**Title**
Computable Continuous Structure Theory



**Author**
Moody, James Gardner

**Publication Date**
2019

Peer reviewed|Thesis/dissertation

# Computable Continuous Structure Theory

by

James Gardner Moody

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Logic and the Methodology of Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Theodore Slaman, Co-chair
Professor Antonio Montalbán, Co-chair
Professor John Steel
Professor Thomas Scanlon
Professor Wesley Holliday
Professor Michael Christ

Spring 2019

**Computable Continuous Structure Theory**

## Abstract

Computable Continuous Structure Theory

by

James Gardner Moody

Doctor of Philosophy in Logic and the Methodology of Science

University of California, Berkeley

Professor Theodore Slaman, Co-chair

Professor Antonio Montalbán, Co-chair

We investigate structures of size at most continuum using various techniques originating from computable structure theory and continuous logic. Our approach, which we are naming "computable continuous structure theory", allows the fine-grained tools of computable structure theory to be generalized to apply to a wide class of separable completely-metrizable structures, such as Hilbert spaces, the p-adic integers, and many others. We can generalize many ideas, such as effective Scott families and effective type-omitting, to this wider class of structures. Since our logic respects the underlying topology of the space under consideration, it is in some sense more natural for structures with a metrizable topology which is not discrete.

To everyone who helped me continue work in times of trouble.

I am greatly indebted to Ted Slaman, whose support went beyond reason.

# Contents

## Acknowledgments

I am using this space to acknowledge the animals who were killed and mistreated as a result of my choices before I decided to stop eating meat and using animals products. They have suffered because of me, and many of their body parts which they could have used better for their own existence are now part of me, being used instead for pursuits such as the academic study of logic which are relatively trivial in comparison. We can hope for a future in which our mathematical results are not powered by exploitation of sentient animals.

# Philosophical Background

## 0.1 A Brief Introduction for a General Audience

The entire edifice of mathematics is rooted in human experience. Mathematical understanding is more than just skill in symbolic manipulation. The languages that mathematicians have developed are tools to expand the reach of the human mind. Far from being universal, mathematical languages are designed fit the cognitive abilities, intuitions, and limitations of the human mind. Mathematical logic studies rigorously how mathematical language and its associated rules carve out mathematical concepts.

It has been considered a mystery why mathematical concepts, apparently objective, unchanging, and non-physical, have proven so useful in understanding reality. In other fields of scientific study, entire theories are regularly discarded and replaced to improve empirical adequacy. We can understand why these theories are useful by analogy to natural selection: only the useful theories survive the brutal process of the scientific method. But in mathematics, theories once established are established forever, barring logical inconsistencies; furthermore, agreement with observation of physical reality is not considered by pure mathematicians to be an important criterion for a mathematical theory. One might expect, then, that mathematics ought to be about as practically useful as theology. Yet somehow, mathematics has been incredibly useful for describing and predicting reality.

The physicist Max Tegmark has proposed one explanation: the physical universe *is* a mathematical structure. There is perhaps a better explanation, however. Mathematicians, in the course of their work, build general purpose cognitive tools which fit together nicely. The tools which work to solve only one particular problem tend to be forgotten and considered not beautiful, while the tools with wide applicability are developed further and refined to be easier to use and more powerful. It doesn't matter that mathematicians are using these tools to solve problems which have almost no connection to physical reality: these are tools to aid in mental work, not physical work. The same human mind which tries to understand the natural numbers is employed when trying to send a rocket to the moon. The constant factor in the application of mathematical thinking to physical problems and to abstract problems in number theory is that there is a human doing the thinking. The cognitive problems which mathematicians learn to overcome are often ubiquitous in human thought in general. From this perspective, the "unreasonable effectiveness of mathematics in the natural sciences", as Eugene Wigner put it, is about as mysterious as the unreasonable effectiveness of English in science fiction novels.

That being said, it is still worth asking for particular mathematical techniques, "why does this work so well?". Far from being blind beneficiaries of the evolution of mathematical thought, we guide it by choosing to care about mathematical virtues like beauty, universality, and simplicity. The areas we are concerned with in this thesis are recursion theory and general topology. Recursion theory can be thought of as the study of the dynamics of information. General topology can be thought of as the study of categorization (in the ordinary, non-mathematical sense of the term). There is a deep connection between the concepts of a recursively enumerable set in recursion theory, an open set in topology, and a verifiable property in philosophy of science. Dually, there is a connection between recursively co-enumerable sets in recursion theory, closed sets in topology, and falsifiable properties in philosophy of science.

This connection has been explored in the field of effective descriptive set theory, but we wish to take an alternative approach using a newly-revived form of logic called "continuous logic", which incorporates topology into the fabric of our mathematical language. In continuous logic, truth values now lie on a continuum, rather than the discrete set $\{True, False\}$ or $\{1, 0\}$, and formulas are understood as continuous functions. The advantage to this approach is that we can, in the same breath, talk about both discrete and continuous structures without sacrificing any naturality of presentation. Our aim is to faithfully generalize results from computable structure theory to this new setting, allowing the tool set of computable structure theory to be transparently applied to well-behaved uncountable structures as well.

## 0.2 Why have a Continuum of Truth Values?

There are a few problems in the philosophy of language known as "vagueness paradoxes". The classical, Aristotelian basis for reasoning is based on assigning statements or propositions truth values in a systematic way. If a statement is well-defined, we expect it to be either true or false, and we can build connections, through logical reasoning, between the truth-values of some sentences and other sentences. A valid deductive argument, with true statements as premises, must have a true conclusion. The problem comes when we try to apply this reasoning to statements involving vague terms, such as "warm", "blue", "huge", etc. Consider the Sorites paradox [7]:

A heap is a large pile of stuff. We seem to understand what a heap of sand would look like. Our intuition also tells us that, since grains of sand are so minuscule, if you have a heap of sand, and remove a single grain, it will still be a heap. On the other hand, we can definitively say that a few grains of sand on the ground do not form a heap: a heap should be large! But then we end up with the following, seemingly valid argument. Our premises are: (i) if we have a billion grains of sand in a pile, it will form a heap; (ii) for any number $n$ of sand grains, if a pile of $n$ grains forms a heap, then a pile of $n-1$ grains also forms a heap; and (iii) a pile of six grains sand on the ground does not form a heap. Using classical reasoning, we start from premise (i), which tells us that for $n = 1000000000$, $n$ grains of sand in a pile forms a heap, and repeatedly apply premise (ii) to get that 999999999, 999999998, 999999997, etc. grains of sand in a pile also would form a heap. After applying premise (ii) 999999993 times, we conclude that six grains of sand in a pile forms a heap, contradicting premise (iii).

Similar arguments can be used to prove absurdities like "red = blue". The idea is that if we take a shade of red, and make it imperceptibly more blue, it will still appear red to us (since the change in hue is imperceptible). On the other hand, repeating such an imperceptible change in hue many, many times could take us from a red hue to a blue hue.

These apparent paradoxes can be resolved without abandoning classical logic, but doing so seems to change the meaning of everyday words. For example, we might redefine "red" to be a precise band of frequencies, but then we are forced to say that it's possible for one hue which is visually indistinguishable from another hue to be red, while the other is not red. Taking this further, the difference in corresponding frequencies could hypothetically be so small that it would be physically impossible to detect a difference in frequency in the lifespan of the universe (an inevitable consequence of Heisenberg's Uncertainty Principle). In other words, if we try to resolve the paradox by creating a precise definition of "red", we end up concluding that whether a given frequency of light is red or not may not be physically observable.

The fundamental problem, from our perspective, is that we, as humans, are forcing a binary distinction on something which does not naturally fit into binary categories but in some kind of continuum. We'll talk a bit more about how physicists solve this problem in the section on Hilbert Spaces and Observations.

For now, we have a somewhat nice resolution to this problem that doesn't require us to completely rewrite mathematical reasoning in the physicists' preferred framework: broaden our perspective to allow for statements to have intermediate truth values. In our color example, the idea would be that as we shift the hue continuously from red to blue, the truth value of "this is red" shifts continuously from true to false, taking on intermediate values in the region of hues which are not clearly red clearly blue. It's convenient here to think of truth values as lying in the interval $[0, 1]$, with 0 corresponding to completely false, and 1 corresponding to completely true. It's important that we *do not* identify truth values with probabilities, as tempting as it may be. Probabilities can certainly be thought of as a sort of generalized truth value, but we don't want to commit ourselves to thinking that if "this is red" is $\frac{3}{4}$ true, then there is a $\frac{3}{4}$ probability it is red, and a $\frac{1}{4}$ probability it is not red.

Rather, we want to eschew our preconceived notion that everything must ultimately either end up completely true or completely false.

Of course, this creates some problems for our simple classical rules of inference. But they can be replaced with approximate rules of inference. For example, we could generalize some of the logical connectives from classical logic by defining the connectives "$\wedge$", "$\vee$" and "$\neg$" to be interpreted as min, max, and $x \mapsto 1-x$ on truth values, and "$\Rightarrow$" to be interpreted as $(x,y) \mapsto 1-(x \dot{-} y)$, where $x \dot{-} y$ is $x-y$ if this is non-negative, and 0 otherwise. Note that with this choice of connectives, the classical equivalence between $P \Rightarrow Q$ and $Q \vee \neg P$ no longer holds (for another choice, this would hold). However, they still retain some of their properties. "$P \Rightarrow Q$" still means something like "Q is at least as true as P". If an implication has truth value, say $\frac{2}{3}$, this means the truth of the consequent is at worst $\frac{1}{3}$ less than the truth of the antecedent. Using this, we can resolve the Sorites paradox as follows. We believe that the implication "n grains of sand in a pile forms a heap $\Rightarrow n-1$ grains of sand in a pile forms a heap" has truth value at least $\frac{999}{1000}$, but not 1. This means if we start with a premise with truth value 1, like "a billion grains of sand in a pile forms a heap", and apply this almost-true implication only a few dozen times, we will still be left with an almost-true conclusion. If we apply it a billion times, however we are no longer guaranteed anything about the truth value of the conclusion. Our new implication is no longer transitive, but it is approximately transitive, in that if $P \Rightarrow Q$ has truth value at least $\tau$, and $Q \Rightarrow R$ has truth value at least $\tau'$ then $P \Rightarrow R$ has truth value at least $1-(1-\tau)-(1-\tau') = \tau+\tau'-1$. When $\tau$ and $\tau'$ are both 1, this gives us transitivity of implication in classical logic.

What we've said so far can be considered one starting point into something like Lotfi Zadeh's "fuzzy logic", which has been successfully applied to the discipline of control theory. See Zadeh's 1965 paper "Fuzzy Sets" [21], for example. So allowing more general truth values is not only of philosophical interest, but also practically useful. You might still be skeptical about whether there are any applications to pure mathematics, however. You might compare some applications of fuzzy logic to the analogue ballistics computers used in old battleships: their design may be inspired by some principled physics, and they may be beautiful machines[1], but it seems unlikely that new physics or mathematics would arise from studying them. If you peruse the literature on fuzzy logic, you will find numerous examples of seemingly arbitrary choices made in the process of comporting fuzzy logic to particular applied problems. One particular mistake is the conflation of intermediate truth values with probabilities.[2]

This is nothing against fuzzy logic itself, it's just that if we want to introduce intermediate

[1]From an engineering perspective, that is, not from the perspective of a scared draftee

[2]Lotfi Zadeh himself has indicated it is important to distinguish between probabilities and fuzziness/vagueness. Some people use the choice of connective $\wedge : (x,y) \mapsto x * y$, which resembles computing the probability of a conjunction of two independent events. It will often be OK in many applications to conflate these two things (since many events are, after all, independent) but it is conceptually wrong.

truth values into mathematics itself in a way mathematicians will actually care about, we should show some tangible benefit and some serious mathematical structure. The direction we are headed in is the direction of Chang and Keisler's "Continuous Model Theory" (1966) [4]. The idea here can be traced back to what many would consider to be normal (even essential, classical) mathematics: studying a space by looking at a ring of functions on it. From a mathematician's perspective, logicians have been perhaps myopically focused only on functions valued in $\mathbb{F}_2 = \{0, 1\}$. This works really well for discrete structures, but could be considered unnatural when applied to continuous structures. For example $\leq$, as a $\{0, 1\}$-valued relation on the reals, does not respect the topology of the reals: it is not continuous. To emphasize that we care more about topology than vagueness (which many would say has no place in mathematics), we'll be using the term "continuous logic" rather than "fuzzy logic" to describe what we are doing. This is the terminology used by Ben Yaacov, Berenstein, Henson, and Usvayatsov, who have recently revived Chang and Keisler's work with their recent "Model Theory for Metric Structures" [19]. Another candidate term would have been "Łukasiewicz logic", after the work of Łukasiewicz and Tarski on many-valued logics in the 1930s.

It could be argued, and we will, that continuous logic with truth values in $[0, 1]$ is in some sense more natural as a general-purpose logical framework for mathematics than classical $\{0, 1\}$-valued logic. We find that using continuous logic greatly expands the reach of computable structure theory to a much larger class of structures, especially those appearing in analysis.

# Chapter 1

# Computable Continuous Structures

We present here many important definitions and conventions we will use to describe uncountable structures. These definitions combine concepts from computable structure theory and the model theory of metric structures. The idea is that uncountable structures are amenable to study from the perspective of recursion theory so long as their features can be controlled by a suitable metric on that structure.

## 1.1 Basic Definitions

**Convention:** Whenever we use the term "function", we allow any acceptable description of a function, either a set-theoretic function (set of ordered pairs) or a Turing machine or Turing functional which computes the function, or any other reasonable mathematical object which has the ability to be evaluated on objects in the domain to obtain objects in the codomain. If we say a function is a computable function, we take this to mean that the function was described by a *particular* Turing machine or Turing functional, meaning we are allowed to ask "What is *the* index of this computable function?", rather than just "What is *an* index of this computable function?". This is at odds with a different convention, where "computable function" just means a function for which there exists some description or other of the function via a Turing machine or Turing functional. The same convention applies to recursively enumerable sets, partial computable functions, etc.

**Convention:** We do not require in general that the domain and codomain of a function are stored as datum in the description of the function. Rather, for any choice of domain and codomain, we have a "type" of functions with that domain and codomain. The same description of a function may pick out functions with many different possible domains. So really, whenever we are using the word "function", that's really shorthand for "function $D \to C$", where D and C are often left implicit. One consequence of this is that when we say two functions are equal, we are always talking about extensional equality of functions with the same domain and codomain. So, functions given by the exact same Turing machine will be considered not equal if they are thought of as operating on different domains, and two computable functions may be equal even if they are given by completely different Turing functionals. The type information about a function (what its domain and codomain are) is thus to some extent extrinsic and stored separately from the description of the function itself.

**Convention:** We assume that we have fixed a standard one-to-one enumeration $(\dot{q}_i)_{i\in\omega}$ of $\mathbb{Q}$ to serve as our means of coding rational numbers as natural numbers. We suppose this coding yields an interpretation of $\mathbb{Q}$ into $\mathbb{N}$ in which the operations on $\mathbb{Q}$ are computable. The dot on top is just to distinguish it from other dense subsets $(q_i)_{i\in\omega}$ of other metric spaces we will consider later.

**Definition 1.** A **lower name** for $r \in \mathbb{R}$ is a non-decreasing function $\lfloor r \rfloor : \omega \to \mathbb{Q}$ such that $\lim_k \lfloor r \rfloor(k) = r$. A lower name $\lfloor r \rfloor$ for $r$ is **computable** if $\lfloor r \rfloor$ is a computable function $\omega \to \mathbb{Q}$.[1]

**Definition 2.** An **upper name** for $r \in \mathbb{R}$ is a non-increasing function $\lceil r \rceil : \omega \to \mathbb{Q}$ such that $\lim_k \lceil r \rceil(k) = r$. An upper name $\lceil r \rceil$ for $r$ is **computable** if $\lceil r \rceil$ is a computable function $\omega \to \mathbb{Q}$.

**Definition 3.** A **name** for $r \in \mathbb{R}$ is a function $[r] : \omega \to \mathbb{Q}$ such that $|[r](k) - r| \leq \frac{1}{2^k}$. A name $[r]$ for $r$ is **computable** if $[r]$ is a computable function $\omega \to \mathbb{Q}$.[2]

**Exercise for the Reader:** There is a computable name for $r \in \mathbb{R}$ if and only if there is a computable lower name for $r$ and a computable upper name for $r$. There are reals which have a computable lower name but no computable upper name, and vice versa (hint: think about a real whose base 2 digits encode the halting set).

**Definition 4.** Let $X$ be a countable space interpreted in $\mathbb{N}$ (for example $\omega^n$, $\omega^{<\omega}$, or $\mathbb{Q}$). Let $A$ be a not-necessarily-definable subset of $X$. A **lower name** for a real-valued function $f : A \to \mathbb{R}$ is a partial function $\lfloor f \rfloor : X \times \omega \to \mathbb{Q}$ such that for all $x \in A$, $\lfloor f \rfloor(x, \cdot)$ is a lower name for $f(x)$. An **upper name** for f is a partial function $\lceil f \rceil : X \times \omega \to \mathbb{Q}$ such that for all $x \in A$, $\lceil f \rceil(x, \cdot)$ is an upper name for $f(x)$. A **name** for f is a partial function $[f] : X \times \omega \to \mathbb{Q}$ such that for all $x \in A$, $[f](x, \cdot)$ is a name for $f(x)$. A lower name $\lfloor f \rfloor$, an upper name $\lceil f \rceil$, or a name $[f]$ for f is **computable** if $\lfloor f \rfloor$, $\lceil f \rceil$, or $[f]$ respectively is partial computable[3].

[1] We have an interpretation of $\mathbb{Q}$ in $\mathbb{N}$, so we can talk about computable functions $\omega \to \mathbb{Q}$

[2] The reason we use non-strict inequality here is motivated by descriptive set theory considerations. One can build many Polish spaces by quotients a $\Pi^0_1$ subset of $\omega^\omega$ (Baire space), by a $\Pi^0_1$ equivalence relation (which you can think of as an equivalence relation defined by $x \sim y \Leftrightarrow f(x) = f(y)$ for some continuous function $f$). This allows us to have the property "$x \in \omega^\omega$ is a name for $p \in \mathcal{M}$", as well as "$x, y \in \omega^\omega$ are names for the same point" (i.e., equality in $\mathcal{M}$ pulled back to names), both be $\Pi^0_1$.

[3] It makes sense to talk about partial computable functions $X \times \omega \to \mathbb{Q}$, since $X$ is interpreted in $\mathbb{N}$.

**Notation:** If $[f]$ is a name for a real-valued function $f : A \to \mathbb{R}$, we will by abuse of notation write $[f(\alpha)]$ to denote the name $[f](\alpha, \cdot) : \omega \to \omega$ of $f(\alpha)$. Note that $f(\alpha) = \beta$ need not imply $[f(\alpha)] = [\beta]$, but the reverse does hold. The square brackets in $[f]$ do not constitute an operation, because there are many possible ways to name the same function. Rather, the removal of brackets to obtain a function from a name for a function is an operation. The same convention holds for upper/lower names and upper/lower square brackets.

**Definition 5.** A **computable metric space** is a quadruple $((M, d), D, (q_i)_{i\in\omega}, [\delta])$, where $(M, d)$ is a complete, separable metric space, D is a countable dense subset of $M$, $(q_i)_{i\in\omega}$ is an enumeration of $D$, and $[\delta] : D^2 \times \omega \to \mathbb{Q}$ is a computable name for the function $\delta = d|_{D^2} : D^2 \to \mathbb{R}$.[4]

**Observation 1.** *Essentially, a computable metric space is just a complete metric space with a countable dense subset, together with a method to calculate distances on that dense subset. We can recover the metric on the whole space by defining the distance between two Cauchy sequences of elements from D to be the limit of the pairwise distances between the elements of those sequences. If we wanted to, we could eschew the metric space entirely, identify the countable dense subset with $\omega$ via the enumeration, and just use $(\omega, [\delta])$ as a computable structure on its own, because we can recover the metric space using just this information. We won't do this, however, as it will be convenient to refer directly to points in the space.*

**Definition 6.** Let $((M, d), D, (q_i)_{i\in\omega}, [\delta])$ be a computable metric space. A **name** for a point $p \in M$ is a function $[p] : \omega \to D$ such that $d([p](k), p) \leq \frac{1}{2^k}$.

