and an $n > 2e$ such that $n \in W_e[s] \setminus W_e[s-1]$ and $\mu(\mathcal{V}_n^m[s]) \leq 2^{-e}$ for all $m < e$. Then the least such n enters A for the sake of e at stage s unless another number has already entered A for the sake of e. It follows that A is promptly simple.

Now suppose that $Y \in \mathcal{D}$. Let the numbers k, m and the 1-random set Z be such that $Z \in P_k(Y) \subseteq \mathcal{S}_m$. Let $B \leq_{\mathrm{T}} Y$ be defined as follows. Given n, let

$$\mathcal{D}_s^n = \{X : (\forall l \leq s)\, \Theta(k, Y \upharpoonright l, X \upharpoonright l)\} \setminus \mathcal{V}_n^m[s].$$

Then $\mathcal{D}_0^n \supseteq \mathcal{D}_1^n \supseteq \cdots$. Furthermore, if $X \in \bigcap_s \mathcal{D}_s^n$ then $P_k(Y, X)$ and $X \notin \mathcal{V}_n^m$. Since $P_k(Y) \subseteq S_m \subseteq \mathcal{V}_n^m$, it follows that $X \notin P_k(Y)$, which is a contradiction. Thus $\bigcap_s \mathcal{D}_s^n = \emptyset$. Since the $\mathcal{D}_s^n$ are nested closed sets, it follows that there is an s such that $\mathcal{D}_s^n = \emptyset$. Let s_n be the least such s (which we can find using Y) and let $B(n) = A(n)[s_n]$. Note that $B \subseteq A$.

Let $T = \{\mathcal{V}_n^m[s] : n \text{ enters } A \text{ at stage } s\}$. We can think of T as a uniform singly-indexed sequence of Σ_1^0 sets since m is fixed and for each n there is at most one s such that $\mathcal{V}_n^m[s] \in T$. For each e, there is at most one n that enters A for the sake of e, and the sum of the measures of the $\mathcal{V}_n^m[s]$ such that n enters A at stage s for the sake of some $e > m$ is bounded by $\sum_e 2^{-e}$, which is finite. Thus T is a Solovay test, and hence Z is in only finitely many elements of T. So for all but finitely many n, if n enters A at stage s then $Z \notin \mathcal{V}_n^m[s]$. Then $Z \in \mathcal{D}_s^n$, so $s_n > s$. Hence, for all such n, we have that $B(n) = A(n)[s_n] = 1$. Thus $B =^* A$, so $A \equiv_{\mathrm{T}} B \leq_{\mathrm{T}} Y$. $\square$

Note that the result of Hirschfeldt and Miller mentioned above follows from this theorem by starting with a Σ_3^0 class $\mathcal{S} = \bigcap_m \mathcal{S}_m$ of measure 0 and letting each P_k be the identity relation.

COROLLARY 7.3.10. *Let $X \leq_{\mathrm{T}} \emptyset'$ be 1-random. There is a promptly simple set A such that if $\overline{\rho}(D \triangle X) < \frac{1}{4}$ then $A \leq_{\mathrm{T}} D$. In particular, X^{c} contains a promptly simple set, and there is a promptly simple set A such that $\mathcal{E}(A) \leq_{\mathrm{nc}} X$.*

PROOF. Say that sets Y and Z are *r-close from m on* if whenever $m < n$, the Hamming distance between $Y \upharpoonright n$ and $Z \upharpoonright n$ (i.e., the number of bits on which these two strings differ) is at most rn.

Let $\mathcal{S}_m$ be the class of all Z such that X and Z are $\frac{1}{2}$-close from m on. Since X is Δ_2^0, the $\mathcal{S}_m$ are uniformly Π_2^0 classes. Furthermore, if X and Z are $\frac{1}{2}$-close from m on for some m, then Z cannot be 1-random relative to X (by the same argument that shows that if C is 1-random then there must be infinitely many n such that $C \upharpoonright n$ has more 1's than 0's), so $\mu(\mathcal{S}_m) = 0$ for all m. Let $P_m(Y, Z)$ hold if and only if Y and Z are $\frac{1}{4}$-close from m on. The P_m are clearly uniformly Π^0_1 relations.

Thus the hypotheses of Theorem 7.3.9 are satisfied. Let A be as in that theorem. Suppose that $\overline{\rho}(D \triangle X) < \frac{1}{4}$. Then there is an m such that D and X are $\frac{1}{4}$-close from m on. If D and Z are $\frac{1}{4}$-close from m on, then by the triangle inequality for Hamming distance, X and Z are $\frac{1}{2}$-close from m on. Thus $X \in P_m(D) \subseteq \mathcal{S}_m$, so $A \leq_{\mathrm{T}} D$. □

After learning about Corollary 7.3.10, Nies [85] gave a different but closely connected proof of this result, which works even for X of positive effective Hausdorff dimension, as long as we sufficiently decrease the bound $\frac{1}{4}$. However, even for X of effective Hausdorff dimension 1 his bound is much worse, namely $\frac{1}{20}$.

Maass, Shore and Stob [75, Corollary 1.6] showed that if A and B are promptly simple then there is a promptly simple set G such that $G \leq_{\mathrm{T}} A$ and $G \leq_{\mathrm{T}} B$. Thus we have the following extension of Kučera's result [61] that two Δ^0_2 1-random sets cannot form a minimal pair, which will also be useful below.

COROLLARY 7.3.11. *Let $X_0, X_1 \leq_{\mathrm{T}} \emptyset'$ be 1-random. There is a promptly simple set A such that if $\overline{\rho}(D \triangle X_i) < \frac{1}{4}$ for some $i \in \{0,1\}$ then $A \leq_{\mathrm{T}} D$.*

It is easy to adapt the proof of Corollary 7.3.10 to give a direct proof of Corollary 7.3.11, and indeed of the fact that for any uniformly $\emptyset'$-computable family $X_0, X_1, \ldots$ of 1-random sets, there is a promptly simple set A such that if $\overline{\rho}(D \triangle X_i) < \frac{1}{4}$ for some i then $A \leq_{\mathrm{T}} D$. (We let $\mathcal{S}_{\langle i,m \rangle}$ be the class of all Z such that X_i and Z are $\frac{1}{2}$-close from m on, and the rest of the proof is essentially as before.)

Given the many (and often surprising) characterisations of K-triviality, it is natural to ask whether there is a converse to Theorem 7.3.2 stating that if A is K-trivial then $A \in X^{\mathfrak{c}}$ for some 1-random X. We now show that is not the case, using a recent result of Bienvenu, Greenberg, Kučera, Nies, and Turetsky [8]. There are many notions of randomness tests in the theory of algorithmic randomness. Some, like Martin-Löf tests, correspond to significant levels of algorithmic randomness, while other, less obviously natural ones have nevertheless become important tools in the development of this theory. Balanced tests belong to the latter class.

DEFINITION 7.3.12. Let $\mathcal{W}_0, \mathcal{W}_1, \ldots \subseteq 2^\omega$ be an effective list of all Σ^0_1 classes. A *balanced test* is a sequence $(\mathcal{U}_n)_{n \in \omega}$ of Σ^0_1 classes such that there is a computable binary function f with the following properties.

1. $|\{s : f(n, s+1) \neq f(n,s)\}| \leq O(2^n)$,
2. $\forall n\, \mathcal{U}_n = \mathcal{W}_{\lim_s f(n,s)}$, and
3. $\forall n\, \forall s\, \mu(\mathcal{W}_{f(n,s)}) \leq 2^{-n}$.

For $\sigma \in 2^{<\omega}$ and $X \in 2^\omega$, we write σX for the element of 2^ω obtained by concatenating σ and X.

THEOREM 7.3.13. (Bienvenu, Greenberg, Kučera, Nies and Turetsky [8]) *There are a K-trivial set A and a balanced test $(\mathcal{U}_n)_{n \in \omega}$ such that if $A \leq_{\mathrm{T}} X$ then there is a string σ with $\sigma X \in \bigcap_n \mathcal{U}_n$.*

We will also use the following measure-theoretic fact.

THEOREM 7.3.14. (Loomis and Whitney [74]) *Let $\mathcal{S} \subseteq 2^\omega$ be open, and let $k \in \omega$. For $i < k$, let $\pi_i(\mathcal{S}) = \{Y^k_{\neq i} : Y \in \mathcal{S}\}$. Then $\mu(\mathcal{S})^{k-1} \leq \mu(\pi_0(\mathcal{S})) \cdots \mu(\pi_{k-1}(\mathcal{S}))$.*

Our result will follow from the following lemma.

LEMMA 7.3.15. *Let X be 1-random, let $k > 1$, and let $(\mathcal{U}_n)_{n\in\omega}$ be a balanced test. There is an $i < k$ such that $X^k_{\neq i} \notin \bigcap_n \mathcal{U}_n$.*

PROOF. Assume for a contradiction that $X^k_{\neq i} \in \bigcap_n \mathcal{U}_n$ for all $i < k$. Let

$$\mathcal{S}_{n,s} = \{Y : \forall i < k\ (Y^k_{\neq i} \in \mathcal{U}_n[s])\}$$

and let $\mathcal{S}_n = \bigcup_s \mathcal{S}_{n,s}$. By Theorem 7.3.14, $\mu(\mathcal{S}_{n,s})^{k-1} \leq \mu(\mathcal{U}_n[s])^k$, so $\mu(\mathcal{S}_n) \leq O(2^n)2^{-\frac{nk}{k-1}} = O(2^{-\frac{n}{k-1}})$, and hence $\sum_n \mu(\mathcal{S}_n) < \infty$. Thus $\{\mathcal{S}_n : n \in \omega\}$ is a Solovay test. However, $X \in \bigcap_n \mathcal{S}_n$, so we have a contradiction. ☐

THEOREM 7.3.16. *There is a K-trivial set A such that $A \notin X^{\mathfrak{c}}$ for all 1-random X.*

PROOF. Let A and $(\mathcal{U}_n)_{n\in\omega}$ be as in Theorem 7.3.13. Let X be 1-random. By Theorem 7.3.7, it is enough to fix $k > 1$ and show that there is an $i < k$ such that $A \nleq_{\mathrm{T}} X^k_{\neq i}$. Assume for a contradiction that $A \leq_{\mathrm{T}} X^k_{\neq i}$ for all $i < k$. Then there are $\sigma_0, \dots, \sigma_{k-1}$ such that $\sigma_i X^k_{\neq i} \in \bigcap_n \mathcal{U}_n$ for all $i < k$. Let $m = \max_{i<k} |\sigma_i|$ and let $\mathcal{V}_n = \{Y : \exists i < k\ (\sigma_i Y \in \mathcal{U}_{n+k+m})\}$. It is easy to check that $(\mathcal{V}_n)_{n\in\omega}$ is a balanced test, and $X^k_{\neq i} \in \bigcap_n \mathcal{V}_n$ for all $i < k$, which contradicts Lemma 7.3.15. ☐

7.4. Further applications of cone-avoiding compactness

We can use Theorem 7.3.7 to give an analogue to Corollary 7.3.3 for effective genericity. In this case, 1-genericity is sufficient, as it is straightforward to show that if X is 1-generic relative to A and A is non-computable, then $A \nleq_{\mathrm{T}} X$ (i.e., unlike the case for 1-randomness, there are no non-computable bases for 1-genericity), and that no 1-generic set can be coarsely computable. The other ingredient we need to replicate the argument we gave in the case of effective randomness is a version of van Lambalgen's Theorem for 1-genericity. This result was established by Yu [126, Proposition 2.2]. Relativising his theorem and applying induction as in the case of Theorem 7.3.5, we obtain the following fact.

THEOREM 7.4.1. (Yu [126]) *The following are equivalent for all sets X and A, and all $k > 1$.*

1. *X is 1-generic relative to A.*
2. *For each $i < k$, the set X^k_i is 1-generic relative to $X^k_{\neq i} \oplus A$.*

Now we can establish the following analogue to Corollary 7.3.3.

THEOREM 7.4.2. *If X is 1-generic then $X^{\mathfrak{c}} = \mathbf{0}$, and hence $\mathcal{E}(A) \nleq_{\mathrm{nc}} X$ for all non-computable A. In particular, in both the uniform and non-uniform coarse degrees, the degree of X is not in the image of the embedding induced by $\mathcal{E}$.*

Proof. Let $A \in X^{\mathfrak{c}}$. As in the proof of Theorem 7.3.2, there is a k such that $A \leq_{\mathrm{T}} X^k_{\neq i}$ for all $i < k$. By the unrelativised form of Theorem 7.4.1, each X^k_i is 1-generic relative to $X^k_{\neq i}$, and hence relative to $X^k_{\neq i} \oplus A \equiv_{\mathrm{T}} X^k_{\neq i}$. Again by Theorem 7.4.1, X is 1-generic relative to A. But $A \leq_{\mathrm{T}} X$, so A is computable. □

Igusa (personal communication) has also found the following application of Theorem 7.3.7. We say that X is *generically computable* if there is a partial computable function φ such that $\varphi(n) = X(n)$ for all n in the domain of φ, and the domain of φ has density 1. Jockusch and Schupp [15, Theorem 2.26] showed that there are generically computable sets that are not coarsely computable, but by Lemma 1.7 in [39], if X is generically computable then $\gamma(X) = 1$, where γ is the coarse computability bound from Definition 7.3.6.

Theorem 7.4.3. (Igusa, personal communication) *If $\gamma(X) = 1$ then $X^{\mathfrak{c}} = \mathbf{0}$, and hence $\mathcal{E}(A) \nleq_{\mathrm{nc}} X$ for all non-computable A. Thus, if $\gamma(X) = 1$ and X is not coarsely computable then in both the uniform and non-uniform coarse degrees, the degree of X is not in the image of the embedding induced by $\mathcal{E}$. In particular, the above holds when X is generically computable but not coarsely computable.*

Proof. Suppose that $\gamma(X) = 1$ and A is not computable. If $\varepsilon > 0$ then there is a computable set C such that $\overline{\rho}(X \triangle C) < \varepsilon$. Since C is computable, $A \nleq_{\mathrm{T}} C$. By Theorem 7.3.7, $A \notin X^{\mathfrak{c}}$. □

7.5. Minimal pairs in the uniform and non-uniform coarse degrees

For any degree structure that acts as a measure of information content, it is reasonable to expect that if two sets are sufficiently random relative to each other, then their degrees form a minimal pair. For the Turing degrees, it is not difficult to show that if Y is not computable and X is weakly 2-random relative to Y, then the degrees of X and Y form a minimal pair. On the other hand, Kučera [61] showed that if $X, Y \leq_{\mathrm{T}} \emptyset'$ are both 1-random, then there is a non-computable set $A \leq_{\mathrm{T}} X, Y$, so there are relatively 1-random sets whose degrees do not form a minimal pair. As we will see, the situation for the non-uniform coarse degrees is similar, but "one jump up".

For an interval I, let $\rho_I(X) = \frac{|X \cap I|}{|I|}$.

Lemma 7.5.1. *Let $J_k = [2^k - 1, 2^{k+1} - 1)$. Then $\rho(X) = 0$ if and only if $\lim_k \rho_{J_k}(X) = 0$.*

Proof. First suppose that $\limsup_k \rho_{J_k}(X) > 0$. Since $|J_k| = 2^k$, we have $\overline{\rho}(X) \geq \limsup_k \rho_{2^{k+1}-1}(X) \geq \limsup_k \frac{\rho_{J_k}(X)}{2} > 0$.

Now suppose that $\limsup_k \rho_{J_k}(X) = 0$. Fix $\varepsilon > 0$. If m is sufficiently large, $k \geq m$, and $n \in J_k$, then

$$|X \cap [0, n)| \leq |X \cap [0, 2^{k+1} - 1)| \leq \sum_{i=0}^{m-1} |J_i| + \sum_{i=m}^{k} \frac{\varepsilon}{2} |J_i|.$$

If k is sufficiently large then this sum is less than $\varepsilon(2^k - 1)$, whence $\rho_n(X) < \frac{\varepsilon(2^k-1)}{n} \leq \frac{\varepsilon n}{n} = \varepsilon$. Thus $\limsup_n \rho_n(X) \leq \varepsilon$. Since ε is arbitrary, $\limsup_n \rho_n(X) = 0$. □

THEOREM 7.5.2. *If A is not coarsely computable and X is weakly* 3*-random relative to A, then there is no X-computable coarse description of A. In particular, $A \not\leq_{\mathrm{nc}} X$.*

PROOF. Suppose that Φ^X is a coarse description of A and let

$$\mathcal{P} = \{Y : \Phi^Y \text{ is a coarse description of } A\}.$$

Then $Y \in \mathcal{P}$ if and only if

1. Φ^Y is total, which is a Π^0_2 property, and
2. for each k there is an m such that, for all $n > m$, we have $\rho_n(\Phi^Y \triangle A) < 2^{-k}$, which is a $\Pi^{0,A}_3$ property.

Thus $\mathcal{P}$ is a $\Pi^{0,A}_3$ class, so it suffices to show that if A is not coarsely computable then $\mu(\mathcal{P}) = 0$.

We prove the contrapositive. Suppose that $\mu(\mathcal{P}) > 0$. Then, by the Lebesgue Density Theorem, there is a σ such that $\mu(\mathcal{P} \cap [\![\sigma]\!]) > \frac{3}{4}2^{-|\sigma|}$. It is now easy to define a Turing functional Ψ such that the measure of the class of Y for which Ψ^Y is a coarse description of A is greater than $\frac{3}{4}$. Define a computable set D as follows. Let $J_k = [2^k - 1, 2^{k+1} - 1)$. For each k, wait until we find a finite set of strings S_k such that $\mu([\![S_k]\!]) > \frac{3}{4}$ and Ψ^σ converges on all of J_k for each $\sigma \in S_k$ (which must happen, by our choice of Ψ). Let n_k be largest such that there is a set $R_k \subseteq S_k$ with $\mu([\![R_k]\!]) > \frac{1}{2}$ and $\rho_{J_k}(\Psi^\sigma \triangle \Psi^\tau) \leq 2^{-n_k}$ for all $\sigma, \tau \in R_k$. Let $\sigma \in R_k$ and define $D \upharpoonright J_k = \Psi^\sigma \upharpoonright J_k$.

We claim that D is a coarse description of A. By Lemma 7.5.1, it is enough to show that $\lim_k \rho_{J_k}(D \triangle A) = 0$. Fix n. Let $\mathcal{B}_k$ be the class of all Y such that Ψ^Y converges on all of J_k and $\rho_{J_k}(\Psi^Y \triangle A) \leq 2^{-n}$. If Ψ^Y is a coarse description of A then, again by Lemma 7.5.1, $\rho_{J_k}(\Psi^Y \triangle A) \leq 2^{-n}$ for all sufficiently large k, so there is an m such that $\mu(\mathcal{B}_k) > \frac{3}{4}$ for each $k > m$, and hence $\mu(\mathcal{B}_k \cap [\![S_k]\!]) > \frac{1}{2}$ for each $k > m$. Let $T_k = \{\sigma \in S_k : \rho_{J_k}(\Psi^\sigma \triangle A) \leq 2^{-n}\}$. Then $[\![T_k]\!] = \mathcal{B}_k \cap [\![S_k]\!]$, so $\mu([\![T_k]\!]) > \frac{1}{2}$ for each $k > m$. Furthermore, by the triangle inequality for Hamming distance, $\rho_{J_k}(\Psi^\sigma \triangle \Psi^\tau) \leq 2^{-(n-1)}$ for all $\sigma, \tau \in T_k$. It follows that, for each $k > m$, we have $n_k \geq n - 1$, and at least one element Y of $\mathcal{B}_k$ is in $[\![R_k]\!]$ (where R_k is as in the definition of D), which implies that

$$\rho_{J_k}(D \triangle A) \leq \rho_{J_k}(D \triangle \Psi^Y) + \rho_{J_k}(\Psi^Y \triangle A) \leq 2^{-n_k} + 2^{-n} < 2^{-n+2}.$$

Since n is arbitrary, $\lim_k \rho_{J_k}(D \triangle A) = 0$. $\square$

COROLLARY 7.5.3. *If Y is not coarsely computable and X is weakly* 3*-random relative to Y, then the non-uniform coarse degrees of X and Y form a minimal pair, and hence so do their uniform coarse degrees.*

PROOF. Let $A \leq_{\mathrm{nc}} X, Y$. Then Y computes a coarse description D of A. We have $D \leq_{\mathrm{nc}} X$, and X is weakly 3-random relative to D, so by the theorem, D is coarsely computable, and hence so is A. $\square$

For the non-uniform coarse degrees at least, this corollary does not hold of 2-randomness in place of weak 3-randomness. To establish this fact, we use the following complementary results. The first was proved by Downey, Jockusch and

Schupp [24, Corollary 3.16] in unrelativised form, but it is easy to check that their proof relativises.

THEOREM 7.5.4. (Downey, Jockusch and Schupp [24]) *If A is c.e., $\rho(A)$ is defined, and $A' \leq_{\mathrm{T}} D'$, then D computes a coarse description of A.*

THEOREM 7.5.5. (Hirschfeldt, Jockusch, McNicholl and Schupp [39]) *Every nonlow c.e. degree contains a c.e. set A such that $\rho(A) = \frac{1}{2}$ and A is not coarsely computable.*

THEOREM 7.5.6. *Let $X, Y \leq_{\mathrm{T}} \emptyset''$ (which is equivalent to $\mathcal{E}(X), \mathcal{E}(Y) \leq_{\mathrm{nc}} \mathcal{E}(\emptyset'')$). If X and Y are both 2-random, then there is an $A \leq_{\mathrm{nc}} X, Y$ such that A is not coarsely computable. In particular, there is a pair of relatively 2-random sets whose non-uniform coarse degrees do not form a minimal pair.*

PROOF. Since X and Y are both 1-random relative to $\emptyset'$, by the relativised form of Corollary 7.3.11 there is an $\emptyset'$-c.e. set $J >_{\mathrm{T}} \emptyset'$ such that for every coarse description D of either X or Y, we have that $D \oplus \emptyset'$ computes J, and hence so does D'. By the Sacks Jump Inversion Theorem [102], there is a c.e. set B such that $B' \equiv_{\mathrm{T}} J$. By Theorem 7.5.5, there is a c.e. set $A \equiv_{\mathrm{T}} B$ such that $\rho(A) = \frac{1}{2}$ and A is not coarsely computable. Let D be a coarse description of either X or Y. Then $D' \geq_{\mathrm{T}} J \equiv_{\mathrm{T}} A'$, so by Theorem 7.5.4, D computes a coarse description of A. $\square$

We do not know whether this theorem holds for uniform coarse reducibility.

7.6. Open questions

We finish with a few questions raised by our results.

OPEN QUESTION 7.6.1. Can the bound $\frac{1}{4}$ in Corollary 7.3.10 be increased?

OPEN QUESTION 7.6.2. Let $X \leq_{\mathrm{T}} \emptyset'$ be 1-random. Must there be a noncomputable (c.e.) set A such that $\mathcal{E}(A) \leq_{\mathrm{uc}} X$? (Recall that Corollary 7.3.10 gives a positive answer to the non-uniform analogue to this question.) If not, then is there any 1-random X for which such an A exists?

OPEN QUESTION 7.6.3. Does Theorem 7.5.6 hold for uniform coarse reducibility?

Part III

ε-Logic

CHAPTER 8

ε-Logic

As discussed in the introduction, regular first-order logic is not *learnable*, in the sense that we cannot decide if a given formula φ holds in a model or not if we are only allowed to take finitely many samples. In this part of the thesis we will study ε-logic, a probability logic introduced by Terwijn [117] which turns out to be learnable in a sense closely related to Valiant's pac-model from computational learning theory.

In this first introductory chapter we will introduce the background of ε-logic and we give all the necessary definitions. Furthermore, we discuss some of the choices we make. This chapter is mostly based on Kuyper and Terwijn [68].

8.1. ε-Logic

In this section, we will repeat the definition of the probabilistic logic from Terwijn [117]. This logic was partly motivated by the idea of what it means to "learn" an ordinary first-order statement φ from a finite amount of data from a model $\mathcal{M}$ of φ, in a way that is similar to learning in Valiant's pac-model [49]. In this setting, atomic data are generated by sampling from an unknown probability distribution $\mathcal{D}$ over $\mathcal{M}$, and the task is to decide with a prescribed amount of certainty whether φ holds in $\mathcal{M}$ or not. On seeing an atomic truth $R(a)$, where R is some relation, one knows with certainty that $\exists x R(x)$, so that the existential quantifier retains its classical interpretation. On the other hand, inducing a universal statement $\forall x R(x)$ can only be done probabilistically. Thus, there is a fundamental asymmetry between the interpretation of the existential quantifier and the interpretation of the universal quantifier. As in the pac-model, it is important that the distribution $\mathcal{D}$ is unknown, which is counterbalanced by the fact that success of the learning task is measured using the same distribution $\mathcal{D}$. (In the pac-setting this is called "distribution-free learning".) In [117] it was shown that ordinary first-order formulas are pac-learnable under the appropriate probabilistic interpretation, given in the definition below. In this thesis the focus will be on the logic, and no background on pac-learning is required any further.

Definition 8.1.1. Let $\mathcal{L}$ be a first-order language, possibly containing equality, of a countable signature. Let $\varphi = \varphi(x_1, \ldots, x_n)$ be a first-order formula in the language $\mathcal{L}$, and let $\varepsilon \in [0, 1]$. Furthermore, let $\mathcal{M}$ be a classical first-order model for $\mathcal{M}$ and let $\mathcal{D}$ be a probability measure on the universe of $\mathcal{M}$. Then we inductively define the notion of *ε-truth*, denoted by $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$, as follows (where we leave the parameters implicit).

(i) For every atomic formula φ:

$$(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi \text{ if } \mathcal{M} \models \varphi.$$

(ii) We treat the logical connectives $\wedge$ and $\vee$ classically, e.g.

$$(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi \wedge \psi \text{ if } (\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi \text{ and } (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi.$$

(iii) The existential quantifier is treated classically as well:

$$(\mathcal{M},\mathcal{D}) \models_\varepsilon \exists x\varphi(x)$$

if there exists an $a \in \mathcal{M}$ such that $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a)$.

(iv) The case of negation is split into sub-cases as follows:
 (a) For φ atomic, $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\varphi$ if $(\mathcal{M},\mathcal{D}) \not\models_\varepsilon \varphi$.
 (b) $\neg$ distributes in the classical way over $\wedge$ and $\vee$, e.g.

$$(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg(\varphi \wedge \psi) \text{ if } (\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\varphi \vee \neg\psi.$$

 (c) $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\neg\varphi$ if $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi$.
 (d) $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg(\varphi \to \psi)$ if $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi \wedge \neg\psi$.
 (e) $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\exists x\varphi(x)$ if $(\mathcal{M},\mathcal{D}) \models_\varepsilon \forall x\neg\varphi(x)$.
 (f) $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\forall x\varphi(x)$ if $(\mathcal{M},\mathcal{D}) \models_\varepsilon \exists x\neg\varphi(x)$.

(v) $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi \to \psi$ if $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\varphi \vee \psi$.

(vi) Finally, we define $(\mathcal{M},\mathcal{D}) \models_\varepsilon \forall x\varphi(x)$ if

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a)\big] \geq 1 - \varepsilon.$$

We make some remarks about the definition of ε-truth. Observe that everything in Definition 8.1.1 is treated classically, except for the interpretation of $\forall x\varphi(x)$ in case (vi). Instead of saying that we have $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a)$ for *all* elements $a \in \mathcal{M}$, we merely say that it holds for "many" of the elements, where "many" depends on the error parameter ε. The treatment of negation requires some care, since we no longer have that $(\mathcal{M},\mathcal{D}) \models_\varepsilon \neg\varphi$ implies that $(\mathcal{M},\mathcal{D}) \not\models_\varepsilon \varphi$ (though the converse still holds, see Terwijn [117, Proposition 3.1]). The clauses for the negation allow us to push the negations down to the atomic formulas.

Note that both $(\mathcal{M},\mathcal{D}) \models_\varepsilon \forall x\varphi(x)$ and $(\mathcal{M},\mathcal{D}) \models_\varepsilon \exists x\neg\varphi(x)$ can hold, for example if $\varphi(x)$ holds on a set of measure one but not for all x. Thus, the logic defined above is *paraconsistent*. This example also shows that for $\varepsilon = 0$ the notion of ε-truth does not coincide with the classical one. Note also that even though both φ and $\neg\varphi$ may be satisfiable, they cannot both be ε-tautologies, as at most one of them can be true in a model with only one point.

We have chosen to define $\varphi \to \psi$ as $\neg\varphi \vee \psi$. We note that this is weaker than the classical implication. The classical definition would say that ψ holds in any model where φ holds. Using an atomic inconsistency as falsum, we would thus obtain a classical negation. Since $\exists$ expresses classical existence, we would then also obtain the classical universal quantifier $\forall$, and our logic would become a strong extension of classical predicate logic, which is not what we are after.

The case for $\varepsilon = 1$ is pathological; for example, all universal statements are always true. We will therefore often exclude this case.

DEFINITION 8.1.2. Let $\mathcal{L}$ be a first-order language of a countable signature, possibly containing equality, and let $\varepsilon \in [0,1]$. Then an *ε-model* $(\mathcal{M},\mathcal{D})$ for the language $\mathcal{L}$ consists of a classical first-order $\mathcal{L}$-model $\mathcal{M}$ together with a probability distribution $\mathcal{D}$ over $\mathcal{M}$ such that:

(15) For all formulas $\varphi = \varphi(x_1,\dots,x_n)$ and all $a_1,\dots,a_{n-1} \in \mathcal{M}$, the set

$$\{a_n \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a_1,\dots,a_n)\}$$

is $\mathcal{D}$-measurable (i.e. all definable sets of dimension 1 are measurable).

(16) All relations of arity n are $\mathcal{D}^n$-measurable (including equality, if it is in $\mathcal{L}$) and all functions of arity n are measurable as functions from $(\mathcal{M}^n,\mathcal{D}^n)$ to $(\mathcal{M},\mathcal{D})$ (where $\mathcal{D}^n$ denotes the n-fold product measure). In particular, constants are $\mathcal{D}$-measurable.

A *probability model* is a pair $(\mathcal{M},\mathcal{D})$ that is an ε-model for every $\varepsilon \in [0,1]$.

We remark that condition (16) does not imply condition (15), because even if a set is measurable, its image under a projection need not be measurable. Nevertheless, the following result holds.

PROPOSITION 8.1.3. *Let $\varphi(x_1,\dots,x_n)$ be a universal formula and let $(\mathcal{M},\mathcal{D})$ be an ε-model. Then*

$$\{(a_1,\dots,a_n) \in \mathcal{M}^n \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a_1,\dots,a_n)\}$$

is $\mathcal{D}^n$-measurable.

PROOF. First, one can use induction to prove that the lemma holds for propositional formulas ψ; the base case is exactly (16). Next, let φ be a universal formula. By Proposition 8.1.7 below we may assume φ to be in prenex normal form; say $\varphi = \forall y_1 \dots \forall y_m \psi(x_1,\dots,x_n,y_1,\dots,y_m)$. We have just argued that the set

$$\{(a_1,\dots,a_n,b_1,\dots,b_m) \in \mathcal{M}^{n+m} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(a_1,\dots,a_n,b_1,\dots,b_m)\}$$

is $\mathcal{D}^{n+m}$-measurable. The result now follows from repeatedly applying the fact that for any $1 \le i \le m$ and any $\mathcal{D}^{n+m-i+1}$-measurable set X the function

$$(a_1,\dots,a_n,b_1,\dots,b_{m-i}) \mapsto \Pr_{\mathcal{D}}\big[b_{m-i+1} \in \mathcal{M} \mid (a_1,\dots,a_n,b_1,\dots,b_{m-i+1}) \in X\big]$$

is a $\mathcal{D}^{n+m-i}$-measurable function, see e.g. Bogachev [10, Theorem 3.4.1]. $\square$

The definition of ε-model is discussed in more detail in section 8.2.

DEFINITION 8.1.4. A formula $\varphi(x_1,\dots,x_n)$ is *ε-satisfiable* if there exists an ε-model $(\mathcal{M},\mathcal{D})$ and there exist $a_1,\dots,a_n \in \mathcal{M}$ such that $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi$. Similarly, we say that φ is an *ε-tautology* or is *ε-valid* (notation: $\models_\varepsilon \varphi$) if for all probability models $(\mathcal{M},\mathcal{D})$ and all $a_1,\dots,a_n \in \mathcal{M}$ it holds that $(\mathcal{M},\mathcal{D}) \models_\varepsilon \varphi(a_1,\dots,a_n)$.

EXAMPLE 8.1.5. Let Q be a unary predicate. Then $\varphi = \forall x Q(x) \vee \forall x \neg Q(x)$ is a $\frac{1}{2}$-tautology. Namely, in every probability model, either the set on which Q holds or its complement has measure at least $\frac{1}{2}$. However, φ is not an ε-tautology for $\varepsilon < \frac{1}{2}$. Furthermore, both φ and $\neg\varphi$ are classically satisfiable and hence ε-satisfiable for every ε; in particular we see that φ can be an ε-tautology while simultaneously $\neg\varphi$ is ε-satisfiable.

One might wonder why for satisfiability we only require ε-models, while for validity we look at the slightly stronger and less elegant probability models. From Proposition 9.2.1 and Theorem 9.2.9 below it will follow that every formula which is ε-satisfiable in an ε-model is also satisfiable in a probability model, while we do not know of a similar result for validity.

EXAMPLE 8.1.6. Let Q be a unary predicate. Then $\varphi = \forall x Q(x) \vee \forall x \neg Q(x)$ is a $\frac{1}{2}$-tautology. Namely, in every $\frac{1}{2}$-model, either the set on which Q holds or its complement has measure at least $\frac{1}{2}$. However, φ is not an ε-tautology for $\varepsilon < \frac{1}{2}$. Furthermore, both φ and $\neg\varphi$ are classically satisfiable and hence ε-satisfiable for every ε; in particular we see that φ can be an ε-tautology while simultaneously $\neg\varphi$ is ε-satisfiable.

In many proofs it will be convenient to work with formulas in prenex normal form. We may assume that formulas are in this form by the following:

PROPOSITION 8.1.7. Terwijn [117] *Every formula φ is semantically equivalent to a formula φ' in prenex normal form; i.e. $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \Leftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi'$ for all $\varepsilon \in [0,1]$ and all ε-models $(\mathcal{M}, \mathcal{D})$.*

8.2. ε-Models

Our definition of ε-model slightly differs from the original definition in [117]. We require more sets to be measurable in our ε-models than in the original definition, where the measurability condition was included in the truth definition. However, we need this stronger requirement on our models to be able to prove anything worthwhile, in fact, (15) in Definition 8.1.2 is already implicit in most proofs published in earlier papers.

We now discuss condition (16) in Definition 8.1.2. This is a natural assumption: When we are talking about probabilities over certain predicates we may as well require that all such probabilities exist, even if in some cases this would not be necessary. To illustrate this point we give an example of what can happen without it.

EXAMPLE 8.2.1. The following example is based on the famous argument of Sierpinski showing that under the continuum hypothesis CH there are unmeasurable subsets of the real plane. Let $\mathcal{D}$ be a measure on the domain ω_1 defined by

$$\mathcal{D}(A) = \begin{cases} 1 & \text{if } A = \omega_1 \text{ with the exception of at most} \\ & \text{countably many elements,} \\ 0 & \text{if A is countable.} \end{cases}$$

It is easy to check that $\mathcal{D}$ is a probability measure. Let $<$ be the usual order relation on ω_1. Then we have

$$(\omega_1, \mathcal{D}) \models_0 \forall x \forall y (x < y)$$

since for every $x \in \omega_1$ the vertical section $\{y \mid x < y\}$ has $\mathcal{D}$-measure 1. Similarly,

$$(\omega_1, \mathcal{D}) \not\models_0 \forall y \forall x (x < y)$$

since for every $y \in \omega_1$ the horizontal section $\{x \mid x < y\}$ has $\mathcal{D}$-measure 0. Note that the relation $\{(x, y) \in {\omega_1}^2 \mid x < y\}$ is not $\mathcal{D}^2$-measurable: Since all its vertical sections $\{y \mid x < y\}$ have $\mathcal{D}$-measure 1, and all its horizontal sections $\{x \mid x < y\}$ have $\mathcal{D}$-measure 0, $\mathcal{D}^2$-measurability of the relation $<$ would contradict Fubini's theorem. That in general universal quantifiers do not commute under a probabilistic interpretation was already remarked in Keisler [51]. In fact, it is easy to give a three-element example of a model $\mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_{\frac{1}{3}} \forall x \forall y R(x, y)$ but not $(\mathcal{M}, \mathcal{D}) \models_{\frac{1}{3}} \forall y \forall x R(x, y)$. So in this respect condition (16) does not help anyway.

We point out that the choice to impose condition (16) or not *does* make a difference for the resulting probability logic: Let

$$\mathsf{lin} = \forall x \forall y (x \le y \vee y \le x) \wedge \forall x \forall y \forall z (x \le y \wedge y \le z \to x \le z)$$

be the sentence saying that $\le$ is a total preorder (not necessarily antisymmetric).