**Observation 2.** *The metric space $\mathbb{R}$ can be turned into a computable metric space by using the rationals as the countable dense subset, $(\dot{q}_i)_{i\in\omega}$ as an enumeration of the rationals, and defining $[\delta] : \mathbb{Q}^2 \times \omega \to \mathbb{R}$ by $[\delta](q, q')(k) = |q - q'|$. A name for a point $r$ in the computable metric space $((\mathbb{R}, d), \mathbb{Q}, (\dot{q}_i)_{i\in\omega}, [\delta])$ is then just a function $[r] : \omega \to \mathbb{Q}$ such that $|[r](k) - r| \leq \frac{1}{2^k}$. This agrees exactly with our original definition of a "name" for $r \in \mathbb{R}$.*

**Definition 7.** Let $((M, d), D, (q_i)_{i\in\omega}, [\delta])$ and $((M', d'), D', (q'_i)_{i\in\omega}, [\delta'])$ be two computable metric spaces. A **name** for a function $\phi : M \to M'$ is a partial function $[\phi] : D^\omega \to D'^\omega$ which sends every name for $p \in M$ to a name for $\phi(p) \in M'$. $[\phi]$ is **computable** if $[\phi]$ can be given by a Turing functional $\Phi$, i.e. for every name $[p]$ for a point in $M$, $\Phi^{[p]}(k) = ([\phi]([p]))(k)$ for all $k \in \omega$.

**Abuse of Notation** Given names $[\phi]$ and $[p]$ for a function $\phi$ and point $p$ respectively, we will write $[\phi(p)]$ as shorthand for $[\phi]([p])$.

[4]A one-to-one enumeration of D allows us to interpret the set $D$ in $\mathbb{N}$, so it makes sense to talk about whether a function $D^2 \times \omega \to \mathbb{R}$ is computable in that context. However, you'll notice we didn't include in our definition that the enumeration is one-to-one. This is not of any essential importance: if we have a computably presented metric space in our sense, we can effectively transform this presentation into a different one with a possibly different dense set with a one-to-one enumeration. We'll go into this more later.

**Note.** The previous definition can be extended in two obvious ways to multivariate functions. One is to give a definition directly of a (computable) name for a multivariate function: just say it sends a *tuple* of names $([p_1], ..., [p_n])$ to a name for $\phi(p_1, ..., p_n)$. Another is to give a definition of product computable metric space (just let the product metric be the max of the distances in each coordinate, and let the countable dense subset be given by the product of the countable dense subsets of each, enumerated in a natural way using a computable pairing/tupling function), and say that a (computable) name for a multivariate function is just a (computable) name for a function on the appropriate product space. These turn out to be equivalent.

**Sanity Check:** With this extended definition, it's easy to verify that if $((M, d), D, (q_i)_{i\in\omega}, [\delta])$ is a computable metric space, then the continuous binary function $d : M^2 \to \mathbb{R}$ has a computable name. Namely, we can set:

$$[d]([a], [b])(k) = [\delta]([a](k + 2), [b](k + 2))(k + 1)$$

.

*Proof.* By definition, $[a](k+2)$ and $[b](k+2)$ are within $2^{-k-2}$ of $a$ and $b$ respectively, so by the triangle inequality $d(a, b)$ is within $2^{-k-1}$ of $d([a](k+2), [b](k+2))$. Since we know the latter is itself within $2^{-k-1}$ of $[\delta]([a](k+2), [b](k+2))(k+1)$, this tells us $[\delta]([a](k+2), [b](k+2))(k+1)$ is within $2^{-k}$ of $d(a, b)$. But $[\delta]([a](k + 2), [b](k + 2))(k + 1)$ is computable uniformly in $k$, an oracle for $[a]$, and an oracle for $[b]$. □

**Definition 8.** A function between two computable metric spaces is **computable** if it is given by a computable name.

In the special case that $\phi$ is a real-valued function on $\mathcal{M}$, where $\mathcal{M}$ is a computable metric space, and $\mathbb{R}$ is given its standard presentation, we can also define upper/lower names for $\phi$:

**Definition 9.** A **lower name** for $\phi$ is a partial function $\lfloor\phi\rfloor : D^\omega \to \mathbb{Q}^\omega$ which sends every name for a point in $p \in M$ to a lower name for $\phi(p)$ in $\mathbb{R}$, i.e. to a non-decreasing sequence of rationals converging to $\phi(p)$ from below. Likewise, an **upper name** for $\phi$ is a partial function $\lceil\phi\rceil : D^\omega \to \mathbb{Q}^\omega$ which sends every name for a point in $p \in M$ to a upper name for $\phi(p)$ in $\mathbb{R}$, i.e. to a non-increasing sequence of rationals converging to $\phi(p)$ from above. A lower/upper name is **computable** if it is given by a Turing functional.

## 1.2 Equalness Relation

We would like to think of metric spaces as being like sets with a continuous notion of equality (close-by points are more equal than far-away points). For this reason (and for some technical reasons that will become clear later), it turns out it is useful to replace the metric $d : M^2 \to \mathbb{R}$ with an **equalness relation** $\varepsilon : M^2 \to [0, 1]$ given by $\varepsilon(x, y) = 2^{-d(x,y)}$. This has the additional advantage of making it easy to accommodate infinite distances (two points are infinite distance apart if and only if $\varepsilon(a, b) = 0$), which means we don't need to talk about $\infty$-metric spaces, and our presentation of a version of the compactness theorem is nicer.[5]

**Definition 10.** An **equalness relation** on a set $M$ is a function $\varepsilon : M^2 \to [0, 1]$ satisfying the following properties:

- $(\forall x, y, z \in M)\big(\varepsilon(x, z) \geq \varepsilon(x, y) * \varepsilon(y, z)\big)$
- $(\forall x, y \in M)\big(\varepsilon(x, y) = \varepsilon(y, x)\big)$
- $(\forall x, y \in M)\big(\varepsilon(x, y) = 1 \Leftrightarrow x = y\big)$

One can easily verify that $d : M^2 \to \mathbb{R}$ is a metric if and only if its corresponding $\varepsilon = 2^{-d} : M^2 \to [0, 1]$ is an equalness relation.

We'll occasionally use the infix notation $x \varepsilon y$ in place of $\varepsilon(x, y)$ when this allows for neater notation, with an order of operations making $\varepsilon$ weaker than every other operation (meaning it is performed last, unless explicitly marked by brackets).

**Definition 11.** We call a set together with an equalness relation a **continuous space**.

All metric spaces can be turned into continuous spaces canonically, but the converse is not true, since a continuous space may have points which are entirely unequal, i.e. have $\varepsilon(a, b) = 0$, which would correspond to an infinite distance.

**Definition 12.** We call a continuous space **discrete** if $\varepsilon$ takes values only in $\{0, 1\}$.

**Observation 3.** *Discrete continuous spaces are just sets, with $\varepsilon$ being identical with equality.*

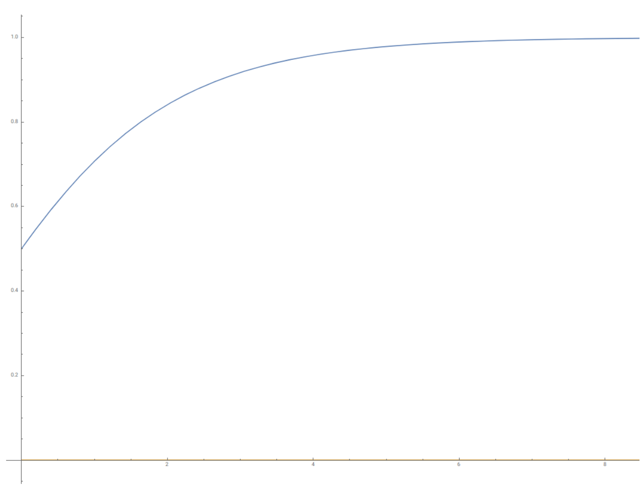
We're now going to replace the definitions in the previous section based on metrics with new definitions based on equalness relations. The reason for giving the previous definitions was to motivate these new definitions, which might seem obtuse otherwise to someone used

[5]The potential issue here is that while any finite subset of the conditions $\{d(a, b) > n : n \in \mathbb{N}\}$ can be realized in a metric space, they cannot all simultaneously be realized.

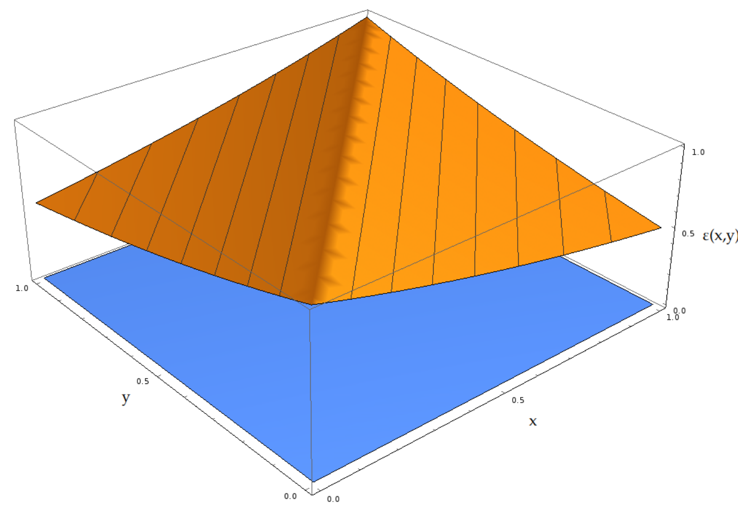
to working with metrics. You can compare some of our definitions with those like that of "effective Polish space" in Moschovakis' "Descriptive Set Theory" [11]. You can also compare the axiom "$(\forall x, y, z \in M)\big(\varepsilon(x, z) \geq \varepsilon(x, y) * \varepsilon(y, z)\big)$" with the triangle inequality: like the ordinary sub-additive triangle inequality for metrics, this super-multiplicative triangle inequality for equalness relations is used to chain together bounds to obtain a common bound.

**Definition 13.** A **name** for $r \in [0, 1]$ is a function $[r] : \omega \to \mathbb{Q} \cap [0, 1]$ with $\varepsilon([r](k), r) \geq 2^{-2^{-k}}$, where $\varepsilon(x, y) = 2^{-|x-y|}$ for $x, y \in [0, 1]$.[6]

The following Mathematica graphs may aid in visualizing this definition:



Graph of $2^{-2^{-k}}$ for $k = 0$ to $10$




Graph of equalness relation on $[0, 1]$

A higher value of $k$ corresponds to a closer-to-1 value of $2^{-2^{-k}}$, which corresponds to $([r](k), r)$ lying closer to the diagonal $x = y$.

**Definition 14.** A **name** for a function $f : A \to [0, 1]$, where $A$ is a not-necessarily-definable subset of some countable $X$ interpretable in $\mathbb{N}$, is a partial function $[f] : X \times \omega \to \mathbb{Q} \cap [0, 1]$ such that for all $x \in A$, $[f](x)$ is a name for $f(x) \in [0, 1]$.

**Definition 15.** A **computable continuous space** is a quadruple $((M, \varepsilon), D, (q_i)_{i \in \omega}, [\hat{\varepsilon}])$, where $(M, \varepsilon)$ is a complete[7], separable[8] continuous space, $D$ is a countable dense subset of

[6] If the typesetting is hard to read here, by $2^{-2^{-k}}$ we mean $2^{(-2^{-k})}$, not $(2^{-2})^{-k}$.

[7] A continuous space is **complete** if whenever $\lim_N \inf_{i,j \geq N} \varepsilon(x_i . x_j) = 1$, there is a limit $x_\infty$ with $\lim_i \varepsilon(x_i, x_\infty) = 1$

[8] A continuous space is **separable** if there is some countable $D \subseteq M$ such that for all $x \in M$ and $\tau < 1$, there is some $q \in D$ with $\varepsilon(q, x) > \tau$

$M$, $(q_i : i \in \omega)$ is an enumeration of $D$, and $[\hat{\varepsilon}] : D^2 \times \omega \to [0,1] \cap \mathbb{Q}$ is a computable name for the function $\hat{\varepsilon} = \varepsilon|_{D^2} : D^2 \to [0,1]$.

**Note.** For continuous spaces arising from metric spaces, we could have used either the old definition of "computable name for a function" here, i.e. $|[\hat{\varepsilon}](q_i, q_j)[k] - \hat{\varepsilon}(q_i, q_j)| \leq 2^{-k}$, or the new one, i.e. $\varepsilon\big([\hat{\varepsilon}](q_i, q_j)[k], \hat{\varepsilon}(q_i, q_j)\big) \geq 2^{-2^{-k}}$. These are equivalent. However, $[\hat{\varepsilon}]$ being a computable name for $\hat{\varepsilon}$ arising from a metric $d$ does not imply that $-log_2([\hat{\varepsilon}])$ is a computable name for $d$ (or rather $\delta = d|_{D^2}$, as we were calling it earlier). The basic reason is that for very large distances, being able to estimate $2^{-d(x,y)}$ to within error $\frac{1}{2^{10000}}$ might not even tell us $d(x,y)$ to within an error 1. In this sense, by moving to an equalness relation rather than a distance function, we are allowing our approximations of large distances to converge more slowly than small distances. This makes sense, though. Imagine trying to name the points of $[0,\infty]$ by branches of an infinite binary tree in an order preserving way. The first split will correspond to breaking up $[0,\infty]$ into two pieces, and the right-most piece will always be bigger, in the sense of Euclidean length, than the left-most piece, meaning the first bit of information gives us less accuracy about the number being described if it is 1 than if it is 0.

**Aside:** You may also noticed, if you read a footnote in the previous section, that we have again cheated slightly here: $D$ is not necessarily interpretable in $\mathbb{N}$ via our enumeration, because we did not require our enumeration to be one-to-one. This is no worry, however: if we have more than one representative for $x, y \in D$, say $q_{i'} = x = q_i$ and $q_{j'} = y = q_j$, then we would find both $|[\hat{\varepsilon}](q_{i'}, q_{j'})(k) - \varepsilon(x,y)| \leq 2^{-k}$ and $|[\hat{\varepsilon}](q_i, q_j)(k) - \varepsilon(x,y)| \leq 2^{-k}$, so our choice of representatives for $p_1, p_2 \in D$ does not really matter, as long as we care only about approximating $\varepsilon(x,y)$ to arbitrarily small error. If we wanted to make a more straightforward (but more notation-heavy) definition, we could have defined $\hat{\varepsilon}$ be a function $\omega^2 \to [0,1]$, given by $\hat{\varepsilon}(i,j) = \varepsilon(q_i, q_j)$, but to simplify our presentation, we have by abuse of notation treated $\hat{\varepsilon}$ as a function $D^2 \to [0,1]$ and swept this issue under the rug. This is unproblematic if we have an injective enumeration of $D$, which is naturally the case for many of the spaces we encounter, such as $\mathbb{R}$. If our enumeration is not injective, we could still identify $D$ with a quotient structure $\mathbb{N}/\sim$ where $i \sim j$ iff $\varepsilon(q_i, q_j) = 1$. In general, the equivalence relation $\sim$ for a computable continuous space would be $\Pi^0_1$ classically, but it turns out this doesn't really matter, since in general strict equality itself of points in $\mathcal{M}$ is $\Pi^0_1$, so using this quotient structure does not increase either the complexity of strict equality on our space $\mathcal{M}$, nor of the equalness relation. In fact, in our context, we don't have strict equality, so what might appear to be a non-computable quotient structure turns out to be computable in the continuous setting. We can think of $\varepsilon(q_i, q_j) : \omega^2 \to [0,1]$ as a continuous generalization of the characteristic function of $\sim$ (which is just another name for classical equality). It's easy to check that for a discrete computable continuous space, $\sim$ will in fact be decidable, meaning we can take the quotient effectively. Hopefully this helps alleviate some worries that our definition of "computable continuous space" might be better described as a "quotient presentation of a continuous space".

**Note.** We prove in Appendix B a theorem that every computable continuous space presented with a non-injective enumeration of $D$ can also be presented with an injective enumeration of $D$. The basic idea is to delay enumerating a point into $D$ (possibly indefinitely) until we have computed the distance from that point to the elements we have already enumerated to within sufficient accuracy to see that it is different from all the points we have already enumerated. This can be done uniformly. Ultimately, this is altogether a boot-strapping problem, not a serious concern. It is just important to keep this in mind, because some model-theoretic constructions more naturally yield a non-injective enumeration of a dense set (such as building function spaces as the completion of a countable algebra of functions on which strict equality is not decidable).

**Definition 16.** Let $((M, \varepsilon), D, (q_i)_{i\in\omega}, [\hat{\varepsilon}])$ be a computable continuous space. A **name** for a point $p \in M$ is a function $[p] : \omega \to D$ such that $\varepsilon([p](k), p) \geq 2^{-2^{-k}}$

**Definition 17.** Let $((M, \varepsilon), D, (q_i)_{i\in\omega}, [\hat{\varepsilon}])$ and $((M', \varepsilon), D', (q'_i)_{i\in\omega}, [\hat{\varepsilon}'])$ be computable continuous spaces. A **name** for a function $\phi : M \to M'$ is a partial function $[\phi] : D^\omega \to D'^\omega$ which sends every name for $p \in M$ to a name for $\phi(p) \in M'$.

The following definition gives us a concept analogous to that of a modulus of continuity for functions between metric spaces:

**Definition 18.** An **n-ary modulus** for functions between continuous spaces is a function $\Lambda : [0, 1]^n \to [0, 1]$ such that:
(i) $\forall \bar{u}, \bar{v} \in [0, 1]^n$, $\Lambda(\bar{u}) \geq \Lambda(\bar{u}\bar{v}) \geq \Lambda(\bar{u})\Lambda(\bar{v})$, where $\bar{u}\bar{v}$ is the coordinate-wise product.
(ii) $\Lambda$ is continuous and $\Lambda(\bar{1}) = \Lambda(1, 1, \ldots, 1) = 1$

**Example 1.** $\Lambda : [0, 1]^n \to [0, 1]$ defined by $\Lambda(v_0, ..., v_{n-1}) = min(v_0, \ldots, v_{n-1})$ is an n-ary modulus. We'll call this the 1-Lipschitz modulus.

**Definition 19.** A function $f : \prod_{i<n} M_i \to M$ between continuity spaces respects the modulus $\Lambda$ if for all $\bar{x}, \bar{y} \in \prod_{i<n} M_i$, $\epsilon(f(\bar{x}), f(\bar{y})) \geq \Lambda(\bar{\epsilon}(\bar{x}, \bar{y}))$, where $\bar{\epsilon}(\bar{x}, \bar{y}) = (d_i(x_i, y_i))_{i<n}$.

If a function respects a modulus, it is not only continuous, but uniformly continuous in a way which can be explicitly controlled. The equalness (think, degree of closeness) of $f(\bar{x})$ and $f(\bar{y})$ can be estimated purely in terms of the degree of equalness of $\bar{x}$ and $\bar{y}$.

**Example 2.** The function $\wedge : [0, 1]^2 \to [0, 1]$ defined by $u \wedge v = min(u, v)$ (with $[0, 1]$ presented in some standard way as a computable continuous space) respects the 1-Lipschitz modulus $\Lambda$. We call such functions **1-Lipschitz**. The functions $\neg : [0, 1] \to [0, 1]$ defined by $\neg v = 1 - v$ and $\vee : [0, 1]^2 \to [0, 1]$ defined by $u \vee v = max(u, v)$ are also 1-Lipschitz.

Intuitively, if a function is 1-Lipschitz, the values of that function on two different inputs are at least as equal as the inputs.

**Observation 4.** *The composition of moduli is a modulus.*

*Proof.* Suppose $\Lambda_i$, $i < n$ are $n_i$-ary moduli, and $\Lambda$ is an $n - ary$ modulus. Since the composition of moduli is continuous, and $\Lambda(\Lambda_0(\bar{1}), \dots, \Lambda_{n-1}(\bar{1})) = \Lambda(1, \dots, 1) = 1$, condition $(ii)$ is satisfied. Condition $(i)$ implies $\Lambda$, $\Lambda_0, \dots, \Lambda_{n-1}$ are non-decreasing in each coordinate, so we can see that $\Lambda_i(\bar{u}_i) \geq \Lambda_i(\bar{u}_i\bar{v}_i)$ for all $i < n$, and thus

$$\Lambda(\Lambda_0(\bar{u}_0), \dots, \Lambda_{n-1}(\bar{u}_{n-1})) \geq \Lambda(\Lambda_0(\bar{u}_0\bar{v}_0), \dots, \Lambda_{n-1}(\bar{u}_{n-1}\bar{v}_{n-1}))$$

. which is the left-hand inequality of condition $(i)$. We can also see

$$\begin{aligned}\Lambda\big(\Lambda_0(\bar{u}_0\bar{v}_0), \dots, \Lambda_{n-1}(\bar{u}_{n-1}\bar{v}_{n-1})\big) &\geq \Lambda\big(\Lambda_0(\bar{u}_0)\Lambda_0(\bar{v}_0), \dots, \Lambda_{n-1}(\bar{u}_{n-1})\Lambda_0(\bar{v}_{n-1})\big) \\ &\geq \Lambda\big(\Lambda_0(\bar{u}_0), \dots, \Lambda_{n-1}(\bar{u}_{n-1})\big)\Lambda\big(\Lambda_0(\bar{v}_0), \dots, \Lambda_{n-1}(\bar{v}_{n-1})\big)\end{aligned}$$

so the right-hand inequality of condition $(i)$ is satisfied. □

**Observation 5.** *If each $f_i : \prod_{j<n_i} M_{ij} \to N_i$ respects modulus $\Lambda_i$ for $i < n$, and $g : \prod_{i<n} N_i \to K$ respects modulus $\Lambda$, then the composition $g \circ (f_0, f_1, \dots, f_{n-1}) : \prod_{j<n_i, i<n} M_{ij} \to K$ respects the composition of their moduli, $\Lambda \circ (\Lambda_0, \Lambda_1, \dots, \Lambda_{n-1})$.*

*Proof.* Exercise for the reader. □

**Observation 6.** *The composition of 1-Lipschitz moduli is itself a 1-Lipschitz modulus, so the composition of 1-Lipschitz functions is 1-Lipschitz.*

*Proof.* $min_{i<n}\big(min_{j<k_i} u_{ij}\big) = \min_{i<n, j<k_i} u_{ij}$. □

**Definition 20.** A modulus $[0,1]^n \to [0,1]$ is **computable** if it is computable as a function $[0,1]^n \to [0,1]$ with $[0,1]$ interpreted as a computable continuous space using the equalness relation $\varepsilon(\tau_1, \tau_2) = 2^{-|\tau_1 - \tau_2|}$.

**Definition 21.** A modulus is **bounded** if $\Lambda(\bar{0}) \neq 0$.

**Note.** A bounded modulus can be everywhere bounded away from 0, i.e. there is some $\delta > 0$, namely $\delta = \Lambda(\bar{0})$, such that $Range(\Lambda) \subseteq [\delta, 1]$. If a function obeys a bounded modulus, it is bounded, as well as uniformly continuous. The problem of determining whether an index of a computable modulus is that of a bounded modulus is $\Sigma_1$: if a computable modulus is bounded, then eventually we know $\Lambda(\bar{0})$ to within sufficient accuracy to conclude $\Lambda(\bar{0}) \neq 0$.

## 1.3 Continuous Logic and Computable Continuous Structures

Now that we've gotten through some of preliminary definitions, we want to talk about continuous spaces with additional constants, functions, and relations on them. A constant is just an element of $M$, which is more-or-less the same thing as a function $\{\emptyset\} \to M$. An n-ary function is a uniformly continuous map $M^n \to M$. A relation (generalizing from first order logic) is a uniformly continuous function $M^n \to [0, 1]$. As mentioned in the introduction, we are using the interval $[0, 1] \subseteq \mathbb{R}$ in place of $\{0, 1\}$ as a set of "truth values". For example, if our continuity space is a Banach space, perhaps we want a constant for the 0 vector, a binary function for addition, etc. Up to this point, our definitions were inspired by mathematical folklore and pedantically refined by us for convenience of exposition. Similar definitions exist scattered throughout the literature, e.g in literature on effective Polish spaces etc., but we made no promise of adhering to any kind of existing standard. In what follows, we are taking definitions almost directly from papers by Ben Yaacov, Berenstein, Henson, & Usvyatsov [19] and Ben Yaacov, Doucha, Nies, & Tsankov [18]. Our contribution is merely to give computable analogues of their definitions, framed using continuous spaces rather than metric spaces, which usually involves just inserting the word "computable" in the right places and adjusting appropriately for an equalness relation rather rather than a distance function. This should make it relatively easy for the reader to relate our results to the work of Ben Yaacov et. al.

To motivate ourselves a bit, consider the classical theorem that total computable functions $2^\omega \to 2^\omega$ (where $2^\omega$ is Cantor Space) are all uniformly continuous, and every uniformly continuous function $2^\omega \to 2^\omega$ is computable relative to some oracle. This is what allows us to make a connection between recursion theory and analysis. Our $n$-ary moduli in continuous spaces play a similar role to moduli of continuity there: they tell you, roughly, how many bits of precision you need on the inputs to a function to compute a certain number of bits of precision on the output. Or more precisely, for us an $n$-ary modulus tells us how equal we can guarantee $f(\bar{x})$ and $f(\bar{y})$ are, given how equal the coordinates of $\bar{x}$ and $\bar{y}$ are. It will be useful to make an analogy to non-standard analysis (and in fact there is a non-trivial connection here, which we'll explain later when we talk about ultraproducts). There, we can equivalently define a function $f : \mathbb{R} \to \mathbb{R}$ to be uniformly continuous if its non-standard extension $f^* : \mathbb{R}^* \to \mathbb{R}^*$ satisfies that if $x \sim y$ then $f(x) \sim f(y)$, where $\sim$ is the relation of being infinitesimally close (which we can think of as being almost equal). In other words, the naive definition of continuity which is painfully scrubbed from undergraduate students minds ("close-by points are sent to close-by points") turns out to be meaningful and formally correct in non-standard analysis.

One can check with our definitions that if $(f_i : i \in \omega)$ converges to $f$ pointwise, and each $f_i$ obeys modulus $\Lambda$, then $f$ obeys modulus $\Lambda$. Likewise if a continuous function $f$ obeys modulus $\Lambda$ on a dense set of points in its domain, it obeys the modulus everywhere. See Appendix A for proofs of these facts. Also recall from the previous section that moduli compose: if $f_0, ..., f_{n-1}$ and $g$ are functions obeying moduli $\Lambda_0, ..., \Lambda_{n-1}$, and $\Lambda$ respectively which can be composed as $g \circ (f_0, ..., f_{n-1})$, then $\Lambda \circ (\Lambda_0, ..., \Lambda_{n-1})$ is a modulus for $g \circ (f_0, ..., f_{n-1})$. This also shows there is a uniform procedure for taking computable moduli for functions and obtaining a computable modulus for their composition, since compositions of computable functions are computable. These facts will enable us to represent uniformly continuous functions and relations on a computable continuous space in a nice way, which paves the way for computable continuous structure theory. But first lets give some of the basic definitions for continuous logic and continuous structures.

**Definition 22** (modified from Ben Yaacov, Henson, et. al.)**.** A (multi-sorted) **continuous signature** $\sigma$ consists of the following:

(a) A set $\mathcal{S}_\sigma$ of sort symbols

(b) a set $\mathcal{R}_\sigma$ of relation symbols

(c) a set $\mathcal{F}_\sigma$ of function symbols

(d) a function $Ari : \mathcal{R}_\sigma \cup \mathcal{F}_\sigma \to \omega$

(e) a function $Dom : \mathcal{R}_\sigma \cup \mathcal{F}_\sigma \to \bigcup_{\bar{S} \in \mathcal{S}_\sigma^{<\omega}} \prod_{j<|\bar{S}|} S_j$ giving the domain

(f) a function $Cod : \mathcal{R}_\sigma \cup \mathcal{F}_\sigma \to (\mathcal{S}_\sigma \cup \{[0,1]\})$ giving the codomain

(g) a function $\Lambda : \mathcal{R}_\sigma \cup \mathcal{F}_\sigma \to \bigcup_{n\in\omega}([0,1]^n \to [0,1])$ giving the modulus

We also assume that the symbols in $\mathcal{R}_\sigma$ and $\mathcal{F}_\sigma$ are assigned arities, domains, and codomains by $Ari$, $Dom$ and $Cod$, consistently with the type of symbol and with the arity of the modulus each is assigned. Constant symbols are just function symbols with arity 0 with domain $\{\emptyset\}$ and codomain $S$ for some sort $S \in \mathcal{S}_\sigma$. Relation symbols should have codomain $[0,1]$, while function symbols should have codomain one of the named sorts.

**Note.** For two continuous signatures $\sigma$ and $\tau$, we say $\sigma \subseteq \tau$ if all of the objects mentioned in (a)-(g) constituting $\sigma$ are contained in the corresponding objects for $\tau$, i.e., $\mathcal{S}_\sigma \subseteq \mathcal{S}_\tau$, the $Dom$ function for $\tau$ is an extension of the $Dom$ function for $\sigma$, etc.[9]

**Definition 23.** A continuous signature is **computable** if it can be interpreted in $\mathbb{N}$ in such a way that the sets of sort, relation, and function symbols are each computable, $Ari$, $Dom$, and $Cod$ are computable functions, and $\Lambda_X$ is a computable function for each symbol $X$, uniformly in the symbol $X$.

[9]If we were more clever with our definition of "continuous signature", we could ensure that extension of signatures is just set containment, but we choose not to because it would greatly complicate the definition.

**Definition 24.** A (multi-sorted) **continuous structure** $\mathcal{M}$ with signature $\sigma$ is a collection of continuous spaces $((M_S, \varepsilon_S) : S \in \mathcal{S}_\sigma)$ together with relations $(R^{\mathcal{M}} \in M_{Dom(R)_0} \times ... \times M_{Dom(R)_{n-1}} \to [0,1] : R \in \mathcal{R}_\sigma$ and $Ari(R) = n)$ and functions $(f^{\mathcal{M}} \in M_{Dom(f)_0} \times ... \times M_{Dom(f)_{n-1}} \to M_{Cod(f)} : f \in \mathcal{F}_\sigma$ and $Ari(f) = n)$ such that $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$ obey the moduli $\Lambda_f$ and $\Lambda_R$ respectively.

**Definition 25.** A (multi-sorted) continuous structure $\mathcal{M}$ with computable signature $\sigma$ is **computable** if the underlying continuous spaces for each sort are computable uniformly and we can uniformly compute a name for each $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$.

It's especially useful to look at continuous structures based off of metric spaces which are bounded or even compact. Ben Yaacov, Henson, et. al. [19] (who have a definition of metric structures) assume all sorts are bounded, and use many sorts to deal with unbounded spaces, but we are hoping to avoid this by using equalness relations. In any case, if we want something like the compactness theorem to be satisfied for continuous structures, the presentation of simple spaces like $\mathbb{R}$ may seem a bit complicated. The use of equalness relations sacrifices the ability to use classical reasoning dealing with subadditivity of metrics and such, but we think it is probably an easier sacrifice than presenting $\mathbb{R}$ as a structure with infinitely many sorts. We'll go into more detail about challenges to presenting $\mathbb{R}$ as a continuous structure later.

In addition to having a continuous version of equality, we also have a continuous version of logical connectives:

**Definition 26.** A **continuous logical connective** is a continuous function $\rho : [0,1]^n \to [0,1]$ together with a modulus of continuity $\Lambda_\rho$ which $\rho$ obeys.

**Definition 27.** A continuous logical connective is **computable** if both $\rho$ and $\Lambda_\rho$ are computable.

**Note.** You might be wondering, if you are thinking ahead, why don't our moduli of continuity themselves need moduli of continuity? After all, a modulus of continuity allows us to compute the values of a continuous function everywhere if we can compute them on a dense set. If we just know a function is continuous, and know its values on a dense set, that doesn't imply we can compute the function everywhere, because we don't know how close we need to approximate our input point to get a given level of closeness to the value of the function at that point. We might worry that we will have an infinite regress, where even if we can compute a modulus of continuity on a dense set, it need not be computable unless we have a further modulus of continuity, and so on. It turns out the reason we don't have an infinite regress is that if you can lower compute $\Lambda$ on a dense set which includes $\bar{0}$, then there is a computable lower name for $\Lambda$, and this is all we need to be able to compute a continuous function which obeys $\Lambda$ everywhere if we can compute it on a dense set. See Appendix C for more details.

**Definition 28.** Let $\sigma$ be a (multi-sorted) continuous signature, and for each $S \in \mathcal{S}_\sigma$, let $\chi_S$ be a set of new symbols for variables of sort S (e.g. $\chi = \{x_i^S : i \in \omega\}$). We will define the **terms** over $\sigma$ with variables $\bar{\chi} = (\chi_S : S \in \mathcal{S}^\sigma)$. $Terms(\sigma, \bar{\chi})$ is the smallest set of strings such that:

(i) If $c \in \mathcal{F}_\sigma$ has arity 0, then $c \in Terms(\sigma, \bar{\chi})$, $Type(c) := \{\emptyset\} \to Cod(c) \cong Cod(c)$, and $\Lambda_c \equiv 1$

(ii) If $x \in \chi_S$, then $x \in Terms(\sigma, \bar{\chi})$, $Type(x) := S \to S$. We set $\Lambda_x := Id_{[0,1]}$, the unary 1-Lipschitz modulus.

(iii) If $t_0, ..., t_{n-1} \in Terms(\sigma, \bar{\chi})$, $f \in \mathcal{F}_\sigma$, and $Type(t_i) = \prod_{j<k_i} S_{ij} \to Dom(f)_i$ for $i < Ari(f)$, then $f(t_0, ..., t_{n-1}) \in Terms(\sigma, \bar{\chi})$, $Type(f(t_0, ..., t_{n-1})) := \prod_{i<k} S_i \to Cod(f)$, and $\Lambda_{f(t_0,...,t_{n-1})} = \Lambda_f \circ (\Lambda_{t_0}, ...., \Lambda_{t_{n-1}}) : [0,1]^k \to [0,1]$. Here $(S_i : i < k)$ is obtained by reading $f(t_0, ..., t_{n-1})$ left-to-right, starting with the empty list, and adding a copy of $S$ to the end of the list whenever we encounter a variable symbol from $\chi_S$ which has not appeared before. We identify variables in the composition of the moduli which correspond to the same variable symbol.[10]

In this definition we also defined the **type** of terms in parallel. We could have written this without (i) and treated constants as a special case of (iii), but instead we separated it out to avoid confusion. We will always make use of the canonical isomorphisms between the space of functions $\{\emptyset\} \to X$ and $X$, as well as the canonical isomorphisms between $\{\emptyset\} \times X$ and $X$. This means we will treat a formula like $R(c, x)$, where $c$ is a constant symbol and $x$ a variable symbol, as denoting a unary [0,1]-valued function rather than a binary [0,1]-valued function with one of the inputs ranging over a domain with only a single point. This means we will always simplify product and function types to exclude irrelevant copies of $\{\emptyset\}$. If we wanted to be really pedantic, we also could have talked about diagonal functions being used to identify variables, but we will leave that to people who are interested in categorical logic.

**Definition 29.** Let $\sigma$ be a (multi-sorted) continuous signature and $\bar{\chi} = (\chi_S : S \in \mathcal{S}^\sigma)$ a set of variable of each sort. Then the **atomic formulas** over $\sigma$ with variables $\chi$, which we will call $Atomic(\sigma, \bar{\chi})$, is the smallest set of strings with the property that if $t_i \in Terms(\sigma, \bar{\chi})$ and $Type(t_i) = Dom(R)_i$ for $i < Ari(R)$ for some $R \in \mathcal{R}_\sigma$, then $R(t_0, ...., t_{n-1}) \in Atomic(\sigma, \bar{\chi})$. Here we treat the equalness relation $\varepsilon_S$ for each sort S as a binary relation symbol with modulus $\Lambda_\varepsilon(\tau_1, \tau_2) = \tau_1 \tau_2$. We also let $Type(R(t_0, ..., t_{n-1})) := \prod_{i<k} S_i \to [0,1]$, where $(S_i : i < k)$ is determined as in part $(iii)$ of the definition of terms. $\Lambda_{R(t_0,...,t_{n-1})} := \Lambda_R \circ (\Lambda_{t_0}, ...., \Lambda_{t_{n-1}})$.

[10] If variables are not repeated, $\prod_{i<k} S_i = \prod_{i<n} \prod_{j<k_i} S_{ij}$. This is just to say what happens when we repeat a variable. For example, suppose our formula is "$f(g(y), s(z), r(2), h(x, y))$". Assuming x ranges over real numbers, y and z over complex numbers, and the codomain of f is the real numbers, this would have type $\mathbb{C} \times \mathbb{C} \times \mathbb{R} \to \mathbb{R}$

**Definition 30.** Let $\sigma$ be a (multi-sorted) continuous signature, and $\nu$ be a collection of continuous logical connectives. Fix $\bar{\chi} = (\chi_S : S \in \mathcal{S}_\sigma)$ a *countable* set of variables for each sort. Then the continuous language with signature $\sigma$ and connectives $\nu$ is called $\mathcal{L}(\sigma, \nu)$ and is defined as the smallest set of formulas satisfying the following:

(i) $Atomic(\sigma, \bar{\chi}) \subseteq \mathcal{L}(\sigma, \nu)$

(ii) If $\rho \in \nu$ is an n-ary logical connective, and $\phi_0, ..., \phi_{n-1} \in \mathcal{L}(\sigma, \nu)$, then $\rho(\phi_0, ..., \phi_{n-1}) \in \mathcal{L}(\sigma, \nu)$. $Type(\rho(\phi_0, ..., \phi_{n-1})) := \prod_{i<k} S_i \to [0,1]$, where $(S_i : i < k)$ is just the list of sorts of the "free" variables appearing in $u(\phi_0, ..., \phi_{n-1})$ from left to right (not counting repeats of a variable). $\Lambda_{\rho(\phi_0,...,\phi_{n-1})} := \Lambda_\rho \circ (\Lambda_{\phi_0}, ...., \Lambda_{\phi_{n-1}})$ (with arguments corresponding to the same variable symbol identified).

(iii) If $\phi \in \mathcal{L}(\sigma, \nu)$, and $x$ is a variable, then $\inf_x \phi \in \mathcal{L}(\sigma, \nu)$ and $\sup_x \phi \in \mathcal{L}(\sigma, \nu)$. $Type(\inf_x \phi) = Type(\sup_x \phi) = \prod_{i<k} S_i \to [0,1]$, where $(S_i : i < k)$ is just the list of sorts of the "free" variables besides $x$ appearing in $\phi$ from left to right (not counting repeats of a variable). $\Lambda_{\inf_x \phi} = \Lambda_{\sup_x \phi}$ is just $\Lambda_\phi$ with 1 plugged in for the coordinate corresponding to $x$ (if it appears). See Appendix A for an explanation of this choice of modulus for $sup_x \phi$ and $\inf_x \phi$.

**Definition 31** (Interpretation and Moduli)**.** These formulas are all to be interpreted in the obvious way in a structure: the interpretation of a complex formula is the interpretation of the outermost symbol used in its construction composed with the interpretations of all the sub-formulas used in the last step of the construction of the term. Additionally, it's possible to extend *Ari*, *Cod*, and *Dom* to both terms and to formulas (as we did with $\Lambda$). *Ari*, *Cod* and *Dom* can be extracted from the type.

**Definition 32.** The language $\mathcal{L}(\sigma, \nu)$ is **computable** if both the signature $\sigma$ is computable (according to the definition given before) and the connective collection $\nu$ is computable, in the sense that we have an interpretation of the set of connectives $\nu$ inside $\mathbb{N}$ and a uniform procedure taking a connective $\rho$ and returning computable names $[\rho] : ([0,1] \cap Q)^{Ari(u)} \times \omega \to [0,1] \cap \mathbb{Q}$ and $[\Lambda_\rho] : ([0,1] \cap \mathbb{Q})^{Ari(u)} \to [0,1] \cap \mathbb{Q}$ for $\rho$ and $\Lambda_\rho$, respectively.[11]

[11] It's fun to note, for those who really enjoy pedantry, that we are using the continuous logical connectives in $\nu$ as symbols for themselves. You might be wondering something like this: what if we have a way of enumerating connectives which is unavoidably non-injective? Then by our convention to use extensional equality rather than intensional equality for functions, when two connectives we enumerate turn out to be extensionally equal, we have committed ourselves to assigning them the same symbol. But it is not decidable whether two connectives are extensionally equal! This seems to exclude perfectly reasonable collections of connectives. We can solve this problem by using a separate set of connective symbols, which are assigned to logical connectives by some assignment function. Another way to solve it would be to consider a kind of effective presentation of a closed collection of connectives, where we have a countable dense subset of connectives approximating the whole collection, and we don't care exactly which connectives are in this dense set, as long as its closure is as desired. In that framework, we could always find a presentation of our closed collection of connectives that has an injective enumeration of a dense subset of that collection, assuming we could find a not-necessarily-injective enumeration of a dense subset of that collection. But we are already so notation heavy, it seemed best to leave this to a footnote.