PROPOSITION 8.2.2. *For every $\varepsilon > 0$, the sentence*

$$\varphi = \neg \mathsf{lin} \vee \exists x \forall y (y \le x)$$

is an ε-tautology if and only if we impose condition (16) *in Definition 8.1.2.*

PROOF. Let $\varepsilon > 0$. Without (16) we can construct a counter-model for φ as follows. Consider ω_1 with the measure $\mathcal{D}$ from Example 8.2.1. Then since every initial segment of ω_1 has measure 0, clearly $(\omega_1, \mathcal{D}) \not\models_\varepsilon \varphi$.

Now suppose (16) holds. Let $(\mathcal{M}, \mathcal{D})$ be an ε-model with $\varepsilon > 0$. We show that φ is an ε-tautology. When $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \neg\mathsf{lin}$ then we are done. When $(\mathcal{M}, \mathcal{D}) \not\models_\varepsilon \neg\mathsf{lin}$ then we have classically $\mathcal{M} \models \mathsf{lin}$ so $\le$ really is a linear preorder in $\mathcal{M}$. Suppose that

$$\forall x \ \Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid x \le y\big] > \varepsilon. \tag{17}$$

Suppose further that

$$\forall x_0 \exists y \ge x_0 \Pr_{\mathcal{D}}\big[[x_0, y]\big] \ge \frac{1}{2}\varepsilon$$

where $[x_0, y]$ is the interval between x_0 and y in $\mathcal{M}$. We may assume that for every $y \in \mathcal{M}$ there exists a $z \in \mathcal{M}$ with $z \not\le y$; otherwise y is clearly a maximal element and we are done. Then we also have

$$\forall x_0 \exists y > x_0 \Pr_{\mathcal{D}}\big[[x_0, y]\big] \ge \frac{1}{2}\varepsilon$$

(where $y > x_0$ denotes $x_0 \le y \wedge y \not\le x_0$). But then we can find infinitely many intervals $[y_0, y_1], [y_1, y_2], \ldots$ with $y_i < y_{i+1}$ of measure at least $\frac{1}{2}\varepsilon$, which are disjoint by the transitivity of $\le$. This is a contradiction. So, choose x_0 such that $\forall y \ge x_0 \ \Pr_{\mathcal{D}}\big[[x_0, y]\big] < \frac{1}{2}\varepsilon$ and consider the set

$$\big\{(x, y) \in \mathcal{M} \times \mathcal{M} \mid x, y \ge x_0 \wedge x \le y\big\}, \tag{18}$$

i.e. the restriction of the relation $\le$ to elements greater than x_0. Then, similarly as in Example 8.2.1, all vertical sections of (18) have measure $> \varepsilon$ and all horizontal sections have measure $< \frac{1}{2}\varepsilon$, so by Fubini's theorem the set (18) is not $\mathcal{D}^2$-measurable. But then, since (18) is the intersection of sets defined using $\le$, the relation $\le$ itself is not measurable, contradicting (16). So (17) is false, and hence there is an $x \in \mathcal{M}$ such that at least $1 - \varepsilon$ of the weight is to the left of x. Hence $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$. $\square$

After these considerations we take the standpoint that it is natural to assume the measurability condition (16) in Definition 8.1.2. As we will see below, for the discussion of compactness it is useful to consider a weaker notion of ε-model, where we drop the condition (16) from the definition:

DEFINITION 8.2.3. If $\mathcal{M}$ is a first-order model and $\mathcal{D}$ is a probability measure on $\mathcal{M}$ such that for all formulas $\varphi = \varphi(x_1, \ldots, x_n)$ and all $a_1, \ldots, a_{n-1} \in \mathcal{M}$, the set

$$\{a_n \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \ldots, a_n)\}$$

is $\mathcal{D}$-measurable, we say that $(\mathcal{M}, \mathcal{D})$ is a *weak ε-model*.

CHAPTER 9

Model Theory of ε-Logic

In this chapter we will study the model theory of ε-logic. First, in section 9.1 we show that there are satisfiable sentences that are not satisfiable in any countable model, but we also prove a downward Löwenheim–Skolem theorem for ε-logic in which "countable" is replaced by "of cardinality of the continuum". In section 9.2 we refine this result and show that every satisfiable sentence is in fact satisfied by a model on the unit interval together with the Lebesgue measure. Section 9.3 continues with the problem discussed in the preceding section, by discussing what the exact value of the Löwenheim number (the smallest cardinality for which every satisfiable sentence has a model of that cardinality) is in ε-logic. Next, in section 9.4 we present a technical result that gives many-one reductions between satisfiability in ε_0-logic and ε_1-logic for different $\varepsilon_0, \varepsilon_1$. Finally, in section 9.5 we show that compactness fails for ε-logic, but that we can recover a weak notion of compactness using an ultraproduct construction.

This chapter is based on Kuyper and Terwijn [68].

9.1. A downward Löwenheim–Skolem theorem

In this section, we will prove a downward Löwenheim–Skolem theorem for ε-logic. We will see that it is not always possible to push the cardinality of a model down to being countable, as in classical logic. In many ways, countable ε-models are analogous to *finite* classical models, as exemplified by the following result:

THEOREM 9.1.1. (Terwijn [119]) *Let φ be a sentence. Then the following are equivalent:*

(i) *φ classically holds in all finite classical models,*
(ii) *$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$ for all $\varepsilon > 0$ and for all ε-models $(\mathcal{M}, \mathcal{D})$ such that $\mathcal{M}$ is countable.*

DEFINITION 9.1.2. We will call a measure ν on a σ-algebra $\mathcal{B}$ of subsets of N a *submeasure* of a measure μ on a σ-algebra $\mathcal{A}$ of subsets of some set $M \supseteq N$ if for every $B \in \mathcal{B}$ there exists an $A_B \in \mathcal{A}$ such that $B = A_B \cap N$ and $\mu(A_B) = \nu(B)$.

To motivate this definition, let us first consider the special case where $A_B = B$ for every $B \in \mathcal{B}$. In this case, we have that $\nu(B) = \mu(B)$ for every $B \in \mathcal{B}$. In other words, ν is just the restriction of the measure ν to the μ-measurable set N. However, requiring N to be a measurable subset of M is too restrictive for the constructions below. Our definition of submeasure also allows us to restrict μ to certain non-μ-measurable sets N, by allowing us some freedom in the choice of the set A_B. The precise method of constructing such submeasures $\mathcal{N}$ will become clear in the proof of Theorem 9.1.6.

DEFINITION 9.1.3. An *ε-submodel* of an ε-model $(\mathcal{M}, \mathcal{D})$ is an ε-model $(\mathcal{N}, \mathcal{E})$ over the same language such that:

- $\mathcal{N}$ is a submodel of $\mathcal{M}$ in the classical sense,
- $\mathcal{E}$ is a submeasure of $\mathcal{D}$.

We will denote this by $(\mathcal{N}, \mathcal{E}) \subset_\varepsilon (\mathcal{M}, \mathcal{D})$.

DEFINITION 9.1.4. An *elementary ε-submodel* of an ε-model $(\mathcal{M}, \mathcal{D})$ is an ε-submodel $(\mathcal{N}, \mathcal{E})$ such that, for all formulas $\varphi = \varphi(x_1, \dots, x_n)$ and sequences $a_1, \dots, a_n \in \mathcal{N}$ we have:

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(a_1, \dots, a_n) \Leftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \dots, a_n).$$

We will denote this by $(\mathcal{N}, \mathcal{E}) \prec_\varepsilon (\mathcal{M}, \mathcal{D})$.

The next example shows that there are satisfiable sentences without any countable model.

EXAMPLE 9.1.5. Let $\varphi = \forall x \forall y (R(x, y) \wedge \neg R(x, x))$. Then φ is 0-satisfiable; for example, take the unit interval $[0, 1]$ equipped with the Lebesgue measure and take $R(x, y)$ to be $x \neq y$. However, φ does not have any countable 0-models. Namely, if $(\mathcal{M}, \mathcal{D}) \models_0 \varphi$ then for almost every $x \in \mathcal{M}$ the set

$$B_x = \{y \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_0 \neg R(x, y) \vee R(x, x)\}$$

has measure zero. Since $x \in B_x$, the set $\bigcup_{x \in \mathcal{M}} B_x$ equals $\mathcal{M}$, and therefore has measure 1. But if $\mathcal{M}$ is countable it is also the union of countable many sets of measure 0 and hence has measure 0, a contradiction.

Note also that φ is *finitely* ε-satisfiable (i.e. ε-satisfiable with a finite model) for every $\varepsilon > 0$.

Using the reduction from Theorem 9.4.1, we now also find for every rational $\varepsilon \in [0, 1)$ a sentence φ_ε which is only ε-satisfiable in uncountable models.

Example 9.1.5 shows that we cannot always find countable elementary submodels. However, we can find such submodels of cardinality 2^ω, as we will show next.

THEOREM 9.1.6. (Downward Löwenheim–Skolem theorem for ε-logic) *Let $\mathcal{L}$ be a countable first-order language, possibly containing equality, but not containing function symbols. Let $(\mathcal{M}, \mathcal{D})$ be an ε-model and let $X \subseteq \mathcal{M}$ be of cardinality at most 2^ω. Then there exists an ε-model*

$$(\mathcal{N}, \mathcal{E}) \prec_\varepsilon (\mathcal{M}, \mathcal{D})$$

such that $X \subseteq \mathcal{N}$ and $\mathcal{N}$ is of cardinality at most 2^ω.

PROOF. We start by fixing some model-theoretic notation. For basics about types we refer the reader to Hodges [41]. For an element $x \in \mathcal{M}$, let $\mathrm{tp}(x/\mathcal{M})$ denote the complete 1-type of x over $\mathcal{M}$, i.e. the set of all formulas $\varphi(z)$ in one free variable and with parameters from $\mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(x)$. Clearly the relation $\mathrm{tp}(x/\mathcal{M}) = \mathrm{tp}(y/\mathcal{M})$ defines an equivalence relation on $\mathcal{M}$. The idea is to construct $\mathcal{N}$ by picking one element from every equivalence class of a finer equivalence relation. We will show that we can do this in such a way that we need at most 2^ω many points. Furthermore, this finer equivalence relation induces a natural submeasure $\mathcal{E}$ of $\mathcal{D}$ on $\mathcal{N}$, which will turn $(\mathcal{N}, \mathcal{E})$ into the desired ε-submodel.

Let $R = R(x_1, \ldots, x_n)$ be a relation. By Definition 8.1.2 we have that $R^{\mathcal{M}}$ is a $\mathcal{D}^n$-measurable set. We can view the construction of the product σ-algebra on $\mathcal{M}^n$ as an inductive process over the countable ordinals: we start with n-fold products or boxes of $\mathcal{D}$-measurable edges, for successor ordinals $\alpha + 1$ we take countable unions and intersections of elements from α (we could also take complements, but this is not necessary) and for limit ordinals λ we take the union of sets constructed in steps $< \lambda$. Now the product σ-algebra is exactly the union of all the sets constructed in this way.

In particular, we see from this construction that $R^{\mathcal{M}}$ can be formed using countable unions and intersections of Cartesian products of at most countably many $\mathcal{D}$-measurable sets. This expression need not be unique — so, for each relation R, pick one such expression t and form the set Δ_R consisting of the $\mathcal{D}$-measurable sets occurring as edges of Cartesian products in this expression. Let Δ be $\bigcup_R \Delta_R$ (where the union includes equality, if it is in the language) together with $\{c^{\mathcal{M}}\}$ for every constant c.

Since Δ is countable, we can fix an enumeration $B_0, B_1, \ldots$ of it. For each $a \in 2^\omega$ define

$$E_a = \bigcap_{a_i = 1} B_i \cap \bigcap_{a_i = 0} (\mathcal{M} \setminus B_i).$$

Then we can check that points in E_a are equivalent, in the sense that for all $x, y \in E_a$ we have that $\mathrm{tp}(x/\mathcal{M}) = \mathrm{tp}(y/\mathcal{M})$. Namely, first we check that for any n-ary relation R and $z_1, \ldots, z_{n-1} \in \mathcal{M}$, $R^{\mathcal{M}}(x, z_1, \ldots, z_{n-1}) \Leftrightarrow R^{\mathcal{M}}(y, z_1, \ldots, z_{n-1})$. This follows by induction on the construction of $R^{\mathcal{M}}$ from Δ. The equivalence for arbitrary formulas from the 1-types then follows by induction over formulas in prenex normal form.

From each non-empty E_a, pick one point x_a, and define

$$\mathcal{N} = X \cup \{x_a \mid a \in 2^\omega\}.$$

Clearly, $\mathcal{N}$ has cardinality at most 2^ω. Finally, for each $\mathcal{D}$-measurable A such that

$$\forall a \in 2^\omega \forall x, y \in E_a (x \in A \Leftrightarrow y \in A), \tag{19}$$

we define $\mathcal{E}(A \cap \mathcal{N}) = \mathcal{D}(A)$. We claim that $(\mathcal{N}, \mathcal{E})$ (with relations restricted to $\mathcal{N}$) satisfies the required properties.

First, observe that $\mathcal{E}$ is well-defined. Namely, let $A \neq C$ be $\mathcal{D}$-measurable sets satisfying (19), say $x \in A$ and $x \notin C$. Let $a \in 2^\omega$ be such that $x \in E_a$. Then $x_a \in A$, but $x_a \notin C$. So $A \cap \mathcal{N} \neq C \cap \mathcal{N}$. Also, $\mathcal{E}$ is a probability measure since $\mathcal{D}$ is.

Next, we prove that $(\mathcal{N}, \mathcal{E}) \prec_\varepsilon (\mathcal{M}, \mathcal{D})$. It is clear that $\mathcal{N}$ is a submodel of $\mathcal{M}$ and that $\mathcal{E}$ is a submeasure of $\mathcal{D}$. We prove that $(\mathcal{N}, \mathcal{E})$ is an elementary ε-submodel. We use formula-induction on formulas in prenex normal form to show that, for all sequences $b_1, \ldots, b_n \in \mathcal{N}$ and for every formula $\varphi = \varphi(x_1, \ldots, x_n)$, we have

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(b_1, \dots, b_n) \Leftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(b_1, \dots, b_n).$$

For propositional formulas, this is clear. For the existential case, observe that

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \exists x \psi(x, b_1, \dots, b_n)$$

clearly implies that this also holds in $(\mathcal{M}, \mathcal{D})$. For the converse, assume

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \exists x \psi(x, b_1, \dots, b_n).$$

Let $x \in \mathcal{M}$ be such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x, b_1, \dots, b_n)$, and let $a \in 2^\omega$ be such that $x \in E_a$. Then, as explained above, x and x_a are equivalent, so we also have $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x_a, b_1, \dots, b_n)$. Using the induction hypothesis, we therefore find $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \psi(x_a, b_1, \dots, b_n)$. Since $x_a \in \mathcal{N}$ this implies that $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \exists x \psi(x, b_1, \dots, b_n)$.

For the universal case, let $\varphi = \forall x \psi(x, x_1, \dots, x_n)$. Let

$$B = \{x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x, b_1, \dots, b_n)\},$$
$$C = \{x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \psi(x, b_1, \dots, b_n)\}.$$

Then by induction hypothesis we have

$$\begin{aligned} C &= \{x \in \mathcal{N} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x, b_1, \dots, b_n)\} \\ &= B \cap \mathcal{N}. \end{aligned}$$

From this and the fact that B satisfies (19), we see that $\mathcal{E}(C) = \mathcal{D}(B)$, and hence

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \psi(x, b_1, \dots, b_n) \Leftrightarrow (\mathcal{N}, \mathcal{E}) \models_\varepsilon \forall x \psi(x, b_1, \dots, b_n).$$

This concludes the induction.

It remains to check that $(\mathcal{N}, \mathcal{E})$ is an ε-model (see Definition 8.1.2). For every formula $\varphi = \varphi(x_1, \dots, x_n)$ and every sequence $a_1, \dots, a_{n-1} \in \mathcal{N}$ we have

$$\begin{aligned} B_\varphi &:= \{a_n \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(a_1, \dots, a_n)\} \\ &= \{a_n \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \dots, a_n)\} \cap \mathcal{N} \end{aligned}$$

and since the right-hand side is the intersection of $\mathcal{N}$ and a $\mathcal{D}$-measurable set satisfying (19), it follows that B_φ is $\mathcal{E}$-measurable. That relations in $\mathcal{N}$ are measurable follows directly from the construction; for constants c use the fact that $\{c\} \in \Delta$ and therefore there exists an $a \in 2^\omega$ such that $E_a = \{c\}$.

Thus, we see that $(\mathcal{N}, \mathcal{E})$ is an elementary ε-submodel of $(\mathcal{M}, \mathcal{D})$. □

Remark 9.1.7. The proof given above uses the full measurability condition (16) from Definition 8.1.2. We remark that one can also prove the theorem without using that the relations are measurable, by following the proof of Keisler [51, Theorem 2.4.4]; in that case the language is also allowed to contain function symbols. However, we need the proof above to be able to derive Theorem 9.2.8 below.

By varying ε, we can easily see that in fact the following strengthening of Theorem 9.1.6 holds.

Theorem 9.1.8. (Downward Löwenheim–Skolem theorem for variable ε) *Let $\mathcal{L}$ be a first-order language as above. Let $A \subseteq [0,1]$, let $(\mathcal{M}, \mathcal{D})$ be an ε-model for all $\varepsilon \in A$ and let $X \subseteq M$ be of cardinality at most 2^ω. Then there exists an ε-model $(\mathcal{N}, \mathcal{E})$ such that $(\mathcal{N}, \mathcal{E}) \prec_\varepsilon (\mathcal{M}, \mathcal{D})$ for all $\varepsilon \in A$, such that $X \subseteq \mathcal{N}$ and such that $\mathcal{N}$ is of cardinality at most 2^ω.*

Proof. This can be shown using the same proof as for Theorem 9.1.6. □

9.2. Satisfiability and Lebesgue measure

The construction from the proof of Theorem 9.1.6 produces an unknown probability measure on 2^ω. However, we can say a bit more about the σ-algebra of measurable sets of $\mathcal{E}$ in that proof: for example, that it is countably generated. We will use this and other facts to show that every ε-satisfiable set Γ of sentences has an ε-model on $[0,1]$ equipped with the Lebesgue measure. This model need not be equivalent to the original model satisfying Γ; the new model will in general satisfy more sentences.[1]

We cannot directly show that the measure space of the given model is isomorphic to $[0,1]$ with the Lebesgue measure — we need to make some modifications to the model first. As a first step, we show that each set ε-satisfiable set Γ of sentences is satisfied in some Borel measure on the Cantor set 2^ω (with the usual topology). For this, we need the following auxiliary result.

Proposition 9.2.1. *Let $\mathcal{M}$ be a first-order model that is a Polish space, and let $\mathcal{D}_0$ be a Borel probability measure on $\mathcal{M}$ such that all relations and functions are $\mathcal{D}_0^n$-measurable. Then all definable sets are analytic. In particular, if we let $\mathcal{D}$ be the completion of $\mathcal{D}_0$, then $(\mathcal{M}, \mathcal{D})$ is an ε-model for every $\varepsilon \in [0,1]$.*

Proof. Since every relation is $\mathcal{D}_0^n$-measurable, it is in particular Borel and therefore analytic. We now verify that every definable set is analytic, using induction over the number of quantifiers of a formula in prenex normal form (see Proposition 8.1.7). Clearly, this holds for propositional formulas. For the existential quantifier, use that projections of analytic sets are analytic (which is clear from the definition of an analytic set, see e.g. Kechris [50, Definition 14.1]), and for the universal quantifier, this fact is expressed by the Kondô-Tugué theorem [60] (see Kechris [50, Theorem 29.26]).

Since all definable sets are analytic, in particular the definable sets of dimension 1 are analytic and hence $\mathcal{D}$-measurable (see e.g. Bogachev [10, Theorem 1.10.5]). This proves the second claim (see Definition 8.1.2). □

Proposition 9.2.2. *Let $\mathcal{L}$ be a countable first-order language not containing function symbols. Let Γ be an ε-satisfiable set of sentences. Then there exists an ε-model $(\mathcal{M}, \mathcal{D})$ on 2^ω which ε-satisfies Γ such that $\mathcal{D}$ is the completion of a Borel measure. Furthermore, the relations in $\mathcal{M}$ can be chosen to be Borel.*

[1]This can happen even if Γ is already complete, i.e. if for every sentence φ at least one of $\varphi \in \Gamma$ and $\neg\varphi \in \Gamma$ holds: because of the paraconsistency of the logic, it could happen that both φ and $\neg\varphi$ hold in our new model, while only one of them is in Γ.

Proof. Fix an ε-model ε-satisfying all sentences from Γ and apply Theorem 9.1.6 (with $X = \emptyset$) to find a model $(\mathcal{N}, \mathcal{E})$. Let $\Delta = \{B_0, B_1, \dots\}$ and E_a be as in the proof of Theorem 9.1.6, that is,

$$E_a = \bigcap_{a_i=1} B_i \cap \bigcap_{a_i=0} (X \setminus B_i).$$

By construction of $\mathcal{N}$, each E_a contains at most one point of $\mathcal{N}$, namely x_a. So, the function $\pi : \mathcal{N} \to 2^\omega$ mapping each $x_a \in \mathcal{N}$ to a is injective.[2]

Now, define the subsets $C_n \subseteq 2^\omega$ by

$$C_n = \{a \in 2^\omega \mid a_n = 1\}.$$

Then $\{C_n \mid n \in \omega\}$ generate the Borel σ-algebra of 2^ω and we have $\pi^{-1}(C_n) = B_n$. Thus, C_n can be seen as an enlargement of B_n.

Next, let $R(x_1, \dots, x_n)$ be an n-ary relation (different from equality). Write $R^{\mathcal{N}}$ as an infinitary expression using countable unions and intersections of Cartesian products of $\mathcal{E}$-measurable sets from $\{B \cap \mathcal{N} \mid B \in \Delta_R\}$ (see the definition of Δ_R and $\mathcal{E}$ in the proof of Theorem 9.1.6); say as the expression $t(B_0, B_1, \dots)$. Then we define $R^{\mathcal{M}}$ by $t(C_0, C_1, \dots)$. Furthermore, we define each constant $c^{\mathcal{M}}$ to be $\pi(c^{\mathcal{N}})$.

Finally, define a probability measure $\mathcal{D}_0$ on the Borel sets of $\mathcal{M} = 2^\omega$ by

$$\mathcal{D}_0 = \mathcal{E} \circ \pi^{-1}.$$

Let $\mathcal{D}$ be the completion of $\mathcal{D}_0$. Then Proposition 9.2.1 tells us that $(\mathcal{M}, \mathcal{D})$ is an ε-model. Now it is easy to see that for all propositional formulas $\varphi(x_1, \dots, x_n)$ and all $x_1, \dots, x_n \in \mathcal{N}$ we have that $\mathcal{N} \models \varphi(x_1, \dots, x_n)$ if and only if $\mathcal{M} \models \varphi(\pi(x_1), \dots, \pi(x_n))$. For the atomic formulas not using equality, this follows from the definition of the relations, and for equality this follows from the injectivity of π. For general formulas in prenex normal form, we can now easily prove the implication from left to right, i.e. that $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(x_1, \dots, x_n)$ implies that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(\pi(x_1), \dots, \pi(x_n))$. Note that for universal quantifiers, going from $\mathcal{N}$ to $\mathcal{M}$, the set on which a formula holds can only increase in measure, and for existential quantifiers, there are at least as many witnesses in $\mathcal{M}$ as in $\mathcal{N}$. It follows from this that in particular $(\mathcal{M}, \mathcal{D})$ satisfies Γ. $\square$

Next, we show that we can eliminate atoms.

Definition 9.2.3. Let μ be a measure and let x be a measurable singleton. We say that x is an *atom* of μ if $\mu(\{x\}) > 0$. The measure μ is called *atomless* if it does not have any atoms.

Often a different notion of atom is used in the literature, in which an atom is a measurable set A of strictly positive measure such that every measurable subset of A either has measure 0 or the same measure as A. However, in Polish spaces such as 2^ω these two notions coincide, as can be seen in e.g. Aliprantis and Border [1, Lemma 12.18].

[2]The idea of sending each x_a to $a \in 2^\omega$ also appears in Bogachev [10, Theorem 9.4.7], albeit in a different context. However, there only the case in which the function π is also surjective is discussed, the non-surjective case being irrelevant in that context.

Note that there are at most countably many atoms, since there can only be finitely many atoms of measure $\geqslant \frac{1}{n}$ for every $n \in \omega$.

DEFINITION 9.2.4. Let $(\mathcal{M}, \mathcal{D})$ and $(\mathcal{N}, \mathcal{E})$ be two ε-models over the same language. Then we say that $(\mathcal{M}, \mathcal{D})$ and $(\mathcal{N}, \mathcal{E})$ are *ε-elementary equivalent*, denoted by $(\mathcal{M}, \mathcal{D}) \equiv_\varepsilon (\mathcal{N}, \mathcal{E})$, if for all sentences φ we have

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi \Leftrightarrow (\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi.$$

LEMMA 9.2.5. *Let $\mathcal{L}$ be a first-order language not containing equality or function symbols. Let $(\mathcal{M}, \mathcal{D})$ be an ε-model such that $\mathcal{D}$ is the completion of a Borel measure $\mathcal{D}_0$. Then there exists an atomless ε-model $(\mathcal{N}, \mathcal{E})$ such that $\mathcal{E}$ is the completion of a Borel measure $\mathcal{E}_0$ and $(\mathcal{N}, \mathcal{E}) \equiv_\varepsilon (\mathcal{M}, \mathcal{D})$. Furthermore, if $\mathcal{M}$ is Polish, then so is $\mathcal{N}$.*

PROOF. We first show how to eliminate a single atom of $\mathcal{D}_0$. Let x_0 be an atom, say of measure r. Let $\mathcal{N}$ be the disjoint union of $\mathcal{M}$ and $[0, r]$. We define a new measure $\mathcal{E}_0$ on $\mathcal{N}$ by setting, for each $\mathcal{D}_0$-measurable $B \subseteq \mathcal{M}$ and Borel $C \subseteq [0, r]$,

$$\mathcal{E}_0(B \cup C) = \mathcal{D}_0(B \setminus \{x_0\}) + \mu(C),$$

where μ denotes the Lebesgue measure, restricted to Borel sets. Then clearly, x_0 is no longer an atom (since it now has measure zero). The interpretation of constants in $\mathcal{N}$ is simply defined by $c^{\mathcal{N}} = c^{\mathcal{M}}$.

Define a function $\pi : \mathcal{N} \to \mathcal{M}$ by

$$\pi(x) = \begin{cases} x & \text{if } x \in M, \\ x_0 & \text{if } x \in [0, r]. \end{cases}$$

We now define the relations on $\mathcal{N}$ by letting

$$R^{\mathcal{N}}(x_1, \ldots, x_n) \Longleftrightarrow R^{\mathcal{M}}(\pi(x_1), \ldots, \pi(x_n))$$

for every relation R. To simplify notation, in the following we write $\vec{x} = x_1, \ldots, x_n$ and $\pi(\vec{x}) = \pi(x_1), \ldots, \pi(x_n)$.

Let $\mathcal{E}$ be the completion of $\mathcal{E}_0$. To show that $(\mathcal{N}, \mathcal{E}) \equiv_\varepsilon (\mathcal{M}, \mathcal{D})$ we prove the stronger assertion that for all formulas φ and for all $\vec{x} \in \mathcal{N}$,

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(\vec{x}) \Longleftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(\pi(\vec{x})), \tag{20}$$

so that this holds in particular for all sentences φ. We prove (20) by formula-induction on φ in prenex normal form. We only prove the universal case; the other cases are easy. So, let $\varphi = \forall y \psi(y, \vec{x})$. To prove (20) it suffices to show that

$$\Pr_{\mathcal{E}}\big[b \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \psi(b, \vec{x})\big] = \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(a, \pi(\vec{x}))\big]$$

for all $\vec{x} \in \mathcal{N}$. Now by induction hypothesis,

$$\begin{aligned}
&\Pr_{\mathcal{E}}\big[b \in \mathcal{N} \mid (\mathcal{N},\mathcal{E}) \models_\varepsilon \psi(b,\vec{x})\big]\\
=&\Pr_{\mathcal{E}}\big[b \in \mathcal{N} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(\pi(b),\pi(\vec{x}))\big]\\
=&\Pr_{\mathcal{E}}\big[b \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(\pi(b),\pi(\vec{x}))\big]\\
&\qquad\qquad + \Pr_{\mathcal{E}}\big[b \in [0,r] \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(\pi(b),\pi(\vec{x}))\big]\\
=&\Pr_{\mathcal{D}}\big[a \in \mathcal{M}\setminus\{x_0\} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(a,\pi(\vec{x}))\big]\\
&\qquad\qquad + \Pr_{\mathcal{D}}\big[x_0 \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(x_0,\pi(\vec{x}))\big]\\
=&\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_\varepsilon \psi(a,\pi(\vec{x}))\big]
\end{aligned}$$

Finally, using the remark below Definition 9.2.3, we can iterate this construction a countable number of times, eliminating the atoms one by one. It should be clear that the limit model exists and satisfies the theorem. □

The next theorem shows the connection between Borel measures on 2^ω and the Lebesgue measure.

Definition 9.2.6. Let $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ be measure spaces. $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ are *isomorphic* if there exists an isomorphism from X to Y, that is, a bijection $f : X \to Y$ such that $f(\mathcal{A}) = \mathcal{B}$ and $\mu \circ f^{-1} = \nu$.

Theorem 9.2.7. (Kechris [50, Theorem 17.41]) *Let $\mathcal{D}$ be an atomless Borel probability measure on a Polish space. Then it is isomorphic to $[0,1]$ with the Lebesgue measure restricted to Borel sets.*

Putting together everything we have found, we reach the theorem announced at the beginning of this section.

Theorem 9.2.8. *Let $\mathcal{L}$ be a countable first-order language not containing equality or function symbols. Let Γ be an ε-satisfiable set of sentences. Then there exists an ε-model on $[0,1]$ with the Lebesgue measure which ε-satisfies Γ. Furthermore, all relations in the new ε-model can be chosen to be Borel.*

Proof. By Proposition 9.2.2 we know that Γ has an ε-model $(\mathcal{M},\mathcal{D})$ on 2^ω with $\mathcal{D}$ the completion of a Borel measure and with Borel relations. By Lemma 9.2.5 we may assume that $\mathcal{M}$ is atomless. By Theorem 9.2.7, $(\mathcal{M},\mathcal{D})$ is isomorphic to $[0,1]$ with the Lebesgue measure, by a Borel isomorphism. Note that isomorphisms preserve ε-truth, so the induced model on $[0,1]$ is ε-elementary equivalent to $\mathcal{M}$. □

Note that Theorem 9.2.8 fails for languages with equality, since then a sentence such as $\exists x\forall y(x=y)$ expresses that there is an atom of measure at least $1-\varepsilon$. For languages with equality we have the following more elaborate result:

Theorem 9.2.9. *Let $\mathcal{L}$ be a countable first-order language not containing function symbols. Let Γ be an ε-satisfiable set of sentences. Then there exists an ε-model $(\mathcal{M},\mathcal{D})$ which ε-satisfies Γ such that:*

(i) *$\mathcal{M}$ is based on $[0,r]\cup X$ for some $r \in [0,1]$ and a countable set X,*

(ii) $\mathcal{D}$ *is the Lebesgue measure on* $[0, r]$,
(iii) *all relations in* $\mathcal{M}$ *are Borel.*

Proof. As in the proof of Theorem 9.2.8, by Proposition 9.2.2 we know that Γ has an ε-model $(\mathcal{N}, \mathcal{E})$ on 2^ω with $\mathcal{E}$ the completion of a Borel measure and with Borel relations. We can separate $\mathcal{N}$ into a countable set of atoms X and an atomless part Y. Let r be the $\mathcal{E}$-measure of the atomless part. If $r = 0$, then observe that we can apply the classical downward Löwenheim–Skolem Theorem to find a countable submodel $\mathcal{M}$ of $\mathcal{N}$ containing X, and then $(\mathcal{M}, \mathcal{E} \upharpoonright \mathcal{M})$ can be easily verified to satisfy the requirements of the theorem. If $r > 0$, then by Theorem 9.2.7 we know that $(Y, \frac{1}{r}\mathcal{E})$ is isomorphic to $[0, 1]$ with the Lebesgue measure, so $(Y, \mathcal{E})$ is isomorphic to $[0, r]$ with the Lebesgue measure. As in Theorem 9.2.8, this induces an ε-elementary equivalent model on $[0, r]$ together with the atoms X. $\square$

9.3. The Löwenheim number

At this point we may ask ourselves how tight Theorem 9.1.6 is. The *Löwenheim number* of a logic is the smallest cardinal λ such that every satisfiable sentence has a model of cardinality at most λ. For every ε, let λ_ε be the Löwenheim number of ε-logic, i.e. the smallest cardinal such that every ε-satisfiable sentence has an ε-model of cardinality at most λ_ε. The next theorem parallels Corollary 2.4.5 in Keisler [51]. MA is Martin's axiom from set theory, see Kunen [62].

Theorem 9.3.1. *Let* $\varepsilon \in [0, 1)$ *be rational. For the Löwenheim number* λ_ε *of* ε*-logic we have*

(i) $\aleph_1 \leqslant \lambda_\varepsilon \leqslant 2^{\aleph_0}$,
(ii) *If Martin's axiom* MA *holds then* $\lambda_\varepsilon = 2^{\aleph_0}$.

Proof. The first part was already proven above, in Example 9.1.5 and Theorem 9.1.6. For the second part, assume that MA holds. Let φ be the sentence from Example 9.1.5. Let $\kappa < 2^\omega$ and assume φ has a model of cardinality κ. We remark that any model of φ has to be atomless. Therefore, if we now use the construction from Proposition 9.2.2, we find a model $(\mathcal{M}, \mathcal{D})$ which ε-satisfies φ and where $\mathcal{D}$ is the completion of an atomless Borel measure. Furthermore, if we let π and $(\mathcal{N}, \mathcal{E})$ be as in the proof of this proposition, the set $\pi(\mathcal{N})$ is a set of cardinality at most κ, so by MA it has measure 0 (see Fremlin [31, p127]). But then

$$\mathcal{E}(\mathcal{N}) = \mathcal{D} \circ \pi(\mathcal{N}) = 0,$$

a contradiction. $\square$

Theorem 9.3.1 shows that we cannot prove in the standard set-theoretic framework of ZFC such statements as $\lambda_\varepsilon = \aleph_1$, because this is independent of ZFC. Namely under CH we have that $\lambda_\varepsilon = \aleph_1$ by item (i), and under MA we have $\lambda_\varepsilon = 2^{\aleph_0}$ by item (ii), and MA is consistent with $2^{\aleph_0} > \aleph_1$, see Kunen [62, p278]. (Note that this does not exclude the possibility that $\lambda_\varepsilon = 2^{\aleph_0}$ could be provable within ZFC.) So in this sense Theorem 9.1.6 is optimal.

9.4. Reductions

In this section we discuss many-one reductions between the sets of ε-satisfiable formulas for various ε. Recall that a *many-one reduction* between two sets A, B of formulas is a computable function f such that for all formulas φ we have that $\varphi \in A$ if and only if $f(\varphi) \in B$. The reductions we present below are useful, e.g. in our discussion of compactness in section 9.5, and also for complexity issues not discussed here.

In what follows we will need to talk about satisfiability of formulas rather than sentences. We will call a formula $\varphi(x_1, \ldots, x_n)$ *ε-satisfiable* if there is an ε-model $(\mathcal{M}, \mathcal{D})$ and elements $a_1, \ldots, a_n \in \mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \ldots, a_n)$.

THEOREM 9.4.1. *Let $\mathcal{L}$ be a countable first-order language not containing function symbols or equality. Then, for all rationals $0 \leqslant \varepsilon_0 \leqslant \varepsilon_1 < 1$, there exists a language $\mathcal{L}'$ (also not containing function symbols or equality) such that ε_0-satisfiability in $\mathcal{L}$ many-one reduces to ε_1-satisfiability in $\mathcal{L}'$.*

PROOF. We can choose integers $a > 0$, n, and $m \leqslant n$ so that $\varepsilon_0 = 1 - \frac{a}{m}$ and $\varepsilon_1 = 1 - \frac{a}{n}$, and hence $\frac{m}{n} = \frac{1-\varepsilon_1}{1-\varepsilon_0}$. Let $\varphi(y_1, \ldots, y_k)$ be a formula in prenex normal form (see Proposition 8.1.7). For simplicity we write $\vec{y} = y_1, \ldots, y_k$. Also, for a function π we let $\pi(\vec{y})$ denote the vector $\pi(y_1), \ldots, \pi(y_k)$. We use formula-induction to define a computable function f such that for every formula φ,

$$\varphi \text{ is } \varepsilon_0\text{-satisfiable if and only if } f(\varphi) \text{ is } \varepsilon_1\text{-satisfiable.} \tag{21}$$

For propositional formulas and existential quantifiers, there is nothing to be done and we use the identity map. Next, we consider the universal quantifiers. Let $\varphi = \forall x \psi(\vec{y}, x)$. The idea is to introduce new unary predicates, that can be used to vary the strength of the universal quantifier. We will make these predicates split the model into disjoint parts. If we split it into just the right number of parts (in this case n), then we can choose m of these parts to get just the right strength.

So, we introduce new unary predicates $X_1, \ldots, X_n$. For $1 \leqslant i \leqslant n$, define

$$Y_i(x) = X_i(x) \wedge \bigwedge_{j \neq i} \neg X_j(x).$$

Then the predicates Y_i define disjoint sets in any model.

We now define the sentence a-n-split by:

$$\bigwedge_{I \subseteq \{1,\ldots,n\}, \#I = a} \forall y \Big(\bigvee_{i \in I} Y_i(y) \Big),$$

where $\#I$ denotes the cardinality of I. Then one can verify that in any model, if the sets X_i are disjoint sets of measure exactly $\frac{1}{n}$ (and hence the same holds for the Y_i), then a-n-split is ε_1-valid. Conversely, if a-n-split holds, then the sets Y_i all have measure $\frac{1}{n}$ by Lemma 9.4.2 below. In particular we see that, if a-n-split holds, then the Y_i together disjointly cover a set of measure 1.

Now define $f(\varphi)$ to be the formula

$$a\text{-}n\text{-split} \wedge \bigwedge_{i_1,\ldots,i_m} \forall x \big((Y_{i_1}(x) \vee \cdots \vee Y_{i_m}(x)) \wedge f(\psi)(\vec{y}, x) \big)$$

where the conjunction is over all subsets of size m from $\{1,\dots,n\}$. (It will be clear from the construction that $f(\psi)$ has the same arity as ψ.) Thus, $f(\varphi)$ expresses that for any choice of m of the n parts, $f(\psi)(x)$ holds often enough when restricted to the resulting part of the model.

We will now prove claim (21) above. For the implication from left to right, we will prove the following strengthening:

For every formula $\varphi(\vec{y})$, if φ is ε_0-satisfied in some ε_0-model $(\mathcal{M},\mathcal{D})$, then there exists an ε_1-model $(\mathcal{N},\mathcal{E})$ together with a measure-preserving surjective measurable function $\pi:\mathcal{N}\to\mathcal{M}$ (i.e. for all $\mathcal{D}$-measurable A we have that $\mathcal{E}(\pi^{-1}(A))=\mathcal{D}(A)$) such that for all $\vec{y}\in\mathcal{N}$ we have that

$$(\mathcal{N},\mathcal{E})\models_{\varepsilon_1} f(\varphi)(\vec{y}) \textit{ if and only if } (\mathcal{M},\mathcal{D})\models_{\varepsilon_0}\varphi(\pi(\vec{y})).$$

We prove this by formula-induction over the formulas in prenex normal form. For propositional formulas, there is nothing to be done (we can simply take the models to be equal and π the identity). For the existential quantifier, let $\varphi=\exists x\psi(x)$ and apply the induction hypothesis to ψ to find a model $(\mathcal{N},\mathcal{E})$ and a mapping π. Then we can take the same model and mapping for φ, as easily follows from the fact that π is surjective.

Next, we consider the universal quantifier. Suppose $\varphi=\forall x\psi(\vec{y},x)$ is ε_0-satisfied in $(\mathcal{M},\mathcal{D})$. Use the induction hypothesis to find a model $(\mathcal{N},\mathcal{E})$ and a measure-preserving surjective measurable function $\pi:\mathcal{N}\to\mathcal{M}$ such that for all $\vec{y},x\in\mathcal{N}$ we have that

$$(\mathcal{N},\mathcal{E})\models_{\varepsilon_1} f(\psi)(\vec{y},x) \text{ if and only if } (\mathcal{M},\mathcal{D})\models_{\varepsilon_0}\psi(\pi(\vec{y}),\pi(x)).$$

Now form the model $(\mathcal{N}',\mathcal{E}')$ consisting of n copies $\mathcal{N}_1,\dots,\mathcal{N}_n$ of $(\mathcal{N},\mathcal{E})$, each with weight $\frac{1}{n}$. That is, $\mathcal{E}'$ is the sum of n copies of $\frac{1}{n}\mathcal{E}$. Let $\pi':\mathcal{N}'\to\mathcal{M}$ be the composition of the projection map $\sigma:\mathcal{N}'\to\mathcal{N}$ with π. Relations in $\mathcal{N}'$ are defined just as on $\mathcal{N}$, that is, for a t-ary relation R we define $R^{\mathcal{N}'}(x_1,\dots,x_t)$ by $R^{\mathcal{N}}(\sigma(x_1),\dots,\sigma(x_t))$. Observe that this is the same as defining $R^{\mathcal{N}'}(x_1,\dots,x_t)$ by $R^{\mathcal{M}}(\pi'(x_1),\dots,\pi'(x_t))$. We interpret constants $c^{\mathcal{N}'}$ by embedding $c^{\mathcal{N}}$ into the first copy $\mathcal{N}_1$. Finally, we let each X_i be true exactly on the copy $\mathcal{N}_i$.

Then π' is clearly surjective. To show that it is measure-preserving, it is enough to show that σ is measure-preserving. If A is $\mathcal{E}$-measurable, then $\sigma^{-1}(A)$ consists of n disjoint copies of A, each having measure $\frac{1}{n}\mathcal{E}(A)$, so $\pi^{-1}(A)$ has $\mathcal{E}'$-measure exactly $\mathcal{E}(A)$.

Now, since $(\mathcal{N}',\mathcal{E}')$ satisfies a-n-split, we see that

$$(\mathcal{N}',\mathcal{E}')\models_{\varepsilon_1} f(\psi)(\vec{y}) \tag{22}$$

is equivalent to the statement that for all $1\leqslant i_1<\cdots<i_m\leqslant n$ we have

$$\Pr_{\mathcal{E}'}\big[x\in\mathcal{N}' \mid (\mathcal{N}',\mathcal{E}')\models_{\varepsilon_1}(Y_{i_1}(x)\vee\cdots\vee Y_{i_m}(x))\wedge f(\psi)(\vec{y},x)\big]\geqslant 1-\varepsilon_1. \tag{23}$$

By Lemma 9.4.3 below we have that $(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\psi)(\vec{y}, x)$ if and only if $(\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), \sigma(x))$. In particular, we see for every $1 \leqslant i \leqslant n$ that

$$\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} Y_i(x) \wedge f(\psi)(\vec{y}, x)\big] = \frac{1}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big]. \tag{24}$$

It follows that (23) is equivalent to

$$\frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big] \geqslant 1 - \varepsilon_1.$$

The induction hypothesis tells us that this is equivalent to

$$\frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), \pi(x))\big] \geqslant 1 - \varepsilon_1$$

and since π is surjective and measure-preserving, this is the same as

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), x)\big] \geqslant \frac{n}{m}(1 - \varepsilon_1) = 1 - \varepsilon_0.$$

This proves that $(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \varphi(\pi'(\vec{y}))$.

We have not yet explained why $(\mathcal{N}', \mathcal{E}')$ is actually an ε_1-model. However, by Theorem 9.2.8 we may assume the relations on the original model $\mathcal{M}$ to be Borel, and it is easily seen that our construction of successively making copies keeps the relations Borel. So, from Proposition 9.2.1 we see that the models $(\mathcal{N}, \mathcal{E})$ and $(\mathcal{N}', \mathcal{E}')$ are in fact ε-models for *every* ε.

To prove the right to left direction of (21) we will use induction to prove the following stronger statement:

If $(\mathcal{M}, \mathcal{D})$ is an ε_1-model and $\vec{y} \in \mathcal{M}$ are such that $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\varphi)(\vec{y})$, then we also have $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$.[3]

In particular, if $f(\varphi)$ is ε_1-satisfiable, then φ is ε_0-satisfiable. The only interesting case is the universal case, so let $\varphi = \forall x \psi(\vec{y}, x)$. Let $\vec{y} \in \mathcal{M}$ be such that $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\varphi)(\vec{y})$. Assume, towards a contradiction, that $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\vec{y})$. Then

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \psi(\vec{y}, x)\big] > \varepsilon_0$$

and by the induction hypothesis we have

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big] > \varepsilon_0. \tag{25}$$

But by taking those m of the Y_i (say $Y_{i_1}, \dots, Y_{i_m}$) which have the largest intersection with this set we find that

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} (Y_{i_1} \vee \dots \vee Y_{i_m}) \wedge f(\psi)(\vec{y}, x)\big] < \frac{m}{n}(1 - \varepsilon_0) = 1 - \varepsilon_1$$

which contradicts our choice of $(\mathcal{M}, \mathcal{D})$. □

LEMMA 9.4.2. *Let $a, n \in \omega$, $a > 0$, and let $\mathcal{D}$ be a probability measure. Let $Y_1, \dots, Y_n$ be disjoint $\mathcal{D}$-measurable sets such that for all subsets $I \subseteq \{1, \dots, n\}$ of cardinality a the union $\bigcup_{i \in I} Y_i$ has measure at least $\frac{a}{n}$. Then all Y_i have measure exactly $\frac{1}{n}$.*

[3]As explained above, we may assume $(\mathcal{M}, \mathcal{D})$ to be an ε_0-model.

PROOF. Assume there exists $1 \leqslant i \leqslant n$ such that Y_i has measure $< \frac{1}{n}$. Determine a sets Y_i with minimal measure, say with indices from the set I. Then, by assumption, $\mathcal{D}(\bigcup_{i \in I} Y_i) \geqslant \frac{a}{n}$. But at least one of the Y_i with $i \in I$ has measure strictly less than $\frac{1}{n}$, so also one of them needs to have measure strictly greater than $\frac{1}{n}$. However, $\mathcal{D}(\bigcup_{i \notin I} Y_i) \leqslant \frac{n-a}{n}$, so there is a Y_j with $j \notin I$ having measure $\leqslant \frac{1}{n}$. This contradicts the minimality. So, all sets Y_i have measure at least $\frac{1}{n}$ and since they are disjoint they therefore have measure exactly $\frac{1}{n}$. □

LEMMA 9.4.3. *Let $(\mathcal{N}', \mathcal{E}')$ and $(\mathcal{N}, \mathcal{E})$ be as in the proof of Theorem 9.4.1 above. Then for every formula $\zeta(x_1, \ldots, x_t)$ in the language of $\mathcal{M}$, for every $\varepsilon \in [0,1]$ and all $x_1, \ldots, x_t \in \mathcal{N}'$ we have that $(\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta(x_1, \ldots, x_t)$ if and only if $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta(\sigma(x_1), \ldots, \sigma(x_t))$.*

PROOF. By induction on the structure of the formulas in prenex normal form. The base case holds by definition of the relations in $\mathcal{N}'$. The only interesting induction step is the one for the universal quantifier. So, let $\zeta = \forall x_0 \zeta'(x_0, \ldots, x_t)$ and let $x_1, \ldots, x_t \in \mathcal{N}'$. Using the induction hypothesis, we find that the set $A = \{x_0 \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta'(x_0, \ldots, x_t)\}$ is equal to the set $\{x_0 \in \mathcal{N}' \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta'(\sigma(x_0), \ldots, \sigma(x_t))\}$, which consists of n disjoint copies of the set $B = \{x_0 \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta'(x_0, \sigma(x_1), \ldots, \sigma(x_t))\}$; denote the copy of B living inside $\mathcal{N}_i$ by B_i. Then

$$\mathcal{D}(A) = \sum_{i=1}^{n} \mathcal{E}'(B_i) = \sum_{i=1}^{n} \tfrac{1}{n}\mathcal{E}(B) = \mathcal{E}(B)$$

from which we directly see that $(\mathcal{N}', \mathcal{E}') \models_\varepsilon \zeta(x_1, \ldots, x_t)$ if and only if $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \zeta(\sigma(x_1), \ldots, \sigma(x_t))$. □

Next we show that we have a similar reduction in the other direction, i.e. from a bigger ε_0 to a smaller ε_1.

THEOREM 9.4.4. *Let $\mathcal{L}$ be a countable first-order language not containing function symbols or equality. Then, for all rationals $0 < \varepsilon_1 \leqslant \varepsilon_0 \leqslant 1$, there exists a language $\mathcal{L}'$ (also not containing function symbols or equality) such that ε_0-satisfiability in $\mathcal{L}$ many-one reduces to ε_1-satisfiability in $\mathcal{L}'$.*

PROOF. The proof is very similar to that of Theorem 9.4.1. There are two main differences: the choice of the integers a, n and m, and a small difference in the construction of f (as we shall see below). We can choose integers a, n and m such that $\varepsilon_1 = 1 - \frac{a}{n}$ and $\frac{m}{n} = \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0}$. The case $a = 0$ is trivial, so we may assume that $a > 0$. We construct a many-one reduction f such that for all formulas φ,

$$\varphi \text{ is } \varepsilon_0\text{-satisfiable if and only if } f(\varphi) \text{ is } \varepsilon_1\text{-satisfiable.}$$

Again, we only consider the nontrivial case where φ is a universal formula $\forall x \psi(\vec{y}, x)$. We define $f(\varphi)$ to be the formula

$$a\text{-}n\text{-split} \wedge \bigwedge_{i_1, \ldots, i_m} \forall x \big(Y_{i_1}(x) \vee \cdots \vee Y_{i_m}(x) \vee f(\psi)(\vec{y}, x)\big)$$

where the conjunction is over all subsets of size m from $\{1, \dots, n\}$. So, the essential change from the proof of Theorem 9.4.1 is that the conjunction of $Y_{i_1}(x) \vee \cdots \vee Y_{i_m}(x)$ and $f(\psi)(\vec{y}, x)$ has become a disjunction.

The remainder of the proof is now almost the same as for Theorem 9.4.1. In the proof for the implication from left to right, follow the proof up to equation (22), i.e.

$$(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\varphi)(\vec{y}).$$

Again this is equivalent to the statement that for all $1 \leqslant i_1 < \cdots < i_m \leqslant n$ we have

$$\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} Y_{i_1}(x) \vee \cdots \vee Y_{i_m}(x) \vee f(\psi)(\vec{y}, x)\big] \geqslant 1 - \varepsilon_1. \tag{26}$$

Similar as before, using Lemma 9.4.3, we find that this is equivalent to

$$\frac{m}{n} + \frac{n-m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big] \geqslant 1 - \varepsilon_1.$$

Again, using the induction hypothesis and the fact that π is measure-preserving we find that this is equivalent to

$$\begin{aligned} &\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), x)\big] \\ &\qquad \geqslant \frac{n}{n-m}\left(1 - \varepsilon_1 - \frac{m}{n}\right) = \frac{\varepsilon_0}{\varepsilon_1}\left(\frac{-\varepsilon_0\varepsilon_1 + \varepsilon_1}{\varepsilon_0}\right) = 1 - \varepsilon_0. \end{aligned}$$

This proves that $(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$.

For the converse implication, we also need to slightly alter the proof of Theorem 9.4.1. Assuming that $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\varphi)(\vec{y})$, follow the proof up to equation (25), where we obtain

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big] > \varepsilon_0. \tag{27}$$

Define

$$\eta = \Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big]$$

and take those m of the Y_i (say $Y_{i_1}, \dots, Y_{i_m}$) which have the smallest intersection with this set. Note that by (27) we have $\eta < 1 - \varepsilon_0$. Then we find that

$$\begin{aligned} &\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} Y_{i_1} \vee \cdots \vee Y_{i_m} \vee f(\psi)(\vec{y}, x)\big] \\ &\leqslant \frac{m}{n} + \left(1 - \frac{m}{n}\right)\eta = \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}\eta \\ &< \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}(1 - \varepsilon_0) = 1 - \varepsilon_1. \end{aligned}$$

which contradicts our choice of $(\mathcal{M}, \mathcal{D})$. Note that the last inequality holds because $\varepsilon_1 \neq 0$. □

9.5. Compactness

We start this section with a negative result, namely that in general ε-logic is not compact. This results holds for rational ε different from 0 and 1. The case $\varepsilon = 1$ is pathological, and in section 11.6 we will show that 0-logic is in fact compact.

First, we prove a technical lemma.

LEMMA 9.5.1. *Let $(\mathcal{M}, \mathcal{D})$ be an ε-model, let $R(x, y)$ be a binary relation and let $\delta > 0$. Then for almost all y there exists a set C_y of strictly positive measure such that for all $y' \in C_y$:*

$$\Pr_{\mathcal{D}}\big[u \in \mathcal{M} \mid R^{\mathcal{M}}(u, y) \leftrightarrow R^{\mathcal{M}}(u, y')\big] \geqslant 1 - \delta.$$

PROOF. It suffices to show this not for almost all y, but instead show that for all $\delta' > 0$ this holds for at least $\mathcal{D}$-measure $1 - \delta'$ many y.

We first show that we can approximate the horizontal sections of $R^{\mathcal{M}}$ in a suitable way, namely that there exist a finite number $U_1, \dots, U_n$ and $V_1, \dots, V_n$ of $\mathcal{D}$-measurable sets such that:

$$\Pr_{\mathcal{D}}\Big[y \in \mathcal{M} \Big| \Pr_{\mathcal{D}}\Big[u \in \mathcal{M} \mid (u, y) \in R^{\mathcal{M}} \triangle \Big(\bigcup_{i=1}^{n} U_i \times V_i\Big)\Big] \leqslant \frac{\delta}{2}\Big] \geqslant 1 - \delta'. \tag{28}$$

(Here $\triangle$ denotes the symmetric difference.) To show that this is possible, determine U_i and V_i such that

$$\Pr_{\mathcal{D}}\Big[R^{\mathcal{M}} \triangle \Big(\bigcup_{i=1}^{n} U_i \times V_i\Big)\Big] < \frac{\delta\delta'}{2}. \tag{29}$$

Observe that such an approximation exists: this obviously holds for relations $R^{\mathcal{M}}$ that are a rectangle $U \times V$, and the existence of an approximation is preserved under countable unions and complements. (This is usually part of a proof of Fubini's theorem.)

Now assume that (28) does not hold for these U_i and V_i. Then we have

$$\Pr_{\mathcal{D}}\Big[y \in \mathcal{M} \Big| \Pr_{\mathcal{D}}\Big[u \in \mathcal{M} \mid (u, y) \in R^{\mathcal{M}} \triangle \Big(\bigcup_{i=1}^{n} U_i \times V_i\Big)\Big] > \frac{\delta}{2}\Big] > \delta'.$$

But then we see that

$$\Pr_{\mathcal{D}}\Big[R^{\mathcal{M}} \triangle \Big(\bigcup_{i=1}^{n} U_i \times V_i\Big)\Big] \geqslant \frac{\delta\delta'}{2}$$

by taking the integral, contradicting (29).

We now show that for almost all y there exists a set C_y of strictly positive measure such that for all $y' \in C_y$ we have that

$$\Pr_{\mathcal{D}}\Big[u \in \mathcal{M} \mid (u, y) \in \bigcup_{i=1}^{n} U_i \times V_i \leftrightarrow (u, y') \in \bigcup_{i=1}^{n} U_i \times V_i\Big] = 1. \tag{30}$$

Note that the V_i induce a partition of $\mathcal{M}$ into at most 2^n many disjoint parts Y_j (by choosing for each $i \leqslant n$ either V_i or its complement, and intersecting these). But each such Y_j has either measure zero (so we can ignore it), or Y_j has strictly positive measure and for all $y \in Y_j$ we can take $C_y = Y_j$. Then it is clear that (30) holds.

Call the set of measure $1 - \delta'$ elements y from (28) $\mathcal{M}'$. The same argument used to prove (30) can be applied to $\mathcal{M}'$, using the partition $Y_j \cap \mathcal{M}'$ of $\mathcal{M}'$. For $y \in \mathcal{M}'$ this gives the same conclusion (30), but with the extra property that $C_y \subseteq \mathcal{M}'$.

Finally, the lemma follows by combining (28) and (30): For every $y \in \mathcal{M}'$, by (28) its sections with $R^{\mathcal{M}}$ and the approximation differ by at most $\frac{\delta}{2}$. By (30) and the previous remark there is $C_y \subseteq \mathcal{M}'$ of positive measure such that for every $y' \in C_y$ the sections of y and y' with the approximation agree almost everywhere. Again by (28) the sections of y' with $R^{\mathcal{M}}$ and the approximation differ by at most $\frac{\delta}{2}$. Hence the sections of y and y' with $R^{\mathcal{M}}$ differ by at most $\frac{\delta}{2} + \frac{\delta}{2}$. $\square$

THEOREM 9.5.2. *For every rational $\varepsilon \in (0,1)$, ε-logic is not compact, i.e. there exists a countable set Γ of sentences such that each finite subset of Γ is ε-satisfiable, but Γ itself is not ε-satisfiable.*

PROOF. The example we use is adapted from Keisler [51, Example 2.6.5]. Let R be a binary relation. Using the reductions from Theorem 9.4.1 (observing, from the proof of that theorem, that we can apply the reduction per quantifier), we can form a sentence φ_n such that φ_n is ε-satisfiable if and only if there is a model satisfying:
For almost all y (i.e. measure 1 many), there exists a set A_y of measure at least $1-\frac{1}{n}$ such that for all $y' \in A_y$ the sets $B_y = \{u \mid R(u,y)\}$ and $B_{y'} = \{u \mid R(u,y')\}$ both have measure $\frac{1}{2}$, while $B_y \cap B_{y'}$ has measure $\frac{1}{4}$ (in other words, the two sets are independent sets of measure $\frac{1}{2}$).

Then each φ_n has a finite ε-model, as illustrated in Figure 1 for $n = 4$: for each x (displayed on the horizontal axis) we let $R(x,y)$ hold exactly for those y (displayed on the vertical axis) where the box has been coloured grey. If we now take for each A_y exactly those three intervals of length $\frac{1}{4}$ of which y is not an element, we can directly verify that φ_n holds.

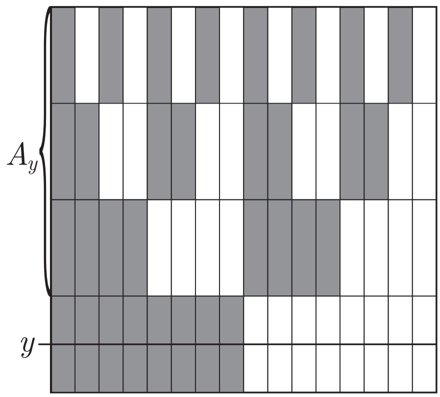


FIGURE 1. A model for φ_4 on $[0,1]$.

However, the set $\{\varphi_n \mid n \in \omega\}$ has no ε-model. Namely, for such a model, we would have that for almost all y, there exists a set A_y of measure 1 such that for all $y' \in A_y$ the sets B_y and $B_{y'}$ (defined above) are independent sets of measure $\frac{1}{2}$. Clearly, such a model would need to be atomless and therefore cannot be countable. But then we would have uncountably many of such independent sets B_y.

Intuitively, this contradicts the fact that R is measurable in the product measure and can therefore be formed using countable unions and countable intersections of Cartesian products.

More formally, Lemma 9.5.1 tells us that in any ε-model with a binary relation R, for almost all y there exists a set C_y of strictly positive measure such that for all $y' \in C_y$ the sets B_y and $B_{y'}$ agree on a set of measure at least $\frac{7}{8}$. Since necessarily $A_y \cap C_y = \emptyset$ this shows that A_y cannot have measure 1. $\square$

Next, we will present an ultraproduct construction that allows us to partially recover compactness, which is due to Hoover and described in Keisler [51]. This construction uses the Loeb measure from nonstandard analysis, which is due to Loeb [73]. The same construction as in Keisler is also described (for a different logic) in Bageri and Pourmahdian [4], however, there the Loeb measure is not explicitly mentioned and the only appearance of nonstandard analysis is in taking the standard part of some element. Below we will describe the construction without resorting to nonstandard analysis. To be able to define the measure, we need the notion of a *limit over an ultrafilter.* This notion corresponds to taking the standard part of a nonstandard real number.

We refer the reader to Hodges [41] for an explanation of the notions of ultrafilter and ultraproduct.

Definition 9.5.3. Let $\mathcal{U}$ be an ultrafilter over ω and let $a_0, a_1, \dots \in \mathbb{R}$. Then a *limit* of the sequence $a_0, a_1, \dots$ over the ultrafilter $\mathcal{U}$, or a *$\mathcal{U}$-limit*, is an $r \in \mathbb{R}$ such that for all $\varepsilon > 0$ we have $\{i \in \omega \mid |a_i - r| < \varepsilon\} \in \mathcal{U}$. We will denote such a limit by $\lim_{\mathcal{U}} a_i$.

Proposition 9.5.4. *Limits over ultrafilters are unique. Furthermore, for any bounded sequence and every ultrafilter $\mathcal{U}$, the limit of the sequence over $\mathcal{U}$ exists.*

Proof. First assume that we have an ultrafilter $\mathcal{U}$ over ω and a sequence $a_0, a_1, \dots$ in $\mathbb{R}$ that has two distinct limits r_0 and r_1. Then the sets

$$\left\{i \in \omega \mid |a_i - r_0| < \tfrac{1}{2}|r_0 - r_1|\right\}$$

and

$$\left\{i \in \omega \mid |a_i - r_1| < \tfrac{1}{2}|r_0 - r_1|\right\}$$

are disjoint elements of $\mathcal{U}$; so, $\emptyset \in \mathcal{U}$, which contradicts $\mathcal{U}$ being a proper filter.

Now, assume the sequence $a_0, a_1, \dots$ is bounded; without loss of generality we may assume that it is a sequence in $[0, 1]$. We will inductively define a decreasing chain $[b_n, c_n]$ of intervals such that for all $n \in \omega$ we have $\{i \in \omega \mid a_i \in [b_n, c_n]\} \in \mathcal{U}$.

First we let $[a_0, b_0] = [0, 1]$. Next, if $\{i \in \omega \mid a_i \in [b_n, c_n]\} \in \mathcal{U}$, then either

$$\left\{i \in \omega \mid a_i \in \left[b_n, \tfrac{b_n + c_n}{2}\right]\right\} \in \mathcal{U}$$

or

$$\left\{i \in \omega \mid a_i \in \left[\tfrac{b_n + c_n}{2}, c_n\right]\right\} \in \mathcal{U}.$$

Choose one of these two intervals to be $[b_{n+1}, c_{n+1}]$.

Now there exists a unique point $r \in \bigcap_{n \in \omega} [b_n, c_n]$, and it is easily verified that this is the limit of the sequence. $\square$

Using these limits over ultrafilters, we show how to define a probability measure on an ultraproduct of measure spaces. As mentioned above, this construction is essentially due to Loeb [73], but we describe it on ultraproducts instead of using nonstandard analysis.

Definition 9.5.5. Let $\mathcal{U}$ be a non-principal ultrafilter on ω and let $\mathcal{D}_0, \mathcal{D}_1, \ldots$ be a sequence of finitely additive probability measures over sets $X_0, X_1, \ldots$. Then we let $\prod_{i\in\omega} X_i/\mathcal{U}$ denote the ultraproduct, and for a sequence $a = a_0, a_1, \ldots$ with $a_i \in X_i$ we let $[a]$ denote the element of $\prod_{i\in\omega} X_i/\mathcal{U}$ corresponding to the equivalence class of the sequence $a_0, a_1, \ldots$.

For each sequence $A = A_0, A_1, \ldots$ with each A_i a $\mathcal{D}_i$-measurable set we will call the set

$$[A] = \left\{[a] \in \prod_{i\in\omega} X_i/\mathcal{U} \mid \{i \in \omega \mid a_i \in A_i\} \in \mathcal{U}\right\}$$

a *basic measurable set.* If we let Δ be the collection of all basic measurable sets, then we define the *ultraproduct measure* to be the unique measure $\mathcal{E}$ on $\sigma(\Delta)$ such that for all basic measurable sets:

$$\Pr_{\mathcal{E}}([A]) = \lim_{\mathcal{U}} \Pr_{\mathcal{D}_i}(A_i).$$

Note that $\mathcal{E}$ is a σ-additive measure, even if the $\mathcal{D}_i$ are only finitely additive.

Proposition 9.5.6. *The ultraproduct measure exists and is well-defined.*

Proof. We need to verify that $\mathcal{E}$, as defined on the Boolean algebra of basic measurable sets, satisfies the conditions of Carathéodory's extension theorem (see e.g. Bogachev [10, Theorem 1.5.6]). Thus, we need to show that, for any disjoint sequence $[A^0], [A^1], \ldots$ of non-empty basic measurable sets such that $\bigcup_{j\in\omega}[A^j]$ is a basic measurable set, we have that

$$\Pr_{\mathcal{E}}\Big(\bigcup_{j\in\omega}[A^j]\Big) = \sum_{j\in\omega} \Pr_{\mathcal{E}}([A^j]).$$

In fact, we will show that if the $[A^j]$ are disjoint and non-empty, then $\bigcup_{j\in\omega}[A^j]$ is never a basic measurable set.

Namely, let $[A^j]$ be as above and assume $\bigcup_{j\in\omega}[A^j]$ is a basic measurable set. Without loss of generality, we may assume that $\bigcup_{j\in\omega}[A^j] = \prod_{i\in\omega} X_i/\mathcal{U}$. We will construct an element of $\prod_{i\in\omega} X_i/\mathcal{U}$ which is not in $\bigcup_{j\in\omega}[A^j]$, which is a contradiction. Observe that, because all the $[A^j]$ are disjoint and non-empty, the set $\bigcup_{j=0}^{n}[A^j]$ will always be a proper subset of $\prod_{i\in\omega} X_i/\mathcal{U}$. So, for each $n \in \omega$, fix $[x^n] \notin \bigcup_{j=0}^{n}[A^j]$. For every $m \in \omega$, let $I_m \in \mathcal{U}$ be the set $\{i \in \omega \mid x_i^m \notin \bigcup_{j=0}^{m} A_i^j\}$. Furthermore, let $k_n \in \omega$ be the biggest $m \leqslant n$ such that $n \in I_m$, and let it be 0 if no such $m \leqslant n$ exists.

Now define x_i as $x_i^{k_i}$. We claim that $[x] \notin \bigcup_{j\in\omega}[A^j]$. Namely, let $m \in \omega$. Then, for every $n \geqslant m$ with $n \in I_m$ we have $k_n \geqslant m$, so we see that $x_n \notin \bigcup_{j=0}^{m} A_n^j$. In particular, we see that $x_n \notin A_n^m$ for every $n \in I_m \setminus \{0, 1, \ldots, n-1\}$. However, since $\mathcal{U}$ is non-principal, this set is in $\mathcal{U}$, so we see $[x] \notin [A^m]$. $\square$

We can now define a model on the ultraproduct in the usual way, however, we cannot guarantee that this is an ε-model, since we will see that we merely know that all definable subsets of $\mathcal{M}$ of arity 1 are measurable. We thus only obtain a *weak ε-model* (see Definition 8.2.3). To even achieve this, we need to extend the measure to all subsets of the model first, at the cost of moving to a finitely additive measure. The final measure that we obtain is still σ-additive though.

THEOREM 9.5.7. (Tarski) *Every finitely additive measure $\mathcal{D}$ on a Boolean algebra of subsets of X can be extended to a finitely additive measure $\mathcal{D}'$ on the power set $\mathcal{P}(X)$.*

PROOF. See Birkhoff [9, p. 185]. □

DEFINITION 9.5.8. Let $\varepsilon \in [0,1]$, let $\mathcal{U}$ be a non-principal ultrafilter over ω and let $(\mathcal{M}_0, \mathcal{D}_0), (\mathcal{M}_1, \mathcal{D}_1), \ldots$ be a sequence of finitely additive weak ε-models, where each $\mathcal{D}_i$ is defined on all of $\mathcal{P}(\mathcal{M}_i)$ (e.g. using Theorem 9.5.7). We then define the *ultraproduct* of this sequence, which we will denote by $\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U}$, to be the classical ultraproduct of the models $\mathcal{M}_i$, equipped with the ultraproduct measure. More precisely, we define it to be the model having as universe $\prod_{i\in\omega}\mathcal{M}_i/\mathcal{U}$, where for each relation $R(x^1, \ldots, x^n)$ we define the relation on $\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U}$ by

$$R([a^1], \ldots, [a^n]) \Leftrightarrow \{i \in \omega \mid R^{\mathcal{M}_i}(a_i^1, \ldots, a_i^n)\} \in \mathcal{U},$$

and we interpret function symbols $f(x_1, \ldots, x_n)$ by

$$f([a^1], \ldots, [a^n]) = [f^{\mathcal{M}_0}(a_0^1, \ldots, a_0^n), f^{\mathcal{M}_1}(a_1^1, \ldots, a_1^n), \ldots].$$

In particular, constants c are interpreted as

$$c = [c^{\mathcal{M}_0}, c^{\mathcal{M}_1}, \ldots].$$

We can now show that a variant of the fundamental theorem of ultraproducts, or Łoś's theorem, holds for this kind of model.

THEOREM 9.5.9. *For every formula $\varphi(x^1, \ldots, x^n)$ and every sequence of elements $[a^1], \ldots, [a^n] \in \prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U}$, the following are equivalent:*

(i) $\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \models_\varepsilon \varphi([a^1], \ldots, [a^n])$,

(ii) *for all $\varepsilon' > \varepsilon$, $\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$,*[4]

(iii) *there exists a sequence $\varepsilon_0, \varepsilon_1, \ldots$ with $\mathcal{U}$-limit ε such that $\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_i} \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$.*

In particular, if $\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_\varepsilon \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$, then we have

$$\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \models_\varepsilon \varphi([a^1], \ldots, [a^n]).$$

PROOF. Before we begin with the proof we note that for any formula ψ, if $\varepsilon_0 \leqslant \varepsilon_1$ and $(\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_0} \psi$, then also $(\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_1} \psi$. This can be directly shown using induction over formulas in prenex normal form, together with the fact that we have assumed every subset of $\mathcal{M}_i$ to be $\mathcal{D}_i$-measurable (see Definition 9.5.8).

[4] Here we also consider $\varepsilon' > 1$, which is interpreted in the same way as in Definition 8.1.1. Of course, this is not necessary if $\varepsilon < 1$, since ε'-truth is equivalent to 1-truth when $\varepsilon' > 1$. However, this way $\varepsilon = 1$ is also included.

We first prove the equivalence of (ii) and (iii). If (ii) holds, let

$$\delta_i = \inf\left\{\varepsilon' > \varepsilon \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \varphi([a^1], \ldots, [a^n])\right\}$$

if this set is non-empty, and 0 otherwise. Note that this set is $\mathcal{U}$-almost always non-empty, as can be seen by applying (ii) with an arbitrary $\varepsilon' > \varepsilon$. Furthermore, (ii) also tells us that the sequence δ_i converges to ε with respect to $\mathcal{U}$. Now, the sequence $\frac{1}{2^i}$ converges to 0, so that the sequence $\varepsilon_i = \min(\delta_i + \frac{1}{2^i}, 1)$ also has $\mathcal{U}$-limit ε. By definition of δ_i, we now have that (iii) holds for the sequence $\varepsilon_0, \varepsilon_1, \ldots$.

Conversely, assume that (iii) holds and fix $\varepsilon' > \varepsilon$. Because the sequence $\varepsilon_0, \varepsilon_1, \ldots$ has $\mathcal{U}$-limit ε, we know that $\{i \in \omega \mid |\varepsilon_i - \varepsilon| < \varepsilon' - \varepsilon\} \in \mathcal{U}$. Using (iii) we therefore see that also

$$\{i \in \omega \mid |\varepsilon_i - \varepsilon| < \varepsilon' - \varepsilon\} \cap \{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_i} \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}$$

and using the observation above we directly see that this set is contained in $\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \varphi(a_i^1, \ldots, a_i^n)\}$, so the latter set is also in $\mathcal{U}$ and we therefore see that (ii) holds.

Next, we simultaneously show the equivalence of (i) with (ii) and (iii) using induction over formulas φ in prenex normal form. For propositional formulas, this proceeds in the same way as the classical proof. For the existential case, we use formulation (iii): using the induction hypothesis, we know that

$$\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \models_\varepsilon \exists x \psi([a^1], \ldots, [a^n], x)$$

is equivalent to saying that there exists an $[a^{n+1}] \in \prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)$ and a sequence $\varepsilon_0, \varepsilon_1, \ldots$ with $\mathcal{U}$-limit ε such that for $\mathcal{U}$-almost all $i \in \omega$ we have that $(\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_i} \psi(a_i^1, \ldots, a_i^n, a_i^{n+1})$, which is in turn equivalent to saying that for the same sequence $\varepsilon_0, \varepsilon_1, \ldots$ we have for $\mathcal{U}$-almost all $i \in \omega$ that $(\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon_i} \exists x \psi(a_i^1, \ldots, a_i^n, x)$.