**Discussion.** There is a sense in which formulas in a computable language can approximate those in an uncountable language. We're not talking about the interpretations of those formulas in a structure, but the formulas themselves. The idea is that a sequence of logical connectives $\rho_n$ might converge to a limit $\rho$. So we should think of the formulas $\rho_n(\phi, \psi)$ as converging to $\rho(\phi, \psi)$, and so too for more complex formulas which differ only by having $\rho_n$ in place of $\rho$. The topology we place on the set of formulas in our language is given as follows: first assign each atomic formula a distinct variable symbol. Next, construct for each complex formula $\phi$ in our language a function $f_\phi : [0,1]^{Atomic(\sigma)} \to [0,1]$ by replacing each occurrence of each atomic formula in $\phi$ with its corresponding variable symbol, and thinking of the result as a $[0,1]$-valued function on $[0,1]^{Atomic(\sigma)}$ with the product topology. $f_\phi$ will depend only on the variables corresponding to the finitely many atomic formulas appearing in $\phi$. Finally, define a metric by

$$d(\phi, \psi) = \max_{\bar{x} \in [0,1]^{Atomic(\sigma)}} |f_\phi(\bar{x}) - f_\psi(\bar{x})|$$

This is a maximum, not just a supremum, because $[0,1]^{Atomic(\sigma)}$ is compact, and each $f_\phi$ is continuous. Essentially, this gives the topology *before* we have a theory relating the values of the atomic formulas. One we have a background theory, it makes sense to take a quotient our space of formulas by equivalence modulo that theory.

## 1.4 Formula Complexity and Definability

Definability is relatively uncomplicated notion in $\mathbb{N}$. Formulas correspond to subsets of (Cartesian powers of) $\mathbb{N}$, and you can gauge the complexity of such a subset by counting alternations of quantifiers in the formula defining it. If you have a computable presentation of $\mathbb{N}$, then there is an algorithm to enumerate the $\Sigma_1$ truths (sentences whose normal forms have only existential quantifiers). $\Pi_1$ truths (universal statements) can be effectively falsified using an algorithm which searches for counter-examples. And more complicated formulas, in general, might require iterates of the halting problem to compute their truth values. The principle here is a syntactic-semantic duality, where syntactic objects (formulas) correspond to semantic objects (subsets), and the complexity of a definable subset can be gauged by the syntactic form of the form of the formula defining it.

It turns out, even for classical computable structures, that in the general case it is more natural to work in a computable infinitary language to describe the complexity of sets. The basic reason for this is that in any computable structure, you can verify a computable infinite disjunction just as easily as you can verify a existential statement. In $\mathbb{Q}$ in the language $(0, 1, +)$, for example, the non-negative dyadic rationals (rationals of the form $\frac{p}{2^k}$) are not definable by an existential formula, but they are definable by a computable infinite disjunction

[12], which is just as easy to verify in any computable copy of $\mathbb{Q}$. See Antonio Montalbán's upcoming book (available as a draft on his website) *Computable Structure Theory* [10] for a nice exposition of this material for classical first-order structures. We will be proving analogues of theorems presented in his book for continuous structures, so it may be useful to compare. It is worth noting that not only do many of the theorems generalize to this setting, but many of their proofs do too, which provides some evidence that our framework for computable continuous structure theory is the right one.

However, care must be taken in laying out the definitions. It is temping to define "definable subset" to mean a set of the form $\{\bar{p} \in M^n : \phi^{\mathcal{M}}(\bar{p}) = 1\}$ (for a single-sorted structure). But we can see this is probably the wrong definition: even if $\phi(x)$ is an atomic formula, and $\mathcal{M}$ is a computable continuous structure, it may not be decidable from an oracle giving a name for $\bar{p}$ whether $\phi^{\mathcal{M}}(\bar{p}) = 1$. In general, we can only evaluate atomic formulas in a computable continuous structure to arbitrary precision. Think, for example, of trying to figure out whether a real number is equal to 0 from its binary expansion. We might read a million zero digits, but never know if later there might be a one. The problem here is essentially that the characteristic function of the set we are trying to define is not continuous. In a discrete structure, this problem does not arise, because every function is continuous. Instead, we want to be able to compute how close elements are to a given set. Specifically:

**Definition 33.** Suppose $A \subseteq M^n$. Let $\chi_A(\bar{x}) := \sup_{\bar{p}\in A} \bigwedge_{i<n} \varepsilon(x_i, p_i)$. This is the **continuous characteristic function** of $A$. For $\tau \in [0, 1]$, $A$ is **$\tau$-definable** (without parameters) if there is a formula $\phi(\bar{x})$ such that

$$\inf_{\bar{x}\in M^n} \varepsilon(\phi(\bar{x}), \chi_A(\bar{x})) \geq \tau$$

The continuous characteristic function tells you how close a point $\bar{x}$ is to $A$. For $A$ to be $\tau$-definable is for there to be a formula which approximates the characteristic function $\chi_A$ to degree $\tau$ of closeness. We are using the symbol "$\tau$" because we might want to think of $[0, 1]$ as a set of truth values. Instead of choosing an error $\epsilon$, and approximating up to an error, we are thinking of choosing a truth value, and requiring that the equality of the functions $\phi(\bar{x})$ and $\chi_A(\bar{x})$ is at least $\tau$-true, or to put it another way: that their values are at least $\tau$-equal. Note that $\tau$-equality is an equivalence relation only when $\tau = 1$ (or 0, in which case it is trivial). In general it is a similarity relation (reflexive and symmetric, but not necessarily transitive). See section 0.2 for some intuitions on how this lack of transitivity plays out. There is still a weakened version of transitivity: if $x$ and $y$ are $\tau$-equal, and $y$ and $z$ are $\tau'$

[12]This formula would just be something like:

$$\phi(x) = \bigvee_{p,k\in\mathbb{N}} x * \big((1+1) * (1+1) * \cdots (1+1) \text{ (k times) } \big) = 1 + 1 + \cdots + 1 \text{ (p times)}$$

equal, then $x$ and $z$ are at least $(\tau * \tau')$-equal. Recall this is just our version of the triangle inequality.

**Note.** In the case that we are working with a discrete continuous structure, where all values lie in $\{0, 1\}$, $\tau$-definability is identical with classical first-order definability for any $\tau > \frac{1}{2}$.

**Definition 34.** $A \subseteq M^n$ is **definable** if for every $0 \leq \tau < 1$, $A$ is $\tau$-definable.

For a long discussion of why this notion of definability is the appropriate one for model-theoretic purposes, see the section titled "Definability in Metric Structures" in [19]. The basic idea is that there is a natural metric on the class of functions $f : M^n \to [0, 1]$ (which we can think of as having all the continuous characteristic functions of sets). This metric roughly corresponds to the Hausdorff distance between the sets so defined. We would like that any of the sub-classes we consider will be closed with respect to this metric, so we define the class of definable functions $M^n \to [0, 1]$ to be the closure of the set of functions given by formulas of our language (rather than just functions given by formulas in our language). If we want to preserve the duality between sets and characteristic functions, we should then say the definable sets are those whose continuous characteristic functions can be uniformly approximated to arbitrary precision by formulas in our language. It's worth noting that this is a faithful generalization of definability in classical first-order model theory, if that makes it more palatable.

However, this may not be the appropriate notion of definability for computable continuous structure theory. One concern is the lack of uniformity in the formulas giving the approximations. For example, what should we say is the complexity of a set which can be approximated to truth value $\tau_n = 2^{-2^{-n}}$ by a formula $\phi_n$ only with $n$ alterations of sups and infs (and no fewer)? Even if each $\phi_n$ is $\Sigma_1$, what if this sequence of formulas is not computable? These worries can be resolved by working in a computable infinitary language. In [19], Ben Yaacov et. al. use the Tietze extension theorem to construct a continuous infinitary connective which essentially takes the limit of a fast Cauchy sequence of formulas, allowing you a formula like $lim_n \phi_n(\bar{x})$ to define a set, where $\phi_n(\bar{x})$ are the successive approximations to its continuous characteristic function. This connective turns out to be computable. We give a version of it here in our formalism, described in a way that make it obvious it is computable:

**Definition 35.** The **forced limit** of $(\tau_n)_{n \in \omega} \in [0, 1]^\omega$, denoted $\lim_n \tau_n$, is computed from an oracle giving a name for $n \mapsto \tau_n$ by $[\lim_n \tau_n](0) = \tau_1(1)$,

$$[\lim_n \tau_n](k+1) := \begin{cases} [\tau_{k+2}](k+2) & \text{if } \varepsilon([\tau_{k+2}](k+2), [\lim_n \tau_n](k)) \geq 2^{-2^{-k}} \\ [\lim_n \tau_n](k) + 2^{-k-1} & \text{if otherwise } [\tau_{k+2}](k+2) > [\lim_n \tau_n](k) \\ [\lim_n \tau_n](k) - 2^{-k-1} & \text{if otherwise } [\tau_{k+2}](k+2) < [\lim_n \tau_n](k) \end{cases}$$

Essentially, if $n \mapsto [\tau_n](n)$ is a name for some $\tau \in [0,1]$, $[\lim_n \tau_n](k)$ is just equal to $[\tau_{k+1}](k+1)$, and $[\lim_n \tau_n]$ is a name for $\tau$. If the terms of $[\tau_{n+1}](n+1)$ don't converge fast enough, the sequence is modified to only change within some predefined bounds at each step, ensuring that $[\lim_n \tau_n]$ always converges fast to *something* (but it won't necessarily be equal to the limit of the $\tau_n$ if the sequence $(\tau_n)_{n\in\omega}$ converges too slowly). All we're doing here is extending the (computable) process of taking the limit of a (quickly) converging sequence of points to the class of all sequences, obtaining a computable connective $[0,1]^\omega \to [0,1]$.

**Observation 7.** *Suppose $R$ is a function $M^n \to [0,1]$. If $\phi_n(\bar{x})$ is a computable sequence of formulas with $\inf_{\bar{x}} \varepsilon(R(\bar{x}), \phi_n(\bar{x})) \geq 2^{-2^{-n}}$, then $\inf_{\bar{x}} \varepsilon(R(\bar{x}), \lim_n \phi_n(\bar{x})) = 1$. In particular, the relation $R(\bar{x})$ is $\tau$-definable for all $\tau \in [0,1)$ by the same formula, $\lim_n \phi_n(\bar{x})$.*

Thus if we allow this connective into our language, every definable set is definable by a single formula, and every set definable by a computable sequence of formulas is definable by a computable formula (although we haven't said what computable formulas are yet).

**Note.** We might consider the addition of the connective lim to be a very weak kind of infinitary connective. While the infinite conjunction and disjunction (inf and sup) of a computable sequence of computable functions need not be computable in a computable structure, the forced limit of computable sequence of computable functions on a computable structure will always be computable.

We now define a computable infinitary language. First we need to port a definition from [18]:

**Definition 36.** A **weak modulus** is a function $\Omega : [0,1]^\omega \to [0,1]$ such that:

(i) $\Omega(\bar{u}) \geq \Omega(\bar{u} * \bar{v}) \geq \Omega(\bar{u})\Omega(\bar{v})$, where $*$ is the coordinate-wise product.

(ii) $\Omega$ is lower semi-continuous in the product topology, and separately continuous in each argument, and $\Omega(\bar{1}) = \Omega(1,1,1,1,....) = 1$

**Example 3.** The connective $\lim_n$ obeys the weak modulus $\Omega(\bar{\tau}) = \prod_{n\in\omega} \max(2^{-2^{-k}}, \tau_i)$, which is not only lower semi-continuous, but continuous.

However, the main reason for using weak moduli is for their truncations, which allow us to simultaneously require an infinite family of formulas, possibly with different free variables, to all obey compatible moduli.

**Definition 37.** The n-th truncation of $\Omega$ is $\Omega|_n : [0,1]^n \to [0,1]$ defined by $\Omega|_n(\tau_0, ...., \tau_{n-1}) = \Omega(\tau_0, ..., \tau_{n-1}, 1, 1, 1, 1, 1, ...)$

**Observation 8.** *The n-th truncation of a weak modulus $\Omega$ is an n-ary modulus.*

An example of a useful weak modulus that is not continuous, but still useful, is the universal 1-Lipschitz modulus:

**Definition 38.** The **universal weak 1-Lipschitz** modulus $\Omega$ is defined by $\Omega(\bar{\tau}) = \inf_i \tau_i$

Its truncations are all the $n$-ary 1-Lipschitz moduli. This is a convenient way say that a class of formulas all with different numbers of free variables are all 1-Lipschitz. Another example is the universal Lipschitz modulus:

**Definition 39.** The **universal weak Lipschitz modulus** is defined by $\Omega(\bar{\tau}) = \prod_{n\in\omega} \tau_n^n$.

Note that if $f$ obeys modulus $\Lambda(\tau) = \tau^n$, and $\varepsilon$ arises from a metric $d(x, y)$, then $2^{-d(f(x),f(y))} = \varepsilon(f(x), f(y)) \geq \varepsilon(x, y)^n = (2^{-|x-y|})^n$, and thus taking the negative base-2 logarithm on both sides, $d(f(x), f(y)) \leq n|x - y|$, which is exactly saying $f$ is $n$-Lipschitz.

**Definition 40.** Fix $\sigma$ a computable signature, and $\nu$ a computable collection of continuous logical connectives, we define the computable infinitary language $\mathcal{L}^c_\Omega(\sigma, \nu)$ as follows:

- The **basic formulas** are formulas $\phi(x_0, ...., x_{n-1}) \in \mathcal{L}(\sigma, \nu)$ that do not make use of the quantifiers $\inf_x$ or $\sup_x$, depend only on the first n variables (but possibly not all of them), and whose modulus is bounded below by the modulus $\Omega|_n$
- If $(\phi_i : i \in \omega)$ is a computable sequence of n-ary formulas in $\mathcal{L}^c_\Omega(\sigma, \nu)$, then $\bigwedge_i \phi_i$ and $\bigvee_i \phi_i$ are n-ary formulas in $\mathcal{L}^c_\Omega(\sigma, \nu)$ (to be interpreted as an infimum and supremum over $i \in \omega$, respectively).
- If $\phi$ is an (n+1)-ary formula in $\mathcal{L}^c_\Omega(\sigma, \nu)$, then $\inf_{x_n} \phi$ and $\sup_{x_n} \phi$ are n-ary formulas in $\mathcal{L}^c_\Omega(\sigma, \nu)$.
- If $(\phi_i : i < \alpha)$, $\alpha \leq \omega$ is a computable sequence of n-ary formulas in $\mathcal{L}^c_\Omega(\sigma, \nu)$, and $u \in \nu$ is a 1-Lipschitz $\alpha$-ary continuous logical connective, then $u((\phi_i : i < \alpha))$ is an n-ary formula in $\mathcal{L}^c_\Omega(\sigma, \nu)$.

Note some of the restrictions we have placed. We are only allowed to quantify over the largest variable in a formula. We are only allowed to apply 1-Lipschitz connectives, aside from the construction of basic formulas. The reason for this is that when we take an infinite conjunction or disjunction, we need all the formulas in that infinite conjunction or disjunction to respect a common modulus, and so we want it to be trivial (computationally speaking) to verify this.

One potential problem with our language, however, is that it may be computationally nontrivial to tell whether a basic formula respects $\Omega|_n$, or verify that a continuous logical connective is 1-Lipschitz. *If* we knew that we could enumerate the basic formulas in our language, and the 1-Lipschitz connectives in $\nu$, then we could enumerate all the formulas of our language by induction. So we will assume from now on that we have chosen a language for which we have an enumeration of the basic formulas in $\mathcal{L}^c_\Omega(\sigma,\nu)$, and can enumerate all the 1-Lipschitz connectives in $\nu$. One way to do this is to choose $\nu$ to consist only of 1-Lipschitz connectives, and to give our relations/functions moduli which are easy to compare to $\Omega|_n$ (for example, maybe we can separately enumerate the $r$-Lipschitz relations/functions for each $r \in \mathbb{Q}_{\geq 0}$). It's worth noting that $\wedge$, $\vee$, $\neg$, and $\lim_n$ are all 1-Lipschitz.

**Definition 41.** We define some complexity classes of formulas in $\mathcal{L}^c_\Omega(\sigma,\nu)$:

- Basic formulas are $\Pi^c_0 = \Sigma^c_0$
- A formula of the form $\bigvee_i \sup_{\bar{x}_i} \phi_i(\bar{x}_i, \bar{y})$ where the $\phi_i$ are uniformly $\Pi^c_n$ is $\Sigma^c_{n+1}$.
- A formula of the form $\bigwedge_i \inf_{\bar{x}_i} \phi_i(\bar{x}_i, \bar{y})$ where the $\phi_i$ are uniformly $\Sigma^c_n$ is $\Pi^c_{n+1}$.
- A formula is $\Delta^c_n$ if it is equivalent to both a $\Pi^c_n$ and a $\Sigma^c_n$ formula.

**Definition 42.** A logical connective $u$ is **order preserving** if whenever $\bar{\tau} \leq \bar{\tau}'$ coordinatewise, $u(\bar{\tau}) \leq u(\bar{\tau}')$

**Observation 9.** *$\wedge$, $\vee$, and $\lim_n$ are all order-preserving, while $\neg$ is order reversing. $\Sigma^c_n$ and $\Pi^c_n$ formulas are thus both closed under order-preserving connectives, in the sense than any formula obtain by applying order-preserving connective like $\wedge$ or $\vee$ to a $\Sigma^c_n$ formula (respectively, $\Pi^c_n$ formula) is equivalent to a $\Sigma^c_n$ formula (respectively, $\Pi^c_n$) formula). This means $\Sigma^c_n$ and $\Pi^n_c$ formulas up to equivalence form a lattice. The negation of a $\Sigma^c_n$ formula is equivalent to $\Pi^c_n$, and vice versa, which tells us that $\Delta^c_n$ formulas form a Boolean algebra. But these lattices and Boolean algebras have many more operations than the classical Boolean connectives, which give us a richer algebraic structure on these classes of formulas in continuous logic than in classical logic.*

**Aside.** We haven't said what it means for two formulas to be equivalent. For us, this will just mean that $\inf_{\bar{x}} \varepsilon(\phi(\bar{x}), \psi(\bar{x})) = 1$ in every model of the theory under consideration. Usually this will be the theory of a particular structure $\mathcal{M}$ whose definable sets we are considering, in which case this is just saying that those two formulas have $\inf_{\bar{x}} \varepsilon(\phi^{\mathcal{M}}(\bar{x}), \psi^{\mathcal{M}}(\bar{x})) = 1$ We haven't said what a theory is in continuous logic, so we'll have to refer back when we do define continuous theories (there are some subtleties). If we are working over the empty theory, however, equivalence can be characterized by the metric on formulas we defined at the end of section 1.3: two formulas are equivalent if $d(\phi, \psi) = 0$ (which is straight-forwardly the same as saying they take on the same values for any interpretation of the function/relation symbols in the language). This also has a syntactic characterization for finitary continuous logic: there is a completeness theorem for continuous logic, proven by Ben Yaacov and Pedersen in [20], but we won't go into this now.

**Definition 43.** A set $A \subseteq M^n$ is $\Sigma_n^c$ (respectively, $\Pi_n^c$) if there is a $\Sigma_n^c$ (respectively, $\Pi_n^c$) formula $\phi(\bar{x}) \in \mathcal{L}_\Omega^c(\sigma, \nu)$ with $\inf_{\bar{x}} \varepsilon(\phi(\bar{x}), \chi_A(\bar{x})) = 1$

From now on when we refer to these complexity classes of formulas, we will always be using equivalence in a given structure, rather than over a class of structures.

**Definition 44.** We say function $f : \mathcal{M} \to [0,1]$ is a $\Gamma$ function for $\Gamma = \Sigma_n^c$ or $\Pi_n^c$ or $\Delta_n^c$ a class of formulas if $f$ is definable by a single formula in $\Gamma$, i.e. $f$ is $\tau$-definable by the same $\Gamma$ formula for each $\tau < 1$.

**Note.** For $\mathbb{N}$ treated as a computable continuous structure with the discrete metric, the subsets whose characteristic functions are $\Sigma_1^c$ are exactly the $\Sigma_1$ (recursively enumerable) sets. This should give us some indication we are using the right definitions.