Finally, consider the universal case. By definition,

$$\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \models_\varepsilon \forall x \psi([a^1], \ldots, [a^n], x)$$

is equivalent to

$$\Pr_{\mathcal{E}}\left(\left\{[a^{n+1}] \,\middle|\, \prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \models_\varepsilon \psi([a^1], \ldots, [a^{n+1}])\right\}\right) \geqslant 1 - \varepsilon.$$

By induction hypothesis, we know that this is equivalent to

$$\Pr_{\mathcal{E}}\left(\bigcap_{\varepsilon' > \varepsilon}\left\{[a^{n+1}] \,\middle|\, \{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \psi(a_i^1, \ldots, a_i^{n+1})\} \in \mathcal{U}\right\}\right) \geqslant 1 - \varepsilon.$$

Because we can restrict ourselves to the (countable) intersection of rational ε', this is the same as having for all $\varepsilon' > \varepsilon$ that

$$\Pr_{\mathcal{E}}\left(\left\{[a^{n+1}] \,\middle|\, \{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \psi(a_i^1, \ldots, a_i^{n+1})\} \in \mathcal{U}\right\}\right) \geqslant 1 - \varepsilon. \tag{31}$$

Observe that the set in (31) is precisely the basic measurable set

$$\left[\{a_i^{n+1} \in \mathcal{M}_i \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \psi(a_i^1, \ldots, a_i^{n+1})\}\right],$$

where we once again use that every subset of $\mathcal{M}_i$ is $\mathcal{D}_i$-measurable. So, using the definition of the ultraproduct measure (Definition 9.5.5), (31) is equivalent to having for every $\delta > 0$ that

$$\bigl\{i \in \omega \bigm| \Pr_{\mathcal{D}_i}\bigl(\bigl\{a_i^{n+1} \in \mathcal{M}_i \bigm| (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \psi(a_i^1, \ldots, a_i^{n+1})\bigr\}\bigr) \geqslant 1 - \varepsilon - \delta\bigr\} \in \mathcal{U}.$$

But this holds for all $\varepsilon' > \varepsilon$ and all $\delta > 0$ if and only if we have for all $\varepsilon' > \varepsilon$ that

$$\bigl\{i \in \omega \bigm| \Pr_{\mathcal{D}_i}\bigl(\bigl\{a_i^{n+1} \in \mathcal{M}_i \bigm| (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \psi(a_i^1, \ldots, a_i^{n+1})\bigr\}\bigr) \geqslant 1 - \varepsilon'\bigr\} \in \mathcal{U}.$$

which is in turn equivalent to having for all $\varepsilon' > \varepsilon$ that

$$\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \forall x \psi(a_i^1, \ldots, a_i^n, x)\} \in \mathcal{U}.$$

This completes the induction. □

COROLLARY 9.5.10. *The ultraproduct is a weak ε-model.*

PROOF. For every formula $\varphi = \varphi(x_1, \ldots, x_n)$ and parameters $[a^1], \ldots, [a^{n-1}]$ in $\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U}$, Theorem 9.5.9 tells us that the subset of $\prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U}$ defined by φ and the parameters $[a^1], \ldots [a^{n-1}]$ is exactly

$$\bigcap_{\varepsilon' > \varepsilon, \varepsilon' \in \mathbb{Q}} \bigl\{[a^n] \in \prod_{i\in\omega}(\mathcal{M}_i, \mathcal{D}_i)/\mathcal{U} \mid \{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_{\varepsilon'} \varphi(a_i^1, \ldots, a_i^n)\} \in \mathcal{U}\bigr\}$$

which is a countable intersection of basic measurable sets, and therefore measurable. □

We remark that this construction, in general, does not yield an ε-model. For example, if we have a binary relation $R(x^1, x^2)$ and on each model $(\mathcal{M}_i, \mathcal{D}_i)$ the relation R consists of the union of two 'boxes' $(X_i \times Y_i) \cup (U_i \times V_i)$, then we would need an uncountable union of boxes of basic measurable sets to form R in the ultraproduct model. This is, of course, not an allowed operation on σ-algebras.

A more formal argument showing that the ultraproduct construction does not necessarily yield ε-models is that this construction allows us to prove a weak compactness result in the usual way. If this would always yield an ε-model, this would contradict Theorem 9.5.2 above.

THEOREM 9.5.11. (Weak compactness theorem) *Let Γ be a countable set of sentences such that each finite subset is satisfied in a weak ε-model. Then there exists a weak ε-model satisfying Γ.*

PROOF. Let $A_0, A_1, \ldots$ be an enumeration of the finite subsets of Γ. For each A_i, fix a weak ε-model $(\mathcal{M}_i, \mathcal{D}_i)$ satisfying all formulas from A_i. Then the filter on ω generated by

$$\bigl\{\{i \in \omega \mid (\mathcal{M}_i, \mathcal{D}_i) \models_\varepsilon \varphi\} \mid \varphi \in \Gamma\bigr\}$$

is a proper filter, so we can use the ultrafilter lemma (see e.g. Hodges [41, Theorem 6.2.1]) to determine an ultrafilter $\mathcal{U}$ on ω containing this filter. If $\mathcal{U}$ is principal, then there exists an $n \in \omega$ with $\{n\} \in \mathcal{U}$ and therefore $(\mathcal{M}_n, \mathcal{D}_n)$ satisfies Γ. Otherwise we form the ultraproduct (where we note that we may assume every subset of $\mathcal{M}_i$ to be $\mathcal{D}_i$-measurable by Theorem 9.5.7, provided we only assume that $\mathcal{D}_i$ is finitely additive). It then follows from Theorem 9.5.9 and Corollary 9.5.10 that this ultraproduct is a weak ε-model that satisfies every $\varphi \in \Gamma$. □

CHAPTER 10

Computational Hardness of Validity in ε-Logic

In this chapter we take a first look at the computational complexity of ε-logic. Terwijn [119] has shown that the set of ε-tautologies is undecidable. This should not be too surprising, because this is of course also the case for classical first-order logic. Terwijn [119] also showed that the 0-tautologies are just the classical tautologies, so 0-validity is in fact Σ^0_1-complete. In this chapter we will show that ε-validity is in fact Π^1_1-hard for $\varepsilon \in (0,1) \cap \mathbb{Q}$. This shows that ε-logic is computationally much harder than first-order logic and that we cannot hope to find an effective calculus for it.

This chapter is based on Kuyper [64].

10.1. Many-one reductions between different ε

In this section we will show that for rational $\varepsilon_0, \varepsilon_1 \in (0,1)$, the ε_0-tautologies many-one reduce to the ε_1-tautologies. Not only does this show that we only need to consider one fixed ε for our hardness results (we will take $\varepsilon = \frac{1}{2}$ below), but in our proof of the Π^1_1-hardness of ε-validity these reductions will also turn out to be useful in a different way.

We will begin with reducing to bigger ε_1. To do this, we refine the argument by Terwijn [119], where it is shown that the 0-tautologies many-one reduce to the ε-tautologies for $\varepsilon \in [0,1)$. Our argument is similar to the one given in section 9.4, where we discussed reductions for satisfiability instead of for validity.

THEOREM 10.1.1. *Let $\mathcal{L}$ be a countable first-order language not containing equality. Then, for all rationals $0 \leq \varepsilon_0 \leq \varepsilon_1 < 1$, the ε_0-tautologies many-one reduce to the ε_1-tautologies.*

PROOF. We can choose integers n and $0 < m \leq n$ so that $\frac{m}{n} = \frac{1-\varepsilon_1}{1-\varepsilon_0}$. Let $\varphi(y_1, \dots, y_k)$ be a formula in prenex normal form (see Proposition 8.1.7). For simplicity we write $\vec{y} = y_1, \dots, y_k$. Also, for a function π we let $\pi(\vec{y})$ denote the vector $\pi(y_1), \dots, \pi(y_k)$. We use formula-induction to define a computable function f such that for every formula φ,

(32) $\qquad \varphi$ is an ε_0-tautology if and only if $f(\varphi)$ is an ε_1-tautology.

For propositional formulas and existential quantifiers, there is nothing to be done and we use the identity map. Next, we consider the universal quantifiers. Let $\varphi = \forall x \psi(\vec{y}, x)$. The idea is to introduce new unary predicates, that can be used to vary the strength of the universal quantifier. We will make these predicates split the model into disjoint parts. If we split it into just the right number of parts (in this case n), then we can choose m of these parts to get just the right strength.

So, we introduce new unary predicates $X_1, \ldots, X_n$. We define the sentence n-split by:

$$\forall x \left((X_1(x) \vee \ldots \vee X_n(x)) \wedge \bigwedge_{1 \leq i < j \leq n} \neg \left(X_i(x) \wedge X_j(x) \right) \right).$$

Then one can verify that in any model, $\neg n$-split does *not* hold if and only if the sets X_i disjointly cover the entire model.

Now define $f(\varphi)$ to be the formula

$$\neg n\text{-split} \vee \bigvee_{i_1, \ldots, i_m} \forall x \big((X_{i_1}(x) \vee \cdots \vee X_{i_m}(x)) \wedge f(\psi)(\vec{y}, x) \big)$$

where the disjunction is over all subsets of size m from $\{1, \ldots, n\}$. (It will be clear from the construction that $f(\psi)$ has the same arity as ψ.) Thus, $f(\varphi)$ expresses that for some choice of m of the n parts, $f(\psi)(x)$ holds often enough when restricted to the resulting part of the model.

We will now prove claim (32) above. For the implication from right to left, we will prove the following strengthening:

For every formula $\varphi(\vec{y})$ and every probability model $(\mathcal{M}, \mathcal{D})$ there exists a probability model $(\mathcal{N}, \mathcal{E})$ together with a measure-preserving surjective measurable function $\pi : \mathcal{N} \to \mathcal{M}$ (i.e. for all $\mathcal{D}$-measurable A we have that $\mathcal{E}(\pi^{-1}(A)) = \mathcal{D}(A)$) such that for all $\vec{y} \in \mathcal{N}$ we have that

$$(\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y}) \text{ if and only if } (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi(\vec{y})).$$

In particular, if $f(\varphi)$ is an ε_1-tautology, then φ is an ε_0-tautology. We prove this by formula-induction over the formulas in prenex normal form. For propositional formulas, there is nothing to be done (we can simply take the models to be equal and π the identity). For the existential quantifier, let $\varphi = \forall x \psi(x)$ and apply the induction hypothesis to ψ to find a model $(\mathcal{N}, \mathcal{E})$ and a mapping π. Then we can take the same model and mapping for φ, as easily follows from the fact that π is surjective.

Next, we consider the universal quantifier. Suppose $\varphi = \forall x \psi(\vec{y}, x)$ and let $(\mathcal{M}, \mathcal{D})$ be a probability model. Use the induction hypothesis to find a model $(\mathcal{N}, \mathcal{E})$ and a measure-preserving surjective measurable function $\pi : \mathcal{N} \to \mathcal{M}$ such that for all $\vec{y}, x \in \mathcal{N}$ we have that

$$(\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\vec{y}, x) \text{ if and only if } (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi(\vec{y}), \pi(x)).$$

Now form the probability model $(\mathcal{N}', \mathcal{E}')$ which consists of n disjoint copies $\mathcal{N}_1, \ldots, \mathcal{N}_n$ of $(\mathcal{N}, \mathcal{E})$, each with weight $\frac{1}{n}$. That is, $\mathcal{E}'$ is the sum of n copies of $\frac{1}{n}\mathcal{E}$. Let $\pi' : \mathcal{N}' \to \mathcal{M}$ be the composition of the projection map $\sigma : \mathcal{N}' \to \mathcal{N}$ with π. Relations in $\mathcal{N}'$ are defined just as on $\mathcal{N}$, that is, for a t-ary relation R we define $R^{\mathcal{N}'}(x_1, \ldots, x_t)$ by $R^{\mathcal{N}}(\sigma(x_1), \ldots, \sigma(x_t))$. Observe that this is the same as defining $R^{\mathcal{N}'}(x_1, \ldots, x_t)$ by $R^{\mathcal{M}}(\pi'(x_1), \ldots, \pi'(x_t))$. We interpret constants $c^{\mathcal{N}'}$ by embedding $c^{\mathcal{N}}$ into the first copy $\mathcal{N}_1$. For functions f of arity t, first note that we can see $f^{\mathcal{N}}$ as a function from $\mathcal{N}^t \to \mathcal{N}'$ by embedding its codomain $\mathcal{N}$ into the first copy $\mathcal{N}_1$. We now interpret $f^{\mathcal{N}'}$ as the composition of this $f^{\mathcal{N}}$ with π'. Finally, we let each X_i be true exactly on the copy $\mathcal{N}_i$.

Then π' is clearly surjective. To show that it is measure-preserving, it is enough to show that σ is measure-preserving. If A is $\mathcal{E}$-measurable, then $\sigma^{-1}(A)$ consists of n disjoint copies of A, each having measure $\frac{1}{n}\mathcal{E}(A)$, so $\pi^{-1}(A)$ has $\mathcal{E}'$-measure exactly $\mathcal{E}(A)$.

Now, since $(\mathcal{N}', \mathcal{E}')$ does not satisfy $\neg n$-split (because the X_i disjointly cover $\mathcal{N}'$), we see that

$$(\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y}) \tag{33}$$

is equivalent to the statement that for all $1 \leq i_1 < \cdots < i_m \leq n$ we have

$$\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} (X_{i_1}(x) \vee \cdots \vee X_{i_m}(x)) \wedge f(\psi)(\vec{y}, x)\big] > \varepsilon_1. \tag{34}$$

By Lemma 9.4.3 above we have that $(\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} f(\psi)(\vec{y}, x)$ holds if and only if $(\mathcal{N}, \mathcal{E}) \models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), \sigma(x))$ holds. In particular, we see for every $1 \leq i \leq n$ that

$$\begin{aligned}
&\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}', \mathcal{E}') \models_{\varepsilon_1} X_i(x) \text{ and } (\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big] \\
&= \frac{1}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big].
\end{aligned}$$

It follows that (34) is equivalent to

$$\frac{n-m}{n} + \frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}), x)\big] > \varepsilon_1.$$

The induction hypothesis tells us that this is equivalent to

$$\frac{n-m}{n} + \frac{m}{n} \Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), \pi(x))\big] > \varepsilon_1$$

and since π is surjective and measure-preserving, this is the same as

$$\begin{aligned}
\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})), x)\big] &> \frac{n}{m}\left(\varepsilon_1 - \frac{n-m}{n}\right) \\
&= \frac{n}{m}(\varepsilon_1 - 1) + 1 = \varepsilon_0.
\end{aligned}$$

This proves that we have $(\mathcal{N}', \mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\pi'(\vec{y}))$.

To prove the left to right direction of (32) we will use induction to prove the following stronger statement:

If $(\mathcal{M}, \mathcal{D})$ is a probability model and $\vec{y} \in \mathcal{M}$ are such that $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$, then we also have $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_0} \varphi(\vec{y})$.

In particular, if φ is an ε_0-tautology, then $f(\varphi)$ is an ε_1-tautology. The only interesting case is the universal case, so let $\varphi = \forall x \psi(\vec{y}, x)$. Let $\vec{y} \in \mathcal{M}$ be such that $(\mathcal{M}, \mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$. Assume, towards a contradiction, that $(\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$. Then

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_0} \psi(\vec{y}, x)\big] \geq 1 - \varepsilon_0$$

and by the induction hypothesis we have

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y}, x)\big] \geq 1 - \varepsilon_0. \tag{35}$$

Because $(\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_1} \neg n$-split, the X_i disjointly cover $\mathcal{M}$, as discussed above. Now, by taking those m of the X_i (say $X_{i_1},\dots,X_{i_m}$) which have the largest intersection with this set we find that

$$\begin{aligned}\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} (X_{i_1} \vee \cdots \vee X_{i_m}) \wedge f(\psi)(\vec{y},x)\big] &\geq \frac{m}{n}(1-\varepsilon_0)\\ &= 1-\varepsilon_1\end{aligned}$$

which contradicts our choice of $(\mathcal{M},\mathcal{D})$. □

Theorem 10.1.2. *Let $\mathcal{L}$ be a countable first-order language not containing equality. Then, for all rationals $0 < \varepsilon_1 \leq \varepsilon_0 \leq 1$, the ε_0-tautologies many-one reduce to the ε_1-tautologies.*

Proof. We can choose integers n and $m < n$ such that $\frac{m}{n} = \frac{\varepsilon_0-\varepsilon_1}{\varepsilon_0}$. We construct a many-one reduction f such that for all formulas φ,

$$\varphi \text{ is an } \varepsilon_0\text{-tautology if and only if } f(\varphi) \text{ is an } \varepsilon_1\text{-tautology.}$$

Again, we only consider the nontrivial case where φ is a universal formula $\forall x\psi(\vec{y},x)$. We define $f(\varphi)$ to be the formula

$$\neg\text{-}n\text{-split} \wedge \bigvee_{i_1,\dots,i_m} \forall x\big(X_{i_1}(x) \vee \cdots \vee X_{i_m}(x) \vee f(\psi)(\vec{y},x)\big)$$

where the disjunction is over all subsets of size m from $\{1,\dots,n\}$.

The proof is almost the same as for Theorem 10.1.1. In the proof for the implication from right to left, follow the proof up to (33), i.e.

$$(\mathcal{N}',\mathcal{E}') \not\models_{\varepsilon_1} f(\varphi)(\vec{y}).$$

This is equivalent to the statement that for all $1 \leq i_1 < \cdots < i_m \leq n$ we have

$$\Pr_{\mathcal{E}'}\big[x \in \mathcal{N}' \mid (\mathcal{N}',\mathcal{E}') \not\models_{\varepsilon_1} X_{i_1}(x) \vee \cdots \vee X_{i_m}(x) \vee f(\psi)(\vec{y},x)\big] > \varepsilon_1.$$

Similar as before, using Lemma 9.4.3, we find that this is equivalent to

$$\frac{n-m}{n}\Pr_{\mathcal{E}}\big[x \in \mathcal{N} \mid (\mathcal{N},\mathcal{E}) \not\models_{\varepsilon_1} f(\psi)(\sigma(\vec{y}),x)\big] > \varepsilon_1.$$

Again, using the induction hypothesis and the fact that π is measure-preserving we find that this is equivalent to

$$\begin{aligned}&\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_0} \psi(\pi(\sigma(\vec{y})),x)\big]\\ &> \frac{n}{n-m}\varepsilon_1 = \frac{\varepsilon_0}{\varepsilon_1}\varepsilon_1 = \varepsilon_0.\end{aligned}$$

This proves that $(\mathcal{N}',\mathcal{E}') \models_{\varepsilon_1} f(\varphi)(\vec{y})$ if and only if $(\mathcal{M},\mathcal{D}) \models_{\varepsilon_0} \varphi(\vec{y})$.

For the converse implication, we also need to slightly alter the proof of Theorem 10.1.1. Assuming that $(\mathcal{M},\mathcal{D}) \not\models_{\varepsilon_1} f(\varphi)(\vec{y})$, follow the proof up to (35), where we obtain

$$\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y},x)\big] \geq 1-\varepsilon_0. \tag{36}$$

Define

$$\eta = \Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\varepsilon_1} f(\psi)(\vec{y},x)\big]$$

and take those m of the X_i (say $X_{i_1}, \ldots, X_{i_m}$) which have the smallest intersection with this set. Note that by (36) we have $\eta \geq 1 - \varepsilon_0$. Then we find that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\big[x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\varepsilon_1} X_{i_1} \vee \cdots \vee X_{i_m} \vee f(\psi)(\vec{y}, x)\big] \\
&\geq \frac{m}{n} + \Big(1 - \frac{m}{n}\Big)\eta = \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}\eta \\
&\geq \frac{\varepsilon_0 - \varepsilon_1}{\varepsilon_0} + \frac{\varepsilon_1}{\varepsilon_0}(1 - \varepsilon_0) = 1 - \varepsilon_1.
\end{aligned}$$

which contradicts our choice of $(\mathcal{M}, \mathcal{D})$. □

Observe that, because of the inductive nature of the reductions above, we can perform these reductions per quantifier. In particular, we can talk about what it means for a formula with variable ε (that is, a separate ε for each quantifier) to be a tautology. This way, we get something like Keisler's probability logic mentioned in the introduction; however, remember that we still have our non-classical negation (unlike Keisler). This idea will be crucial in our hardness proof.

10.2. Validity is Π^1_1-hard

To show that the set of ε-tautologies is indeed Π^1_1-hard, we adapt a proof by Hoover [42] which shows that $\mathcal{L}_{\omega P}$ is Π^1_1-complete. We will show that, to a certain extent, we can define the natural numbers within probability logic.

Definition 10.2.1. Let φ be a formula in prenex normal form and N a unary predicate. Then φ^N, or φ *relativised to* N, is defined as the formula where each $\forall x \psi(x)$ is replaced by $\forall x(N(x) \to \psi(x))$ and each $\exists x \psi(x)$ is replaced by $\exists x(N(x) \wedge \psi(x))$.

Theorem 10.2.2. *Let $\mathcal{L}$ be the language consisting of a constant symbol* 0, *a unary relation $N(x)$, binary relations $x = y$,*[1] *$S(x) = y$ and $R(x, y)$, and ternary relations $x + y = z$ and $x \cdot y = z$. Furthermore, let f be the reduction from* 0-*tautologies to $\frac{1}{2}$-tautologies from Proposition 10.1.1. Then there exists finite theories T, T' in the language $\mathcal{L}$ such that, for every first-order sentence φ containing a new predicate symbol Q, the following are equivalent:*

(i) $\models_{\frac{1}{2}} f(\neg(\bigwedge T)) \vee \neg\left(\bigwedge T'\right) \vee f\left(\neg\varphi^N\right)$;
(ii) $\mathbb{N} \models \forall Q \neg\varphi(Q)$.[2]

Proof. We will prove the contrapositives of the implications (i) $\to$ (ii) and (ii) $\to$ (i). During the entire proof, one should mainly think about what it means for a formula ψ that its negation $\neg\psi$ does not hold. Note that we have that $(\mathcal{M}, \mathcal{D}) \not\models_0 \neg\psi$ if and only if all universal quantifiers hold classically and all existential quantifiers hold on a set of strictly positive measure. Likewise, $(\mathcal{M}, \mathcal{D}) \not\models_{\frac{1}{2}} \neg\psi$ holds if and only if all universal quantifiers hold classically and all existential quantifiers hold on a set of measure strictly greater than $\frac{1}{2}$.

---

[1] Here we do not mean true equality, but rather a binary relation that we will use to represent equality.

[2] We denote by $\forall Q \neg\varphi(Q)$ the second-order formula $\forall X \neg\varphi(X/Q)$, where $\varphi(X/Q)$ is the formula where the predicate symbol Q is replaced by a second-order variable X.

Inspired by this, we form the theories T and T'. T consists of Robinson's Q relativised to N, axioms specifying that the arithmetical relations only hold on N, and some special axioms for N and R. That is, we put the following axioms in T (keeping in mind that we are mostly interested in what happens when the negation of these formulas does not hold, i.e. one should read the $\forall$ as a classical universal quantifier and the $\exists$ as saying that the statement holds on a set of strictly positive measure):

All equality axioms. For example:

$$\forall x(x = x)$$
$$\forall x \forall y((N(x) \wedge x = y) \to N(y))$$

We should guarantee that 0 is in N:

$$N(0)$$

We now give the axioms for the successor function:

$$\forall x \forall y(S(x) = y \to (N(x) \wedge N(y)))$$
$$(\forall x \exists y S(x) = y)^N$$
$$(\forall x \forall y \forall u \forall v((S(x) = y \wedge S(u) = v \wedge x = u) \to y = v))^N$$
$$(\forall x \neg S(x) = 0)^N$$
$$(\forall x(x = 0 \vee \exists y S(y) = x))^N.$$[3]

In the axioms below, we will leisurely denote by $\psi(S(x))$ the formula $\forall y(S(x) = y \to \psi(y))$ and similarly for $x + y$ and $x \cdot y$. We proceed with the inductive definitions of $+$ and $\cdot$:

$$(\forall x \forall y \forall z(x + y = z \to (N(x) \wedge N(y) \wedge N(z))))$$
$$(\forall x(x + 0 = x))^N$$
$$(\forall x \forall y(x + S(y) = S(x + y)))^N$$
$$(\forall x \forall y \forall z(x \cdot y = z \to (N(x) \wedge N(y) \wedge N(z))))^N$$
$$(\forall x(x \cdot 0 = 0))^N$$
$$(\forall x \forall y(x \cdot S(y) = (x \cdot y) + x))^N.$$

Finally, we introduce a predicate R. This predicate is meant to function as a sort of 'padding'. The goal of this predicate is to force the measure of a point $S^n(0)$ to be larger than the measure of $\{x \mid N(x) \wedge x > S^n(0)\}$ (the precise use will be made clear in the proof below).

$$(\forall x \forall y \neg R(x, y))^N$$

[3] We do not really need this last axiom, but we have added it anyway so that all axioms of Robinson's Q are in T.

The last two axioms will be in T' instead of in T, because these need to be evaluated for $\varepsilon = \frac{1}{2}$ while the rest will be evaluated for $\varepsilon = 0$. So, because we will be looking at when the negation does not hold, the existential quantifier should be read as "strictly more than measure $\frac{1}{2}$ many".

$$\begin{aligned}
&\forall x(N(x) \to \exists y(R(x,y) \vee x = y)) \\
&\forall x(N(x) \to \exists y \neg(R(x,y) \vee x < y))
\end{aligned}$$

Here, $x < y$ is short for $f(\exists z(x + S(z) = y))$, i.e. the usual definition of $x < y$ evaluated for $\varepsilon = 0$.

Note that for universal formulas it does not matter if they are in T or T' because in both cases the negation of the formula does not hold if and only if the formula holds classically.

We will now show that these axioms indeed do what we promised. First, we show that (i) implies (ii). So, assume $\mathbb{N} \not\models \forall Q(\neg\varphi(Q))$. Fix a predicate $Q^{\mathbb{N}}$ such that $\mathbb{N} \not\models \neg\varphi(Q)$. Now take the model $\mathcal{M} = \omega \times \{0,1\}$ to be the disjoint union of two copies of ω, where we define $S, +, \cdot, \leq, 0$ on the first copy $\omega \times \{0\}$ of ω as usual, and let these be undefined elsewhere. Let

$$N := \omega \times \{0\} \text{ and } R := \left\{((a,0),(b,1)) \mid \mu k \left[2^{k+1} > 3^{a+1}\right] \neq b\right\}.$$

We let $Q^{\mathcal{M}}(a,0)$ hold if $Q^{\mathbb{N}}(a)$ and we never let it hold on the second copy of ω. Finally, define $\mathcal{D}$ by

$$\mathcal{D}(a,0) = \mathcal{D}(a,1) := \frac{1}{3^{a+1}}.$$

Then it is directly verified that

$$(\mathcal{M}, \mathcal{D}) \not\models_0 \neg\Big(\bigwedge T\Big) \vee \neg\varphi^N,$$

i.e. all formulas in $T \cup \left\{\varphi^N\right\}$ hold in $(\mathcal{M}, \mathcal{D})$ if universal quantifiers are interpreted classically and existential quantifiers as expressing that there exists a set of positive measure. Note that because all points have positive measure this is equivalent to the classical existential quantifier, so all we are really saying is that T and φ^N hold classically in $\mathcal{M}$.

Furthermore, if we let $a \in \omega$ and denote b for $\mu k[2^{k+1} > 3^{a+1}]$ then we have that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\left[y \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\frac{1}{2}} R((a,0),y) \vee (a,0) = y\right] \\
&= \frac{1}{2} - \frac{1}{2^{b+1}} + \frac{1}{3^{a+1}} \\
&> \frac{1}{2}
\end{aligned}$$

while we also have that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\frac{1}{2}} \neg(R((a,0),y) \vee (a,0) < y)\big] \\
&= 1 - \left(\frac{1}{2} - \frac{1}{2^{b+1}} + \sum_{i=a+2}^{\infty} 3^{-i}\right) \\
&= 1 - \frac{1}{2}\left(1 - \frac{1}{2^b} + \frac{1}{3^{a+1}}\right) \\
&> \frac{1}{2}
\end{aligned}$$

where the last inequality follows from the fact that b is the smallest $k \in \omega$ such that $2^{k+1} > 3^{a+1}$, so that $2^b \leq 3^{a+1}$. Thus, we see that $(\mathcal{M}, \mathcal{D}) \not\models_{\frac{1}{2}} \neg(\bigwedge T')$. But then we see from (the proof of) Theorem 10.1.1, together with the remark below Theorem 10.1.2 that there is a probability model $(\mathcal{N}, \mathcal{E})$ such that

$$(\mathcal{N}, \mathcal{E}) \not\models_{\frac{1}{2}} f\left(\neg\Big(\bigwedge T\Big)\right) \vee \neg\Big(\bigwedge T'\Big) \vee f\left(\neg\varphi^N\right),$$

i.e. (i) does not hold.

Conversely, assume that statement (i) does not hold. Without loss of generality, we may assume the equality relation on $\mathcal{M}$ to be true equality; otherwise, because (i) does not hold and all equality axioms are in T we could look at $\mathcal{M}/=$ instead.

Again, from (the proof of) Theorem 10.1.1 we see that

$$(\mathcal{M}, \mathcal{D}) \not\models_0 \neg\Big(\bigwedge T\Big) \vee \neg\varphi^N \text{ and } (\mathcal{M}, \mathcal{D}) \not\models_{\frac{1}{2}} \neg\Big(\bigwedge T'\Big).$$

We will now use the three axioms involving R. Let $m \in \mathcal{M}$ with $\mathcal{M} \models N(m)$. Then $\{a \in \mathcal{M} \mid \mathcal{M} \models a = m\} \subseteq N^{\mathcal{M}}$ by the equality axioms, and similarly $\{a \in \mathcal{M} \mid m < a\} \subseteq N^{\mathcal{M}}$. So the axiom $(\forall x \forall y \neg R(x,y))^N$ tells us that these two sets are disjoint from $\{a \in \mathcal{M} \mid \mathcal{M} \models R(m,a)\}$. Therefore, from the two axioms in T' it now follows that

$$\begin{aligned}
\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models m = a\big] &> \frac{1}{2} - \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models R(m,a)\big] \\
&> \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models m < a\big].
\end{aligned}$$

Thus,

$$\Pr_{\mathcal{D}}\big[\mathcal{M}\big] > \frac{1}{2}\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid m \leq a\big] \tag{37}$$

We now claim that, if we denote $S(x)$ for the unique y such that $S(x) = y$ (as guaranteed to exist and be unique by T):

$$\Pr_{\mathcal{D}}\left[\big\{0, \ldots, S^k(0)\big\}\right] > \left(1 - \frac{1}{2^{k+1}}\right)\Pr_{\mathcal{D}}\big[N\big].$$

For $k = 0$ this is clear: from the axioms in T it follows that for all elements $a \in N$ different from 0 we have $a > 0$, and therefore $\Pr_{\mathcal{D}}\big[\{0\}\big] > \frac{1}{2}\Pr_{\mathcal{D}}\big[N\big]$ by

(37). Similarly, assume this holds for $k \in \omega$. Then we have by (37):

$$\begin{aligned}
&\Pr_{\mathcal{D}}\left[\{0,\ldots,S^{k+1}(0)\}\right]\\
&> \Pr_{\mathcal{D}}\left[\{0,\ldots,S^{k}(0)\}\right] + \frac{1}{2}\left(\Pr_{\mathcal{D}}[N] - \Pr_{\mathcal{D}}\left[\{0,\ldots,S^{k}(0)\}\right]\right)
\end{aligned}$$

so from the induction hypothesis we obtain

$$\begin{aligned}
&> \left(1 - \frac{1}{2^{k+1}}\right)\Pr_{\mathcal{D}}[N] + \frac{1}{2^{k+2}}\Pr_{\mathcal{D}}[N]\\
&= \left(1 - \frac{1}{2^{k+2}}\right)\Pr_{\mathcal{D}}[N].
\end{aligned}$$

Because this converges to $\Pr_{\mathcal{D}}[N]$ if k goes to infinity, we see that all weight of N rests on $X := \{S^n(0) \mid n \in \omega\} \subseteq N$. Now, if some universal quantifier holds when relativised to N, it certainly holds when restricted to X. Furthermore, if some existential quantifier holds with positive measure in N, then it also has to hold with positive measure in X because $X \subseteq N$ has the same measure as N. Therefore, we see that $(\mathcal{M}, \mathcal{D}) \not\models_0 \neg\varphi^N$ implies that also $(\mathcal{M}\restriction X, \mathcal{D}\restriction X) \not\models_0 \neg\varphi^X$ (see the discussion at the beginning of the proof about what it means for the negation of a formula to not hold).

However, we can directly verify that $\mathcal{M}\restriction X$ is isomorphic to the standard natural numbers $\mathbb{N} = (\omega, S, +, \cdot, 0)$. So, by transferring the predicate Q from $\mathcal{M}$ to $\mathbb{N}$ (i.e. letting $Q^{\mathbb{N}}(k)$ hold if $Q^{\mathcal{M}}(S^k(0))$ holds) we find that indeed $\mathbb{N} \not\models \forall Q \neg\varphi(Q)$. □

Putting this together, we reach our conclusion.

THEOREM 10.2.3. *For rational $\varepsilon \in (0,1)$, the set of ε-tautologies is Π^1_1-hard.*

PROOF. From Theorem 10.1.1, Theorem 10.1.2 and Theorem 10.2.2. □

In fact, we have shown that even for languages not containing function symbols or equality, ε-validity is already Π^1_1-hard. Our proof above uses one constant: 0. However, we could also replace 0 by a unary relation representing $0 = x$ and modify the proof to show that the relational fragment of ε-validity is Π^1_1-hard.

We do not yet know of an upper bound for the complexity of ε-validity. While we have developed methods for proving upper bounds for ε-satisfiability, which will be discussed in the next chapter, these methods do not seem to work for proving any results about ε-validity. Thus, the exact complexity of ε-validity is still an open problem.

CHAPTER 11

Computational Hardness of Satisfiability in ε-Logic

In this chapter we study the computational complexity of satisfiability in ε-logic. Unfortunately, it turns out that ε-logic is computationally quite hard: as we already saw in the previous chapter, ε-validity is Π^1_1-hard.

In this chapter, we will mainly study the fragment of ε-logic not containing equality or function symbols, i.e. containing only relation and constant symbols. For this fragment, we will show that ε-satisfiability is, in general, Σ^1_1-complete, refuting a conjecture by Terwijn [119, Conjecture 5.3]. At first one might think that the Σ^1_1-hardness of ε-satisfiability already follows from the Π^1_1-hardness of ε-validity mentioned above: indeed, in the classical case a formula φ is satisfiable if and only if its negation $\neg\varphi$ is not valid. However, because our logic is paraconsistent (i.e. both a formula φ and its negation $\neg\varphi$ can hold at the same time) this complementarity does not hold for ε-satisfiability and ε-validity. Therefore, we need to consider the complexity of these two problems separately.

For $\varepsilon = 0$ the results are vastly different: as mentioned above, 0-validity coincides with classical validity, as shown in Terwijn [119], so it is Σ^0_1-complete. In this chapter we show that 0-satisfiability is decidable, so 0-satisfiability is even easier than classical satisfiability. This also clearly shows that the complexities of 0-validity and 0-satisfiability are not complementary, as argued above. The different complexities for ε-satisfiability and ε-validity are summarised in Table 1 below. Note that the exact complexity of ε-validity is still open, as there is no known matching upper bound for the Π^1_1-hardness.

	$\varepsilon \in (0,1) \cap \mathbb{Q}$	$\varepsilon = 0$
ε-satisfiability	Σ^1_1-complete	decidable
ε-validity	Π^1_1-hard	Σ^0_1-complete

TABLE 1. Complexity of validity and satisfiability in ε-logic.

The structure of this chapter is as follows. First, in section 11.1 we show that a certain weak version of ε-satisfiability is Σ^1_1. In section 11.2 we briefly turn towards the problem of Skolemisation in ε-logic, which we will need for the later results in this chapter. After that, in section 11.3 we show that the relational fragment of ε-satisfiability is Σ^1_1. Next, in section 11.4 we turn to 0-satisfiability and we show that this problem is, quite surprisingly, decidable. This contrasts our result in section 11.5 that ε-satisfiability is Σ^1_1-hard for rational $\varepsilon \in (0,1)$, which completes our proof that ε-satisfiability is Σ^1_1-complete for such ε. Finally, in section 11.6 we use the results from section 11.4 to show that 0-logic is compact, contrasting Theorem 9.5.2 which says that ε-logic is not compact for rational $\varepsilon \in (0,1)$.

This chapter is based on Kuyper [66].

11.1. Towards an upper bound for ε-satisfiability

Our first goal is to show that (the relational fragment of) ε-satisfiability is Σ^1_1. Together with the proof in section 11.5 that ε-satisfiability is Σ^1_1-hard for rational $\varepsilon \in (0,1)$ this will show that ε-satisfiability is Σ^1_1-complete. There are multiple ways one could go about proving this. The first would be to reduce ε-satisfiability to Keisler's logic $\mathcal{L}_{\omega P}$. In Hoover [42], it is shown that validity for this logic is Π^1_1-complete, hence satisfiability for $\mathcal{L}_{\omega P}$ (which is dual to validity in the same way as for classical logic) is Σ^1_1-complete. Keisler proves this through the use of a deduction system with infinitary deduction rules.

We will take a different, more direct approach. We will prove that there is a natural, equivalent formulation of ε-satisfiability of which we can directly see that it is Σ^1_1. This equivalent formulation is the hidden heart of the completeness proof for $\mathcal{L}_{\omega P}$, and our method allows one to grasp the true essence of the proof. Of course, this comes at the price of not having a deduction system, but it is questionable how useful a deduction system with infinitary deduction rules is in the first place. Furthermore, because of the reduction to Keisler's logic, this deduction system would talk about formulas containing Keisler's quantifiers ($Px \geq r$). So, we would lose the advantage of using only the language of first-order logic. We will briefly come back to this point in Remark 11.1.6 and Remark 11.3.6.

As a first step, we will show that a certain weaker form of ε-satisfiability is Σ^1_1.

Definition 11.1.1. A *weak ε-model* is a pair $(\mathcal{M}, \mathcal{D})$ which satisfies the conditions of Definition 8.1.2, except possibly for condition (16).

A *finitely additive model* is a pair $(\mathcal{M}, \mathcal{D})$ consisting of a first-order model $\mathcal{M}$ and a finitely additive measure $\mathcal{D}$ over $\mathcal{M}$.

We say that $\varphi(x_1, \dots, x_n)$ is *weakly ε-satisfiable* (respectively *finite additively ε-satisfiable*) if there exist a weak ε-model (respectively finitely additive model) $(\mathcal{M}, \mathcal{D})$ and $a_1, \dots, a_n \in \mathcal{M}$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \dots, a_n)$.[1]

Note that, by Theorem 9.5.7, every finitely additive measure $\mathcal{D}$ on a set X can be extended to a finitely additive measure $\mathcal{D}'$ on the power set $\mathcal{P}(X)$. This explains why we did not impose any measurability conditions on our finitely additive models: if φ is satisfied in some finitely additive model $(\mathcal{M}, \mathcal{D})$, then it is also satisfied in the model $(\mathcal{M}, \mathcal{D}')$ in which every set is measurable.

[1] Because we did not impose any measurability conditions on our finitely additive model, it could be the case that the set occurring in case (vi) of Definition 8.1.1 is not measurable. For finitely additive models we therefore say that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \varphi(x)$ if $\{a \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a)\}$ is $\mathcal{D}$-measurable and

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a)\big] \geq 1 - \varepsilon,$$

i.e. we let a universal quantifier be false if the corresponding set is not measurable.

At first sight, finitely additive ε-satisfiability might seem to be much weaker than weak ε-satisfiability. Surprisingly, it turns out that they are equivalent.

THEOREM 11.1.2. *A formula φ is weakly ε-satisfiable if and only if it is finite additively ε-satisfiable.*

PROOF. Clearly any weak ε-model is also a finitely additive model, so if φ is weakly ε-satisfiable it is certainly finite additively ε-satisfiable.

For the converse, assume $(\mathcal{M}, \mathcal{D})$ is a finitely additive model ε-satisfying φ. Extend $\mathcal{D}$ to a finitely additive measure $\mathcal{D}'$ on $\mathcal{P}(X)$ using Theorem 9.5.7 and take an ultrapower $(\mathcal{N}, \mathcal{E})$ of $(\mathcal{M}, \mathcal{D}')$ (as defined in Definition 9.5.8). Then $(\mathcal{N}, \mathcal{E})$ is a weak ε-model by Corollary 9.5.10 and it ε-satisfies φ by Theorem 9.5.9. $\square$

In Example 9.1.5 it was shown that, in general, not every ε-satisfiable sentence has a countable ε-model (i.e. the Downward Löwenheim–Skolem theorem does not hold in the usual sense). In contrast, this does hold if we look at finitely additive models.

DEFINITION 11.1.3. Let $\mathcal{D}$ be a probability measure. We say that $\mathcal{D}$ is *countable* if the Boolean algebra on which $\mathcal{D}$ is defined is countable.

THEOREM 11.1.4. (Downward Löwenheim–Skolem theorem for finitely additive ε-satisfiability) *Let $\mathcal{L}$ be a countable first-order language. Let $(\mathcal{M}, \mathcal{D})$ be a finitely additive model and let $X \subseteq \mathcal{M}$ be countable. Then there exists a finitely additive model*

$$(\mathcal{N}, \mathcal{E}) \prec_\varepsilon (\mathcal{M}, \mathcal{D})$$

such that $X \subseteq \mathcal{N}$ and such that $\mathcal{N}$ and $\mathcal{E}$ are countable.

PROOF. By Theorem 9.5.7, we may without loss of generality assume that the domain of $\mathcal{D}$ is $\mathcal{P}(\mathcal{M})$. We will define a sequence $X_0 \subseteq X_1 \subseteq \dots$ of countable subsets of $\mathcal{M}$ and let $\mathcal{N}$ be the restriction of $\mathcal{M}$ to $\bigcup_{n \in \omega} X_n$.

Let X_0 consist of X together with the interpretation $c^{\mathcal{M}}$ of all constants. Next, given X_n, we show how to define X_{n+1}. Let $\mathcal{B}_{n+1}$ be the Boolean algebra generated by the ε-definable subsets of $(\mathcal{M}, \mathcal{D})$ using parameters from X_n. Since X_n is countable, $\mathcal{B}_{n+1}$ will also be countable. Fix an element $x_B \in B$ for every non-empty $B \in \mathcal{B}_{n+1}$. Now let

$$\begin{aligned} X_{n+1} = & X_n \cup \{x_B \mid B \in \mathcal{B}_{n+1}\} \cup \\ & \{f^{\mathcal{M}}(a_1, \dots, a_m) \mid f \in \mathcal{L}, f \text{ is } m\text{-ary and } a_1, \dots, a_m \in X_n\}. \end{aligned}$$

Then X_{n+1} is countable and $X_n \subseteq X_{n+1} \subseteq \mathcal{M}$.

As announced above, we let $\mathcal{N}$ be $\mathcal{M} \restriction \bigcup_{n \in \omega} X_n$. Note that if $a_1, \dots, a_m \in \mathcal{N}$, then there is some $n \in \omega$ such that $a_1, \dots, a_m \in X_n$. Then $f^{\mathcal{M}}(a_1, \dots, a_m) \in X_{n+1} \subseteq \mathcal{N}$ so the functions on $\mathcal{N}$ are well-defined.

For every B which is in the Boolean algebra generated by the ε-definable subsets of $(\mathcal{M}, \mathcal{D})$ using parameters from $\mathcal{N}$, let $\mathcal{E}(B \cap \mathcal{N}) = \mathcal{D}(B)$. Note that such a B uses only finitely parameters and hence is already ε-definable using parameters from some X_n. We show that the finitely additive measure $\mathcal{E}$ is well-defined. To this end, assume that B_1, B_2 in the Boolean algebra described above are such that $B_1 \cap \mathcal{N} = B_2 \cap \mathcal{N}$. Let X_n be such that B_1, B_2 can be ε-defined using parameters from X_n. If $B_1 \neq B_2$, then either $B_1 \cap (\mathcal{M} \setminus B_2) \neq \emptyset$ or $(\mathcal{M} \setminus B_1) \cap B_2 \neq \emptyset$. In both cases, one of these points gets added to X_{n+1}, hence $B_1 \cap \mathcal{N} \neq B_2 \cap \mathcal{N}$, a contradiction. So, $B_1 = B_2$ and therefore $\mathcal{D}(B_1) = \mathcal{D}(B_2)$, as desired.

It remains to show that $(\mathcal{N}, \mathcal{E})$ is an elementary finitely additive ε-submodel of $(\mathcal{M}, \mathcal{D})$, i.e. that for all sequences $a_1, \ldots, a_n \in \mathcal{N}$ and for all formulas $\varphi(x_1, \ldots, x_n)$ we have

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \varphi(a_1, \ldots, a_n) \Leftrightarrow (\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi(a_1, \ldots, a_n).$$

We prove this using induction over the formulas in prenex normal form. For propositional formulas, this is directly clear. For the existential case, note that

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \exists x \psi(x, a_1, \ldots, a_n)$$

clearly implies that this also holds in $(\mathcal{M}, \mathcal{D})$. For the converse, assume

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \exists x \psi(x, a_1, \ldots, a_n).$$

Let X_m be such that $a_1, \ldots, a_n \in X_m$. By construction, we have added a point from the non-empty set

$$\{b \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(b, a_1, \ldots, a_n)\}$$

to $\mathcal{N}$. So, there exists some point $b \in \mathcal{N}$ such that $(\mathcal{N}, \mathcal{E}) \models_\varepsilon \psi(b, a_1, \ldots, a_n)$. Using the induction hypothesis, this directly implies that

$$(\mathcal{N}, \mathcal{E}) \models_\varepsilon \exists x \psi(x, a_1, \ldots, a_n).$$

For the universal case, let $\varphi = \forall x \psi(x, x_1, \ldots, x_n)$. Let

$$\begin{aligned} B &= \{x \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x, b_1, \ldots, b_n)\}, \\ C &= \{x \in \mathcal{N} \mid (\mathcal{N}, \mathcal{E}) \models_\varepsilon \psi(x, b_1, \ldots, b_n)\}. \end{aligned}$$

Then by induction hypothesis we have

$$\begin{aligned} C &= \{x \in \mathcal{N} \mid (\mathcal{M}, \mathcal{D}) \models_\varepsilon \psi(x, b_1, \ldots, b_n)\} \\ &= B \cap \mathcal{N}. \end{aligned}$$

From this we see that $\mathcal{E}(C) = \mathcal{D}(B)$, and hence

$$(\mathcal{M}, \mathcal{D}) \models_\varepsilon \forall x \psi(x, b_1, \ldots, b_n) \Leftrightarrow (\mathcal{N}, \mathcal{E}) \models_\varepsilon \forall x \psi(x, b_1, \ldots, b_n).$$

This concludes the induction. □

THEOREM 11.1.5. *Let $\mathcal{L}$ be a countable language and let ε be rational. Then weak ε-satisfiability is Σ^1_1.*

PROOF. From Theorem 11.1.2 and Theorem 11.1.4 we see that a formula φ is weakly ε-satisfiable if and only if it is ε-satisfied in a countable finitely additive model. We sketch how to express the latter as a Σ^1_1-formula. Namely, we can express this by saying:

- there exist interpretations $R^{\mathcal{M}} \subseteq \omega^n$ for the relations R and interpretations for the constants (note that this can be done through one second-order existential quantifiers, by using a pairing function);
- there exist $Q \subseteq \omega$ (representing the Boolean algebra) and $\mathcal{D} : \omega \to \mathbb{R}$;

such that:

- the sets $Q_n := \{m \in \omega \mid \langle n, m \rangle \in Q\}$ for $n \in \omega$ form a Boolean algebra;

- the function mapping Q_n to $\mathcal{D}(n)$ is a finitely additive measure;
- φ is ε-satisfied in $(\mathcal{M}, \mathcal{D})$.

It can be directly verified that all the items in the second list can be expressed as arithmetical formulas. Combining this with the existential second-order quantifiers from the first list, we obtain the desired result. □

REMARK 11.1.6. In Keisler's logic $\mathcal{L}_{\omega P}$, finitely additive models can be obtained by taking a maximal consistent set of formulas in the calculus he calls *weak* $\mathcal{L}_{\omega P}$ (see Keisler [51, Definition 1.4.1 and Theorem 1.5.3]). By extending his logic with rules for the existential quantifier, we would be able to obtain something similar: note that $\varphi = \forall x_1 \exists x_2 \ldots \forall x_n \psi(x_1, \ldots, x_n)$ is ε-satisfiable if and only if $(Px_1 \geq 1 - \varepsilon)\exists x_2 \ldots (Px_n \geq 1 - \varepsilon)\psi(x_1, \ldots, x_n)$ is satisfiable. In fact, we can base our system on the finitary fragment of weak $\mathcal{L}_{\omega P}$ (i.e. we only need Keisler's axioms A1-A5) since our formulas correspond to formulas in Keisler's positive fragment. From this we find that weak ε-satisfiability is in fact even Π^0_1. Since we are not interested in weak ε-satisfiability and merely see it as a stepping stone for our results on (regular) ε-satisfiability, which we have already shown to be Σ^1_1-hard, we will not pursue this matter any further.

Our next goal is to extend Theorem 11.1.5 to regular ε-satisfiability, instead of just weak ε-satisfiability. Before we do so, we will first look at some results concerning Skolemisation in ε-logic.

11.2. Skolemisation in ε-logic

For classical satisfiability, we can eliminate existential quantifiers by introducing Skolem functions. Clearly, we can also do something similar here: for example, if a statement $\forall x \exists y \varphi(x, y)$ is 0-true in some model $(\mathcal{M}, \mathcal{D})$, we can extend $\mathcal{M}$ to a model $\mathcal{N}$ such that $(\mathcal{N}, \mathcal{D}) \models_0 \forall x \varphi(x, f(x))$ holds. We do not even need that $\mathcal{M} \models \varphi(x, f(x))$ holds for every $x \in \mathcal{M}$ for which $\mathcal{M} \models \exists y \varphi(x, y)$ holds — it is enough if this is true for almost all such x.

However, we have required our functions to be measurable. It is not directly clear that we can also pick our Skolem functions in a measurable way. Our next result shows that this is possible.

THEOREM 11.2.1. *Let φ be a formula not containing function symbols. Then φ is ε-satisfiable if and only if its Skolemisation is ε-satisfiable.*

PROOF. We need to show that we can pick the Skolem functions in a measurable way. First, assume φ does not contain the equality symbol. Now, if φ is ε-satisfiable, then by Theorem 9.2.9 we may assume it is ε-satisfied in a model $(\mathcal{M}, \lambda)$ on $[0, 1]$ with the Lebesgue measure, with Borel relations. Then all definable sets are analytic by Proposition 9.2.1.

Let $\varphi = \forall x_1 \exists y_1 \ldots \forall x_n \exists y_n \psi(x_1, \ldots, x_n, y_1, \ldots, y_n)$ be in prenex normal form. Now, we use recursion over $1 \leq i \leq n$ to define Borel measurable Skolem functions $f_1, f_2, \ldots, f_n$, by performing the following steps:

(1) Use the Jankov and von Neumann Uniformisation Theorem (see e.g. Kechris [50, Theorem 18.1][2]) to find a Lebesgue-measurable *uniformising function* g_i: that is, for all $x_1, x_2, \dots, x_i \in [0,1]$, if there exists an $y_i \in [0,1]$ such that

$$(\mathcal{M}, \lambda) \models_\varepsilon \forall x_{i+1} \dots \exists y_n \psi(x_1, \dots, x_n, \\ f_1(x_1), \dots, f_{i-1}(x_1, \dots, x_{i-1}), y_i, \dots, y_n),$$

then we have, if we denote $y = g_i(x_1, x_2, \dots, x_i)$:

$$(\mathcal{M}, \lambda) \models_\varepsilon \forall x_{i+1} \dots \exists y_n \psi(x_1, \dots, x_n, \\ f_1(x_1), \dots, f_{i-1}(x_1, \dots, x_{i-1}), y, y_{i+1}, \dots, y_n).$$

(2) Let f_i be a Borel function which is equal to g_i almost everywhere; that such functions exist is shown in Bogachev [10, Proposition 2.1.11], or can easily be proven using Lusin's Theorem on measurable functions (see e.g. Kechris [50, Theorem 17.12][3]).

In the case that φ does contain the equality symbol, we may assume it is satisfied by a model on $[0, r]$ plus atoms by Theorem 9.2.9. We can apply the construction above on $[0, r]$ and use AC to choose the values f_i takes in the atoms; since there are only countably many atoms, this does not alter the measurability of the f_i. □

It turns out we can do even more. We assumed that our formula φ did not contain any function symbols, but now its Skolemisation does contain function symbols. The next proposition shows that we can also construct a kind of Skolemisation which does not need function symbols.

THEOREM 11.2.2. *Let $\mathcal{L}$ be a language not containing equality or function symbols (but it may contain constant symbols). Then there exists a language $\mathcal{L}'$ only containing relation symbols and a computable function mapping each formula φ in the language $\mathcal{L}$ to a universal formula φ' in the language $\mathcal{L}'$ such that for every $\varepsilon \in [0,1]$: φ is ε-satisfiable if and only if φ' is ε-satisfiable.*

PROOF. The idea of the proof is roughly as follows: because universal quantifiers only talk about measure, we can always change the interpretation of a relation on a set of measure zero without affecting the truth of the universal quantifiers. Thus, we can always add witnesses of measure zero, as long as there are no contradicting statements about a witness. Let R be a relation symbol, say of arity k. To express which statements hold about which witnesses, we add a new copy of R for every atomic formula of the form $R(t_1, \dots, t_k)$ up to a permutation of the universally bound variables. We then form the formula φ' by replacing the atomic formula by this new copy of R. If φ is satisfiable, then φ' is satisfiable just by interpreting the new relation symbol by the interpretation of $R(t_1, \dots, t_k)$.

[2]Kechris only states that we get a partial function with as domain exactly the (analytic) set of those $x_1, \dots, x_i$ for which an y_i as above exists, but we can easily extend this to a total Lebesgue measurable function by letting its value be 0 outside this set.

[3]Kechris only states this theorem for Borel measures, but it holds for the Lebesgue measure with exactly the same proof.

Conversely, if φ' is satisfiable, we pick our witnesses in a set of measure zero (say, the Cantor set) and recombine all the relations into one relation R. We will now give the proof in full detail.

Let R be a relation symbol, say of arity k. First, we enlarge our language by adding a copy of R for every atomic formula of the form $R(t_1, \ldots, t_k)$, up to a permutation of the universally bound variables. To this end, fix an enumeration $c_0, c_1, \ldots$ of the constant symbols in $\mathcal{L}$. Let A_R be the set of functions from $\{1, \ldots, k\}$ to $\{0\} \cup (\{1,2\} \times \omega)$. To every atomic formula $R(t_1, \ldots, t_k)$ we assign a function $\alpha \in A_R$ by letting $\alpha(i) = 0$ if t_i equals some (universally bound) x_j, by letting $\alpha(i) = (1, j)$ if t_i equals (the existentially bound) y_j and by letting $\alpha(i) = (2, j)$ if t_i equals the constant symbol c_j. During the proof, we will say that the α constructed in this way is the α *corresponding to the atomic formula* $R(t_1, \ldots, t_k)$. Now, for every relation symbol R and every $\alpha \in A_R$, we introduce a new relation symbol R_α, where the arity of R_α equals $|\{1 \leq i \leq k \mid \alpha(i) = 0\}| + \max\{j \in \omega \mid (1, j) \in \operatorname{ran}(\alpha)\}$.

Let $\varphi = \forall x_1 \exists y_1 \ldots \exists y_n \forall x_n \psi(x_1, \ldots, x_n, y_1, \ldots, y_n)$ be a formula in $\mathcal{L}$ in prenex normal form. We describe how to form $\varphi' = \forall x_1 \ldots \forall x_n \psi'(x_1, \ldots, x_n)$ from φ. Let $R(t_1, \ldots, t_k)$ be an atomic formula occurring in φ and let α be the function corresponding to this atomic formula, as described above. Let $m = \max\{j \in \omega \mid (1, j) \in \operatorname{ran}(\alpha)\}$ and let $s_1, \ldots, s_{k'}$ be the subsequence of $t_1, \ldots, t_k$ consisting of just those t_i which equal a universally bound variable x_j (during the proof, we will say that $s_1, \ldots, s_{k'}$ is the *universal subsequence of* $t_1, \ldots, t_k$). Then we form ψ' from ψ by replacing each $R(t_1, \ldots, t_k)$ with $R_\alpha(s_1, \ldots, s_{k'}, x_1, \ldots, x_m)$.

We show that φ' is as desired. First, assume φ is ε-satisfiable. So, by Theorem 9.2.9 there exists an ε-model $(\mathcal{M}, \lambda)$ with the Lebesgue measure λ and Borel relations ε-satisfying φ. By the previous proposition, for each y_i we can find a Borel Skolem function $f_i : [0,1]^i \to [0,1]$. Let R be a relation symbol and let $\alpha \in A_R$. We interpret $R_\alpha^{\mathcal{N}}(s_1, \ldots, s_{k'}, x_1, \ldots, x_m)$ as $R^{\mathcal{M}}(v_1, \ldots, v_k)$, where

$$v_i := \begin{cases} s_j & \text{if } \alpha(i) = 0 \text{ and } |\{s \leq j \mid \alpha(s) = 0\}| = i \\ f_j(x_1, \ldots, x_j) & \text{if } \alpha(i) = (1, j) \\ c_j & \text{if } \alpha(i) = (2, j). \end{cases}$$

We claim that $R^{\mathcal{N}}$ is both analytic and co-analytic; then it is Borel by Souslin's theorem (see e.g. Kechris [50, Theorem 14.11]). Let β be the function such that

$$\beta_i(s_1, \ldots, s_{k'}, y_1, \ldots, y_m) := \begin{cases} s_j & \text{if } \alpha(i) = 0 \text{ and } |\{s \leq j \mid \alpha(s) = 0\}| = i \\ y_j & \text{if } \alpha(i) = (1, j) \\ c_j & \text{if } \alpha(i) = (2, j), \end{cases}$$

i.e. β 'restores' the atomic formula $R(t_1, \ldots, t_k)$ in the sense that $R(t_1, \ldots, t_k)$ is the same as $R(\beta(s_1, \ldots, s_{k'}, x_1, \ldots, x_m))$. Finally, let B_j be the graph of f_j. Then we have

$$\begin{aligned} R_\alpha^{\mathcal{N}} = \{(s_1, \ldots, s_{k'}, x_1, \ldots, x_m) \mid \exists y_1 \ldots y_m (&\beta(s_1, \ldots, s_{k'}, y_1, \ldots, y_m) \in R^{\mathcal{M}} \wedge \\ &(x_1, y_1) \in B_1 \wedge \cdots \wedge (x_1, \ldots, x_m, y_m) \in B_m)\} \end{aligned}$$

so it is analytic, and

$$\begin{aligned} R_\alpha^{\mathcal{N}} = \{(s_1, \dots, s_{k'}, x_1, \dots, x_m) \mid \forall y_1 \dots y_m \\ ((x_1, y_1) \in B_1 \wedge \dots \wedge (x_1, \dots, x_m, y_m) \in B_m) \\ \to \beta(s_1, \dots, s_{k'}, y_1, \dots, y_m) \in R^{\mathcal{M}}\} \end{aligned}$$

so it is also co-analytic. Therefore, $(\mathcal{N}, \lambda)$ is an ε-model by Proposition 9.2.1.

We know that

$$(\mathcal{M}, \lambda) \models_\varepsilon \forall x_1 \dots \forall x_n \psi(x_1, \dots, x_n, f_1(x_1), \dots, f_n(x_1, \dots, x_n))$$

because the f_i are Skolem functions. Let $R(t_1, \dots, t_k)$ be an atomic formula with corresponding α and let $s_1, \dots, s_{k'}$ be the universal subsequence of $t_1, \dots, t_k$. Let $R(u_1, \dots, u_k)$ be the atomic formula where $f_i(x_1, \dots, x_i)$ is substituted for each y_i. Then

$$\begin{aligned} &\{(x_1, \dots, x_n) \in \mathcal{M}^n \mid R^{\mathcal{M}}(u_1, \dots, u_k)\} \\ &= \{(x_1, \dots, x_n) \in \mathcal{N}^n \mid R_\alpha^{\mathcal{N}}(s_1, \dots, s_{k'}, x_1, \dots, x_m)\} \end{aligned} \tag{38}$$

by the definition of $R_\alpha^{\mathcal{N}}$. Then it follows using formula induction that

$$\begin{aligned} &\{(x_1, \dots, x_n) \in \mathcal{M}^n \mid \mathcal{M} \models \psi(x_1, \dots, x_m, f_1(x_1), \dots, f_n(x_1, \dots, x_n))\} \\ &= \{(x_1, \dots, x_n) \in \mathcal{N}^n \mid \mathcal{N} \models \psi'(x_1, \dots, x_m)\}, \end{aligned}$$

where the only interesting case is the atomic case, which is exactly (38). So, we see that also $(\mathcal{N}, \lambda) \models_\varepsilon \varphi'$ holds.

Conversely, assume we have a model $(\mathcal{N}, \lambda)$ with the Lebesgue measure and Borel relations satisfying φ'. We now need to put all the $R_\alpha^{\mathcal{N}}$ together into one $R^{\mathcal{M}}$. The basic idea is that, because universal quantifiers only talk about measure, we can glue the relations together on a set of measure zero. We will pick all of our witnesses inside Cantor space, which has measure 0 inside $[0, 1]$. To this end, let $\zeta : \mathcal{C} \to \bigcup_{1 \le i \le n} [0, 1]^i$ be a Borel isomorphism of Cantor space (as subset of $[0, 1]$) with copies of unit boxes of increasing dimension (that such an isomorphism exists follows from Kechris [50, Theorem 15.6]). Also fix an injective function $\eta : \omega \hookrightarrow [0, 1] \setminus \mathcal{C}$, which we will use to interpret the constant symbols.

We construct a model $\mathcal{M}$ on $[0, 1]$. For any constant symbol c_i, let $c_i^{\mathcal{M}} = \eta(i)$. Next we show how to define $R^{\mathcal{M}}$ of arity k. Let $a_1, \dots, a_k \in [0, 1]$. Let $\beta(i) = (1, j)$ if $a_i \in \mathcal{C}$ and $\zeta(a_i)$ has length j, let $\beta(i) = (2, j)$ if $a_i \in \operatorname{ran}(\eta)$ and $\eta^{-1}(a_i) = j$ and finally let $\beta(i) = 0$ if neither of these cases hold. We say that this β *corresponds to the sequence* $a_1, \dots, a_k$. Let $b_1, \dots, b_{k'}$ be the subsequence of $a_1, \dots, a_k$ obtained by taking just those a_i satisfying $\beta(i) = 0$. Let $m = \max\{j \in \omega \mid (1, j) \in \operatorname{ran}(\beta)\}$ and let i be the least $1 \le j \le k$ such that $\beta(j) = (1, m)$. Now let $R^{\mathcal{M}}(a_1, \dots, a_k)$ be defined as $R_\beta^{\mathcal{N}}(b_1, \dots, b_{k'}, \zeta_1(a_i), \dots, \zeta_m(a_i))$. We can prove that $R^{\mathcal{N}}$ is Borel using Souslin's theorem, in the same way as above.

Next, let $R(t_1, \dots, t_k)$ be an atomic formula and let α correspond to this atomic formula . Furthermore, let $s_1, \dots, s_{k'}$ be the universal subsequence of $t_1, \dots, t_k$. Let $m = \max\{j \in \omega \mid (1, j) \in \operatorname{ran}(\alpha)\}$. Let $b_1, \dots, b_k \in [0, 1] \setminus (\mathcal{C} \cup \operatorname{ran}(\eta))$ and consider the sequence $a_i = t_i^{\mathcal{M}}[x_j := b_j, y_j := \zeta^{-1}(b_1, \dots, b_j)]$. Then it is directly verified that the β corresponding to the sequence $a_1, \ldots, a_k$ is equal to α. Thus, by the definition of $R^{\mathcal{N}}$ we see that

$$R^{\mathcal{M}}(t_1, \ldots, t_k)[x_i := b_i, y_i := \zeta^{-1}(b_1, \ldots, b_i)]$$

holds if and only if

$$R^{\mathcal{N}}_\alpha(s_1, \ldots, s_{k'}, x_1, \ldots, x_m)[x_i := b_i]$$

holds. So, from the construction of ψ' we see that

$$\{\vec{x} \in ([0,1] \setminus (\mathcal{C} \cup \mathrm{ran}(\eta)))^n \mid \mathcal{M} \models \psi(x_1, \ldots, x_n, \zeta^{-1}(x_1), \ldots, \zeta^{-1}(x_1, \ldots, x_n))\}$$
$$= \{\vec{x} \in ([0,1] \setminus (\mathcal{C} \cup \mathrm{ran}(\eta)))^n \mid \mathcal{N} \models \psi'(x_1, \ldots, x_n)\}.$$

Because $\mathrm{ran}(\eta)$ is countable we know that $\mathcal{C} \cup \mathrm{ran}(\eta)$ has Lebesgue-measure 0, so using the fact that φ' holds in $(\mathcal{N}, \lambda)$ we then directly see that

$$(\mathcal{M}, \lambda) \models_\varepsilon \forall x_1 \forall x_2 \ldots \forall x_n \psi(x_1, \ldots, x_n, \zeta^{-1}(x_1), \ldots, \zeta^{-1}(x_1, \ldots, x_n))$$

holds and therefore $(\mathcal{M}, \lambda) \models_\varepsilon \varphi$ holds, as desired. $\square$

11.3. Satisfiability is Σ^1_1

In the previous section we saw that every formula in a language not containing equality or function symbols is equisatisfiable to a universal relational formula, through a computable transformation. Therefore, in our proof of the Σ^1_1 upper bound for ε-satisfiability without equality or function symbols, we will only need to consider the universal relational fragment.

In Section 11.1 we already showed that weak ε-satisfiability is Σ^1_1. However, it is not directly clear how to extend Theorem 11.1.5 to regular ε-satisfiability: the relations in the ε-model we built there using an ultrapower will, in general, not be measurable. Thus, if we want a result similar to Theorem 11.1.2 for ε-satisfiability, we will need to impose extra conditions on our finitely additive models. Our extra conditions will be motivated by the next lemma.

LEMMA 11.3.1. *Let $\mathcal{D}$ be a probability measure on a set X and let R be a $\mathcal{D}^n$-measurable set. Then for every $m \in \omega$ there exists $k \in \omega$ and $\mathcal{D}$-measurable sets $X_{i,j}$ for $1 \leq i \leq k$, $1 \leq j \leq n$ such that for every function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$:*

$$\Pr_{\mathcal{D}^n}\big[(a_1, \ldots, a_n) \in \mathcal{M}^n \mid (a_{f(1)}, \ldots, a_{f(n)}) \in (R^{\mathcal{M}}$$
$$\triangle((X_{1,1} \times \cdots \times X_{1,n}) \cup \cdots \cup (X_{k,1} \times \cdots \times X_{k,n})))\big] \leq \frac{1}{m}.$$

PROOF. Fix an enumeration $f_1, \ldots, f_s$ of all functions from $\{1, \ldots, n\}$ to $\{1, \ldots, n\}$. Let $H(R)$ be the statement of the lemma, i.e. $H(R)$ states that: *for every $m \in \omega$ there exists $k \in \omega$ and $\mathcal{D}$-measurable sets $X_{i,j}$ for $1 \leq i \leq k$, $1 \leq j \leq n$ such that for every function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$:*

$$\Pr_{\mathcal{D}^n}\big[(a_1, \ldots, a_n) \in \mathcal{M}^n \mid (a_{f(1)}, \ldots, a_{f(n)}) \in (R^{\mathcal{M}}$$
$$\triangle((X_{1,1} \times \cdots \times X_{1,n}) \cup \cdots \cup (X_{k,1} \times \cdots \times X_{k,n})))\big] \leq \frac{1}{m}.$$

We show that the class of subsets $R \subseteq X^n$ for which $H(R)$ holds is a monotone class containing all finite unions of boxes, from which then follows that $H(R)$ holds for all $\mathcal{D}^n$-measurable sets R by the monotone class theorem (see e.g. Bogachev [10, Theorem 1.9.3]).

Clearly, $H(R)$ holds if R is a finite union of boxes. Next, let $R = \bigcap_{j \in \omega} R_j$ with $R_0 \supseteq R_1 \supseteq \dots$ be a monotone decreasing sequence and assume that $H(R_j)$ holds for all these R_j. Fix $m \in \omega$. By countable additivity we can determine, for each $1 \leq i \leq s$ an $l_i \in \omega$ such that

$$\Pr_{\mathcal{D}^n}\left[(a_1, \dots, a_n) \in \mathcal{M}^n \mid (a_{f_i(1)}, \dots, a_{f_i(n)}) \in \left(\bigcap_{j \leq l_i} R_j\right) \triangle R\right] \leq \frac{1}{2m}.$$

Let l be the maximum of these l_i. By our hypotheses $H(R_j)$ we can determine for each $j \in \omega$ a $k_j \in \omega$ and measurable sets $X^j_{u,v}$ such that for every $1 \leq i \leq s$ we have

$$\begin{aligned}\Pr_{\mathcal{D}^n}\big[&(a_1, \dots, a_n) \in \mathcal{M}^n \mid (a_{f_i(1)}, \dots, a_{f_i(n)}) \in (R_j^{\mathcal{M}} \\ &\triangle((X^j_{1,1} \times \dots \times X^j_{1,n}) \cup \dots \cup (X^j_{k_j,1} \times \dots \times X^j_{k_j,n}))\big] \leq \frac{1}{2^{j+1}m}.\end{aligned}$$

Now consider the set

$$A = \bigcap_{j \leq l}((X^j_{1,1} \times \dots \times X^j_{1,n}) \cup \dots \cup (X^j_{k_j,1} \times \dots \times X^j_{k_j,n})).$$

Then A is of the desired form (i.e. it can be written as a finite union of Cartesian products): by distributivity, A is equal to a finite union of expressions of the form

$$(X^1_{u_1,1} \times \dots \times X^1_{u_1,k_1}) \cap \dots \cap (X^l_{u_l,1} \times \dots \times X^l_{u_l,k_l})$$

and this expression is equal to

$$(X^1_{u_1,1} \cap \dots \cap X^l_{u_l,1}) \times \dots \times (X^1_{u_1,n} \cap \dots \cap X^l_{u_l,n}).$$

Furthermore, it is also directly verified that for every $1 \leq i \leq s$ we have

$$\Pr_{\mathcal{D}^n}\big[(a_1, \dots, a_n) \in \mathcal{M}^n \mid (a_{f_i(1)}, \dots, a_{f_i(n)}) \in R \triangle A\big] \leq \frac{1}{2m} + \sum_{j=1}^{l} \frac{1}{2^{j+1}m} \leq \frac{1}{m}.$$

So, this shows that $H(R)$ holds.

The case in which R is the union of a monotone increasing sequence can be proven in a similar way, which completes the proof. $\square$

Now, the important idea is that this property of 'having finite approximations' can be expressed in the language of ε-logic, and is hence preserved under taking ultrapowers. Therefore, if we take an ultrapower as in the proof of Theorem 11.1.2, this ultrapower will also possess these finite approximations. By taking a suitable limit of a sequence approximating the relation R we will obtain a *measurable* relation S that coincides with R almost everywhere. The next lemma expresses that such an approximation is good enough for our purposes.

LEMMA 11.3.2. *Let $(\mathcal{M},\mathcal{D})$ be an ε-model and let R,S be two n-ary relation symbols such that for every function $f:\{1,\dots,n\}\to\{1,\dots,n\}$ we have:*

$$(\mathcal{M},\mathcal{D})\models_0 \forall x_1\dots\forall x_n(R(x_{f(1)},\dots,x_{f(n)})\leftrightarrow S(x_{f(1)},\dots,x_{f(n)})).$$

Then for every universal formula φ not containing function symbols or constant symbols, if we let φ' be the formula where every occurrence of R is replaced by S: $(\mathcal{M},\mathcal{D})\models_\varepsilon \varphi$ if and only if $(\mathcal{M},\mathcal{D})\models_\varepsilon \varphi'$.

PROOF. First we prove using formula induction that for every propositional formula $\psi(x_1,\dots,x_m)$:

$$(\mathcal{M},\mathcal{D})\models_0 \forall x_1\dots\forall x_n(\psi(x_1,\dots,x_n)\leftrightarrow \psi'(x_1,\dots,x_n)).$$

For atomic formulas this follows from our assumption: since our language does not contain function or constant symbols, every atomic subformula of φ is of the form $R(y_1,\dots,y_n)$ where $y_1,\dots,y_n\in\{x_1,\dots,x_n\}$. Now let f be the function such that $y_i=x_{f(i)}$ and apply the hypothesis. The other cases in the formula induction are direct.