**Observation 10.** *$\Sigma_1^c$ functions $M^n \to [0,1]$ on a computable continuous structure have computable lower names (i.e., they can be computably approximated from below) $\Pi_1^c$ functions have a computable upper names (i.e. they can be computably approximated from above). $\Delta_1^c$ functions have computable names.*

What's interesting is that the syntactically defined class of $\Sigma_1^c$ functions $M^n \to [0,1]$, which we will call "$\Sigma_1^c$ relations", corresponds to a semantic class of relations: the "uniformly relatively intrinsically computably enumerable", or "u.r.i.c.e.", relations. We extend some definitions from Antonio Montalbán's "Computable Structure Theory [10] to continuous structures:

**Definition 45.** An **$\omega$-presentation** (a.k.a. a **copy**) of a continuous structure $\mathcal{M}$ over a computable language is an isomorphic continuous structure $\mathcal{A}$ whose dense set $D = \omega$, whose elements are equivalence classes of fast Cauchy sequences of elements of $D = \omega$ under the equivalence relation $(n_i)_{i\in\omega} \sim (m_i)_{i\in\omega}$ iff $\lim_i \varepsilon(n_i, m_i) = 1$, and whose constants, functions, and relations are given uniformly by names (not necessarily computable names, but the data should be stored as a mapping from symbols to names for functions/relations/constants). The **named atomic diagram** of an $\omega$-presentation $\mathcal{A}$ of $\mathcal{M}$, denoted $\mathcal{D}(\mathcal{A})$, is exactly this datum: the mapping from the set of function/relation/constant symbols to names for the corresponding functions/relations/constants.

**Definition 46.** A **computable copy** of a continuous structure $\mathcal{M}$ is an $\omega$-presentation of $\mathcal{M}$ whose named atomic diagram is computable.

An $\omega$-presentation of a continuous structure is essentially just a computable continuous structure relative to some oracle, with the constraint that the dense set is $\omega$. The idea behind restricting the dense set to be $\omega$ rather than some arbitrary countable dense set $D$ is that it makes it easier to consider the class of all separable continuous structures in a given language. Likewise, a computable copy of a continuous structure $\mathcal{M}$ is just a computable structure $\mathcal{A}$ isomorphic to $\mathcal{M}$ whose dense set is $\omega$, so this definition is not really anything new. The main difference here is that it's possible to describe these structures completely (up to equality, not just isomorphism) with only a countable amount of information.

**Definition 47.** A relation $R : M^n \to [0,1]$ obeying weak modulus $\Omega$ is **u.r.i.c.e.** if there is a Turing functional $\Phi$ such that $\Phi^{\mathcal{D}(\mathcal{A})}$ is a computable lower name for $R$ for every copy $\mathcal{A}$ of $\mathcal{M}$.

**Theorem 11.** *Let $\mathcal{M}$ be a computable continuous structure with a given signature, and $\Omega$ is the universal 1-Lipschitz modulus, $\Omega(\bar{\tau}) := \inf_i \tau_i$, and our collection $\nu$ of logical connectives contains all rational piecewise linear connectives, together with* $\lim_n$. *Then the $\Sigma_1^c$ in relations over $\mathcal{L}^c_\Omega(\sigma,\nu)$ are exactly the* 1*-Lipschitz u.r.i.c.e. relations.*

*Proof.* One direction is straightforward: if a 1-Lipschitz relation $R$ is $\Sigma_1^c$, say by a formula $\bigvee_i \sup_{\bar{y}_i} \varphi_i(\bar{x},\bar{y}_i)$ we can approximate $R(\bar{p})$ from below using an oracle for the diagram of any copy $\mathcal{A}$ of $\mathcal{M}$, and an oracle for a name $[\bar{p}]$ for $\bar{p} \in \mathcal{A}^{|\bar{x}|}$ by setting

$$[R(\bar{p})](k) := \bigvee_{i<k} \bigvee_{\bar{q}_i \in \{q_0,\ldots,q_{k-1}\}^{|\bar{y}_i|}} [\phi_i^{\mathcal{A}}]([\bar{x}](k+1),\bar{q}_i)(k+1) - 2^{-k}$$

if this is not less than $[R(\bar{p})](k-1)$, and $[R](\bar{p})(k-1)$ otherwise.

The reason this is *u.r.i.c.e.* is because, given the diagram of $\mathcal{A}$, we can uniformly in $i$ compute a name for $\phi_i^{\mathcal{A}}$, and thus approximate to arbitrary precision a finite maximum of such functions. By subtracting $2^{-k}$, we ensure that we are approximating $R(\bar{p})$ from below: this is to account for the fact that our estimate for $\varphi_i^{\mathcal{A}}(\bar{x},\bar{q}_i)$ (namely $[\varphi_i^{\mathcal{A}}]([\bar{x}](k+1),\bar{q}_i)(k+1)$ may be too high by an error of $\frac{1}{2^k}$, accounting for our error in approximating $\bar{x}$, and our error in approximating the value of $\phi$ on that approximation of $\bar{x}$. To see that this converges to $R(\bar{p})$, just note that if $\bigvee_i \sup_{\bar{y}_i} \varphi(\bar{x},\bar{y}_i) > \tau$, then $\sup_{\bar{y}_{i_0}} \varphi(\bar{x},\bar{y}_{i_0}) > \tau$ for some $i_0 \in \omega$, and if $\sup_{\bar{y}_{i_0}} \varphi(\bar{x},\bar{y}_{i_0}) > \tau$, then for some $\bar{q} \in D^{|\bar{y}_i|} = \omega^{|\bar{y}_i|}$, we have $\varphi(\bar{x},\bar{q}) > \tau$. We only need to evaluate the supremum ever a dense set of tuples, because $\varphi$ is continuous, and the supremum of a continuous function over some domain is the same as the supremum of that continuous function over any dense subset of that domain.

For the converse, suppose the $n$-ary relation $R$ is *u.r.i.c.e.* by some Turing functional $\Phi$. That is, given the diagram $\mathcal{D}(\mathcal{A})$ of a copy $\mathcal{A}$ of $\mathcal{M}$, and any $\bar{n} \in \omega^n \subseteq \mathcal{A}^n$, $\Phi^{\mathcal{D}(\mathcal{A})}(\bar{n})(k)$, thought of as a function of $k$, is a lower name for $R^{\mathcal{A}}(\bar{n})$. We want to obtain from this a $\Sigma_c^1$ formula which defines $R$.

The main observation is that if $\Phi^{\mathcal{D}(\mathcal{A})}(\bar{n})(k) \downarrow = q_k \in \mathbb{Q} \cap [0,1]$ for a particular $k$ in, say, $s$ steps of computation, it could only access the first $s$ bits of $\mathcal{D}(\mathcal{A})$. This represents a finite amount of information about some tuple $\bar{m} \supseteq \bar{n}$, and can be expressed by a finite collection of closed conditions of the form $[\varepsilon(\phi_{i_l}^{at,\mathcal{A}}(\bar{m}), \dot{q}_{j_l}) \geq 2^{-2^{-k_l}}]$ for $l < s$. For each $l$, using a computable 1-Lipschitz connective $u_l$ (obtainable uniformly from $\dot{q}_{j_l}$ and $k_l$), we can re-express the $l$-th condition as: $[u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{m})) = 1]$, so the entire collection of conditions can be expressed as:

$$[\bigwedge_{l<s} u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{m})) = 1]$$

Given $\bar{n}$ and $\mathcal{D}(\mathcal{A})$, and $k$, there is a uniform way of obtaining this formula $\bigwedge_{l<s} u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{m}))$ and $q_k$. This formula is 1-Lipschitz, and any $\bar{q} \in \mathcal{A}^{|\bar{m}|}$ meeting this condition will necessarily have $R^{\mathcal{A}}(\bar{p}) \geq \bar{q}_k$ for $\bar{p} = \bar{q} \upharpoonright n$, because if $\bar{q}$ meets this condition, then there is an alternate isomorphic copy $\mathcal{B}$ of $\mathcal{M}$ for which $\bar{q} \in \mathcal{A}^{|\bar{m}|}$ can be carried by an isomorphism to $\bar{m}$ interpreted as an element of $\mathcal{B}$, and for which the first $s$ bits of $\mathcal{D}(\mathcal{B})$ are equal to the first $s$ bits of $\mathcal{D}(\mathcal{A})$ (we obtain this copy $\mathcal{B}$ of $\mathcal{M}$ by simply choosing an enumeration of a dense set in $\mathcal{A}$ so that the elements corresponding to $\bar{q}$ are labeled with the labels $\bar{m}$). But then we can see that:

$$\mathcal{A} \models [\sup_{\bar{y}} \bigwedge_{l<s} u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{x}^\frown\bar{y})) = 1] \Longrightarrow \mathcal{A} \models [R(\bar{x}) \geq q_k]$$

Without loss of generality, we may suppose $\mathcal{A} = \mathcal{M}$, which tells us:

$$\mathcal{M} \models [\sup_{\bar{y}} \bigwedge_{l<s} u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{x}^\frown\bar{y})) = 1] \Longrightarrow \mathcal{M} \models [R(\bar{x}) \geq q_k]$$

The problem, of course, is that we cannot necessarily reverse this implication, because there may be many different ways of making $R(\bar{x}) \geq q_k$. The first thing we can do is quantify over points nearby points meeting this condition, to get an open neighborhood (with a slightly weaker result):

$$\mathcal{M} \models \left[\sup_{\bar{z}} \left(\varepsilon(\bar{x}, \bar{z}) + 2^{-k} \wedge \sup_{\bar{y}} \bigwedge_{l<s} u_l(\phi_{i_l}^{at,\mathcal{A}}(\bar{z}^\frown\bar{y}))\right) = 1\right] \Longrightarrow \mathcal{M} \models [R(\bar{x}) > q_k - 2^{-k}]$$

We are taking advantage of the fact that our language is 1-Lipschitz, here. Now we just want to take an infinite disjunction over all possible ways to to witness $R(\bar{x}) \geq q$ for some rational $q$. For $\bar{n} \in \omega^n \subseteq \mathcal{M}^n$, and $k \in \omega$, define:

$$\psi_{\bar{n},k}(\bar{x}) := \left(q_{\bar{n},k} - 2^{-k}\right) - \left(1 - \sup_{\bar{z}} \left(\varepsilon(\bar{x}, \bar{z}) + 2^{-k} \wedge \sup_{\bar{y}} \bigwedge_{l<s} u_{\bar{n},l}(\phi_{i_{\bar{n},l}}^{at,\mathcal{A}}(\bar{z}^\frown\bar{y}))\right)\right)$$

where for any $\bar{n}$ and $k$, $q_{\bar{n},k}$, $u_{\bar{n},l}$, and $i_{\bar{n},l}$, are chosen according to our uniform construction above, and we take this formula to have value 0 if it evaluates to a negative number (to keep its values inside $[0, 1]$). This construction of $\psi_{\bar{n},k}$ is uniform in $\bar{n}$ and $\mathcal{D}(\mathcal{M})$, and $\mathcal{D}(\mathcal{M})$ is computable, so this family of formulas for $\bar{n} \in \mathcal{M}^n$ is *c.e.* even without an oracle. Note also that by construction, for any $\bar{n}, k$, and any $\bar{x} \in \mathcal{M}^n$, we always have $\psi^{\mathcal{M}}_{\bar{n},k}(\bar{x}) \leq R^{\mathcal{M}}(\bar{x})$, i.e. we are never overestimating $R(\bar{x})$. But then we can take a compute disjunction over $\bar{n} \in \mathcal{M}^n$ and $k \in \omega$ to obtain a $\Sigma^c_1$ formula:

$$\varphi(\bar{x}) := \bigvee_{\bar{n} \in \mathcal{M}^n, k \in \omega} \psi_{\bar{n},k}(\bar{x})$$

We claim this $\Sigma^c_1$ formula defines $\mathcal{R}^{\mathcal{M}}$. We have already shown one direction (namely that it never overestimates $R(\bar{x})$. We need to show that it never underestimates $R(\bar{x})$ either:

Let $\bar{x} \in \mathcal{M}^n$, and fix $k' \in \omega$. Pick a point $\bar{n} \in \mathcal{M}$ with $\varepsilon(\bar{n}, \bar{x}) \geq 2^{-2^{-(k'+2)}}$. Now pick $k$ so that $\Phi^{\mathcal{D}(\mathcal{M})}(\bar{n})(k) \geq R(\bar{n}) - 2^{-(k'+2)}$. Then

$$\begin{aligned} \psi_{\bar{n},k}(\bar{x}) &\geq \psi_{\bar{n},k}(\bar{n}) - 2^{-(k'+2)} \geq R(\bar{n}) - 2^{-(k'-2)} - 2^{-(k'-2)} \\ &\geq R(\bar{x}) - 2^{-(k'-2)} - 2^{-(k'-2)} - 2^{-(k'-2)} > R(\bar{x}) - 2^{-k'} \end{aligned}$$

We are using 1-Lipschitzness of $\psi_{\bar{n},k}$ and $R$ in the first and second-to-last inequalities. As $k'$ was arbitrary, $\varphi(\bar{x}) \geq R(\bar{x})$ for any $\bar{x} \in \mathcal{M}^n$. □

## 1.5 Examples of Computable Continuous Structures

Before we go further, it's worthwhile to present a few structures of general interest to mathematicians as computable continuous structures. The main obstacle to presenting a structure is possibly needing to modify it so that all the functions and relations are uniformly continuous. We of course have that complete, separable metric spaces whose metric can be computed on a dense set can all be represented as computable continuous structure in the language of equalness by defining $\varepsilon(x, y) = 2^{-d(x,y)}$, but this is sort of trivial. Let's start with a non-trivia structure which doesn't need any modification to present it as a computable continuous structure: the p-adic integers, which are the completion of the integers under a different norm (or valuation, if that is your preference).

### The $p$-adic Integers

The p-adic integers $\mathbb{Z}_p$ can be defined in a number of equivalent ways, but the quickest way to see that they form a continuous structure is to define $\mathbb{Z}_p$ as the completion of the integers under the norm $|x|_p = p^{-\max\{k\,:\,p^k|x\}}$.[13] We can then put a metric (in fact, an ultrametric) in $\mathbb{Z}_p$ by defining $d(x, y) = |x - y|_p$. We can extend the ring structure of $\mathbb{Z}$ to $\mathbb{Z}_p$ by continuity. As a topological space, $\mathbb{Z}_p$ can be thought of as $p^\omega$, which is homeomorphic to Cantor space, and it is useful think of it as an infinite $p$-branching tree. Choosing a branch along this tree can be thought of as choosing a sequence of numbers $(a_0, a_1, a_2, \dots)$ with each $a_n \in \mathbb{Z}/p^n\mathbb{Z}$, and $a_n \equiv a_{n+1} \pmod{p^n}$. Just as finite binary strings correspond to basic open sets in Cantor space, an element of $\mathbb{Z}/p^n\mathbb{Z}$ corresponds to a basic open subset $[a] = \{x \in \mathbb{Z}_p : x \equiv a \pmod{p^n}\}$. Here's our main observation:

**Observation 12.** *The ring $\mathbb{Z}_p$ has a computable presentation as a continuous structure.*

*Proof.* We identify $\mathbb{Z}$ with a dense subset of $\mathbb{Z}_p$ by identifying $n \in \mathbb{Z}$ with

$$\big(n \bmod p^0, n \bmod p^1, n \bmod p^2, \dots\big) \in \mathbb{Z}_p$$

It's clear that addition and multiplication are computable on this dense subset, since it's the same as multiplication on $\mathbb{Z}$. Using the equalness relation $\varepsilon(x, y) = 2^{-|x-y|_p}$, we can verify that:

$$-\log_p(-\log_2 \varepsilon(x \cdot y, x' \cdot y')) = -\log_p |xy - x'y'|_p = \max\{k : p^k|(xy - x'y')\}$$
$$\geq \max\big\{k : p^k|\big((x-x')y+x'(y-y')\big)\big\} \geq \min\big(\ \max\{k : p^k|(x-x')y\},\ \max\big\{k : p^k|x'(y-y')\big\}\ \big)$$
$$\geq \min\big(\ \max\{k : p^k|(x-x')\},\ \max\big\{k : p^k|(y-y')\big\}\ \big) = \min\big(-\log_p |x-x'|_p, -\log_p |y-y'|\big)$$
$$= -\max(\log_p |x - x'|_p, \log_p |y - y'|_p)$$

[13] We define $|0|_0 = p^{-\infty} = 0$

Taking $p$ to the negative of both sides, we find

$$-\log_2 \varepsilon(x \cdot y, x' \cdot y') \leq \max(|x - x'|_p, |y - y'|_p) = -\min(\log_2 \varepsilon(x, x'), \log_2 \varepsilon(y, y'))$$

Taking 2 to the negative of both sides of this, we find

$$\varepsilon(x \cdot y, x' \cdot y')) \geq \min(\varepsilon(x, x'), \varepsilon(y, y'))$$

So multiplication is 1-Lipschitz. Addition is also 1-Lipschitz, although we leave this as an exercise to the reader. This means we can compute addition and multiplication on $\mathbb{Z}_p$ with a Turing functional if we represent its elements as fast Cauchy sequences of elements of $\mathbb{Z}$ identified as a dense subset of $\mathbb{Z}_p$. We can assign every formula in our language the 1-Lipschitz modulus. □

Note that it's probably easier in this case to work with the metric rather than the equalness relation (as is often the case when you are given a space as a metric space). It's even easier to work with the valuation $\nu(x) = \max\{k : p^k | x\}$ valued in $\{0, 1, 2, 3, \ldots, \infty\}$. Once you've proven the corresponding inequality in terms of valuation, the corresponding inequality in terms of $\varepsilon$ follows immediately.

$\mathbb{Z}_p$ is an example of a compact continuous structure, and we will see later that compact continuous structures have special properties which make them behave like finite structures do in classical model theory.

## Addition in Hilbert Spaces

If we want to do analysis over a continuous logic base, we should probably say how to present Hilbert spaces. The first thing to notice is that addition in a Hilbert space (or more generally a Banach space) will be uniformly continuous with respect to the equalness relation $\varepsilon(x, y) = 2^{-||x-y||}$, with modulus of continuity $\Lambda_+(\tau_1, \tau_2) = \tau_1\tau_2$. Let's verify this:

$$\varepsilon(x + y, x' + y') = 2^{-||(x+y)-(x'+y')||} = 2^{-||(x-x')+(y-y')||} \geq 2^{-||x-x'||}2^{-||y-y'||} = \varepsilon(x, x')\varepsilon(y, y')$$

. We should also verify that $\Lambda_+$ is in fact a binary modulus:

$$\Lambda_+(\tau_1, \tau_2) = \tau_1\tau_2 \geq (\tau_1\tau_1')(\tau_2\tau_2') = \Lambda_+(\tau_1\tau_1', \tau_2\tau_2')$$

$$\Lambda_+(\tau_1\tau_1', \tau_2\tau_2') = (\tau_1\tau_1')(\tau_2\tau_2') = (\tau_1\tau_2)(\tau_1'\tau_2') \geq \Lambda_+(\tau_1, \tau_2)\Lambda_+(\tau_1', \tau_2')$$

We can see $\Lambda_+$ is continuous, and $\Lambda_+(1, 1) = 1^2 = 1$.

The classical way to present a vector space over a field $\mathbb{F}$ as a first-order structure is to include a constant symbol for the zero vector, a binary operation for addition, and unary function symbol for scalar multiplication by $c$ for each $c \in \mathbb{F}$. One reason for not using a two-sorted structure with a sort for the field and a sort for the vector space is that you don't have to deal with field enlarging if you move to an elementary extension. This isn't too much of a worry, though, since any vector space over an elementary extension $\mathbb{R}^*$ of $\mathbb{R}$, for example, will also be a vector space over $\mathbb{R}$, so classically there is not much harm in including a sort for our field classically. In continuous logic, however, there is an additional problem in including the scalars as a sort: scalar multiplication is not uniformly continuous as a function $\mathbb{R} \times \mathbb{H} \to \mathbb{H}$ for $\mathbb{H}$ a real Hilbert space. If we try to make a naive bound, using the metric for convenience, we might write:

$$||cx - c'x'|| = ||(cx - cx') + (cx' - c'x')|| \leq |c|||x - x'|| + |c - c'|||x'||$$

We cannot express this bound as a function just of $|c - c'|$ and $||x - x'||$. If we fix $c = c'$, then we see that scalar multiplication by a fixed scalar $c \in \mathbb{R}$ is $|c|$-Lipschitz, so if we include a function symbol for multiplication by $c$ for each $c \in \mathbb{R}$, then we don't have a problem. But we have an issue here if we want to include a sort for $\mathbb{R}$, for example to allow an easy presentation of linear functionals and inner products, or to express linear dependence in a first-order way.

But even though scalar multiplication is not uniformly continuous, it seems to be computable in many natural presentations of Hilbert spaces, such as a presentation of $\mathbb{H} = \mathcal{L}^2([0,1], \mathbb{R})$ with the metric $||f - g|| = \sqrt{\int_0^1 fg dx}$, where our dense set consists of piecewise-linear functions. The reason for this is that for every bounded region $U \subseteq \mathbb{H}$ and every bounded region in $V \subseteq \mathbb{R}$, scalar multiplication *is* uniformly continuous on the product $V \times U$ of those regions. More generally, if we have different moduli for different regions covering our continuous structure, and a uniform procedure for selecting an appropriate computable modulus for the restriction of a given function to each of those regions, then we can compute the function given its values on a dense set. The way Ben Yaacov et. al. handle this problem is to make a sort for each region, and then build into the language the inclusion maps between these various regions. The problem with this is, of course, that even to present just the one-dimensional Hilbert space $\mathbb{R}$, you need infinitely many sorts, because you need infinitely many bounded regions to cover $\mathbb{R}$, and this makes the language tedious to work with.

To understand why Ben Yaacov makes this choice, we should consider an alternative and see what goes wrong.

## Local Moduli

The idea is that as long as we can locally compute a modulus for a function $f$ from other functions/relations in our language that we already know how to compute, then we can compute $f$ given its values on a dense set. These other functions and relations will serve as parameters which tell us which modulus to use in any given region of our structure. Our main challenge is that we want this computation of a local modulus to be uniform, and we want this local modulus to be properly a part of the language itself, rather than a part of the theory of a given structure. We'll see we also will need a kind of well-foundedness in what functions local moduli are allowed to depend on: we don't want the modulus of $R_0$ to depend on $R_1$, whose modulus depends on $R_2$, whose modulus depends on $R_0$. So we will assume that we have well-order of all the symbols in our language (which we will always have if our language is computable).

Consider the following inequality for scalar multiplication:

$$||cx-c'x'|| = ||c(x-x')+(c'-c)x+(c-c')(x-x')|| \leq |c|||x-x'||+|c'-c|||x||+|c-c'|||x-x'||$$

If we think of $c$ and $x$ as being fixed, and thus $|c|$ and $||x||$ as constants, this gives us a bound only in terms of $||x-x'||$ and $|c-c'|$. Re-expressing this in terms of equalness relations, we have:

$$\varepsilon(cx, c'x') = 2^{-||cx-c'x'||} \geq 2^{-|c|||x-x'||}2^{-|c'-c|||x||}2^{-|c-c'|||x-x'||}$$

$$= 2^{-\left(-\log_2 \varepsilon(0,c)\right)\left(-\log_2 \varepsilon(x,x')\right)} * 2^{-\left(-log_2\varepsilon(c',c)\right)\left(-log_2\varepsilon(x,0)\right)} * 2^{-\left(-log_2\varepsilon(c,c')\right)\left(-log_2\varepsilon(x,x')\right)}$$

We can see that $\big(-\log_2 \varepsilon(0,c)\big)\big(-\log_2(\varepsilon(x,x')\big)$, $\big(-\log_2(\varepsilon(x,0)\big)\big(-log_2\varepsilon(c,c')\big)$, and $\big(-log_2\varepsilon(x,x')\big)\big(-log_2(\varepsilon(c,c')\big)$ are all upper computable (the trick here is that each of the terms of each of these products must be a finite number, since $\varepsilon$ will never take on the value 0, so we can just wait until we see that $\varepsilon$ is bounded away from zero to get an upper bound on the $-\log_2$ of each, which gives us an upper bound on the product). This tells us that:

$$2^{-\left(-\log_2 \varepsilon(0,c)\right)\left(-\log_2 \varepsilon(x,x')\right)} * 2^{-\left(-log_2\varepsilon(c',c)\right)\left(-log_2\varepsilon(x,0)\right)} * 2^{-\left(-log_2\varepsilon(c,c')\right)\left(-log_2\varepsilon(x,x')\right)}$$

is lower computable, if equalness is computable.

To translate this into something like a modulus, define

$$\beta(\rho_1, \rho_2, \tau_1, \tau_2) := 2^{-\log_2 \rho_1 \log_2 \tau_2}2^{-\log_2 \rho_2 \log_2 \tau_1}2^{-\log_2 \tau_1 \log_2 \tau_2}$$

And note that our inequality can be translated as: $\varepsilon(cx, c'x') \geq \beta(\varepsilon(c,0), \varepsilon(x,0), \varepsilon(c,c'), \varepsilon(x,x'))$ We can think of $\rho_1, \rho_2$ as parameters, and we get a different "modulus" for each choice of parameters. The function $\beta$ is lower computable, which suffices to allow us to approximate $\varepsilon(cx, c'x')$ from below in terms of approximations for $c$, $x$, $c'$, and $x'$, which is all we need to be able compute scalar multiplication given its values on a dense set.

Note that for fixed $\rho_1, \rho_2 \in [0,1)$, we have that:

$$\beta(\rho_1, \rho_2, \tau_1, \tau_2) \geq \beta(\rho_1, \rho_2, \tau_1\tau_1', \tau_2\tau_2')$$

and

$$\beta(\rho_1, \rho_2, \tau_1\tau_1', \tau_2\tau_2')$$

$$= 2^{-\log_2 \rho_1 (\log_2 \tau_2 + \log_2 \tau_2')} 2^{-\log_2 \rho_2 (\log_2 \tau_1 + \log_2 \tau_1')} 2^{-\log_2 \tau_1 \log_2 \tau_2 - \log_2 \tau_1 \log_2 \tau_2' - \log_2 \tau_1' \log_2 \tau_2 - \log_2 \tau_1' \log_2 \tau_2'}$$

$$= \beta(\rho_1, \rho_2, \tau_1, \tau_2)\beta(\rho_1, \rho_2, \tau_1', \tau_2') 2^{-\log_2 \tau_1 \log_2 \tau_2' - \log_2 \tau_1' \log_2 \tau_2}$$

So $\beta$ with $\rho_1, \rho_2$ fixed is not quite a binary modulus in $\tau_1$, $\tau_2$, but it has similar enough properties to a modulus to be used for the same purposes in this application.

We can do a similar trick with the inner product, and end up with a computable presentation with local moduli of a separable Hilbert space, with sorts for scalars and vectors, and where addition, scalar multiplication, and the inner product are all computable. This is a fairly standard practice in the field of computable analysis. But we are not really trying to do computable analysis here: we're trying to do computable structure theory for continuous structures. We don't care just about computing particular functions in particular presentations of a structure, but about classes of structures, and the general model theory of such structures.

We could give a general definition of a "local modulus", generalizing the idea in this example with Hilbert spaces, and thus expand our notion of what a "continuous structure" is to those which have local moduli instead of just moduli. A local modulus would be something with properties similar to a modulus which depends continuously on some parameters given by other functions and relations in the language, in a well-founded way. But we've already said enough in this example with scalar multiplication in Hilbert spaces to notice a problem with this approach from the logical point of view: a version of continuous logic with local moduli will not satisfy the compactness theorem! You can see Appendix D for a discussion of how the compactness theorem can be proven in continuous logic when we restrict to moduli and uniformly continuous functions. You might want to read this appendix first so the following discussion makes sense.

The basic problem here is that while $\beta$ is continuous on $(0, 1]^4$, which are the only values $\varepsilon$ takes on standard Hilbert spaces, it cannot be extended to a continuous function on $[0, 1]^4$. The idea here is to look at what happens in an elementary extension of our Hilbert space given by an ultraproduct which has "infinite" elements. If you multiply infinite elements by infinite scalars, even if those infinite elements and infinite scalars are very close to each other, the scalar products may not be close to each other.

**Example 4.** Here's an explicit example. Let $c_i = i$, $c_i' = i + 2^{-i}$, $x_0$ be some non-zero vector, $x_{i+1} = 2x$, and $x_i' = (1 + 2^{-2i})x_i$. Consider the sequences $\bar{c} = (c_i : i \in \omega)$, $\bar{c}' = (c_i' : i \in \omega)$ and $\bar{x} = (x_i : i \in \omega)$, $\bar{x}' = (x_i' : i \in \omega)$, thought of as representatives of elements of the classical ultraproduct (before quotienting by the "infinitesimally close" relation to obtain the continuous ultraproduct). Notice that

$$\varepsilon(c_i, c_i') = 2^{-|i-(i+2^{-i})|} = 2^{-2^{-i}} \to 1 \text{ as } i \to \infty$$

and

$$\varepsilon(x_i, x_i') = 2^{-||2^i x_0 - (1+2^{-2i})2^i x_0||} = 2^{-2^{-i}||x_0||} \to 1 \text{ as } i \to \infty$$

but $\varepsilon(c_i, 0) \to 0$, $\varepsilon(c_i', 0) \to 0$, $\varepsilon(x_i, 0) \to 0$, and $\varepsilon(x_i', 0) \to 0$, so in the ultraproduct $\bar{c}, \bar{c}', \bar{x}$, and $\bar{x}'$ are infinite elements. We can calculate that:

$$\varepsilon(c_i x_i, c_i' x_i') = 2^{-||i2^i x_0 - (i+2^{-i})(1+2^{-2i})2^i x_0||} = 2^{-||i2^i x_0 - (i+i2^{-2i}+2^i+2^{-3i})2^i x_0||}$$
$$= 2^{-||-(i2^{-i}+1+2^{-2i})x_0||} = 2^{-|1+i2^{-i}+2^{-2i}|||x_0||} \to 2^{-||x_0||} \neq 1$$

which means that scalar multiplication does not respect the "infinitesimally close" relation, so we cannot quotient by it to obtain the continuous ultraproduct. That is, the pairs $(\bar{c}, \bar{x})$ and $(\bar{c}', \bar{x}')$ are infinitesimally close, so should correspond to the same element in the continuous ultraproduct, but $\bar{c}\bar{x}$ and $\bar{c}'\bar{x}$ are not infinitesimally close, so they would have to correspond to different elements in the continuous ultraproduct, meaning scalar multiplication does not extend to a well-defined function (never mind a continuous function) on the continuous ultraproduct.

This tells us that expanding our notion of continuous structure to have local moduli loses some important logical features, so we will not pursue this line. However, it would still be nice if we could present bilinear maps in a natural and non-tedious way.

## Making Bilinear Maps Uniformly Continuous

Note that to be able to express that a vector $u$ is linearly dependent on vectors $v_1, \ldots, v_k$ in a vector space over $\mathbb{R}$, it suffices to be able to express

$$(\exists \lambda \in (0,1))(\exists c_1, \ldots c_k \in [0,1)) \text{ such that } \lambda u = c_1 v_1 + \cdots + c_k v_k$$

So in fact for most purposes, it is not necessary for scalar multiplication to be defined for large scalars, which should be able to help us find a version of scalar multiplication which is both useful in practice and uniformly continuous.

There are some geometric tricks we can do in Hilbert spaces which will help us.

**Fact 1.** The function $x \mapsto \frac{x}{1+||x||}$ is Lipschitz in any Hilbert space.

This fact does not immediately follow from the fact that we are scaling by a positive scalar less than 1, because it is possible in general for any $K$ to have $||c_1 x - c_2 y|| > K||x - y||$ for some $x, y$, and scalars $0 < c_1, c_2 < 1$.

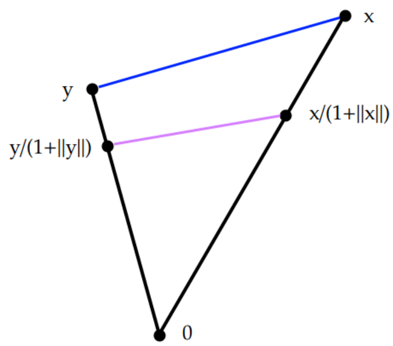



Diagram for Reference

*Proof of Fact 1.* Let $x, y \in \mathbb{H}$, a Hilbert space. Without loss of generality, we may suppose that $\mathbb{H}$ is two-dimensional by considering a two-dimensional subspace containing the span of $x$ and $y$. Write $x = x_1 u + x_2 v$, $y = y_1 u + y_2 v$ for $\{u, v\}$ an orthonormal basis.

Then we can see:

$$||\frac{x}{1+||x||} - \frac{y}{1+||y||}||^2 = ||(\frac{x_1}{1+||x||}u + \frac{x_2}{1+||x||}v) - (\frac{y_1}{1+||y||}u + \frac{y_2}{1+||y||}v)||^2$$

$$= (\frac{x_1}{1+||x||} - \frac{y_1}{1+||y||})^2 + (\frac{x_2}{1+||x||} - \frac{y_2}{1+||y||})^2$$

Now notice that

$$\left|\frac{\frac{x_i}{1+||x||} - \frac{y_i}{1+||y||}}{||x-y||}\right| = \left|\frac{(1 - \frac{1+||x||-x_i}{1+||x||}) - (1 - \frac{1+||y||-y_i}{1+||y||})}{||x-y||}\right| = \left|\frac{\frac{1+||y||-y_i}{1+||y||} - \frac{1+||x||-x_i}{1+||x||}}{||x-y||}\right|$$

$$= \left|\frac{\frac{(1+||y||-y_i)(1+||x||)}{(1+||x||)(1+||y||)} - \frac{(1+||x||-x_i)(1+||y||)}{(1+||x||)(1+||y||)}}{||x-y||}\right|$$

$$= \left|\frac{\frac{1+||x||+(||y||-y_i)||x||+||y||-y_i}{(1+||x||)(1+||y||)} - \frac{1+||y||+(||x||-x_i)||y||+||x||-x_i}{(1+||x||)(1+||y||)}}{||x-y||}\right|$$

$$= \left|\frac{\left(\frac{x_i-y_i}{(1+||x||)(1+||y||)}\right) + \frac{(||y||-y_i)||x||-(||x||-x_i)||y||}{(1+||x||)(1+||y||)}}{||x-y||}\right|$$

$$= \left|\frac{\frac{x_i-y_i}{||x-y||}}{(1+||x||)(1+||y||)} + \frac{\left(\frac{(||y||-y_i)(||x||-||y||)+\big((||y||-y_i)-(||x||-x_i)\big)||y||}{(1+||x||)(1+||y||)}\right)}{||x-y||}\right|$$

$$= \left| \frac{\frac{x_i - y_i}{||x-y||}}{(1+||x||)(1+||y||)} + \frac{(||x|| - ||y||) \cdot \frac{||y|| - y_i}{(1+||y||)} \cdot \frac{1}{1+||x||}}{||x-y||} + \frac{\left(\frac{||y|| - ||x||}{1+||x||} - \frac{x_i - y_i}{1+||x||}\right)\frac{||y||}{1+||y||}}{||x-y||} \right|$$

$$\leq \left| \frac{x_i - y_i}{||x-y||} \right| + \left| \frac{(||x|| - ||y||) \cdot \frac{||y|| - y_i}{(1+||y||)}}{||x-y||} \right| + \left| \frac{(||y|| - ||x||) - (x_i - y_i)}{||x-y||} \right|$$

$$\leq \left| \frac{x_i - y_i}{||x-y||} \right| + \left| \frac{||x|| - ||y||}{||x-y||} \right| + \left| \frac{||y|| - ||x||}{||x-y||} \right| + \left| \frac{x_i - y_i}{||x-y||} \right| \leq 1+1+1+1 = 4$$

Thus $x \mapsto \frac{x}{1+||x||}$ is $\sqrt{32}$-Lipschitz[14], so in particular uniformly continuous.

□

**Aside:** We can also generalize this argument to Banach spaces by considering in a given two-dimensional subspace an inner product whose corresponding norm can be bounded multiplicatively both above and below the restriction of the Banach space norm to that two-dimensional subspace. Of course different two-dimensional subspaces of a Banach space might have wildly different norms, but for an appropriate choice of coordinates on that subspace, the restriction of the Banach space norm can be bounded above and below by multiples of the Euclidean norm on that subspace, i.e. $c||x||_2 \leq ||x|| \leq C||x||_2$ for some $0 < c < C$, where $c, C$ can be chosen independently of the subspace. To choose the coordinate system, consider $B := \{x : ||x|| \leq 1\}$, and choose linear functionals $l_1, l_2$ such that $l_1(B)$ and $l_2(B)$ are bounded by 1, and the maximum of $l_1$ and $l_2$ are achieved at some $v_1$ and $v_2$ respectively on the boundary of $B$ with $l_i(v_{1-i}) = 0$. We can see that using the linear coordinate system with basis $\{v_1, v_2\}$, the unit ball in the $\infty$-norm with respect to that coordinate system contains $B$ by construction, and the unit ball in the 1-norm with respect to that coordinate system is contained in $B$ by convexity of $B$. We can then see that $c = \frac{1}{\sqrt{2}}$ and $C = \sqrt{2}$ suffices.

**Fact 2.** The map $(c, x) \mapsto \frac{1}{1+|c|}x$ is uniformly continuous when restricted to vectors of norm at most 1.

*Proof of Fact 2.* Let $x \in \mathbb{H}$ with $||x|| \leq 1$. Then we can compute:

$$\left|\left| \frac{1}{1+|c|}x - \frac{1}{1+|c'|}x' \right|\right| = \left|\left| \left(\frac{1}{1+|c|}x - \frac{1}{1+|c'|}x\right) + \left(\frac{1}{1+|c'|}x - \frac{1}{1+|c'|}x'\right) \right|\right|$$

$$\leq \left| \frac{1}{1+|c|} - \frac{1}{1+|c'|} \right| \cdot ||x|| + \frac{||x-x'||}{1+|c'|} = \frac{(1+|c'|) - (1+|c|)}{(1+|c|)(1+|c'|)}||x|| + \frac{||x-x'||}{1+|c'|}$$

[14] $\sqrt{32} = \sqrt{4^2+4^2}$

$$= \frac{|c'| - |c|}{(1+|c|)(1+|c'|)}||x|| + \frac{||x-x'||}{1+|c'|} \leq |c-c'| + ||x-x'||$$

□

Combining facts 1 & 2, we find that the composition $(c, x) \mapsto \frac{1}{1+|c|}\frac{x}{1+|x|}$ is uniformly continuous $\mathbb{R} \times \mathbb{H} \to \mathbb{H}$. It is a total function, but it is a deviant version of scalar multiplication which only allows you to scale by small positive amounts.

**Definition 48.** The **small scalar multiplication** on a real Hilbert space, $\odot : \mathbb{R} \times \mathbb{H} \to \mathbb{H}$, is defined by $c \odot x = \frac{1}{1+|c|}\frac{x}{1+||x||}$.

**Definition 49.** The **small inner product** on a real Hilbert space, $\langle\langle \cdot, \cdot \rangle\rangle : \mathbb{H} \times \mathbb{H} \to \mathbb{R}$, is defined by $\langle\langle x, y \rangle\rangle = \langle \frac{x}{1+||x||}, \frac{y}{1+||y\}} \rangle$.

Using fact 1, we can see that the small inner product is uniformly continuous, because for $||x||, ||y|| \leq 1$, we have:

$$||\langle x,y\rangle - \langle x',y'\rangle|| = ||\langle x-x',y\rangle + \langle x',y-y'\rangle|| \leq ||x-x'|| \cdot ||y|| + ||x'|| \cdot ||y-y'|| \leq ||x-x'|| + ||y-y'||$$

It's useful to note that $\langle\langle x, y \rangle\rangle = 0$ if and only if $\langle x, y \rangle = 0$ in a real Hilbert space.

This basic idea of replacing one function which isn't uniformly continuous with a another function which is is quite useful. Another idea we can use is that even if an inevitable map has no uniformly continuous inverse, we may be able to find an analogous function to its inverse in the set of truth values. The map $||x|| \mapsto \left|\left|\frac{x}{1+||x||}\right|\right|$ has a continuous inverse, but it's not uniformly continuous. However, we can, in some sense, build an inverse to this with the following logical connective:

**Definition 50.** We define the unary continuous logical connective $\eta(\tau) := 2^{\frac{\log_2 \tau}{1+\log_2 \tau}}$ if $\frac{1}{2} < \tau \leq 1$, and 0 if $0 \leq \tau \leq \frac{1}{2}$.