Because of our definition of $\to$, which is only classical on the propositional level, we see that $\alpha\leftrightarrow\beta$ which is defined as $(\alpha\to\beta)\wedge(\beta\to\alpha)$ is also only classical on the propositional level. To ease our notation, for the rest of the proof we let $\Leftrightarrow$ be the connective expressing ε-equivalence, i.e. we define $(\mathcal{M},\mathcal{D})\models_\varepsilon \alpha\Leftrightarrow\beta$ if either $(\mathcal{M},\mathcal{D})\models_\varepsilon \alpha\wedge\beta$ or $(\mathcal{M},\mathcal{D})\not\models_\varepsilon \alpha\vee\beta$. In particular $\Leftrightarrow$ coincides with $\leftrightarrow$ on the propositional level, so we see that

$$(\mathcal{M},\mathcal{D})\models_0 \forall x_1\dots\forall x_n(\psi(x_1,\dots,x_n)\Leftrightarrow \psi'(x_1,\dots,x_n)).$$

Next, we note that universal quantifiers distribute over $\Leftrightarrow$ in 0-logic, i.e. we have that $(\mathcal{M},\mathcal{D})\models_0 \forall x(\alpha(x)\Leftrightarrow\beta(x))$ if and only if $(\mathcal{M},\mathcal{D})\models_0 (\forall x\alpha(x))\Leftrightarrow(\forall x\beta(x))$. From this we see that for every propositional formula ψ we have

$$(\mathcal{M},\mathcal{D})\models_0 (\forall x_1\dots\forall x_n\psi(x_1,\dots,x_n))\Leftrightarrow(\forall x_1\dots\forall x_n\psi'(x_1,\dots,x_n)),$$

i.e. for every universal formula φ we have that $(\mathcal{M},\mathcal{D})\models_0 \varphi\Leftrightarrow\varphi'$. From this fact the statement of the theorem directly follows. □

Before we continue with the completeness proof, we note the following independently interesting corollary of the previous two lemmas.

COROLLARY 11.3.3. *If a formula φ not containing equality, function and constant symbols is ε-satisfiable, then it is also ε-satisfiable in a model on $[0,1]$ with the Lebesgue measure where we can choose each relation to be either $\mathbf{\Pi^0_2}$ or $\mathbf{\Sigma^0_2}$.*

PROOF. Let φ be ε-satisfiable; for now we assume that φ is a universal relational formula. By Theorem 9.2.9 it is then also ε-satisfied in a model $(\mathcal{M},\lambda)$ on $[0,1]$ with the Lebesgue measure. For every relation R of arity n and every $m\in\omega$ use Lemma 11.3.1 to determine $k_m\in\omega$ and Lebesgue-measurable sets $X_{m,i,j}$ such that for all functions $f:\{1,\dots,n\}\to\{1,\dots,n\}$ we have

$$\Pr_{\mathcal{D}^n}\big[(a_1,\dots,a_n)\in\mathcal{M}^n \mid (a_{f(1)},\dots,a_{f(n)})\in(R^{\mathcal{M}}$$
$$\triangle((X_{m,1,1}\times\dots\times X_{m,1,n})\cup\dots\cup(X_{m,k_m,1}\times\dots\times X_{m,k_m,n})))\big]\le\frac{1}{2^m}.$$

For every such $X_{m,i,j}$, let $Y_{m,i,j}$ be a $\mathbf{\Sigma^0_1}$ set such that $X_{m,i,j} \triangle Y_{m,i,j}$ has Lebesgue-measure 0 (see e.g. Kechris [50, Theorem 17.10]). Now let $\mathcal{N}$ be the model where the relations are taken to be the lim sup of these approximations; i.e. let

$$R^{\mathcal{N}} = \bigcap_{M \in \omega} \bigcup_{m \geq M} ((Y_{m,1,1} \times \cdots \times Y_{m,1,n}) \cup \cdots \cup (Y_{m,k_m,1} \times \cdots \times Y_{m,k_m,n})).$$

Then $\Pr_{\mathcal{D}^n}\left[R^{\mathcal{N}} \triangle R^{\mathcal{M}}\right] = 0$, see e.g. Bogachev [10, Theorem 1.12.6], so $(\mathcal{M}, \mathcal{D})$ and $(\mathcal{N}, \mathcal{D})$ satisfy the same universal formulas by Lemma 11.3.2. Furthermore, the relations in $\mathcal{N}$ are clearly $\mathbf{\Pi^0_2}$.

In the general case, let φ be any formula and let φ' be the universal relational formula from Proposition 11.2.2. Then φ' is satisfiable in a model with $\mathbf{\Pi^0_2}$ relations by our argument above, and from the proof of Proposition 11.2.2 we directly see that in fact then also φ is satisfiable in a model with $\mathbf{\Pi^0_2}$ relations.

Finally, if we want our relations to be $\mathbf{\Sigma^0_2}$, we can take the lim inf of $\mathbf{\Pi^0_1}$ sets $Y_{m,i,j}$ instead of the lim sup of $\mathbf{\Sigma^0_1}$ sets $Y_{m,i,j}$ we took above. □

We will now formalise the ideas discussed above to obtain an analogue of Theorem 11.1.2 for regular ε-satisfiability.

THEOREM 11.3.4. *Let φ be a universal formula in a relational language $\mathcal{L}$. Then the following are equivalent:*

(i) *φ is ε-satisfiable;*
(ii) *there exists a finitely additive model $(\mathcal{M}, \mathcal{D})$ such that $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$, and such that for every $m \in \omega$ and every relation R occurring in φ there exists $k_m \in \omega$ and interpretations $X^{\mathcal{M}}_{m,i,j} \subseteq \mathcal{M}$ for $1 \leq i \leq k_m$, $1 \leq j \leq n$ which satisfy that for every function $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ we have that:*

$$(\mathcal{M}, \mathcal{D}) \models_{\frac{1}{m}} \forall x_1 \ldots \forall x_n \left(R(x_1, \ldots, x_n) \leftrightarrow \bigvee_{i=1}^{k_m} \bigwedge_{j=1}^{n} X_{m,i,j}\left(x_{f(j)}\right) \right). \tag{39}$$

PROOF. The implication from (i) to (ii) directly follows from Lemma 11.3.1. For the converse, we note that (39) can be expressed in the language $\mathcal{L}'$ which consists of $\mathcal{L}$ enlarged with countably many unary predicate symbols $X_{m,i,j}$. Thus, if we have some finitely additive model satisfying (ii), then by Theorem 11.1.2 we can find a weak ε-model $(\mathcal{M}', \mathcal{D}')$ which also satisfies (ii). For every $m \in \omega$, fix $k_m \in \omega$ and interpretations $X^{\mathcal{M}'}_{m,i,j} \subseteq \mathcal{M}$ as in (ii). For ease of notation, write

$$Z_m = \left(\left(X^{\mathcal{M}'}_{m,1,1} \times \cdots \times X^{\mathcal{M}'}_{m,1,n} \right) \cup \cdots \cup \left(X^{\mathcal{M}'}_{m,k_m,1} \times \cdots \times X^{\mathcal{M}'}_{m,k_m,n} \right) \right).$$

From (39) we directly see that

$$\Pr_{\mathcal{D}^n}\left[Z_m \triangle Z_{m'}\right] \leq \left(\frac{1}{m} + \frac{1}{m'} \right)^n.$$

In particular, we see that $Z_0, Z_1, \ldots$ is a Cauchy sequence in the pseudometric $d(X, Y) = \Pr_{\mathcal{D}^n}\left[X \triangle Y\right]$. Since this metric is complete (see e.g. Bogachev [10, Theorem 1.12.6]) we can determine a $\mathcal{D}^n$-measurable set Z_R to which this sequence converges.

Now, let $\varepsilon' > 0$. Fix m such that $\frac{1}{m} < \frac{\varepsilon'}{2}$ and such that $d(Z_m, Z_R) < \frac{\varepsilon'}{2}$. Then from (39) it easily follows that

$$(\mathcal{M}', \mathcal{D}) \models_{\varepsilon'} \forall x_1 \ldots \forall x_n (R(x_1, \ldots, x_n) \leftrightarrow Z_R(x_1, \ldots, x_n)).$$

So, from the proof of Terwijn [119, Proposition 3.4] it then follows that

$$(\mathcal{M}', \mathcal{D}) \models_0 \forall x_1 \ldots \forall x_n (R(x_1, \ldots, x_n) \leftrightarrow Z_R(x_1, \ldots, x_n)).$$

Now, if we let $\mathcal{N}$ be the model where $R^{\mathcal{N}} = Z_R$, then $(\mathcal{M}', \mathcal{D})$ and $(\mathcal{N}, \mathcal{D})$ ε-satisfy the same universal formulas by Lemma 11.3.2. In particular, $(\mathcal{N}, \mathcal{D})$ is an ε-model which ε-satisfies φ. $\square$

THEOREM 11.3.5. *Let $\mathcal{L}$ be a countable language not containing equality and function symbols. Then ε-satisfiability is Σ^1_1.*

PROOF. We only need to consider universal relational formulas by Theorem 11.2.2. From Theorem 11.1.4 and Theorem 11.3.4 we see that a universal relational formula φ is ε-satisfiable if and only if there is a countable finitely additive model as in (ii) of Theorem 11.3.4. However, this last statement can be expressed as a Σ^1_1-formula, in a similar way as explained in the proof of Theorem 11.1.5. $\square$

REMARK 11.3.6. As in Remark 11.1.6, we can actually make finitely additive models as in Theorem 11.3.4 (ii) by using a variant of Keisler's calculus. The extra requirement (39) corresponds to adding an extra deduction rule with countable infinitely many hypotheses. We will not go into more detail here, for the reasons explained in the beginning of section 11.1.

11.4. Decidability of 0-satisfiability

In this section we will prove that 0-satisfiability (for languages not containing equality and function symbols) is not only Σ^1_1, but is in fact decidable. This stands in stark contrast to the fact that for $\varepsilon \in (0, 1)$ we have that ε-satisfiability is Σ^1_1-complete, as will be shown in section 11.5. It also contrasts the fact that 0-validity is undecidable: the 0-tautologies are exactly the classical tautologies, as shown in Terwijn [119].

As in the previous section, we only need to consider universal relational formulas because of Theorem 11.2.2. First, we show that 0-satisfiability corresponds to classical satisfiability in a natural way.

PROPOSITION 11.4.1. *Let $m \in \omega$. If a universal relational formula*

$$\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$$

with $n \leq m$ has an ε-model $(\mathcal{M}, \mathcal{D})$, and $\frac{m!}{(m-n)!}(1 - (1 - \varepsilon)^n) < 1$, then there exists a finite, classical model $\mathcal{N}$ of size m which satisfies

$$\tilde{\varphi} = \forall x_1 \ldots \forall x_n (\bigwedge_{i<j} x_i \neq x_j \to \psi(x_1, \ldots, x_n))$$

classically. In particular, this holds if $\varepsilon = 0$.

PROOF. By Theorem 9.2.9 we may assume $\mathcal{D}$ to be the Lebesgue measure λ. In particular the diagonals have measure 0. Since $(\mathcal{M}, \mathcal{D}) \models_\varepsilon \varphi$ we therefore have that, for each injective function $f : \{1, \ldots, n\} \hookrightarrow \{1, \ldots, m\}$, the set

$$A_f := \{(a_1, \ldots, a_m) \in \mathcal{M}^n \mid a_i \neq a_j \text{ for } i \neq j \text{ and } \mathcal{M} \models \psi(a_{f(1)}, \ldots, a_{f(n)})\}$$

has $\mathcal{D}^m$-measure at least $(1 - \varepsilon)^n$. Let

$$B = \bigcap_{f:\{1,\ldots,n\}\hookrightarrow\{1,\ldots,m\}} A_f.$$

Then we see that the set

$$\mathcal{M} \setminus B = \bigcup_{f:\{1,\ldots,n\}\hookrightarrow\{1,\ldots,m\}} (\mathcal{M} \setminus A_f)$$

has measure at most $\frac{m!}{(m-n)!}(1 - (1 - \varepsilon)^n)$, which is strictly smaller than 1 by assumption. So, we see that B has positive measure. In particular B is non-empty, so choose any $(a_1, \ldots, a_m) \in B$. Then $\mathcal{M}$ restricted to $\{a_1, \ldots, a_m\}$ classically satisfies $\tilde{\varphi}$: indeed, if $b_1, \ldots, b_n \in \{a_1, \ldots, a_m\}$ are distinct elements, and we let f be the function sending each $1 \leq i \leq n$ to the unique $1 \leq j \leq m$ such that $b_i = a_j$, then $(a_1, \ldots, a_m) \in A_f$ implies that $\mathcal{M} \models \psi(b_1, \ldots, b_n)$. $\square$

We would also like to have a converse to this proposition, i.e. we would like to know if there exists an $m \in \omega$ such that if $\forall x_1 \ldots \forall x_n (\bigwedge_{i<j} x_i \neq x_j \to \psi(x_1, \ldots, x_n))$ has a finite classical model of size m, then φ also has a 0-model. It turns out that we can do this if we choose m big enough, by using Ramsey's theorem. In fact, the way we will use Ramsey's theorem is very similar to the original use of this theorem by Ramsey in [97], where Ramsey proved his combinatorial theorem in order to prove that the variant of the Entscheidungsproblem asking if a universal relational formula has an infinite model is decidable. First, we need a definition.

DEFINITION 11.4.2. Let $\mathcal{M}$ be a first-order model, let $X \subseteq \mathcal{M}$ and let $<$ be a linear ordering of X. Then we call $(X, <)$ a *sequence of indiscernibles* if for every $n \in \omega$ and all sequences $a_1 < a_2 < \cdots < a_n$, $b_1 < b_2 < \cdots < b_n$ both in X we have for every formula $\varphi(x_1, \ldots, x_n)$ that $\mathcal{M} \models \varphi(a_1, \ldots, a_n)$ if and only if $\mathcal{M} \models \varphi(b_1, \ldots, b_n)$.

THEOREM 11.4.3. (Ramsey [97, page 279]) *Let $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$ be a universal formula not containing function and constant symbols. Let $\mathcal{M}$ be a classical model of size n which satisfies φ and for which there exists a linear order $<$ on $\mathcal{M}$ which turns $(\mathcal{M}, <)$ into a sequence of indiscernibles. Let $(X, \prec)$ be any linearly ordered set. Then there exists a model $\mathcal{N}$ for φ on X which has $(X, \prec)$ as a sequence of indiscernibles.*[4]

COROLLARY 11.4.4. *If for a propositional relational formula ψ the formula*

$$\tilde{\varphi} = \forall x_1 \ldots \forall x_n (\bigwedge_{i<j} x_i \neq x_j \to \psi(x_1, \ldots, x_n))$$

[4]This is not entirely the way in which Ramsey formulated his theorem, but in fact his proof directly yields us the result we stated here.

has a model $\mathcal{M}$ of size n for which there exists a linear order $<$ on $\mathcal{M}$ which turns $(\mathcal{M}, <)$ into a sequence of indiscernibles, then

$$\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$$

has a 0-model.

PROOF. By the previous theorem, there exists a model on $[0,1]$ satisfying $\tilde{\varphi}$ which has $([0,1], <)$ (where $<$ is the usual ordering on $[0,1]$) as a sequence of indiscernibles. Because the diagonal has Lebesgue-measure 0, we then directly see that also $(\mathcal{N}, \lambda) \models_0 \varphi$. The only thing we still need to verify is that $(\mathcal{N}, \lambda)$ is a 0-model. By Proposition 9.2.1 it is enough to prove that the relations $R^{\mathcal{N}}$ are Borel. Because $([0,1], <)$ is a sequence of indiscernibles we have for every relation R of arity k:

$$R^{\mathcal{N}} = \bigcup \big\{ \{(a_1, \ldots, a_k) \in [0,1]^k \mid \forall 1 \leq i,j \leq k (a_i < a_j \leftrightarrow f(i) < f(j))\} \\ \mid f : \{1, \ldots, k\} \to \{1, \ldots, k\} \text{ and } R^{\mathcal{M}}(b_{f(1)}, \ldots, b_{f(k)}) \big\}.$$

Then $R^{\mathcal{N}}$ is equal to a finite union of sets of the form

$$\{(a_1, \ldots, a_k) \in [0,1]^k \mid \forall 1 \leq i,j \leq k (a_i < a_j \leftrightarrow f(i) < f(j))\}$$

which are Borel because the ordering $<$ on $[0,1]$ is Borel. So, $R^{\mathcal{N}}$ is Borel, as desired. □

The following result follows from the finite Ramsey theorem, together with the fact that there are computable upper bounds for the Ramsey numbers.

THEOREM 11.4.5. (Ramsey [97, page 279]) *There exists a computable $f : \omega^3 \to \omega$ such that for every universal relational formula $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$ containing k relation symbols of arity at most m, if φ has a classical model of size at least $f(n,k,m)$, then it also has a model containing a sequence of indiscernibles of size n.*

Putting all things together, we obtain:

THEOREM 11.4.6. *There exists a computable function $f : \mathrm{Form} \to \omega$ such that for every universal relational formula $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$ the following are equivalent:*

(i) *φ is 0-satisfiable;*
(ii) *φ is ε-satisfiable for some $\varepsilon \in [0,1]$ satisfying $\frac{f(\varphi)!}{(f(\varphi)-n)!}(1-(1-\varepsilon)^k) < 1$;*
(iii) *$\tilde{\varphi} = \forall x_1 \ldots \forall x_n \bigwedge_{i<j}(x_i \neq x_j \to \psi(x_1, \ldots, x_n))$ has a classical model of size $f(\varphi)$;*
(iv) *$\tilde{\varphi}$ has a classical model containing a sequence of indiscernibles of size n.*

PROOF. Let $\tilde{f}$ be the computable function from Proposition 11.4.5; this function directly induces a computable function $f : \mathrm{Form} \to \omega$. We prove the equivalences.

(i) $\to$ (ii): This is directly clear.
(ii) $\to$ (iii): This follows from Proposition 11.4.1.
(iii) $\to$ (iv): This was shown in Theorem 11.4.5.
(iv) $\to$ (i): Finally, this was shown in Corollary 11.4.4. □

In particular, we see:

THEOREM 11.4.7. *0-satisfiability is decidable for languages not containing equality and function symbols.*

PROOF. By Theorem 11.2.2, we only need to consider universal relational formulas φ. Let f be the computable function as in Theorem 11.4.6. If we want to check if φ is 0-satisfiable, then by Theorem 11.4.6 (iii) we only need to check if

$$\tilde{\varphi} = \forall x_1 \dots \forall x_n \bigwedge_{i<j} (x_i \neq x_j \to \psi(x_1, \dots, x_n))$$

has a classical model of size $f(\varphi)$, which is a decidable property. $\square$

There is another interesting fact which follows from Theorem 11.4.6: the equivalence of (ii) and (i) says that a formula φ which is ε-satisfiable for small enough $\varepsilon > 0$ is in fact also 0-satisfiable. Thus, one can see 0-satisfiability as the limit of ε-satisfiability for $\varepsilon > 0$, in the sense given below.

COROLLARY 11.4.8. $\bigcap_{\varepsilon>0} \varepsilon\text{-SAT} = 0\text{-SAT}$.

PROOF. From Theorem 11.4.6. $\square$

Note that we cannot have in general that $\bigcap_{\varepsilon'>\varepsilon} \varepsilon'\text{-SAT} = \varepsilon\text{-SAT}$. Namely, in that case, one could prove using Lemma 11.3.1 that a universal relational sentence φ is ε-satisfiable if and only if it has a finite ε'-model for every rational $\varepsilon' > \varepsilon$. However, the latter can be expressed as a first-order arithmetical sentence, while the universal relational fragment of ε-satisfiability will be shown to be Σ^1_1-hard in Theorem 11.5.13.

11.5. Satisfiability is Σ^1_1-hard

In this section we will show that ε-satisfiability is Σ^1_1-hard for rational $\varepsilon \in (0,1)$. Together with the result from section 11.3 this will show that ε-satisfiability is Σ^1_1-complete. We will prove this hardness step by step, interpreting more and more of arithmetic within ε-logic as we go. As a first step, we will look at sentences of the form $\exists Q \varphi(Q)$ in the language of arithmetic (that is, the language consisting of $S, +, \cdot, 0$ and $=$) where φ is universal, in the sense the only quantifiers occurring in φ are first-order universal quantifiers. Equivalently, we can see $\varphi(Q)$ as a first-order universal sentence in φ in the language of arithmetic enlarged with a unary predicate Q, and $\exists Q \varphi(Q)$ is satisfiable in second-order arithmetic if and only if φ is satisfiable in first-order arithmetic (i.e. there exists some interpretation $Q^{\mathbb{N}}$ such that $\mathbb{N} \models \varphi$ under this interpretation for Q). We will implicitly use this equivalence throughout this section.

Furthermore, to optimise our result and show that we do not need function or constant symbols in our language to prove hardness (i.e. to show that the relational fragment is already Σ^1_1-hard), we will not look at $S, +, \cdot$ and 0 as functions or constants, but instead as relations $S(x) = y, x + y = z, x \cdot y = z$ and $0 = x$. It is easy to see that any universal sentence φ in the language with functions $S, +, \cdot$ and 0 can be transformed into a universal sentence in the language with relations for $S, +, \cdot$ and 0. Henceforth, when we talk about the language of arithmetic, we will mean the language with relation symbols for $S, +, \cdot, 0$ and $=$.

DEFINITION 11.5.1. In this chapter, the *language of arithmetic* is the language consisting of relation symbols $S(x) = y, x + y = z, x \cdot y = z, 0 = x$ and $x = y$. For a formula φ in the language of arithmetic, we say that $\mathbb{N} \models \varphi$ if φ holds in the natural numbers ω together with the usual interpretations for $S, +, \cdot, 0$ and $=$.

To formulate the first step in our interpretation, we need a few more tools. Remember that φ^N is φ relativised to N, see Definition 10.2.1.

DEFINITION 11.5.2. Let φ be a formula in prenex normal form and N a unary predicate. Then φ^N, or φ *relativised to* N, is defined as the formula where each $\forall x \psi(x)$ is replaced by $\forall x(N(x) \to \psi(x))$ and each $\exists x \psi(x)$ is replaced by $\exists x(N(x) \wedge \psi(x))$.

DEFINITION 11.5.3. Let $\psi(x_1, \ldots, x_n)$ be a propositional formula. Then we denote by $\psi^\#(x_1, \ldots, x_n)$ the formula

$$\bigwedge_{f:\{1,\ldots,n\}\to\{1,\ldots,n\}} \psi(x_{f(1)}, \ldots, x_{f(n)}).$$

If $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$ with ψ a propositional formula, then we denote by $\varphi^\#$ the formula $\forall x_1 \ldots \forall x_n \psi^\#(x_1, \ldots, x_n)$.

Note that for every universal sentence φ, if $\mathcal{M} \models \varphi$ holds classically, then also $\mathcal{M} \models \varphi^\#$. It is the following converse which we will need.

LEMMA 11.5.4. *For every propositional formula $\psi(x_1, \ldots, x_n)$ and every countable model $\mathcal{M}$: if there exists an enumeration $\{a_0, a_1, \ldots\}$ of $\mathcal{M}$ such that for all injective functions $\pi : \{1, \ldots, n\} \hookrightarrow \omega$ we have $\mathcal{M} \models \psi^\#(a_{\pi(1)}, \ldots, a_{\pi(n)})$, then $\mathcal{M} \models \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$.*

PROOF. To derive a contradiction, assume $\mathcal{M} \not\models \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$. Let $a_{i_1}, \ldots, a_{i_n} \in \mathcal{M}$ (not necessarily distinct) be such that $\mathcal{M} \not\models \psi(a_{i_1}, \ldots, a_{i_n})$. Fix a subset $A \subseteq \omega$ of size n such that $\{i_1, \ldots, i_n\} \subseteq A$ and fix a bijection $\pi : \{1, \ldots, n\} \to A$. Finally, let $f : \{1, \ldots, n\} \to \{1, \ldots, n\}$ be the function mapping $1 \leq j \leq n$ to $\pi^{-1}(i_j)$. Then $i_j = \pi(f(j))$ so $\mathcal{M} \not\models \psi(a_{\pi(f(1))}, \ldots, a_{\pi(f(n))})$. Therefore $\mathcal{M} \not\models \psi^\#(a_{\pi(1)}, \ldots, a_{\pi(n)})$, a contradiction. $\square$

PROPOSITION 11.5.5. *Let $\mathcal{L}$ be a relational language, let φ be a universal formula in the language $\mathcal{L}$ and let $(\mathcal{M}, \mathcal{D}) \models_0 \varphi^\#$. Then there exists a countable submodel $\mathcal{N} \subseteq \mathcal{M}$ such that $\mathcal{N} \models \varphi$ classically. In fact, for $\mathcal{D}^\omega$-almost all $(a_0, a_1, \ldots) \in \mathcal{M}^\omega$ we have that $\mathcal{M} \upharpoonright \{a_0, a_1, \ldots\} \models \varphi$ classically.*

PROOF. Let $\varphi = \forall x_1 \ldots \forall x_n \psi(x_1, \ldots, x_n)$. For each injective function $\pi : \{1, \ldots, n\} \hookrightarrow \omega$, the set

$$\{(a_0, a_1, \ldots) \in \mathcal{M}^\omega \mid (\mathcal{M}, \mathcal{D}) \models_0 \psi^\#(a_{\pi(1)}, \ldots, a_{\pi(n)})\}$$

is measurable (by Proposition 8.1.3). Furthermore, one can easily see that it in fact has measure 1, by using Fubini's theorem and the fact that $(\mathcal{M}, \mathcal{D}) \models_0 \varphi^\#$; for this it is essential that π is injective.

Therefore the set

$$B = \bigcap_{\pi:\{1,\ldots,n\}\hookrightarrow\omega} \{(a_0, a_1, \ldots) \in \mathcal{M}^\omega \mid (\mathcal{M}, \mathcal{D}) \models_0 \psi^\#(a_{\pi(1)}, \ldots, a_{\pi(n)})\}$$

is a countable intersection of sets of measure 1, and therefore has measure 1 itself. From Lemma 11.5.4 we know that for every $(a_0, a_1, \dots) \in B$ we have that φ holds classically in $\mathcal{M} \restriction \{a_0, a_1, \dots\}$, which completes our proof. $\square$

THEOREM 11.5.6. *Let $\mathcal{L}$ be the language consisting of unary relation symbols $0 = x$ and $N(x)$ (which will represent our set of natural numbers), binary relation symbols $S(x) = y$, $x = y$,*[5] *$x < y$, $x \prec y$ (which will represent $x < S(y)$) and $R(x, y)$*[6]*, ternary relation symbols $x + y = z$ and $x \cdot y = z$, and a unary relation symbol Q. Furthermore, let f_0 be the reduction from 0-satisfiability to $\frac{1}{2}$-satisfiability from Theorem 9.4.1, and let $f_{\frac{1}{4}}$ and $f_{\frac{3}{4}}$ be similar reductions for $\frac{1}{4}$- and $\frac{3}{4}$-satisfiability. Then there exists finite theories $T_0, T_{\frac{1}{4}}, T_{\frac{1}{2}}$ and $T_{\frac{3}{4}}$ in the language $\mathcal{L}$, containing only universal sentences, such that for every first-order universal sentence φ in the language of arithmetic enlarged with Q,*[7] *the following are equivalent:*

(i) *$f_0((\bigwedge T_0 \wedge \varphi^N)^\#) \wedge f_{\frac{1}{4}}(\bigwedge T_{\frac{1}{4}}) \wedge f_{\frac{3}{4}}(\bigwedge T_{\frac{3}{4}}) \wedge \bigwedge T_{\frac{1}{2}}$ is $\frac{1}{2}$-satisfiable;*[8]

(ii) *$\mathbb{N} \models \exists Q \varphi(Q)$.*

PROOF. For T_0 we take the universal axioms of Robinson's Q relativised to N, axioms specifying that our relations only hold on N, axioms for $<$ and $\prec$, and some special axioms for N and R. Some of the axioms will turn out to be redundant, but we have added them anyway so that all axioms of Robinson's Q are in T_0. The reason we only add the universal axioms and we avoid the axioms involving existential quantifiers is that, while we can say that something exists (using the $\exists$ quantifier), we cannot say that something holds with strictly positive measure. Our relation R is meant as a trick to work around this problem.

Thus, we will add the following axioms to T_0.

All equality axioms. For example:

$$\forall x(x = x)$$

$$\forall x \forall y((N(x) \wedge x = y) \to N(y))$$

We should guarantee that 0 is in N:

$$\forall x(0 = x \to N(x))$$

[5] Here we do not mean true equality, but rather a binary relation that we will use to represent equality.

[6] The intended interpretation of $R(x, y)$ is quite technical. It is best to think of our model $\mathcal{M}$ as consisting of two copies of ω: one living inside N, one living outside N. However, all the operations will only be defined on the elements of N; there is no extra structure on $\mathcal{M} \setminus N^{\mathcal{M}}$. Our relation R will then be a subset of $N^{\mathcal{M}} \times (\mathcal{M} \setminus N^{\mathcal{M}})$; the intended interpretation is then that $R(a, b)$ holds if $a \in N^{\mathcal{M}}$, $b \in \mathcal{M} \setminus N^{\mathcal{M}}$ and $b \neq S(a)$.

[7] Remember, as defined above this is the language consisting of (relation symbols for) $S, +, \cdot, 0, =$ and Q.

[8] Strictly speaking, $(\bigwedge T_0 \wedge \varphi^N)^\#$ is undefined because it is not in prenex normal form. To avoid the problem that prenex normal forms are not unique, we assume it has been transformed into prenex normal form using some fixed algorithm (for example, the algorithm arising from the proof of Proposition 8.1.7).

We now give the axioms for the successor function:

$$\forall x \forall y (S(x) = y \to N(x) \wedge N(y))$$
$$(\forall x \forall y \neg (S(x) = y \wedge 0 = y))^N$$
$$(\forall x \forall y \forall u \forall v ((S(x) = u \wedge S(y) = v) \to (x = y \leftrightarrow u = v)))^N$$

We proceed with the inductive definitions of $+$ and $\cdot$:

$$(\forall x \forall y \forall z (x + y = z \to (N(x) \wedge N(y) \wedge N(z))))$$
$$(\forall x \forall y \forall z (0 = y \to (x + y = z \leftrightarrow x = z)))^N$$
$$(\forall x \forall y \forall u \forall v \forall w ((S(y) = u \wedge x + y = w) \to (x + u = v \leftrightarrow S(w) = v)))^N$$
$$(\forall x \forall y \forall z (x \cdot y = z \to (N(x) \wedge N(y) \wedge N(z))))^N$$
$$(\forall x \forall y \forall z (0 = y \to (x \cdot y = z \leftrightarrow 0 = z)))^N$$
$$(\forall x \forall y \forall u \forall v \forall w ((S(y) = u \wedge x \cdot y = w) \to (x \cdot u = v \leftrightarrow w + x = v)))^N.$$

Normally, we would be able to use an existential quantifier and the other relations to define $x < y$ and $x \prec y$ (the latter standing for $S(x) < y$). However, because we are trying to avoid the existential quantifier in order to make the proof work, we give axioms for $<$ and $\prec$ which we need in our proof. It is easy to see that these axioms are sound. Because our formula φ does not contain $<$ or $\prec$, it is irrelevant if they completely define $<$ and $\prec$ or not, so we only add those axioms which we need in our proof.

$$\forall x \forall y (x < y \to N(x) \wedge N(y))$$
$$\forall x \forall y (x \prec y \to N(x) \wedge N(y))$$
$$(\forall x \forall y \neg (S(x) = y \wedge x \prec y))^N$$
$$(\forall x \forall y (x < y \leftrightarrow (x \prec y \vee S(x) = y)))^N$$
$$(\forall x \forall y \neg (0 = y \wedge x < y))^N$$
$$(\forall x \forall y (0 = x \vee (0 = y \to y < x)))^N$$
$$(\forall x \forall y \forall z (S(x) = z \to (x \prec y \leftrightarrow z < y)))^N$$

Finally, we introduce a predicate R, with axioms which use $<$ and $\prec$. This predicate is meant to function as a sort of 'padding'. The goal of this predicate is to force the measure of $\{x \mid x = S(y)\}$ to be exactly half of $\{x \mid x > y\}$. Since we will also add an axiom saying that the set of points equal to 0 has measure $\frac{1}{4}$, this means the set of points equal to $S^n(0)$ will have measure 2^{-n-2}. We will use this to show that the collection of 'standard submodels' of $\mathcal{M}$ has measure 1. The intended interpretation was described above. First, we want the second coordinate of R to only hold outside of N.

$$\forall x \forall y (R(x, y) \to \neg N(y))$$

The next axioms for R will be in $T_{\frac{1}{2}}$ instead of in T_0, because these need to be evaluated for $\varepsilon = \frac{1}{2}$ while the rest will be evaluated for $\varepsilon = 0$. So, this means in these axioms the universal quantifier will mean "for measure at least $\frac{1}{2}$ many" instead of the interpretation "for almost all" in T_0.

$$\begin{aligned}
&\forall x(N(x) \wedge \forall y(R(x,y) \vee S(x) = y))\\
&\forall x(N(x) \wedge \forall y\neg(R(x,y) \vee S(x) = y))\\
&\forall x(N(x) \wedge \forall y(R(x,y) \vee x \prec y))\\
&\forall x(N(x) \wedge \forall y\neg(R(x,y) \vee x \prec y)).
\end{aligned}$$

We also add axioms to $T_{\frac{1}{2}}$ to make sure that N has measure $\frac{1}{2}$:

$$\begin{aligned}
&\forall x N(x)\\
&\forall x \neg N(x).
\end{aligned}$$

Finally, we add the next two axioms to $T_{\frac{1}{4}}$ respectively $T_{\frac{3}{4}}$ to ensure that the points equal to 0 together have measure $\frac{1}{4}$:

$$\begin{aligned}
&(\forall x(0 = x))^N\\
&(\forall x\neg(0 = x))^N
\end{aligned}$$

We will now show that these axioms indeed do what we promised.

(ii) $\rightarrow$ (i): Assume $\mathbb{N} \models \exists Q\varphi(Q)$. Let $Q^{\mathbb{N}}$ witness this fact. Now take the model $\mathcal{M} := \omega \times \{0,1\}$ to be the disjoint union of two copies of ω, where we define $S, +, \cdot, \leq, \prec, 0$ on the first copy $\omega \times \{0\}$ of ω as in $\mathbb{N}$ (remembering that $x \prec y$ should mean $S(x) < y$) and let them be false elsewhere. For example, $(a,0)+(b,0) = (c,0)$ holds if and only if $a + b = c$, and it is false in all other cases. Let $Q^{\mathcal{M}}((a,0))$ hold if $Q^{\mathbb{N}}(a)$ holds and let it be false elsewhere. Let

$$N = \omega \times \{0\} \text{ and } R = \{((a,0),(b,1)) \mid b \neq S(a)\}.$$

Finally, define $\mathcal{D}$ by

$$\mathcal{D}(a,0) = \mathcal{D}(a,1) := \frac{1}{2^{a+2}}.$$

Then it is directly verified that

$$(\mathcal{M},\mathcal{D}) \models_0 \left(\bigwedge T_0 \wedge \varphi^N\right)^{\#},$$

because all formulas in T_0 even hold classically in $\mathcal{M}$.

Furthermore, N clearly has measure $\frac{1}{2}$ and the points equal to $0 = (a_0)$ have measure $\frac{1}{4}$, so the axioms expressing this fact hold; in particular $(\mathcal{M},\mathcal{D}) \models_{\frac{1}{4}} \bigwedge T_{\frac{1}{4}}$ and $(\mathcal{M},\mathcal{D}) \models_{\frac{3}{4}} \bigwedge T_{\frac{3}{4}}$. Next, if we let $a \in \omega$ then we have that

$$\begin{aligned}
&\Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\frac{1}{2}} R((a,0),y) \vee S(a,0) = y\big]\\
&= \frac{1}{2} - \frac{1}{2^{a+3}} + \frac{1}{2^{a+3}}\\
&= \frac{1}{2}
\end{aligned}$$

while we also have that

$$\begin{aligned}&\Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid (\mathcal{M},\mathcal{D}) \models_{\frac{1}{2}} R((a,0),y) \vee (a,0) \prec y\big]\\ &= \frac{1}{2} - \frac{1}{2^{a+3}} + \sum_{i=a+4}^{\infty} 2^{-i}\\ &= \frac{1}{2}.\end{aligned}$$

Thus, we also see that the axioms expressing that these two sets have measure $\frac{1}{2}$ hold, i.e. that $(\mathcal{M},\mathcal{D}) \models_{\frac{1}{2}} \bigwedge T_{\frac{1}{2}}$. But then we see that there is a $\frac{1}{2}$-model $(\mathcal{N},\mathcal{E})$ such that

$$(\mathcal{N},\mathcal{E}) \models_{\frac{1}{2}} f_0\left(\left(\bigwedge T_0 \wedge \varphi^N\right)^{\#}\right) \wedge f_{\frac{1}{4}}\left(\bigwedge T_{\frac{1}{4}}\right) \wedge f_{\frac{3}{4}}\left(\bigwedge T_{\frac{3}{4}}\right) \wedge \bigwedge T_{\frac{1}{2}}$$

i.e. that (i) holds.

(i) $\to$ (ii): Assume (i) holds. Then there exists a probability model $(\mathcal{M},\mathcal{D})$ such that $(\mathcal{M},\mathcal{D}) \models_\varepsilon T_\varepsilon$ for $\varepsilon \in \left\{\frac{1}{4},\frac{1}{2},\frac{3}{4}\right\}$ and $(\mathcal{M},\mathcal{D}) \models_0 (\bigwedge T_0 \wedge \varphi^N)^{\#}$.