*Verification of Uniform Continuity.* Let's try to bound $\eta'(t)$ for $t > \frac{1}{2}$:

$$\eta'(t) = \frac{\frac{1+log_2 t}{t \ln 2} - \frac{\log_2 t}{t \ln 2}}{(1+\log_2 t)^2} \cdot \ln 2 \cdot 2^{\frac{\log_2 t}{1+\log_2 t}} = \frac{\frac{1}{t}}{(1+\log_2 t)^2} 2^{\frac{\log_2 t}{1+\log_2 t}} = \frac{2^{-\log_2 t} \cdot 2^{\frac{\log_2 t}{1+\log_2 t}}}{(1+\log_2 t)^2}$$

$$= \frac{2^{\frac{\log_2 t}{1+\log_2 t} - \log_2 t}}{(1+\log_2 t)^2} = \frac{2^{\frac{\log_2 t}{1+\log_2 t} - \frac{\log_2 t + (\log_2 t)^2}{1+\log_2 t}}}{(1+\log_2 t)^2} = \frac{2^{-\frac{(\log_2 t)^2}{1+\log_2 t}}}{(1+\log_2 t)^2}$$

$$\eta''(t) = \frac{\frac{-\frac{1}{t \ln 2} 2(\log_2 t)(1+\log_2 t) + (log_2 t)^2 \frac{1}{t \ln 2}}{(1+\log_2 t)^2} \cdot \ln 2 \cdot 2^{-\frac{(\log_2 t)^2}{1+\log_2 t}} \cdot (1+log_2 t)^2 - 2^{-\frac{(\log_2 t)^2}{1+\log_2 t}} \frac{1}{t \ln 2} 2(1+\log_2 t)}{(1+\log_2 t)^4}$$

$$= \frac{\frac{1}{t} \cdot 2^{-\frac{(\log_2 t)^2}{1+\log_2 t}}}{(1+\log_2 t)^4 \ln 2}\Big( -2\log_2 t(1+\log_2 t) + (log_2 t)^2 - \frac{2}{\ln 2}(1+\log_2 t)\Big)$$

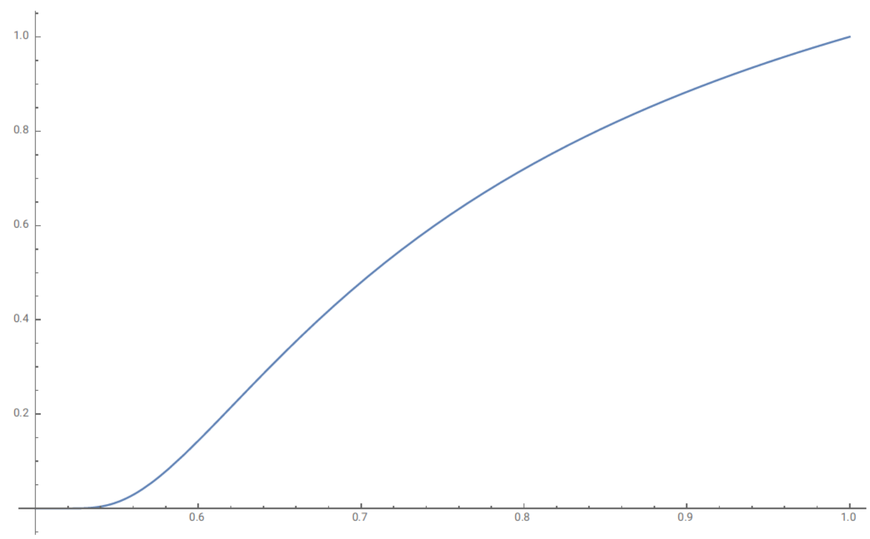
For $t > \frac{1}{2}$, we can see $\eta''(t) = 0$ iff $-2\log_2 t(1+\log_2 t) + (log_2 t)^2 - \frac{2}{\ln 2}(1+\log_2 t) = 0$. This is a quadratic polynomial in $\log_2 t$, so substitute $s = \log_2 t$ to solve:

$$0 = -2s(1+s) + s^2 - \frac{2}{\ln 2}(1+s) = -s^2 - (2+\frac{2}{\ln 2})s - \frac{2}{\ln 2}$$

We can now apply a technique developed by the Indian mathematician Brahmagupta in the seventh century A.D.[15]:

$$s = \frac{-2(1+\frac{1}{\ln 2}) \pm \sqrt{\frac{4}{(\ln 2)^2} + \frac{8}{\ln 2} + 4 - \frac{8}{\ln_2}}}{2} = -1 + -\frac{1}{\ln 2} \pm \sqrt{\frac{1}{(\ln_2)^2} + 1} = -1 - \frac{1}{\ln 2}(1 \pm \sqrt{1+(\ln 2)^2})$$

So our solutions are $t = \frac{1}{2}(2^{-\frac{1}{\ln 2}(1\pm\sqrt{1+(\ln 2)^2})}) = \frac{1}{2}(e^{-(1\pm\sqrt{1+(\ln 2)^2})}) = (\frac{1}{2e})e^{\pm\sqrt{1+(\ln 2)^2}}$. Only the solution $t = (\frac{1}{2e})e^{\sqrt{1+(\ln 2)^2}}$ will correspond to $t \in (\frac{1}{2}, 1]$. A numerical calculation shows that $\eta'((\frac{1}{2e})e^{\sqrt{1+(\ln 2)^2}}) \leq 4$. We can also check that $\lim_{t\to\frac{1}{2}} \eta'(t) = 0$, and $\eta'(1) = 1$, so $\eta'$ is uniformly bounded by 4 on $(\frac{1}{2}, 1]$, and it is by definition 0 on $[0, \frac{1}{2}]$. We can conclude that $\eta$ is 4-Lipschitz. □




Graph of $\eta(t)$ for $\frac{1}{2} \leq t \leq 1$

Let's check that $\eta$ actually does what it is supposed to do:

[15]This obscure technique is known as the quadratic formula.

$$\eta(\varepsilon(\frac{x}{1+||x||},0)) = \eta(2^{-\left|\left|\frac{x}{1+||x||}\right|\right|}) = 2^{\left(\frac{\log_2 2^{-\left|\left|\frac{x}{1+||x||}\right|\right|}}{1+\log_2 2^{-\left|\left|\frac{x}{1+||x||}\right|\right|}}\right)} = 2^{\left(\frac{-\left|\left|\frac{x}{1+||x||}\right|\right|}{1-\left|\left|\frac{x}{1+||x||}\right|\right|}\right)}$$

$$= 2^{\left(\frac{-\left|\left|\frac{x}{1+||x||}(1+||x||)\right|\right|}{(1+||x||)-\left|\left|\frac{x}{1+||x||}(1+||x||)\right|\right|}\right)} = 2^{\left(\frac{-||x||}{(1+||x||)-||x||}\right)} = 2^{-||x||} = \varepsilon(x,0)$$

Now let's give an example of how to express the relation that one vector $x$ is a scalar multiple of another vector $y$. There is a slight subtlety here, because we are not *defining* this relation, in the strict sense of definability of a relation in continuous logic. We cannot define non-continuous relations, and the relation $R(x,y) = \exists c(x = cy)$ is not continuous. Nor can we define the continuous analog of the classical relation, i.e. the distance/equalness function to the set of pairs $S := \{x, y : \exists c(x - cy)\}$ given by $\varepsilon(\bar{x}, S) = \sup_{\bar{s}\in S} \varepsilon(\bar{s}, \bar{x})$. We are rather expressing it in a much weaker sense by a continuous relation $R(x,y)$ for which $R(x,y) = 1 \Leftrightarrow \exists c(x = cy)$. Such things are useful, but it is important to remember that this is a much weaker notion of definability than the standard notion in continuous logic. For the sake of clarity, we will give the following definition:

**Definition 51.** A subset $X$ of (a Cartesian power of) a continuous structure $\mathbb{M}$ is **weakly definable** if there is a formula $\phi$ such that $\mathcal{M} \models [\phi(\bar{x}) = 1]$ if and only if $\bar{x} \in X$.

The important thing to think about here is: what happens to a weakly definable set in an elementary extension of a structure? We'll give some examples later that show weakly definable sets are not the appropriate notion of definability over classes of structures. But if we have a particular structure in mind, it can be useful, but we should really think of ourselves as no longer working inside continuous logic when we do so. We need to be especially careful when quantifying over weakly defined sets.

**Example 5.** Let's give an example of how quantifying over weakly defined sets can lead to mistakes. Suppose we want to say that $x, y \in \mathbb{H}$ are linearly dependent if and only if $\varepsilon(x,0) = 1$, $\varepsilon(y,0) = 1$, or there exist $c, d \in \mathbb{R}$ and such that $\varepsilon(c \odot x, d \odot y) = 1$ or $\varepsilon(c \odot x, d \odot (-y)) = 1$. This is certainly true, and we might be tempted to replace this existential quantifier with a supremum to a get a formula in continuous logic. But unfortunately, when we do so, our definition no longer works, because the limit as $c \to \infty$ and $d \to \infty$ of $\varepsilon(c \odot x, d \odot y) = 1$ for *any* $x, y \in \mathbb{H}$. Essentially the problem is that the relation $R(c,x,d,y) := \varepsilon(c \odot x, d \odot y)$ only weakly defines the set $\{(c,x,d,y) : c \odot x = d \odot y\}$. Tuples $(c,x,d,y) \in (\mathbb{R} \times \mathbb{H})^2$ that are closeness $\frac{1}{2}$ to this set (i.e., distance 1 from this set) can have $R(c,x,d,y)$ arbitrarily close to 1: fix $x, y$ linearly independent and choose $c = 2^n$, $d = 2^n + 1$ for $n$ sufficiently large.

This explains why Ben Yaacov et. al. chose the definition they did for a set being definable in continuous logic (the distance function from that set being definable). That being said, it is still possible to weakly define the linear dependence of $x, y \in \mathbb{H}$ for standard Hilbert spaces. The trick is to bound at least one of $c, d$ in $\varepsilon(c \odot x, d \odot y)$ away from $\infty$. A similar trick will work for linear dependence of $n$ vectors for any $n \in \omega$.

The case $n = 2$ is simpler, however, so we will start with this. We know either $0 < ||x|| \leq ||y||$, or $0 < ||y|| \leq ||x||$, or $x = 0$, or $y = 0$. In the case $0 < ||x|| \leq ||y||$, we can without loss of generality assume $d = 0$, and in the case $0 < ||y|| \leq ||x||$, we can assume $c = 0$. So the following formula weakly defines $\{(x, y) \in \mathbb{H}^2 : x, y \text{ are linearly dependent}\}$:

$$\varepsilon(x,0)\vee\varepsilon(y,0)\vee\sup_d(\varepsilon(0\odot x, d\odot y)\vee\varepsilon(0\odot x, d\odot(-y)))\vee\sup_c(\varepsilon(c\odot x, 0\odot y))\vee(\varepsilon(c\odot(-x), 0\odot y))$$

Recall $0 \odot x = \frac{x}{1+||x||}$, so assuming $0 < ||y|| \leq ||x||$, we have:

$$\sup_d \varepsilon(0 \odot x, d \odot y) = \sup_d \varepsilon(\frac{x}{1+||x||}, \frac{1}{1+|d|}\frac{y}{1+||y||}) = \sup_{0<\alpha\leq 1} \varepsilon(\frac{x}{1+||x||}, \alpha\frac{y}{1+||y||})$$

$$= \sup_{0<\alpha\leq 1} \varepsilon(x, \alpha\frac{1+||x||}{1+||y||}y)^{\frac{1}{1+||x||}} \geq \sup_{0<\beta\leq 1} \varepsilon(x, \beta y) \qquad (\text{since } \frac{1+||x||}{1+||y||} \geq 1\ )$$

So $x = \beta y$ has a solution $0 < \beta \leq 1$ (in this case where $0 < ||y|| \leq ||x||$) if and only if $\sup_d \varepsilon(0 \odot x, d \odot y) = 1$. We can simplify our formula slightly by noticing that if $x = 0$, then $\sup_d \varepsilon(0 \odot x, d \odot y) = \sup_d \varepsilon(0, d \odot y) = 1$ for any $y$ in a standard Hilbert space, so we don't need to include the $\varepsilon(x, 0)$ and $\varepsilon(y, 0)$ disjuncts.

We unfortunately cannot get rid of the disjunction over the signs of the vectors in the linear combination. Let's now write our general formula weakly defining linear dependence of $n$ vectors in a (standard) real Hilbert space:

$$\bigvee_{i<n} \bigvee_{\bar{\sigma}\in\{+,-\}^n} \sup_{\bar{c}} \varepsilon\big(0 \odot x_i, \sum_{j<n, j\neq i} c_j \odot \sigma_j x_j\big)$$

The only trick we are using here is that if $\sum_{i<n} \alpha_i v_i = 0$ for some $\alpha_i \in \mathbb{R}$ not all 0, then we can rescale this equation so one of the coefficients has absolute value 1, and all other coefficients have absolute value at most 1. To weakly define linear dependence of $n$ vectors in a Hilbert space over $\mathbb{C}$, we would need to replace the second disjunction over signs with a supremum over multiplication by complex scalars of absolute value 1.

**Observation 13.** *If $X$ is weakly definable by a formula $\phi$, this does not imply $X^c$ is weakly definable by $\neg\phi$. In fact, in most cases the complement of a weakly-definable set is not weakly definable. In particular, linear dependence is weakly definable, but linear independence is not.*

*Proof.* A weakly definable set is always closed, since the inverse image of $\{1\}$ under any continuous function must be closed. But this implies that the complement of a weakly definable set is always open. Thus the complement of a weakly definable set can be weakly definable only if it is clopen. The set of $n$-tuples of linearly independent vectors is not closed, so cannot be weakly-definable. $\square$

Of course, we could always give a definition of a **weakly codefinable** set as the complement of a weakly definable set, and even create a Borel hierarchy on a continuous structure by considering countable unions of definable sets, countable intersections of codefinable sets, countable unions of countable intersections of codefinable sets, countable intersections of countable unions of definable sets, and so on. In a computable continuous structure, we could define an analogue of the effective Borel hierarchy, but we'll discuss why this is not necessarily a good idea in a later section.

**Summary:** We can present standard Hilbert spaces with a two-sorted continuous structure $\mathbb{H}, \mathbb{R}, +, \otimes, \odot)$, and in this presentation linear dependence is weakly definable.

**Note:** Our approach, defining small scalar multiplication and the small inner product as uniformly continuous replacements for the ordinary scalar and inner products, is useful for using the language of computable continuous structure theory to do computable analysis, since in that case we are only concerned with standard Hilbert spaces. However, there is an alternate approach favored by Ben Yaacov et. al. which can be used, at the expense of some increased verbosity.

The observation is that if we restrict to bounded subsets of $\mathbb{R} \times \mathbb{H}$ and $\mathbb{H} \times \mathbb{H}$, the scalar and inner product *are* uniformly continuous. If we write both $\mathbb{R}$ and $\mathbb{H}$ as countable unions of bounded subsets (say as an increasing union of balls centered around the origin), we can present a Hilbert space as a multi-sorted structure whose sorts correspond to these bounded subsets. If we add inclusion maps between their overlaps to our langauge, this gives us a reasonable presentation of Hilbert spaces with the usual scalar and inner products. However, the limitation here is that we can only quantify over a priori bounded subsets, so there is some tradeoff.

# Chapter 2

# Computable Continuous Structure Theory

"Computable continuous structure theory", a term we are coining, is a generalization of computable structure theory to continuous structures. Computable structure theory as a sub-field of logic combines aspects of recursion theory and model theory. There are a few directions you might approach computable structure theory from. One is that you would like to generalize the fine-grained analysis of $\mathbb{N}$ and its subsets central to recursion theory to other structures. Another direction is from a concern about the effectiveness of various model-theoretic constructions and a desire for more delicate tools for classifying structures.

In this chapter, we'll try to generalize a representative sample of theorems from computable structure theory to continuous structures. We hope to demonstrate that computable continuous structure theory is a natural generalization of computable structure theory, and that the benefit of moving to this new framework is worth the additional effort required to use these new definitions by default.

## 2.1 Quasi Back-and-Forth Arguments

We would like to give a technique for producing isomorphisms between continuous structures analogous to the classical back-and-forth method for countable structures. In the classical back-and-forth method, you construct an isomorphism from $\mathcal{A}$ to $\mathcal{B}$ by finding a sequence of finite partial isomorphims whose limit, in some sense, is an isomorphism. This sequence can be thought of as given by a winning strategy to a particular infinite game. The game board starts with $\bar{a}_0 = \emptyset \in \mathcal{A}^0$ and $\bar{b}_0 = \emptyset \in \mathcal{B}^0$. At the start of stage $n+1$ of the game, we will have $|\bar{a}_n| = |\bar{b}_n| = n$. An antagonistic Player 1 then either extends $\bar{a}_n$ to $\bar{a}_{n+1}$ by adding a single element from $\mathcal{A}$ to the end, or else extends $\bar{b}_n$ to $\bar{b}_{n+1}$ by adding a single element from $\mathcal{B}$ to the end. Player 2 must then extend the other tuple that Player 1 didn't choose to also have length $n+1$, and tries (if they can) to ensure that $\mathcal{D}_{\mathcal{A}}(\bar{a}_{n+1}) = \mathcal{D}_{\mathcal{B}}(\bar{b}_{n+1})$, i.e. the atomic diagrams[1] of these finite structures are the same. After $\omega$ stages have been played, the result is a pair of infinite sequences $\bar{a}$ and $\bar{b}$. Player I, the antagonist, wins if $\bar{a}$ enumerates $\mathcal{A}$, $\bar{b}$ enumerates $\mathcal{B}$, but $\mathcal{D}_{\mathcal{A}}(\bar{a} \upharpoonright n) \neq \mathcal{D}_{\mathcal{B}}(\bar{b} \upharpoonright n)$ for some $n \in \omega$. Otherwise, Player II wins. If Player II has a winning strategy, then there is an isomorphism between the two structures given by sending $\bar{a}$ to $\bar{b}$. A classical application of this is showing that any two countable dense linear orders without endpoints are isomorphic.

We can extend this idea to continuous structures by weakening the requirement that the diagrams of the finite tuples match exactly, and slightly modifying the back and forth steps. Back-and-forth style arguments for computable metric structures were explored by Melnikov in his proof of the computable categoricity of Urysohn space in "Computably Isometric Spaces"[9]. We describe how to faithfully generalize classical back-and-forth arguments to continuous structures. But first we need to give a definition of the atomic diagram of a tuple which works well for our purposes.

**Definition 52.** Suppose $\bar{p} \in \mathcal{M}^k$. Then the **atomic diagram**[2] of $\bar{p}$ is a function $k! \to [0,1] \cap \mathbb{Q}$ such that

$$\mathcal{D}^{\mathcal{M}}(\bar{p})(i) = \frac{m}{2^{k+2}} \text{ only if } \varepsilon(\phi_i^{at,\mathcal{M}}[x_j \mapsto p_j,\, j < k], \frac{m}{2^{k+2}}) > 2^{-2^{-(k+2)}}$$

where "$[x_j \mapsto p_j,\, j < k]$" refers to substituting variables with corresponding points in $\mathcal{M}$, and $(\phi_i^{at} : i \in \omega)$ is a computable enumeration of the atomic $\mathcal{L}(\sigma, \nu)$ formulas such that

- The free variables of $\phi_i^{at}$ are contained in $\{x_j : j < k\}$, where $(x_j : j \in \omega)$ is a fixed one-to-one enumeration of the variable symbols.
- If $j < j' \leq k$, $\varepsilon(x_j, x_{j'})$ is among $\{\phi_i^{at} : i < k!\}$.
- If $\pi$ is a permutation of $k$, and $\phi(x_{i_1}, x_{i_2}, ...., x_{i_s})$ is among $\{\phi_i^{at} : i < k!\}$, then so is $\phi(x_{\pi(i_1)}, x_{\pi(i_2)}, ..., x_{\pi(i_s)})$.

Notice that the atomic diagram is not uniquely defined, because it's possible for the truth value of $\phi$ to be both in the interval $(\frac{m-1}{2^{k+2}}, \frac{m+1}{2^{k+2}})$ and $(\frac{m}{2^{k+2}}, \frac{m+2}{2^{k+2}})$. However, it is well defined up to a difference of at most $\frac{1}{2^{m+2}}$. Moreover if $\mathcal{M}$ is computable, it is possible to make a unique choice of atomic diagram by evaluating $\phi_i^{at}[x_j \mapsto p_j,\, j < k]$ to higher and higher accuracy,

[1] If we are working in a language with infinitely many atomic formulas of some arity, we define the atomic diagram $\mathcal{D}_{\mathcal{A}}(\bar{a})$ to be restricted to the first $|\bar{a}|$-many atomic formulas, so that this can be computed. See Antonio Montalbán's "Computable Structure Theory" (Draft)[10] for more details.

[2] From now on, we assume for convenience that all structures are single-sorted. All of the following results generalize in an obvious way to multi-sorted structures, but we choose not to present this because it would be easier for a reader who wishes to apply these ideas to multi-sorted structures to read the argument for single-sorted structures, and generalize it themself to multi-sorted structures, rather than to sift through the many arbitrary conventions we would inevitably use in making our arguments more general.

and choosing the first value of $\frac{m}{2^{k+2}}$ for which we are sure $\frac{m-1}{2^{k+2}} < \phi_i^{at}[x_j \mapsto p_j,\ j < k] < \frac{m+1}{2^{k+2}}$. This can be done uniformly in a name for $\bar{p}$.