By Proposition 11.5.5 we have that the set

$$B = \{(a_0,a_1,\dots) \in \mathcal{M}^\omega \mid \mathcal{M} \restriction \{a_0,a_1,\dots\} \models \bigwedge T_0 \wedge \varphi^N\}$$

has $\mathcal{D}^\omega$-measure 1. In particular, because all universal Robinson axioms are in T_0 we see that for every $(a_0,a_1,\dots) \in B$ these universal Robinson axioms hold classically in $\mathcal{M} \restriction \left(\{a_0,a_1,\dots\} \cap N^{\mathcal{M}}\right)$. The only problem is that we do not know if the two missing non-universal axioms $\forall x(0 = x \vee \exists y S(y) = x)$ and $\forall x \exists y S(x) = y$ also hold. However, if we were able to find a subsequence $(b_0,b_1,\dots)$ of a permutation of some $(a_0,a_1,\dots) \in B$ satisfying $0 = b_0$ and $S(a_i) = a_{i+1}$ for all $i \in \omega$, these two axioms would also hold in $\mathcal{M} \restriction (\{b_0,b_1,\dots\})$. Furthermore, note that T_0 guarantees that $\{b_0,b_1,\dots\} \subseteq N^{\mathcal{M}}$, so we then have that $\mathcal{M} \restriction \{b_0,b_1,\dots\}$ models *all* the Robinson axioms. Even stronger: we have that this model is a model isomorphic to $\mathbb{N}$ in which φ holds (note that φ is universal, so it holds in any submodel of $\mathcal{M} \restriction \{a_0,a_1,\dots\}$), showing that (ii) holds. The rest of the proof will therefore consist of showing that we can find such a sequence in B.

First, we restrict to a subset B'' of B which still has measure 1. For every sentence $\alpha \in T_0$, say $\forall x_1 \dots \forall x_t \beta(x_1,\dots,x_t)$ (so $t \leq 3$), let $\tilde{\alpha}(y_1,y_2)$ be the formula $\forall x_t \beta(y_1,\dots,y_{t-1},x_t)$. Then

$$B' := B \cap \bigcap_{\pi:\{1,2\}\hookrightarrow\omega} \{(a_0,a_1,\dots) \in \mathcal{M}^\omega \mid (\mathcal{M},\mathcal{D}) \models_0 \bigwedge_{\alpha\in T_0} \tilde{\alpha}(a_{\pi(1)},a_{\pi(2)})\}$$

is still a set of measure 1: it is an intersection of measurable sets by Proposition 8.1.3 and all of these sets have measure 1 because $(\mathcal{M},\mathcal{D}) \models_0 (\bigwedge T_0 \wedge \varphi^N)^{\#}$. We make one more restriction. Let X be the set

$$\begin{aligned}&\{x \in \mathcal{M} \mid \mathcal{M} \models \neg N(x)\} \cup\\ &\left\{x \in N^{\mathcal{M}} \mid \Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid R(x,y) \vee S(x) = y\big] = \Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid R(x,y) \vee x \prec y\big] = \frac{1}{2}\right\}.\end{aligned}$$

Then X has measure 1: the union above is disjoint, $\neg N$ has measure exactly $\frac{1}{2}$ and by the axioms for R it follows that the second set has measure $\frac{1}{2}$. Now let $B'' = B' \cap X^\omega$. Then our axioms hold inside B'' in a strong sense (that is, the axioms of T_0 hold classically and the axioms about R hold classically in the first coordinate), which is a fact we will use shortly.

For each $n \in \omega$, let D_n be the set

$$\bigcup_{(m_0,\dots,m_n)\in\omega^{n+1}} \{(a_0, a_1, \dots) \in B'' \mid 0 = a_{m_0}, S(a_{m_0}) = a_{m_1}, \dots, S(a_{m_{n-1}}) = a_{m_n}\}.$$

Then $D_0 \supseteq D_1 \supseteq \dots$. By Lemma 11.5.7 below, each D_n has measure 1. Therefore, $\bigcap_{n\in\omega} D_n$ has measure 1 and in particular it is non-empty. Let $(a_0, a_1, \dots) \in \bigcap_{n\in\omega} D_n$. For each $n \in \omega$, fix an α_n such that

$$\exists i_0, \dots, i_{n-1}(0 = a_0 \wedge \dots \wedge S(a_{i_{n-2}}) = a_{i_{n-1}} \wedge S(a_{i_{n-1}}) = a_{\alpha_n}),$$

which exists because $(a_0, a_1, \dots) \in D_n$. Now let $(b_0, b_1, \dots)$ be the sequence $(a_{\alpha_0}, a_{\alpha_1}, \dots)$. Then $0 = b_0$. Also, we claim that $S(b_0) = b_1$. We know that there exists some i_0 such that $a_{i_0} = 0$ and $S(a_{i_0}) = b_1$. Now, because $\mathcal{M} \restriction \{a_0, a_1, \dots\} \models \bigwedge T_0$ classically, we know in particular that the equality axioms hold classically in this model. Thus, $a_{i_0} = b_0$ and $S(b_0) = b_1$, as desired. In the same way, we can show that $S(b_i) = b_{i+1}$ holds for any $i \in \omega$. Therefore, as discussed above the model $\mathcal{M} \restriction \{b_0, b_1, \dots\}$ is isomorphic to $\mathbb{N}$ and $\varphi(Q)$ holds in it, which shows that (ii) holds. □

LEMMA 11.5.7. *Let D_n be as in the proof of Theorem 11.5.6. Then* $\Pr_{\mathcal{D}}[D_n] = 1$.

PROOF. Fix $n \in \omega$. For $k, i \in \omega$ with $0 \le i < k$, denote by $D_n^{i,k}$ the set

$$\{(a_0, a_1, \dots) \in B'' \mid 0 = a_i \wedge S(a_i) = a_{i+k} \wedge \dots \wedge S(a_{i+(n-1)k}) = a_{i+nk}\}.$$

Then $D_n^{i,k} \subseteq D_n$. From $(\mathcal{M}, \mathcal{D}) \models_{\frac{1}{4}} T_{\frac{1}{4}}$ and $(\mathcal{M}, \mathcal{D}) \models_{\frac{3}{4}} T_{\frac{3}{4}}$ we see that $\Pr_{\mathcal{D}}[a \in \mathcal{M} \mid \mathcal{M} \models 0 = a] = \frac{1}{4}$, so together with Lemma 11.5.8 below this gives us

$$\Pr_{\mathcal{D}}[D_n^{i,k}] = \frac{1}{4}\frac{1}{8}\cdots\frac{1}{2^{n+2}} = \frac{1}{2^{c_n}}$$

where $c_n = \frac{(n+2)(n+3)}{2} - 1 > 0$. Clearly, for $i \neq j$ the sets $D_n^{i,k}$ and $D_n^{j,k}$ are independent. So, we have

$$\Pr_{\mathcal{D}}[\bigcup_{i\le k} D_n^{i,k}] = 1 - (1 - \frac{1}{2^{c_n}})^k.$$

Combining this with the fact that $\bigcup_{i\le k} D_n^{i,k} \subseteq D_n$ for every $k \in \omega$ we thus have

$$\Pr_{\mathcal{D}}[D_n] \ge \lim_{k\to\infty} 1 - (1 - \frac{1}{2^{c_n}})^k = 1. \qquad \square$$

LEMMA 11.5.8. *Let B'' be as in the proof of Theorem 11.5.6. Let $(a_0, \dots, a_n) \in \mathcal{M}^n$ be any subsequence of a sequence in B'' such that $0 = a_0$ and $S(a_i) = a_{i+1}$. Then*

$$\Pr_{\mathcal{D}}[a \in \mathcal{M} \mid \mathcal{M} \models S(a_n) = a] = \frac{1}{2^{n+3}}.$$

PROOF. Using induction over n. First, let $n = 0$. Then by the definition of B'' we have that

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_0) = a\big] = \frac{1}{2} - \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models R(a_0, a)\big] = \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_0 \prec a\big]$$

(where we use that R only holds outside N while S and $\prec$ only hold inside N, so R is disjoint from $<$ and $\prec$). We then have that

$$2\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_0) = a\big] = \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_0) = a\big] + \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_0 \prec a\big] = \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_0 < a\big].$$

(For this, we use the definition of B' together with the axiom $(\forall x \forall y(x < y \leftrightarrow (x \prec y \vee S(x) = y)))^N$.) Now, because $0 = a$ this is equal to

$$\Pr_{\mathcal{D}}[N] - \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models 0 = a\big] = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$

So, $\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_0) = a\big]$ is one half of that, i.e. $\frac{1}{8}$.

Next, assume

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_n) = a\big] = \frac{1}{2^{n+3}}.$$

In the same way as above, we find that

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_{n+1}) = a\big] = \Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_{n+1} \prec a\big].$$

Again, the sum of these two probabilities is

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_{n+1} < a\big].$$

This is equal to

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models a_n \prec a\big],$$

so by the induction hypothesis we now see that

$$\Pr_{\mathcal{D}}\big[a \in \mathcal{M} \mid \mathcal{M} \models S(a_{n+1}) = a\big] = \frac{1}{2}\frac{1}{2^{n+3}} = \frac{1}{2^{n+4}}.$$

□

Next, we turn to the language of arithmetic enlarged with relation symbols representing finitely many primitive recursive functions — say $g_1, \ldots, g_n$. So, for each of these primitive recursive function $g_i(x_1, \ldots, x_{m_i})$ we add a relation symbol $g_i(x_1, \ldots, x_{m_i}) = y$ to our language which represents this function. We next show that Theorem 11.5.6 still holds if we add these finitely many primitive recursive functions to our language.

THEOREM 11.5.9. *Let $g_1, \ldots, g_n$ be primitive recursive functions. Let $f_0, f_{\frac{1}{4}}, f_{\frac{1}{2}}$ and $f_{\frac{3}{4}}$ be as in Theorem 11.5.6. Then there exists finite theories $T_0, T_{\frac{1}{4}}, T_{\frac{1}{2}}$ and $T_{\frac{3}{4}}$, containing only universal relational sentences, such that for each universal first-order sentence φ in the language of arithmetic enlarged with relation symbols $Q, g_1, \ldots, g_n$, the following are equivalent:*

(i) $f_0((\bigwedge T_0 \wedge \varphi^N)^\#) \wedge f_{\frac{1}{4}}(\bigwedge T_{\frac{1}{4}}) \wedge f_{\frac{3}{4}}(\bigwedge T_{\frac{3}{4}}) \wedge \bigwedge T_{\frac{1}{2}}$ *is* $\frac{1}{2}$*-satisfiable;*
(ii) $\mathbb{N} \models \exists Q \varphi(Q)$.

Proof. The proof is very similar to that of Theorem 11.5.6. We let $T_{\frac{1}{4}}, T_{\frac{1}{2}}$ and $T_{\frac{3}{4}}$ be as in the proof of that theorem and we extend T_0 with axioms for every g_i.

So, let $1 \leq i \leq n$. Fix any sequence $h_0, \dots, h_k$ of primitive recursive functions in such that for each $1 \leq j \leq k$, either:

- $h_j = 0$, $h_j = S$, or h_j is a projection.
- h_j is the composition of h_s with $h_{t_1}, \dots, h_{t_m}$ for some $1 \leq s, t_1, \dots, t_m < j$.
- h_j is defined by primitive recursion from h_s and h_t for some $1 \leq s, t < j$.
- $h_k = g_i$.

We add axioms to define each such h_j. It is clear what we need to do in the first case. If $h_j(x_1, \dots, x_m)$ is the composition of h_s with $h_{t_1}, \dots, h_{t_m}$, we take the relativisation to N of the universal closure of

$$\begin{aligned}&(h_{t_1}(x_1, \dots, x_n) = z_1 \wedge \dots \wedge h_{t_m}(x_1, \dots, x_n) = z_m)\\ &\to (h_j(x_1, \dots, x_m) = y \leftrightarrow h_s(z_1, \dots, z_m) = y).\end{aligned}$$

Similarly, if h_j is defined by primitive recursion over h_s and h_t, then we take the relativisation to N of the universal closure of

$$0 = x \to (h_j(x, y_1, \dots, y_n) = z \leftrightarrow h_s(y_1, \dots, y_n) = z)$$

and of

$$\begin{aligned}&(h_j(x, y_1, \dots, y_n) = z \wedge S(x) = u)\\ &\to (h_t(x, z, y_1, \dots, y_n) = v \leftrightarrow h_j(u, y_1, \dots, y_n) = v).\end{aligned}$$

This completes our description of the construction of T_0.

To prove that (ii) implies (i), we can follow the proof of the previous theorem, giving every primitive recursive function g is usual interpretation on $\omega \times \{0\}$ and letting it be undefined elsewhere.

Conversely, in our proof that (i) implies (ii), we define B for our new formula $\bigwedge T_0 \wedge \varphi^N$. It will still have measure 1. Then for any $(a_0, a_1, \dots) \in B$, the axioms given above hold classically in $\mathcal{M} \upharpoonright (\{a_0, a_1, \dots\} \cap N^{\mathcal{M}})$ and therefore the primitive recursive functions h_i, so in particular $g_1, \dots, g_n$, have their usual or true interpretation in this restricted model. So, if we perform the construction above with this new B, then we get a model $\mathcal{M} \upharpoonright \{b_0, b_1, \dots\}$ in which φ holds which is not just isomorphic to $\mathbb{N}$, but is even isomorphic to $\mathbb{N}$ enlarged with the usual interpretations of $g_1, \dots, g_n$. Therefore we still see that (i) implies (ii) even with the addition of these primitive recursive functions. □

Next, we will want to include relation symbols of the form $\{e\}^Q(x_1, x_2) = y$. Here, $\{e\}^Q(x_1, x_2) = y$ means that the eth partial computable functional halts with oracle Q and input x_1, x_2 and outputs y, or in the notation of Kleene's normal form theorem for partial computable restricted functionals [54] (see e.g. Odifreddi [87, Theorem II.3.11]), this means that both $\exists z \mathcal{T}_{1,2}(e, x_1, x_2, \hat{Q}(z), z)$ and $\mathcal{U}(\mu z[\mathcal{T}_{1,2}(e, x_1, x_2, \hat{Q}(z), z)]) = y$.

To this end, we add ternary relation symbols $\{e\}^Q(x_1, x_2) = y$ to our language for all $e \in \omega$ (so, we do not see e or Q as a variable here; we add a relation symbol for every e and the Q is purely notational). Using the normal form of partial computable restricted functionals we can now easily expand our result to include these functionals.

THEOREM 11.5.10. *Let $f_0, f_{\frac{1}{4}}, f_{\frac{1}{2}}$ and $f_{\frac{3}{4}}$ be as in Theorem 11.5.6. Then there exists finite theories $T_0, T_{\frac{1}{4}}, T_{\frac{1}{2}}$ and $T_{\frac{3}{4}}$, containing only universal relational sentences, and a computable function $e \mapsto \psi_e$ mapping each $e \in \omega$ to a universal relational sentence, such that for each universal sentence φ in the language of arithmetic enlarged with relation symbols $Q, \{0\}^Q, \{1\}^Q, \dots,$ the following are equivalent:*

(i) $f_0((\bigwedge T_0 \wedge \bigwedge_{\{e\}^Q \in \varphi} \psi_e \wedge \varphi^N)^\#) \wedge f_{\frac{1}{4}}(\bigwedge T_{\frac{1}{4}}) \wedge f_{\frac{3}{4}}(\bigwedge T_{\frac{3}{4}}) \wedge \bigwedge T_{\frac{1}{2}}$ *is $\frac{1}{2}$-satisfiable;*[9]

(ii) $\mathbb{N} \models \exists Q(\varphi)$, *i.e. there exists some $Q^{\mathbb{N}} \subseteq \omega$ such that $\mathbb{N} \models \varphi$ if Q is interpreted as $Q^{\mathbb{N}}$ and $\{e\}^Q(x_1, x_2) = y$ is interpreted as $\{e\}^{Q^{\mathbb{N}}}(x_1, x_2) = y$.*

PROOF. We let $\mathcal{U}$ and $\mathcal{T}_{1,2}$ represent the primitive recursive functions from the normal form theorem for partial computable restricted functionals, as discussed above. Furthermore, we let $\mathcal{V}(x, y)$ represent the primitive recursive function which outputs $z + 1$ if y has length at least x and the xth digit equals z, and which outputs 0 otherwise. Now let the T_ε be as in Theorem 11.5.9, applied to $\mathcal{U}$, $\mathcal{T}_{1,2}$ and $\mathcal{V}$.

We let $\mathcal{W}(x, y)$ represent "the sequence x is an approximation to Q up to digit y", so we add to T_0 the relativisation to N of the universal closure of

$$\begin{aligned}&(0 = y \wedge 0 = u \wedge S(u) = v \wedge S(v) = w)\\ &\to (\mathcal{W}(x, y) \leftrightarrow ((\mathcal{V}(y, x) = v \leftrightarrow \neg Q(y)) \wedge (\mathcal{V}(y, x) = w \leftrightarrow Q(y)))),\end{aligned}$$

and of

$$\begin{aligned}&(S(z) = y \wedge 0 = u \wedge S(u) = v \wedge S(v) = w)\\ &\to (\mathcal{W}(x, y) \leftrightarrow (\mathcal{W}(x, z) \wedge (\mathcal{V}(y, x) = v \leftrightarrow \neg Q(y)) \wedge (\mathcal{V}(y, x) = w \leftrightarrow Q(y)))).\end{aligned}$$

Finally, for each $e \in \omega$, let ψ_e be the relativisation to N of the universal closure of

$$\mathcal{W}(u, z) \wedge \mathcal{T}_{1,2}(e, x_1, x_2, u, z) \to (\{e\}^Q(x_1, x_2) = y \leftrightarrow \mathcal{U}(z) = y)).$$

The proof now proceeds in the same way as for Theorem 11.5.9, observing that our new axioms ensure that $\{e\}^Q$ gets the right interpretation. □

So, now we have Π^0_1-hardness. To get to Σ^1_1-hardness, we first put the Σ^1_1-formulas into a more appropriate form.

LEMMA 11.5.11. *Let $\mathcal{A}$ be the set of those universal (relational) formulas φ in the language of arithmetic enlarged with relation symbols $Q, \{0\}^Q, \{1\}^Q, \dots$ for which there exists some interpretation $Q^{\mathbb{N}} \subseteq \omega$ satisfying:*

- $\mathbb{N} \models \varphi$ *if Q is interpreted as $Q^{\mathbb{N}}$ and $\{e\}^Q(x_1, x_2) = y$ is interpreted as $\{e\}^{Q^{\mathbb{N}}}(x_1, x_2) = y$;*
- *For all n there is some $m \in Q^{\mathbb{N}}$ coding a string of length n.*

[9]By $\{e\}^Q \in \varphi$ we mean that the relation symbol $\{e\}^Q$ occurs in the formula φ.

Then A is Σ^1_1-hard.

PROOF. The set of indices of non-well-founded computable trees is Σ^1_1-hard (for example, see Odifreddi [87, Corollary IV.2.16]). This set many-one reduces to $\mathcal{A}$: for every index $e \in \omega$, the expression "$\{e\}$ computes a tree and $Q \subseteq \{e\}$" can be written as a universal formula $\varphi(Q)$ in the language as above. Now $\{e\}$ computes a non-well-founded tree if and only if there exists a Q as above, completing our proof. □

So, we need to somehow expand our theories in such a way that it only allows those Q which contain a string of every possible length. To achieve this, we will add even more relation symbols to our language, and we will use some of these as additional oracles. Let $A \subseteq \omega \times X$ for some set X. We view elements from the set X as parameters: i.e. for all $p \in X$ we will write $\{e\}^{Q \oplus A_p}(x_1, x_2) = y$ if the eth partial computable function halts with oracle $Q \oplus \{n \in \omega \mid (n,p) \in A\}$ and input x_1, x_2, and outputs y. We will add a binary relation symbol A, a ternary relation symbol B and a 5-ary relation symbol $\{e\}^{Q \oplus A_p}(x_1, x_2) = y$ for every $e \in \omega$ to our language. Using the same method as in the proof of Theorem 11.5.10 above, we will guarantee that for every model $(\mathcal{M}, \mathcal{D})$ satisfying our theory we have for almost all $p \in \mathcal{M}$ that our interpretation of $\{e\}^{Q \oplus A_p}(x_1, x_2) = y$ in the model $\mathcal{M}$ agrees with its true interpretation in $\mathbb{N}$. We will combine this with the construction used in the proof of Theorem 9.5.2 where we showed that compactness does not hold for ε-logic; in some sense, compactness failing is the same as forcing infinite objects to exist, which explains the correlation with our current goal.

THEOREM 11.5.12. *Let $f_0, f_{\frac{1}{4}}, f_{\frac{1}{2}}$ and $f_{\frac{3}{4}}$ be as in Theorem 11.5.6. Then there exists finite theories $T_0, T_{\frac{1}{4}}, T_{\frac{1}{2}}$ and $T_{\frac{3}{4}}$, containing only universal relational sentences, and a computable function $e \mapsto \psi_e$ mapping each $e \in \omega$ to a universal relational sentence, such that for each universal sentence φ in the language of arithmetic enlarged with relation symbols $Q, \{0\}^Q, \{1\}^Q, \dots,$ the following are equivalent:*

(i) $f_0((\bigwedge T_0 \wedge \bigwedge_{\{e\}^Q \in \varphi} \psi_e \wedge \varphi^N)^{\#}) \wedge f_{\frac{1}{4}}(\bigwedge T_{\frac{1}{4}}) \wedge f_{\frac{3}{4}}(\bigwedge T_{\frac{3}{4}}) \wedge \bigwedge T_{\frac{1}{2}}$ *is $\frac{1}{2}$-satisfiable;*
(ii) *φ is in the set $\mathcal{A}$ of Lemma 11.5.11.*

PROOF. We start with the theories T_ε and the universal sentences ψ_e as in Theorem 11.5.10. We need to expand our theory in such a way that case (i) can only hold if Q contains a string of length n for every $n \in \omega$, using the relation symbols A and B.

For this, we will make use of the construction which was used to show that ε-logic is not compact in the proof of Theorem 9.5.2. There, sentences α_n were defined such that for any model $(\mathcal{M}, \mathcal{D})$ for some language containing a binary relation symbols symbol E we have that $(\mathcal{M}, \mathcal{D}) \models_{\frac{1}{2}} \alpha_n$ if and only if we have that: *For almost all y (i.e. measure 1 many), there exists a set C_y of measure at least $1 - \frac{1}{n}$ such that for all $y' \in C_y$ the sets $D_y = \{u \mid E(u, y)\}$ and $D_{y'} = \{u \mid E(u, y')\}$ both have measure $\frac{1}{2}$, while $D_y \cap D_{y'}$ has measure $\frac{1}{4}$ (in other words, the two sets are independent sets of measure $\frac{1}{2}$).* In the proof of Theorem 9.5.2, it was shown that every finite subset of $\{\alpha_1, \alpha_2, \dots\}$ is satisfiable, but no infinite subset is satisfiable.

Fix an index e_1 such that $\{e_1\}^Q(n,m)$ computes the indicator function of "there exists some $n' \leq n$ such that n' codes a string of length m and $n' \in Q$". Our idea is to let α_n hold, with E interpreted as $\{(x,y) \mid B(m,x,y)\}$, if and only $\{e_1\}^Q(n,m) = 0$. Then, as discussed above, only finitely many of these α_n are allowed to hold for any fixed m, so we see that there has to be some $n \in \omega$ such that $\{e\}^Q(n,m) = 1$, i.e. Q contains a string of length m. This is precisely what we want to achieve.

To this end we first need to change the definition of α_n from Theorem 9.5.2 into a more uniform one. In that proof, we have a different formula for every n (which in fact grows exponentially in n). We wish to change it into one fixed formula $\tilde{\alpha}$ which has n as a parameter, i.e. we want one formula $\tilde{\alpha}$ such that α_n is equivalent to $\tilde{\alpha}(n)$ for every $n \in \omega$. This is what we will use A for. If we unfold the first two reductions in the definition of α_n for $n \geq 2$, and rename some of the relation symbols X_i, Y_i, we obtain the following formula:

$$
\begin{aligned}
&\forall y Y_1(y) \wedge \forall y Y_2(y)\\
&\wedge \bigwedge_{I \subseteq \{3,\ldots,2n\}, |I| = n-1} \forall y \left(\bigvee_{i \in I} Y_i(y) \right)\\
&\wedge \bigwedge_{1 \leq i \leq 2} \forall y \left(Y_i(y) \right.\\
&\left. \wedge \bigwedge_{3 \leq i_1 < i_2 < \cdots < i_n \leq 2n} \forall y' ((Y_{i_1}(y) \vee \cdots \vee Y_{i_m}(y)) \wedge \alpha'(y,y')) \right)
\end{aligned}
$$

where α' is a formula that does not depend on n. Here, we denoted by $Y_i(x)$ the formula $X_i(x) \wedge \bigwedge_{1 \leq j \leq 2, j \neq i} \neg X_j(x)$ for $i \in \{1,2\}$ and $X_i(x) \wedge \bigwedge_{3 \leq j \leq 2n, j \neq i} \neg X_j(x)$ for $i \in \{3,\ldots,2n\}$, and the X_i are new predicates.

Because we want to make a more uniform version of this formula, we want to use the single binary predicate A instead of these predicates X_i and Y_i. To this end, fix an index e_2 such that for all oracles Q, C and all binary sequences σ we have that $\{e_2\}^{Q \oplus C}(\sigma, \langle n,m \rangle)$ computes the indicator function of

$$
\exists i \in \sigma \left(C(\langle \langle n,m \rangle, i \rangle) \wedge \bigwedge_{3 \leq j \leq 2n, j \neq i} \neg C(\langle \langle n,m \rangle, j \rangle) \right).
$$

Next, fix an index e_3 such that $\{e_3\}^Q(\sigma, n)$ computes if σ is a subset of $\{3,\ldots,2n\}$ of size $n-1$, and similarly an index e_4 which computes if σ is a subset of $\{3,\ldots,2n\}$ of size n. Finally, fix an index e_5 such that $\{e_5\}^Q(n,m)$ computes the pairing function $\langle x, y\rangle$. We can now define our uniform version $\tilde{\alpha}(n,m)$ of α_n:

$$\begin{aligned}
&\forall y A(\langle\langle n,m\rangle,1\rangle,y) \wedge \forall y A(\langle\langle n,m\rangle,2\rangle,y)\\
&\wedge \forall u(\{e_3\}^Q(u,n) \to \forall y \{e_2\}^{Q,A_y}(u,\langle n,m\rangle))\\
&\wedge \bigwedge_{1\leq i\leq 2} \forall y\Bigg(A(\langle\langle n,m\rangle,i\rangle,y)\\
&\wedge \forall u(\{e_4\}^Q(u,n) \to \forall y'(\{e_2\}^{Q,A_{y'}}(u,\langle n,m\rangle) \wedge \tilde{\alpha}'(m,y,y')))\Bigg).
\end{aligned}$$

Here the formula $\tilde{\alpha}'(m,y,y')$ is the formula obtained by replacing each occurrence of $E(x,y)$ by $B(m,x,y)$ in α'. We freely wrote $\alpha(0)$ for $\forall x(N(x) \wedge (0 = x \to \alpha(x)))$; we will interpret this formula for $\varepsilon = \frac{1}{2}$, and because N has measure $\frac{1}{2}$ this essentially means that for almost all $x \in N$ we have that $0 = x \to \alpha(x)$. We did similar things for 1, for 2 and for the pairing function $\langle x, y\rangle$, using the index e_5 for the latter.

We now add the following formula to $T_{\frac{1}{2}}$:

$$\forall n \forall m(N(n) \wedge N(m) \wedge (\{e_1\}^Q(n,m) = 0 \to \tilde{\alpha}(n,m))).$$

Furthermore, we add ψ_e to T_0 for $e \in \{e_1, e_3, e_4, e_5\}$. Like in the proof of Theorem 11.5.10, we add an axiom for $\mathcal{W}'(x,y,z)$ specifying that $\mathcal{W}'(x,y,z)$ holds if and only if the sequence x is an approximation to $Q \oplus A_z$, where $A_z = \{n \mid (n,z) \in A\}$, up to digit y. Finally, we add a rule for e_2 to T_0, namely the universal closure of

$$\mathcal{W}'(u,z,v) \wedge \mathcal{T}_{1,2}(e_2,x_1,x_2,u,z) \to (\{e_2\}^{Q\oplus A_v}(x_1,x_2) = y \leftrightarrow \mathcal{U}(z) = y),$$

where all quantifiers except the one for v are relativised to N.

We prove that the theories defined above are as desired. First, assume (ii) holds. We define a model $\mathcal{M}$ in a similar way as in the proof of Theorem 11.5.10. However, this time we do not take the universe to be two copies of ω, but instead we take our universe to be the unit interval. We let our measure be the Lebesgue measure. Fix a partition $\{U_0, U_1, \dots\} \cup \{V_0, V_1, \dots\}$ of $[0,1]$ such that all U_n, V_n are measurable with measure $\frac{1}{2^{n+2}}$. Looking back at our proof of Theorem 11.5.10 (and Theorem 11.5.6), we identify the elements of each U_n with the element n in the first copy of ω and each V_n with the element n in the second copy of ω. This allows us to define the relations in the language of arithmetic in the same way as before: for example, we let $x + y = z$ hold precisely if $x \in U_n$, $y \in U_m$ and $z \in U_{n+m}$ for some $n, m \in \omega$. We let $N(x)$ hold if $x \in U_n$ for some $n \in \omega$. With this in mind, it should be clear from the proof of Theorem 11.5.10 how to define R and all the other relations already appearing in the proof of that theorem.

The only relations of which it might not be directly clear how to define them are A and B. How to define B is explained in the proof of Theorem 9.5.2, while how to define A follows from the proof of Theorem 9.4.1.

For the converse, assume (i) holds. Then, as in the proof of Theorem 11.5.10, we can find a model $\mathcal{M} \restriction \{b_0, b_1, \dots\}$ which is isomorphic to $\mathbb{N}$ and a relation Q on $\{b_0, b_1, \dots\}$, such that $\{e_i\}^Q, \{e_2\}^{Q\oplus A_p}$ have their usual interpretation and such that φ holds in this model. Using similar ideas as in the proof of that theorem, we

can even assume that our model satisfies our new axioms in the following way:
If $\{e_1\}^Q(b_n, b_m)$ does not hold, then $\tilde{\alpha}(b_n, b_m)$ holds in the sense that all of the following hold:

$$\Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid \mathcal{M} \models A(\langle\langle b_n, b_m\rangle, b_1\rangle, y)\big] \geq \frac{1}{2}$$

$$\Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid \mathcal{M} \models A(\langle\langle b_n, b_m\rangle, b_2\rangle, y)\big] \geq \frac{1}{2}$$

for all $u \in \{b_0, b_1, \dots\}$, if $\mathcal{M} \models \{e_3\}^Q(u, n)$

$$\text{then } \Pr_{\mathcal{D}}\big[y \in \mathcal{M} \mid \mathcal{M} \models \{e_2\}^{Q \oplus A_y}(u, \langle b_m, b_m\rangle)\big] \geq \frac{1}{2}.$$

$$\text{for } 1 \leq i \leq 2,\ \Pr_{\mathcal{D}}\Bigg[y \in \mathcal{M} \mid \mathcal{M} \models A(\langle\langle b_n, b_m\rangle, b_i\rangle, y) \text{ and for all } u \in \{b_0, b_1, \dots\}, \text{ if } \mathcal{M} \models \{e_4\}^Q(u, n) \text{ then } \Pr_{\mathcal{D}}\big[y' \in \mathcal{M} \mid (\mathcal{M}, \mathcal{D}) \models_{\frac{1}{2}} \{e_2\}^{Q \oplus A_{y'}}(u, \langle n, m\rangle) \wedge \tilde{\alpha}'(m, y, y')\big] \geq \frac{1}{2}\Bigg] \geq \frac{1}{2}.$$

Now one can directly verify (using the proof of Theorem 9.4.1) that under this interpretation, $\tilde{\alpha}(b_n, b_m)$ is equivalent to the formula α_n with $E(x, y)$ replaced by $\{(x, y) \mid B(m, x, y)\}$. However, we also have that if Q does not contain a string of length m, then $\tilde{\alpha}(b_n, b_m)$ holds for all $n \in \omega$, which is a contradiction (as explained above). So, we see that Q contains a string of every length and therefore (ii) holds. □

THEOREM 11.5.13. *For rational $\varepsilon \in (0, 1)$ we have that ε-satisfiability is Σ^1_1-hard.*

PROOF. From Lemma 11.5.7 and Theorem 11.5.12. □

THEOREM 11.5.14. *Let $\mathcal{L}$ be a countable language not containing equality and function symbols and let $\varepsilon \in (0, 1)$ be rational. Then ε-satisfiability is Σ^1_1-complete.*

PROOF. The Σ^1_1-hardness was shown in Theorem 11.5.13, while Theorem 11.3.5 gives the matching upper bound. □

11.6. Compactness of 0-logic

We conclude this chapter by showing that 0-logic is compact (when considering languages not containing equality and function symbols), which will quite directly follow from the results from the previous section. This contrasts the fact that for rational $\varepsilon \in (0, 1)$ we have that ε-logic is not compact, as shown in Theorem 9.5.2.

First, we need the following strengthening of Theorem 11.2.2.

THEOREM 11.6.1. *Let $\mathcal{L}$ be a language not containing equality and function symbols (but it may contain constant symbols) and let Γ be a countable set of formulas in $\mathcal{L}$. Then there exists a language $\mathcal{L}'$ only containing relation symbols and a computable function mapping each formula $\varphi \in \Gamma$ in the language $\mathcal{L}$ to a universal formula φ' in the language $\mathcal{L}'$ such that for every $\varepsilon \in [0, 1]$ and every subset $\Delta \subseteq \Gamma$: Δ is ε-satisfiable if and only if $\Delta' = \{\varphi' \mid \varphi \in \Delta\}$ is ε-satisfiable.*

PROOF. We prove this in almost the same way as Theorem 11.2.2. We construct φ' as in that theorem, but instead of using relations R_α to define φ' we use relations $R_{\alpha,\varphi}$. That is, we introduce new relation symbols for every formula φ, instead of reusing the same ones. There is one exception to this rule: when $\alpha = 0$ (i.e. the function which is constantly 0) we always use the relation R_α, which does not depend on φ.

Now let $\Delta \subseteq \Gamma$. For ease of notation we assume Δ is infinite, say $\Delta = \{\varphi_0, \varphi_1, \dots\}$; the finite case follows in the same way. If Δ is ε-satisfiable, then one can prove that Δ' is ε-satisfiable in the same way as in the proof of Theorem 11.2.2. For the converse we also use a similar proof, but we need to make some slight modifications. Note that, in the proof of Theorem 11.2.2, we used the measure 0 Cantor set $\mathcal{C} \subseteq [0,1]$ to provide us with witnesses. Now we partition $\mathcal{C}$ into countably many uncountable Borel measure 0 sets. For example, we can take $\mathcal{C}_i \subseteq \mathcal{C}$ to be those $x \in \mathcal{C}$ which correspond to a sequence starting with 0^i1. We can then take Borel isomorphisms $\zeta^i : \mathcal{C}_i \to \bigcup_{1 \leq i \leq n} [0,1]^i$ and use ζ^i to provide the witnesses for φ_i.

More precisely, we show how to modify the definition of $R^{\mathcal{M}}$ in the proof of Theorem 11.2.2. Let $a_1, \dots, a_k \in [0,1]$. If there are $1 \leq i_1 < i_2 \leq k$ such that $a_{i_1} \in \mathcal{C}_{j_1}, a_{i_2} \in \mathcal{C}_{j_2}$ but $j_1 \neq j_2$, we let $R^{\mathcal{M}}(a_1, \dots, a_k)$ be false (but it does not really matter how we define $R^{\mathcal{M}}$ in this case, as long as we make sure that the resulting relation is Borel; for example, we could also define it to always be true in this case). Otherwise, we let $r \in \omega$ be such that for all $1 \leq i \leq k$ we have that $a_i \in \mathcal{C}$ implies that in fact $a_i \in \mathcal{C}_r$. Let $\alpha(i) = (1,j)$ if $a_i \in \mathcal{C}$ and $\zeta(a_i)$ has length j, let $\alpha(i) = (2,j)$ if $a_i \in \operatorname{ran}(\eta)$ and $\eta^{-1}(a_i) = j$ and finally let $\alpha(i) = 0$ if neither of these cases hold. Let $b_1, \dots, b_{k'}$ be the subsequence of $a_1, \dots, a_k$ obtained by taking just those a_i satisfying $\alpha(a_i) = 0$. Let $m = \max\{j \in \omega \mid (1,j) \in \operatorname{ran}(\alpha)\}$ and let i be the least $1 \leq j \leq k$ such that $\alpha(j) = (1,m)$. Now let $R^{\mathcal{M}}(a_1, \dots, a_k)$ be defined as $R^{\mathcal{N}}_{\alpha,\varphi_r}(b_1, \dots, b_{k'}, (\zeta^r)_1(a_i), \dots, (\zeta^r)_m(a_i))$. The rest of the proof proceeds in the same way as for Theorem 11.2.2. $\square$

THEOREM 11.6.2. *For countable languages $\mathcal{L}$ not containing equality and function symbols, 0-logic is compact. That is, if Γ is any (countable) set of sentences such that any finite subset of Γ is 0-satisfiable, then Γ is 0-satisfiable.*

PROOF. By Theorem 11.6.1 we may assume all formulas in Γ to be universal relational formulas. For every formula $\varphi = \forall x_1 \dots \forall x_n \psi(x_1, \dots, x_n)$, let

$$\tilde{\varphi} = \forall x_1 \dots \forall x_n \bigwedge_{i<j} (x_i \neq x_j \to \psi(x_1, \dots, x_n)).$$

Let $\tilde{\Gamma} = \{\tilde{\varphi} \mid \varphi \in \Gamma\}$. We claim: $\tilde{\Gamma}$ has an infinite classical model.

By classical compactness, it is enough to show that for every $m \in \omega$ and every finite subset $\tilde{\Delta} \subseteq \tilde{\Gamma}$ there is a model for $\tilde{\Delta}$ of size at least m. Let $\Delta = \{\varphi \mid \tilde{\varphi} \in \tilde{\Delta}\}$. Then by assumption we have that $\bigwedge \Delta$ is 0-satisfiable. So, by Proposition 11.4.1 we see that $\widetilde{\bigwedge \Delta}$ has a classical model $\mathcal{M}$ of size at least m. However, it is directly seen that then in fact $\mathcal{M} \models \tilde{\Delta}$, so $\mathcal{M}$ is a classical model for $\tilde{\Delta}$ of size at least m.

So, $\tilde{\Gamma}$ has an infinite classical model. Using the Ehrenfeucht–Mostowski theorem [28] we then know that $\tilde{\Gamma}$ has a model on $[0,1]$ with $([0,1],<)$ as a sequence of indiscernibles, see e.g. Chang and Keisler [18, Theorem 3.3.10]. Because $([0,1],<)$ is a sequence of indiscernibles we can directly see that the relations on this model are Borel (in a similar way as in the proof of Corollary 11.4.4), and because the diagonal has Lebesgue-measure 0 we see that $(\mathcal{M},\lambda) \models_0 \Gamma$, as desired. $\square$

Bibliography

[1] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis: A hitchhiker's guide*, Springer, 2006. [112]

[2] S. N. Artemov, *Explicit provability and constructive semantics*, The Bulletin of Symbolic Logic **7** (2001), no. 1, 1–36. [2]

[3] E. Astor, *Asymptotic density, immunity, and randomness*, to appear. [84]

[4] S.-M. Bageri and M. Pourmahdian, *The logic of integration*, Archive for Mathematical Logic **48** (2009), 465–492. [123]

[5] R. Balbes and P. Dwinger, *Distributive lattices*, University of Missouri Press, 1975. [8, 14]

[6] J. Barwise and S. Feferman (eds.), *Model-theoretic logics*, Perspectives in Mathematical Logic, vol. 8, Springer-Verlag, 1985. [173, 176]

[7] S. S. Basu and S. G. Simpson, *Mass problems and higher-order intuitionistic logic*, submitted. [43]

[8] L. Bienvenu, N. Greenberg, A. Kučera, A. Nies and D. Turetsky, *Coherent randomness tests and computing the K-trivial sets*, to appear. [94]

[9] G. Birkhoff, *Lattice theory*, revised ed., American Mathematical Society, 1948. [125]

[10] V. I. Bogachev, *Measure theory*, vol. 1, Springer, 2007. [9, 103, 111, 112, 124, 144, 148, 150]

[11] R. Brattka, J. S. Miller and A. Nies, *Randomness and differentiability*, to appear in Transactions of the American Mathematical Society. [5, 71]

[12] R. P. Brent, *Fast multiple-precision evaluation of elementary functions*, Journal of the ACM **23** (1976), no. 2, 242–251. [80]

[13] A. M. Bruckner and J. L. Leonard, *Derivatives*, The American Mathematical Monthly **73** (1966), no. 4, 24–56. [71]

[14] C. G. Jockusch, Jr., *Degrees of generic sets*, Recursion theory: its generalisations and applications (F. R. Drake and S. S. Wainer, eds.), London Mathematical Society Lecture Note Series, Cambridge University Press, 1980, pp. 110–139. [19, 71, 72, 79, 87]

[15] C. G. Jockusch, Jr. and P. E. Schupp, *Generic computability, Turing degrees, and asymptotic density*, Journal of the London Mathematical Society **85** (2012), 472–490. [83, 84, 85, 87, 96]

[16] R. Carnap, *On inductive logic*, Philosophy of Science **12** (1945), 72–97. [6]

[17] A. Chagrov and M. Zakharyaschev, *Modal logic*, Clarendon Press, 1997. [8, 15, 21, 22, 28]

[18] C. C. Chang and H. J. Keisler, *Model theory*, third ed., Studies in Logic and the Foundations of Mathematics, vol. 73, North Holland, 1990. [9, 169]

[19] A. Church, *Review: A. M. Turing, on computable numbers, with an application to the Entscheidungsproblem*, The Journal of Symbolic Logic **2** (1937), no. 1, 42–43. [1]

[20] H. B. Curry and R. Feys, *Combinatory logic*, North-Holland, 1958. [2]

[21] D. H. J. de Jongh and A. S. Troelstra, *On the connection of partially ordered sets with some pseudo-Boolean algebras*, Indagationes Mathematicae **28** (1966), 317–329. [21]

[22] O. Demuth, *The differentiability of constructive functions of weakly bounded variation on pseudo numbers*, Commentationes Mathematicae Universitatis Carolinae **16** (1975), no. 3, 583–599, in Russian. [5, 71]

[23] R. G. Downey, C. G. Jockusch, Jr., T. H. McNicholl and P. E. Schupp, *Asymptotic density and the Ershov hierarchy*, to appear in Mathematical Logic Quarterly. [84]

[24] R. G. Downey, C. G. Jockusch, Jr. and P. E. Schupp, *Asymptotic density and computably enumerable sets*, Journal of Mathematical Logic **13** (2013), no. 2, 1350005, 43 pp. [84, 98]

[25] R. G. Downey and D. R. Hirschfeldt, *Algorithmic randomness and complexity*, Springer, 2010. [3, 8, 37, 71, 72, 79, 89, 92]

[26] R. G. Downey, A. Nies, R. Weber and L. Yu, *Lowness and* Π^0_2 *nullsets*, The Journal of Symbolic Logic **71** (2006), no. 3, 1044–1052. [89, 92]

[27] D. D. Dzhafarov and G. Igusa, *Notions of robust information coding*, to appear. [83, 84, 85]

[28] A. Ehrenfeucht and A. Mostowski, *Models of axiomatic theories admitting automorphisms*, Fundamenta Mathematicae **43** (1956), no. 1, 50–68. [168]

[29] S. Figueira, J. S. Miller and A. Nies, *Indifferent sets*, Journal of Logic and Computation **19** (2009), no. 2, 425–443. [87]

[30] C. Freer, B. Kjos-Hanssen, A. Nies and F. Stephan, *Algorithmic aspects of Lipschitz functions*, Computability **3** (2014), no. 1, 45–61. [5]

[31] D. H. Fremlin, *Consequences of Martin's axiom*, Cambridge Tracts in Mathematics, vol. 84, Cambridge University Press, 1984. [115]

[32] D. M. Gabbay, *Properties of Heyting's predicate calculus with respect to r.e. models*, The Journal of Symbolic Logic **41** (1976), no. 1, 81–94. [60]

[33] ———, *Semantical investigations in Heyting's intuitionistic logic*, Springer, 1981. [67]

[34] I. Goldbring and H. Towsner, *An approximate logic for measures*, Israel Journal of Mathematics **199** (2014), no. 2, 867–913. [7]

[35] C. A. Haught, *The degrees below a 1-generic degree* $< 0'$, The Journal of Symbolic Logic **51** (1986), no. 3, 770–777. [19]

[36] P. G. Hinman, *A survey of Mučnik and Medvedev degrees*, The Bulletin of Symbolic Logic **18** (2012), no. 2, 161–229. [8]

[37] D. R. Hirschfeldt and C. G. Jockusch, Jr., *On notions of computability theoretic reduction between* Π^1_2 *principles*, to appear. [83, 84]

[38] D. R. Hirschfeldt, C. G. Jockusch, Jr., R. Kuyper and P. E. Schupp, *Coarse reducibility and algorithmic randomness*, submitted. [83]

[39] D. R. Hirschfeldt, C. G. Jockusch, Jr., T. H. McNicholl and P. E. Schupp, *Asymptotic density and the coarse computability bound*, to appear. [90, 96, 98]

[40] D. R. Hirschfeldt, A. Nies and F. Stephan, *Using random sets as oracles*, Journal of the London Mathematical Society **75** (2007), no. 3, 610–622. [89]

[41] W. Hodges, *Model theory*, Cambridge University Press, 1993. [9, 108, 123, 127]

[42] D. N. Hoover, *Probability logic*, Annals of Mathematical Logic **14** (1978), 287–313. [7, 133, 140]

[43] W. A. Howard, *The formulae-as-types notion of construction*, To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism (J. P. Seldin and J. R. Hindley, eds.), Academic Press, 1980, pp. 479–490. [2]

[44] H. Ishihara, B. Khoussainov and A. Nerode, *Computable Kripke models and intermediate logics*, Information and Computation **143** (1998), 205–230. [60]

[45] ———, *Decidable Kripke models of intuitionistic theories*, Annals of Pure and Applied Logic **93** (1998), 115–123. [60]

[46] V. A. Jankov, *The calculus of the weak "law of excluded middle"*, Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya **32** (1968), no. 5, 1044–1051. [40]

[47] ———, *Three sequences of formulas with two variables in positive propositional logic*, Izvestiya Akademii Nauk SSSR Seriya Matematicheskaya **32** (1968), no. 4, 880–883. [34]

[48] I. Kapovich, A. Myasnikov, P. Schupp and V. Shpilrain, *Generic-case complexity, decision problems in group theory and random walks*, Journal of Algebra **264** (2003), no. 2, 665–694. [6, 83]

[49] M. J. Kearns and U. V. Vazirani, *An introduction to computational learning theory*, MIT Press, 1994. [101]

[50] A. S. Kechris, *Classical descriptive set theory*, Springer-Verlag, 1995. [8, 74, 75, 111, 114, 144, 145, 146, 150]

[51] H. J. Keisler, *Probability quantifiers*, in Barwise and Feferman [6]. [6, 105, 110, 115, 122, 123, 143]

[52] B. Kjos-Hanssen, A. Nies and F. Stephan, *Lowness for the class of Schnorr random sets*, SIAM Journal on Computing **35** (2005), 647–657. [29]

[53] S. C. Kleene, *On the interpretation of intuitionistic number theory*, The Journal of Symbolic Logic **10** (1945), no. 4, 109–124. [2]

[54] ———, *Introduction to metamathematics*, North Holland, 1952. [162]

[55] S. C. Kleene and E. L. Post, *The upper semi-lattice of degrees of recursive unsolvability*, Annals of Mathematics, Second Series **59** (1954), no. 3, 379–407. [3]

[56] S. C. Kleene and R. E. Vesley, *The foundations of intuitionistic mathematics*, Studies in Logic and the Foundations of Mathematics, North-Holland Publishing Company, 1965. [50]

[57] K. Ko, *Complexity theory of real functions*, Birkhäuser, 1991. [80]

[58] A. Kolmogorov, *Zur Deutung der intuitionistischen Logik*, Mathematische Zeitschrift **35** (1932), no. 1, 58–65. [4, 13]

[59] ______, *On the interpretation of intuitionistic logic*, Selected works of A. N. Kolmogorov, Volume I: Mathematics and Mechanics (V. M. Tikhomirov, ed.), Kluwer, 1991, pp. 151–158. [4]

[60] M. Kondô and T. Tugué, *Quelques cribles par rapport aux mesures*, Journal of Mathematics **1** (1952), 55–61. [111]

[61] A. Kučera, *An alternative priority-free solution to Post's problem*, Mathematical Foundations of Computer Science 1986 (J. Gruska, B. Rovan and J. Wiederman, eds.), Lecture Notes in Computer Science, vol. 233, Springer, 1986, pp. 493–500. [94, 96]

[62] K. Kunen, *Set theory: An introduction to independence proofs*, Elsevier, 1980. [115]

[63] R. Kuyper, *First-order logic in the Medvedev lattice*, submitted. [41]

[64] ______, *Computational hardness of validity in probability logic*, Proceedings of Logical Foundations of Computer Science (S. Artemov and A. Nerode, eds.), Lecture Notes in Computer Science, vol. 7734, Springer, 2013, pp. 252–265. [129]

[65] ______, *Natural factors of the Muchnik lattice capturing IPC*, Annals of Pure and Applied Logic **164** (2013), no. 10, 1025–1036. [17]

[66] ______, *Computational aspects of satisfiability in probability logic*, Mathematical Logic Quarterly **60** (2014), no. 6, 444–470. [140]

[67] ______, *Natural factors of the Medvedev lattice capturing IPC*, Archive for Mathematical Logic **53** (2014), no. 7, 865–879. [31]

[68] R. Kuyper and S. A. Terwijn, *Model theory of measure spaces and probability logic*, The Review of Symbolic Logic **6** (2013), no. 3, 367–393. [6, 101, 107]

[69] ______, *Effective genericity and differentiability*, Journal of Logic & Analysis **6** (2014), no. 4, 1–14. [71]

[70] A. H. Lachlan, *Distributive initial segments of the degrees of unsolvability*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik **14** (1968), no. 5, 457–472. [32]

[71] M. van Lambalgen, *The axiomatization of randomness*, The Journal of Symbolic Logic **55** (1990), no. 3, 1143–1167. [3, 90]

[72] F. W. Lawvere, *Adjointness in foundations*, Dialectica **23** (1969), 281–296. [41]

[73] P. A. Loeb, *Conversion from nonstandard to standard measure spaces and applications in probability theory*, Transactions of the American Mathematical Society **211** (1975), 113–122. [123, 124]

[74] L. H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bulletin of the American Mathematical Society **55** (1949), no. 10, 961–962. [95]

[75] W. Maass, R. A. Shore and M. Stob, *Splitting properties and jump classes*, Israel Journal of Mathematics **39** (1981), no. 3, 210–224. [94]

[76] D. A. Martin, *Classes of recursively enumerable sets and degrees of unsolvability*, Mathematical Logic Quarterly **12** (1966), no. 1, 295–310. [25]

[77] J. C. C. McKinsey and A. Tarski, *On closed elements in closure algebras*, The Annals of Mathematics Second Series **47** (1946), no. 1, 122–162. [13]

[78] ———, *Some theorems about the sentential calculi of Lewis and Heyting*, The Journal of Symbolic Logic **13** (1948), no. 1, 1–15. [14, 15]

[79] Yu. T. Medvedev, *Degrees of difficulty of the mass problems*, Doklady Akademii Nauk SSSR, (NS) **104** (1955), no. 4, 501–504. [2, 13, 14]

[80] ———, *Finite problems*, Doklady Akademii Nauk SSSR, (NS) **142** (1962), no. 5, 1015–1018. [34]

[81] W. Miller and D. A. Martin, *The degrees of hyperimmune sets*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik **14** (1968), 159–166. [24, 25, 28]

[82] B. Monin, *Higher computability and randomness*, Ph.D. thesis, Université Paris Diderot–Paris 7, 2014. [87]

[83] Y. N. Moschovakis, *Descriptive set theory*, second ed., Mathematical Surveys and Monographs, vol. 155, American Mathematical Society, 2009. [8, 73, 74]

[84] A. A. Muchnik, *On strong and weak reducibilities of algorithmic problems*, Sibirskii Matematicheskii Zhurnal **4** (1963), 1328–1341. [3, 13, 14]

[85] A. Nies, *Notes on a theorem of Hirschfeldt, Jockusch, Kuyper and Schupp regarding coarse computation and K-triviality*, in A. Nies, ed., Logic Blog 2013, Part 5, Section 24, available at http://arxiv.org/abs/1403.5719. [94]

[86] ———, *Computability and randomness*, Oxford University Press, 2008. [89]

[87] P. G. Odifreddi, *Classical recursion theory*, Studies in Logic and the Foundations of Mathematics, vol. 125, North-Holland, 1989. [8, 24, 25, 26, 162, 164]

[88] ———, *Classical recursion theory: Volume II*, Studies in Logic and the Foundations of Mathematics, vol. 143, North-Holland, 1999. [71]

[89] J. van Oosten, *Realizability: an introduction to its categorical side*, Studies in Logic and the Foundations of Mathematics, vol. 152, Elsevier, 2008. [2, 50]

[90] J. C. Oxtoby, *Measure and category*, second ed., Graduate Texts in Mathematics, vol. 2, Springer-Verlag, 1980. [8, 75]

[91] N. Pathak, C. Rojas and S. G. Simpson, *Schnorr randomness and the Lebesgue differentiation theorem*, Proceedings of the American Mathematical Society **142** (2014), 335–349. [5]

[92] A. M. Pitts, *Notes on categorical logic*, Computer Laboratory, University of Cambridge, Lent Term, 1989. [41, 43, 44, 45]

[93] ———, *Categorical logic*, Logic and Algebraic Methods (S. Abramsky, D. M. Gabbay and T. S. E. Maibaum, eds.), Handbook of Logic in Computer Science, vol. 5, Clarendon Press, 2000, pp. 39–128. [45]

[94] ———, *Tripos theory in retrospect*, Mathematical Structures in Computer Science **12** (2002), 265–279. [41, 42]

[95] D. B. Posner and R. W. Robinson, *Degrees joining to* $0'$, The Journal of Symbolic Logic **46** (1981), no. 4, 714–722. [20, 35]

[96] M. B. Pour-El and J. I. Richards, *Computability in analysis and physics*, Perspectives in Mathematical Logic, vol. 1, Springer-Verlag, 1989. [74, 80]

[97] F. P. Ramsey, *On a problem of formal logic*, Proceedings of the London Mathematical Society, Series 2 **30** (1930), 264–286. [152, 153]

[98] H. Rogers, Jr., *Theory of recursive functions and effective computability*, MIT Press, 1987. [3]

[99] W. Rudin, *Principles of mathematical analysis*, third ed., McGraw-Hill Book Company, 1976. [80]

[100] J. Rute, *Algorithmic randomness, martingales and differentiability*, preprint. [5]

[101] G. E. Sacks, *Degrees of unsolvability*, Princeton University Press, 1963. [37]

[102] ———, *Recursive enumerability and the jump operator*, Transactions of the American Mathematical Society **108** (1963), no. 2, 223–239. [98]

[103] D. S. Scott and P. H. Krauss, *Assigning probabilities to logical formulas*, Aspects of Inductive Logic (K. J. J. Hintikka and P. Suppes, eds.), North-Holland Publishing Company, 1966, pp. 219–264. [6]

[104] P. Shafer, *Characterizing the join-irreducible Medvedev degrees*, Notre Dame Journal of Formal Logic **52** (2011), no. 1, 21–38. [4]

[105] J. Shoenfield, *A theorem on minimal degrees*, The Journal of Symbolic Logic **31** (1966), no. 4, 539–544. [24]

[106] E. Z. Skvortsova, *A faithful interpretation of the intuitionistic propositional calculus by means of an initial segment of the Medvedev lattice*, Sibirskii Matematicheskii Zhurnal **29** (1988), no. 1, 171–178. [3, 15, 41, 61, 176]

[107] ———, *A faithful interpretation of the intuitionistic propositional calculus by means of an initial segment of the Medvedev lattice*, Siberian Mathematical Journal **29** (1988), no. 1, 133–139, translation of [106]. [32, 33, 34, 35]

[108] T. A. Slaman and R. M. Solovay, *When oracles do not help*, COLT '91: Proceedings of the fourth annual workshop on Computational learning theory (M. K. Warmuth and L. G. Valiant, eds.), Morgan Kaufmann Publishers Inc., 1991, pp. 379–383. [18]

[109] C. Smoryński, *Investigations of intuitionistic formal systems by means of Kripke frames*, Ph.D. thesis, University of Illinois, 1973. [22]

[110] A. Sorbi, *Embedding Brouwer algebras in the Medvedev lattice*, Notre Dame Journal of Formal Logic **32** (1991), no. 2, 266–275. [3, 15]

[111] ———, *The Medvedev lattice of degrees of difficulty*, Computability, Enumerability, Unsolvability: Directions in Recursion Theory (S. B. Cooper, T. A. Slaman and S. S. Wainer, eds.), London Mathematical Society Lecture Notes, vol. 224, Cambridge University Press, 1996, pp. 289–312. [8]

[112] A. Sorbi and S. A. Terwijn, *Intermediate logics and factors of the Medvedev lattice*, Annals of Pure and Applied Logic **155** (2008), 69–85. [3, 31]

[113] ———, *Intuitionistic logic and Muchnik degrees*, Algebra Universalis **67** (2012), 175–188. [3, 16, 17]

[114] C. Spector, *On degrees of recursive unsolvability*, Annals of Mathematics **64** (1956), no. 3, 581–592. [26]

[115] C. I. Steinhorn, *Borel structures and measure and category logics*, in Barwise and Feferman [6]. [7]

[116] S. Tennenbaum, *Non-Archimedean models for arithmetic*, Notices of the American Mathematical Society **6** (1959), 270. [5, 41, 55]

[117] S. A. Terwijn, *Probabilistic logic and induction*, Journal of Logic and Computation **15** (2005), no. 4, 507–515. [6, 101, 102, 104]

[118] ———, *The Medvedev lattice of computably closed sets*, Archive for Mathematical Logic **45** (2006), no. 2, 179–190. [3]

[119] ———, *Decidability and undecidability in probability logic*, Proceedings of Logical Foundations of Computer Science (S. Artemov and A. Nerode, eds.), Lecture Notes in Computer Science, vol. 5407, Springer, 2009, pp. 441–450. [7, 107, 129, 139, 151]

[120] S. A. Terwijn and D. Zambella, *Computational randomness and lowness*, The Journal of Symbolic Logic **66** (2001), no. 3, 1199–1205. [29]

[121] A. S. Troelstra and D. van Dalen, *Constructivism in mathematics*, vol. 1, North Holland, 1988. [8, 60]

[122] L. G. Valiant, *A theory of the learnable*, Communications of the ACM **27** (1984), no. 11, 1134–1142. [6]

[123] ———, *Robust logics*, Artificial Intelligence **117** (2000), 231–253. [6]

[124] K. Weihrauch, *Computable analysis: an introduction*, Springer, 2000. [80]

[125] G. Yang, *Computabilities of validity and satisfiability in probability logics over finite and countable models*, arXiv:1410.3059. [7]

[126] L. Yu, *Lowness for genericity*, Archive for Mathematical Logic **45** (2006), no. 2, 233–238. [95]

Index

Samenvatting

In dit proefschrift wordt het samenspel van drie gebieden uit de wiskunde bekeken, namelijk *berekenbaarheid* (Engels: computability), *waarschijnlijkheid* (probability) en *logica* (logic). Voordat we het hierover hebben, is het van belang iets meer te zeggen over wat wiskunde exact inhoudt. Velen zijn bekend met de rekenkant van de wiskunde, getuige de vele malen dat een wiskundige de opmerking "dus jij bent goed met cijfertjes" naar zijn of haar hoofd geworpen krijgt. Dat is echter niet de kern van de wiskunde.

Zoals het woord *wiskunde*, van *wisconst*, al zegt, gaat wiskunde over de de kunst van het gewisse; met andere woorden, de kunst van het zeker weten. De wiskunde gaat over het *bewijzen* van *stellingen*: door te beginnen met eenvoudige basisaannames, de *axioma's*, leidt een wiskundige met behulp van logisch redeneren nieuwe, complexere feiten af, de zogenaamde stellingen. Het meest bekende voorbeeld van een stelling is waarschijnlijk de stelling van Pythagoras over driehoeken. Ondanks dat je uiteraard nooit alle driehoeken kunt controleren, kan een wiskundige door logisch te redeneren toch bewijzen dat deze stelling voor alle driehoeken geldt en daarin ligt de kracht van de wiskunde. Een stelling kan in het bijzonder gerelateerd zijn aan berekeningen; in dit geval zegt de stelling dat een bepaald rekentrucje altijd werkt.

In de logica bestuderen we de logische denkstappen van een wiskundige. Met andere woorden, we bekijken de bewijzen die wiskundigen opschrijven. Dit kan vanuit filosofische hoek, door te bekijken wat de geldige axioma's zijn van waaruit een wiskundige kan redeneren. De moderne *wiskundige logica*, waar dit proefschrift over gaat, bekijkt de logica echter vanuit een andere hoek. Het cruciale idee is dat een wiskundig bewijs zelf een wiskundig object is, waarover wiskundig geredeneerd kan worden. Daarom kunnen we stellingen bewijzen over bewijzen. We weten bijvoorbeeld, dankzij Gödel, dat er *onafhankelijke* vragen bestaan: dat zijn vragen waarvan we kunnen laten zien dat een wiskundige niet kan bewijzen dat het antwoord 'ja' is, noch dat het antwoord 'nee' is. De kracht van de wiskunde is dus niet onbegrensd.

In de *recursietheorie*, het deelgebied van de wiskunde dat *berekenbaarheid* bestudeert, bekijken we het concept van een *algoritme*. Volgens Van Dale is een algoritme een "systematisch stelsel voor het uitvoeren van rekenkundige bewerkingen en de volgorde daarvan". Bekijk bijvoorbeeld de opgave "gegeven drie getallen a, b en c, geef een nulpunt x van $ax^2 + bx + c = 0$". Zoals iedereen als het goed is bekend is, is er een systematische manier om een nulpunt uit te rekenen, namelijk met de *abc*-formule; dit is een algoritme voor deze opgave. We noemen dit probleem

daarom *berekenbaar*, wat wil zeggen dat er een algoritme bestaat dat dit probleem oplost.

De notie van een algoritme is informeel, maar is geformaliseerd door zowel Church als Turing in 1936. Hoewel hun definities op het eerste gezicht lijken te verschillen, blijken zij *equivalent* aan elkaar te zijn; dat wil zeggen, je kunt wiskundig bewijzen dat ze allebei hetzelfde concept karakteriseren. De definitie van Turing is wellicht het meest intuïtief. Volgens deze definitie is een een probleem berekenbaar precies als er een *Turingmachine* bestaat die het probleem oplost. Ruwweg is een Turingmachine een abstracte vorm van de computer zoals wij die kennen, met in het bijzonder de abstracte eigenschap dat de machine onbegrensd veel geheugen heeft.

Het is interessant om te kijken naar de problemen die berekenbaar zijn, maar in de recursietheorie zijn we vooral geïnteresseerd in de problemen die niet berekenbaar zijn. Dat zulke problemen bestaan werd als eerste opgemerkt door Turing, en hij gaf zelfs een expliciet voorbeeld: het zogenaamde *Stopprobleem*. Als we dit probleem terugvertalen naar de concrete computers zoals wij die kennen, zegt het feit dat het Stopprobleem niet berekenbaar is dat er geen computerprogramma bestaat dat detecteert wanneer een ander programma vastloopt. Er zijn echter vele onberekenbare problemen en het is zelfs mogelijk om een stratificatie aan te brengen die uitdrukt hoe onberekenbaar een probleem precies is.

Het eerste deel van dit proefschrift gaat over de combinatie van berekenbaarheid en logica in de vorm van de *Medvedev* en *Muchnik tralies*. Dit zijn twee wiskundige structuren, geïntroduceerd door Medvedev en Muchnik, die gedefinieerd zijn met behulp van concepten uit de recursietheorie. Deze structuren zijn formaliseringen van een eerder idee van Kolmogorov, die probeerde een *klassieke* interpretatie te geven aan de *intuïtionistische logica*.

Het verschil tussen de klassieke logica, de logica waarbinnen bijna alle wiskundigen werken, en de intuïtionistische logica, zit in de exacte elementaire denkstappen die een wiskundige kan zetten. Een voorbeeld van zo een elementaire denkstap, die zowel in de klassieke als in de intuïtionistische logica geldt, is "als de uitspraak A waar is, dan is de uitspraak 'A of B' ook waar". Het verschil tussen de twee logica's zit echter in de zogenaamde *wet van de uitgesloten derde*: in de klassieke logica nemen we aan dat voor elke uitspraak A de uitspraak 'A is waar of A is niet waar' een ware uitspraak is, terwijl we dit in de intuïtionistische logica niet doen.

Hoewel de klassieke logica de standaard is in de wiskunde, is het niettemin interessant om de intuïtionistische logica te bekijken vanuit een klassiek standpunt. Daarmee bedoelen we dat we intuïtionistische bewijzen bekijken als wiskundige objecten, en over deze objecten vervolgens klassiek redeneren. Eén van de mogelijkheden om dit te doen is met behulp van de Medvedev en Muchnik tralies, die laten zien dat er een sterk verband is tussen redeneren in de intuïtionistische logica aan de ene kant en berekenen aan de andere kant. In dit proefschrift bekijken we dit verband nader.

In het tweede deel van dit proefschrift bekijken we enkele onderwerpen in de *algoritmische waarschijnlijkheid*, een gebied dat een combinatie is van berekenbaarheid en waarschijnlijkheid. In de algoritmische waarschijnlijkheid bestuderen we de vraag: wanneer is een oneindige reeks *willekeurig*? Bekijk bijvoorbeeld de volgende rij:

$$01010101010101010101\ldots.$$

In deze rij is een duidelijk patroon te herkennen, dus we zouden deze rij niet snel als willekeurig bestempelen. Een rij als

$$00100100001111110110\ldots$$

lijkt al een stuk willekeuriger, maar is dat in feite niet. Het zijn namelijk de decimalen van π, maar dan in het binaire getallenstelsel zoals dat onder andere door computers gebruikt wordt. Deze rij is exact uit te rekenen en daarom dus voorspelbaar. Bekijken we echter een rij als

$$00100101110101010101\ldots$$

dan lijkt deze een stuk willekeuriger (deze rij heeft de auteur verkregen door 20 keer een munt op te gooien). Wat bedoelen we echter precies wanneer we zeggen dat een rij willekeurig is?

Gezien de informele aard van dit begrip is het niet verrassend dat er meerdere antwoorden mogelijk zijn. Er blijkt echter één notie te zijn die met kop en schouders boven de rest uitsteekt om de beste kandidaat te zijn: Martin-Löf-willekeurigheid. De natuurlijkheid van deze notie wordt versterkt doordat er drie, op het eerste gezicht compleet verschillende, maar equivalente definities blijken te zijn.

De *onvoorspelbaarheidsdefinitie* is het makkelijkste om uit te leggen. Stel je voor dat je in een casino bent, en je mag je geld inzetten op een zekere onbekende oneindige rij x van nullen en enen, zoals de drie rijen hierboven. Dat wil het volgende zeggen: we beginnen met 10 euro, en mogen dit bedrag verdelen over de optie 0 en de optie 1. Zet bijvoorbeeld 3,49 euro in op 0 en 6,51 euro op 1. Als de rij x begint met een 0 krijgen we twee keer onze inleg op 0 terug, oftewel 6,98 euro, en als de rij begint met een 1 krijgen we twee keer onze inleg op 1 terug, oftewel 13,02 euro. Dit geld zetten we vervolgens opnieuw in, maar nu op het tweede getal uit de rij x. We krijgen opnieuw uitbetaald en kunnen nu inzetten op het derde getal uit de rij, enzovoort. We gaan zo net zo lang door als we willen, totdat we besluiten om te stoppen. Het geld dat we op dat moment hebben, nemen we mee naar huis.

Als we volgens deze regels spelen, dan willen we graag inzetten op de eerste rij die we hierboven noemden. We weten immers elke keer precies wat het volgende getal wordt, dus we kunnen elke keer al ons geld op dat getal inzetten en op die manier zoveel geld verdienen als we willen, mits we maar lang genoeg doorgaan. Hetzelfde geldt voor de tweede rij. Voor de derde rij wordt het echter een ander verhaal, want we kunnen niet anders doen dan willekeurig gokken en we kunnen dus niet garanderen dat we met winst thuiskomen. Dit is de kern van de onvoorspelbaarheidsdefinitie van Martin-Löf willekeurigheid: een rij x is willekeurig als er geen algoritme bestaat waarmee je willekeurig veel geld kan verdienen door lang genoeg op x in te zetten, waarbij we met een algoritme in abstracte zin een Turing machine zoals hierboven beschreven bedoelen, of in concrete zin een computerprogramma.

Voor de lezer die iets meer bekend is in de informatica is de tweede definitie ook goed uit te leggen. Een oneindige rij x is Martin-Löf-willekeurig als er geen compressiealgoritme (zoals gzip, RAR of ZIP) bestaat dat x comprimeert. Dus, de oneindige rijen x zijn de rijen die zoveel informatie bevatten, dat ze niet op een eenvoudige manier te comprimeren of samen te vatten zijn. Het feit dat de eerder genoemde algoritmes toch in het dagelijks gebruik nuttig blijken te zijn berust dan ook vooral op het feit dat de meeste computerbestanden die wij produceren niet willekeurig zijn. De derde definitie is lastiger uit te leggen aan niet-wiskundigen en noemen we daarom hier niet.

In dit proefschrift bestuderen we een verband tussen het begrip *1-genericiteit* (een begrip dat nauw verwant is aan Martin-Löf-willekeurigheid) aan de ene kant, en differentieerbaarheid van functies (een ander wiskundig begrip) aan de andere kant. Verder bestuderen we verbanden tussen Martin-Löf-willekeurigheid en *grove berekenbaarheid*, een variant van berekenbaarheid die een mogelijke formalisering is van 'bijna berekenbaar'.

Het derde en laatste deel van dit proefschrift gaat over het combineren van logica met waarschijnlijkheid, wat leidt tot zogenaamde *probabilistische logica's*. We bekijken een specifiek voorbeeld van een probabilistische logica, genaamd *ε-logica*. Ter motivatie van deze logica bekijken we eerst de volgende situatie: stel, we lopen langs een meer en zien daar een grote groep zwanen zwemmen. Het valt ons op dat deze allemaal wit zijn. We vermoeden daarom dat alle zwanen wit zijn. Echter, om de uitspraak 'alle zwanen zijn wit' te controleren, moeten we óf alle zwanen controleren en zien dat deze inderdaad allemaal wit zijn, óf we moeten een zwaan van een andere kleur vinden. We weten echter nooit wanneer we álle zwanen gezien hebben, dus als we na het controleren van duizenden zwanen alleen nog maar witte zwanen hebben gezien, zijn we geneigd om te concluderen dat inderdaad alle zwanen wit zijn. Dit biedt echter geen garantie, want we kunnen vijf minuten later bij toeval alsnog een zwarte zwaan tegenkomen. Met andere woorden, we kunnen een 'voor alle'-uitspraak niet empirisch bewijzen.

In ε-logica vervangen we daarom 'voor alle' door 'voor bijna alle', waarbij een 'voor bijna alle'-uitspraak nog steeds waar kan zijn als er tegenvoorbeelden zijn, zolang de kans maar klein is dat we een dergelijk tegenvoorbeeld tegenkomen. Om terug te komen op de zwanen, als we duizenden witte zwanen hebben gezien zonder ook maar een enkele niet-witte zwaan tegen te komen, kunnen we zoals hierboven omschreven niet concluderen dat alle zwanen wit zijn, maar we kunnen wél met hoge waarschijnlijkheid concluderen dat bijna alle zwanen wit zijn. Aan de andere kant, om te concluderen dat niet alle zwanen wit zijn is het genoeg om slechts één zwaan tegen te komen die niet wit is. We noemen ε-logica daarom een *leerbare* logica: gegeven een claim kunnen we empirisch controleren of de claim of zijn ontkenning waar is.

We bekijken in dit proefschrift enkele wiskundige basiseigenschappen van deze logica. Zo bewijzen we varianten van enkele belangrijke stellingen uit de klassieke logica, zoals de neerwaartse Löwenheim–Skolem-stelling en de stelling van Łoś. Verder bekijken we hoe lastig het voor een computer is om over deze logica te redeneren, waarmee de cirkel tussen *Computability, Probability and Logic* rond is.

Curriculum Vitae

Rutger Kuyper was born in 1989 in Haarlem. After attending Atheneum College Hageveld, graduating *cum laude* in 2007, he studied mathematics at the Radboud University Nijmegen. From this university he obtained a Bachelor of Science in Mathematics in 2010 and a Master of Science in Mathematics with the specialisation Mathematical Foundations of Computer Science in 2011, both *summa cum laude*. In 2012 he started his Ph.D. research under supervision of dr. Sebastiaan A. Terwijn, funded by an NWO/DIAMANT grant and the University of Leiden.

In 2013 he obtained a "Turing Centenary Research Scholarship" funded by the John Templeton Foundation, which he used to visit many experts in the field. In particular he visited the Victoria University of Wellington in New Zealand for two months.