**Exercise for the Reader.** Let $\mathcal{M}, \mathcal{N}$ be $\sigma$ structures and $\bar{p}, \bar{q}$ $k$-tuples from each. Show that if $\varepsilon(D^{\mathcal{N}}(\bar{p}), D^{\mathcal{M}}(\bar{q})) > 2^{-2^{-(k+1)}}$ then $\varepsilon(\phi^{\mathcal{M}}(\bar{p}), \phi^{\mathcal{N}}(\bar{q})) > 2^{-2^{-k}}$ for any $\phi \in \{\phi_i^{at} : i < k!\}$.

Our definition of the atomic diagram of a tuple allows us to give a more canonical form for the atomic diagram of a copy ($\omega$-presentation) of a structure. By restricting to only certain dyadic rational values with bounded denominators at each stage of the approximation, we can make the possible atomic diagrams of $\omega$-presentations of structures in a given computable language into the set of branches of a finitely-branching tree.

**Definition 53.** "The" **atomic diagram** of an $\omega$-presentation (copy) $\mathcal{A}$ of a continuous structure is $\bigoplus_{n\in\omega} \mathcal{D}((0,1,...,n-1))$.[3] Recall that for an $\omega$-presentation, we identify the dense set $D$ with $\omega$, so $(0,1,...,n-1)$ is the tuple consisting of the first $k$ elements of $D$.[4]

**Exercise for the Reader.** Given a copy of a structure with named atomic diagram, we can compute "the" atomic diagram by using the values of the names for its atomic formulas and modifying them to be nearby dyadic rationals with appropriately bounded denominators. Hint: if you know a rational number $r$ with $\varepsilon(\phi(\bar{p}), r) \geq 2^{2^{-(k+5)}}$, you should be able to find (not necessarily uniquely) a dyadic rational $\frac{m}{2^{k+2}}$ with $\varepsilon(\phi(\bar{p}), r) > 2^{2^{-(k+2)}}$. You can also make this choice canonical, by choosing the lower of the two possibilities, if there is more than one choice. For this reason, we will now be justified in saying "the" atomic diagram of a copy of a structure (or tuple from a copy of structure), because given the named atomic diagram of a copy of a structure, we have defined a canonical way of obtaining a named atomic diagram for it.

**Exercise for the Reader.** Suppose we identify copies of $\sigma$-structures, for a fixed computable language $\sigma$, with their atomic diagrams (instead of using *named* atomic diagrams). Show that the space of copies of $\sigma$-structures is a $\Pi_1^0$ subset of $2^\omega$.[5]. Hint: the main idea is to build a finite-branching tree of initial segments of atomic diagrams. Stop a branch if you see what you've said so far about the atomic formulas violates a condition (like one of the moduli of continuity).

---

[3] $\mathcal{D}((0,1,...,n-1))$ is a length $k!$ string of of rationals of the form $\frac{m}{2^{k+2}}$, which can be coded in a canonical way as a finite binary string. Given an infinite sequence of finite or infinite binary strings $(\sigma_i : i \in \omega)$, we can code these into a single infinite binary string $\bigoplus_{i\in\omega} \sigma_i$ in such a way that it is possible to compute from $\bigoplus_{i\in\omega} \sigma_i$ the length (finite or infinite) of $\sigma_i$ and its $k$-th digit, uniformly in $i$ and $k$. We assume we have fixed a canonical way of doing this.

[4] Again, recall that our enumeration of the dense set may have repeats, so by "first $k$ elements of $D$", we are not implying that these $k$ elements are pairwise non-equal.

[5] We mean here the space of all copies of all $\sigma$-structures, not the space of copies of a particular $\sigma$ structure

**Definition 54.** Let $\mathcal{Q} \subseteq \mathcal{M}^{<\omega} \times \mathcal{N}^{<\omega}$. $\mathcal{Q}$ is **quasi back-and-forth set** if the following hold:

- **[FORTH]** Suppose $\bar{p} \in \mathcal{M}^k$ and $\bar{q} \in \mathcal{N}^k$. If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, and $c \in \mathcal{M}$, then there is some $\bar{q}' \in \mathcal{N}^k$ and $d \in \mathcal{N}$ with $\varepsilon(q_i, q_i') > 2^{-2^{-(k+1)}}$ for $i < k$ and $(\bar{p}c, \bar{q}'d) \in \mathcal{Q}$.
- **[BACK]** Suppose $\bar{p} \in \mathcal{M}^k$ and $\bar{q} \in \mathcal{N}^k$. If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, and $d \in \mathcal{N}$, then there is some $\bar{p}' \in \mathcal{M}^k$ and $c \in \mathcal{M}$ with $\varepsilon(p_i, p_i') > 2^{-2^{-(k+1)}}$ for $i < k$ and $(\bar{p}'c, \bar{q}d) \in \mathcal{Q}$.
- **[QUASI ISOMORPHIC DIAGRAMS]** If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, then $|\bar{p}| = |\bar{q}| =: k$ and

$$\varepsilon(\mathcal{D}^{\mathcal{M}}(\bar{p}), \mathcal{D}^{\mathcal{N}}(\bar{q})) > 2^{-2^{-(k+1)}}$$

**Observation 14.** *If $\mathcal{M}$ and $\mathcal{N}$ both have the discrete equalness relation[6], then [BACK] and [FORTH] are equivalent to the much simpler classical back-and-forth conditions:*

- ***[FORTH*]*** *If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, and $c \in \mathcal{M}$, then $\exists d \in \mathcal{N}$ with $(\bar{p}c, \bar{q}d) \in \mathcal{Q}$.*
- ***[BACK*]*** *If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, and $d \in \mathcal{N}$, then $\exists c \in \mathcal{M}$ with $(\bar{p}c, \bar{q}d) \in \mathcal{Q}$.*

**Observation 15.** *$\mathcal{M}$ and $\mathcal{N}$ both have the discrete equalness relation, all non-logical[7] relations are $\{0,1\}$-valued, and all logical connectives $u$ are $\{0,1\}$-valued when restricted to $\{0,1\}^{Ari(u)}$ then the condition [QUASI ISOMORPHIC DIAGRAMS] can be replaced by the following:*

- ***[ISOMORPHIC DIAGRAMS]*** *If $(\bar{p}, \bar{q}) \in \mathcal{Q}$, then $|\bar{p}| = |\bar{q}|$ and $D^{\mathcal{M}}(\bar{p}) = D^{\mathcal{N}}(\bar{q})$.*

The reason for this is that in this situation, if we know $\phi_i^M(\bar{p}) > \frac{1}{2}$, we know its equal to 1, and if we know it's $< \frac{1}{2}$, we know it's equal to 0, so more precise information is not necessary.

**Definition 55.** A (multi-sorted) continuous structure is **discrete** if it has the discrete equalness relation (on each sort), and all relations are $\{0,1\}$ valued.

**Definition 56.** A collection of continuous logical connectives $\nu$ is **classical** if it includes $\wedge : (x,y) \mapsto min(x,y)$, $\vee : (x,y) \mapsto max(x,y)$, and $\neg : x \mapsto 1-x$, and for all $u \in \nu$, we have $u(\{0,1\}^{Ari(u)}) \subseteq \{0,1\}$.

[6] This is the metric defined by $\varepsilon(x,y) = 1$ if $x = y$, and 0 otherwise. Any countable first-order structure augmented with the discrete equalness relation becomes a continuous structure in a canonical way. There is no constraint on functions and relations, because every function on a discrete space is continuous.

[7] This just means we allow the equalness relation to take on values other than 0 and 1. We could change this to allow *all* relations by using a "normalized" equalness relation $\hat{\varepsilon}(x,y) = \frac{2^{-d(x,y)} - 2^{-\partial}}{1-2^{-\partial}}$, where $\partial$ is the diameter of the structure. But this would have negative effects elsewhere, because it is possible to have a computable metric structure whose diameter is not computable (imagine constructing a space whose diameter is some lower computable but not computable real)

You can probably guess what's coming, from these definitions:

**Observation 16.** *Assuming we work over a classical set of connectives, if $\mathcal{M}$ and $\mathcal{N}$ are discrete structures, then $\mathcal{Q} \subseteq \mathcal{M}^{<\omega} \times \mathcal{N}^{<\omega}$ is a quasi back-and-forth set if and only if it is a back-and-forth set.*[8]

**Theorem 17.** *Let $\mathcal{M}$ and $\mathcal{N}$ be two complete, separable continuous structures. There is a quasi back-and-forth set for $\mathcal{M}$ and $\mathcal{N}$ if and only if $\mathcal{M} \cong \mathcal{N}$*

This theorem would be interesting in it's own right. But what makes it beautiful is that not only is the theorem a generalization of the classical theorem about the back-and-forth property, but its proof itself is a natural generalization of the proof of the classical theorem. This is evidence for the following heuristic conjecture: *every theorem from classical (first-order) structure theory has a natural analogue in continuous structure theory, which reduces to the original theorem when restricted to discrete structures.*

*Proof.* [$\Longleftarrow$] If $f : \mathcal{M} \cong \mathcal{N}$, $\{(\bar{p}, \bar{q}) : f(\bar{p}) = \bar{q}\}$ is a quasi back-and-forth set.

[$\Longrightarrow$] $\mathcal{Q}$ is a partial order under the relation $\leq^*$ defined to be the reflexive and transitive closure of the relation: $(\bar{p}, \bar{q}) \leq (\bar{p}'c, \bar{q}'d)$ if $(\bar{p}'c, \bar{q}'d)$ could be obtained from $(\bar{p}, \bar{q})$ by a [FORTH] step, or by a [BACK] step. $(\mathcal{Q}, \leq^*)$ is a well-founded partial order of height $\omega$ and no dead ends.

A branch of $\mathcal{Q}$ is a sequence of pairs of tuples $((p_0^k, p_1^k, ..., p_{k-1}^k), (q_0^k, q_1^k, ..., q_{k-1}^k)) = (\bar{p}^k, \bar{q}^k)$. Define the limit[9] of a branch by $p_i^\infty = \lim_k p_i^k$, which exists because

$$\varepsilon(p_i^{k+j}, p_i^k) \geq \prod_{l<j} \varepsilon(p_i^{k+l+1}, p_i^{k+l}) \geq \prod_{l<j} 2^{-2^{-(k+l+1)}} > 2^{-2^{-k}}$$

and $q_i^\infty = \lim_k q_i^k$, which exists because

$$\varepsilon(q_i^{k+j}, q_i^k) \geq \prod_{l<j} \varepsilon(q_i^{k+l+1}, q_i^{k+l}) \geq \prod_{l<j} 2^{-2^{-(k+l+1)}} > 2^{-2^{-k}}$$

By [QUASI ISOMORPHIC DIAGRAMS] and continuity of all relation and function symbols, we have that for any atomic formula $\phi(x_{i_0}, ...., x_{i_{s-1}}) = \phi_m^{at}$, we have

$$\varepsilon\big(\phi^{\mathcal{M}}(p_{i_0}^\infty, ..., p_{i_{s-1}}^\infty), \phi^{\mathcal{N}}(q_{i_0}^\infty, ..., q_{i_{s-1}}^\infty)\big) = \lim_k \varepsilon\big(\phi^{\mathcal{M}}(p_{i_0}^k, ..., p_{i_{s-1}}^k), \phi^{\mathcal{N}}(q_{i_0}^k, ..., q_{i_{s-1}}^k)\big)$$

$$\geq \lim_k \varepsilon\big(\phi^{\mathcal{M}}(p_{i_0}^k, ..., p_{i_{s-1}}^k), D^{\mathcal{M}}(\bar{p}^k)(m)\big)\varepsilon\big(D^{\mathcal{M}}(\bar{p}^k)(m), D^{\mathcal{N}}(\bar{q}^k)(m)\big)\varepsilon\big(D^{\mathcal{N}}(\bar{q}^k)(m), \phi^{\mathcal{N}}(q_{i_0}^k, ..., q_{i_{s-1}}^k)\big)$$

$$\geq \lim_k \left(2^{-2^{-(k+2)}} 2^{-2^{-(k+1)}} 2^{-2^{-(k+2)}}\right) = 1$$

[8] A (classical) back-and-forth set is a set with the back-and-forth property. See Definition III.3.1 of Montalbán's "Computable Structure Theory (Draft)[10]

[9] By assumption $\mathcal{M}$ and $\mathcal{N}$ are complete metric spaces, so Cauchy sequences converge.

From this we see that for any branch of $\mathcal{Q}$, we can define a map $\tilde{f} : p_i^\infty \mapsto q_i^\infty$ which preserves the truth of all atomic formulas. This map will necessarily be 1-Lipschitz as a map $\mathcal{M} \supseteq \{p_i^\infty : i \in \omega\} \to \{q_i^\infty : i \in \omega\} \subseteq \mathcal{N}$ (as the equalness relation, our proxy for distance, is among the atomic formulas), and has 1-Lipschitz inverse (by symmetry). If $\{p_i^\infty : i \in \omega\}$, $\{q_i^\infty : i \in \omega\}$ are dense in $\mathcal{M}$, $\mathcal{N}$ respectively, a basic theorem of analysis tells us $\tilde{f}$ extends uniquely to a bi-Lipschitz map $f : \mathcal{M} \to \mathcal{N}$, and since $f$ preserves the truth values of all atomic formulas on a dense set of points, and atomic formulas are uniformly continuous, $f$ preserves the truth values of all atomic formulas everywhere. That is, $f$ is an isomorphism.

It thus suffices to show $\mathcal{Q}$ has a branch whose limit has $\bar{p}^\infty$ and $\bar{q}^\infty$ dense in $\mathcal{M}$ and $\mathcal{N}$ respectively. But we can do this: pick countable bases for $\mathcal{M}$ and $\mathcal{N}$ and enumerate them by $(U_i : i \in \omega)$ and $(V_i : i \in \omega)$, where each open set in each base is repeated infinitely many times[10] in the enumeration of the base. Construct a branch of $\mathcal{Q}$ as follows: start with $\bar{p}^0 = \bar{q}^0 = \emptyset$; at stage $2n$, extend $\bar{p}^{2n-1}$ to $\bar{p}^{2n}$ by adding point in $U_n$, and then choose $\bar{q}^{2n}$ by [FORTH]; at stage $2n+1$, extend $\bar{q}^{2n}$ to $\bar{q}^{2n+1}$ by adding point in $V_n$, and then choose $\bar{p}^{2n+1}$ by [BACK]. □

Our concern now is what is needed to perform this back-and-forth construction *effectively*. This will be useful for establishing the computable categoricity of spaces like Cantor space. Constructing the isomorphism, given a path through a back-and-forth set $\mathcal{Q}$ which is dense in both $\mathcal{M}$ and $\mathcal{N}$, is fairly easy: look along the path until you find a $p_i^k$ sufficiently equal to the $p \in \mathcal{M}$ you want to map, and then approximate the value of $f(p)$ by $q_i^k$. So the difficulty in effectivizing this construction of an isomorphism lies with being able to construct a quasi back-and-forth set $\mathcal{Q}$ for which we can effectively find a path.

In the last chapter, we gave a definition of u.r.i.c.e. relations on $\mathcal{M}^n$, which we can think of as presentation-independent effectively enumerable relations. A presentation-independent effective open set, then, ought to be a set whose continuous characteristic function is u.r.i.c.e.. In this case, we have two particular copies of $\mathcal{M}$ and $\mathcal{N}$, say $\mathcal{A}$ and $\mathcal{B}$, and we are concerned with a presentation-dependent problem (finding $\mathcal{Q} \subseteq \mathcal{A}^{<\omega} \times \mathcal{B}^{<\omega}$ with certain properties). So we should define presentation-dependent effective open sets.

**Definition 57.** Given copies $\mathcal{A}$ and $\mathcal{B}$ of two continuous structures $\mathcal{M}$ and $\mathcal{N}$ (possibly in two different languages), and an oracle $z \in 2^\omega$, a relation $R \subseteq \mathcal{A}^{<\omega} \times \mathcal{B}^{<\omega} := \sqcup_{n<\omega}\mathcal{A}^n \times \sqcup_{m<\omega}\mathcal{B}^m$ is **z-effectively open** if, we have, uniformly in $n, m$, a $z$-computable lower name for the continuous characteristic function $\chi_{R|_{\mathcal{A}^n \times \mathcal{B}^m}}$ of the restriction $R|_{\mathcal{A}^n \times \mathcal{B}^m}$ of $R$ to $\mathcal{A}^n \times \mathcal{B}^m$.

**Theorem 18.** *Let $\mathcal{A}$ and $\mathcal{B}$ be two copies of the same continuous structure. Then an isomorphism $f : \mathcal{M} \to \mathcal{N}$ is computable from oracles for $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$ if there is a $\big(\mathcal{D}(\mathcal{A}) \oplus \mathcal{D}(\mathcal{B})\big)$-effectively open quasi back-and-forth set $\mathcal{Q} \subseteq \mathcal{A}^{<\omega} \times \mathcal{B}^{<\omega}$.*

[10] Essentially, this boils down to the $p_i^k$ and $q_i^k$ being "tail-dense" in $\mathcal{M}$ and $\mathcal{N}$ respectively. We need this to ensure denseness of the limit sequences $p_i^\infty$ and $q_i^\infty$.

*Proof.* By the above, we just need to show that, with oracle $z = \big(\mathcal{D}(\mathcal{A}) \oplus \mathcal{D}(\mathcal{B})\big)$, we can construct a path through $\mathcal{Q}$ whose points in the respective tuples of the sequence are tail-dense in $\mathcal{A}$ and $\mathcal{B}$. Since $\mathcal{Q}$ is $z$-effectively open, we can easily find a tail-dense enumeration of points in $\mathcal{M}$ and $\mathcal{N}$. Proceed as follows: given $\bar{p}^k$ and $\bar{q}^k$, if $k = 2j$, let $p_k^{k+1} = q_j^{\mathcal{A}}$, the $j$-th point in our tail-dense sequence $(q_j^{\mathcal{A}} : j \in \omega)$ in $\mathcal{A}$, and search for a $\bar{q}^{k+1}$ with $\varepsilon(q_i^{k+1}, q_i^k) > 2^{-2^{-(k+1)}}$ for all $i < k$ for which $(\bar{p}^k p_k^{k+1}, \bar{q}^{k+1}) \in \mathcal{Q}$. We can do this, because $\mathcal{Q}$ is $z$-effectively open, as is the set of $\bar{q}^{k+1}$ with $\varepsilon(q_i^{k+1}, q_i^k) > 2^{-2^{-(k+1)}}$ for all $i < k$. For $k = 2k + 1$, interchange the roles of $\mathcal{A}$ and $\mathcal{B}$. Since $\mathcal{Q}$ is a quasi back-and-forth set, this procedure will build an infinite path, and we have constructed it so that its limit is dense in $\mathcal{A}$ and $\mathcal{B}$. By our previous argument, we can use this path to build a $z$-computable isomorphism from $\mathcal{A}$ to $\mathcal{B}$. ☐

**Definition 58.** A complete, separable continuous structure $\mathcal{M}$ is **computably categorical** if all of its computable copies are computably isomorphic (i.e., there exists a computable isomorphism between them).

**Example 6** (Application)**.** Cantor space, $2^\omega$, as a continuous space with only the equalness relation $\varepsilon(x, y) = 2^{-d(x,y)}$ (where $d(x, y) = 2^{-min\{n: x(n) \neq y(n)\}}$) is computably categorical.

*Proof.* Let $\mathcal{A}$ and $\mathcal{B}$ be two computable copies of Cantor space. By the previous theorem (specialized to the case where $\mathcal{D}(\mathcal{A})$ and $\mathcal{D}(\mathcal{B})$ are both computable), it suffices to find an effectively-open back-and-forth set $\mathcal{Q}$. Let $\mathcal{Q} := \{(\bar{p}, \bar{q}) \in \mathcal{A}^n \times \mathcal{B}^n : \varepsilon(\varepsilon(p_i, p_j), \varepsilon(q_i, q_j)) > 2^{-2^{-(n+3)}}$ for all $i, j < n\}$. It's easy to see $\mathcal{Q}$ this satisfies [QUASI ISOMORPHIC DIAGRAMS], since our language only has $\varepsilon$ as a relation symbol. To see it satisfies [FORTH], suppose $\varepsilon(\varepsilon(p_i, p_j), \varepsilon(q_i, q_j)) > 2^{-2^{-(n+3)}}$ for all $i, j < n$, and let $p_n \in \mathcal{A}$. First, note that by moving the $q_i$ to some $q_i'$ with $\varepsilon(q_i, q_i') > 2^{-2^{-(n+1)}}$ (think about only modifying the digits of each $q_i$ after the first $n+1$ digits), we can ensure that $\varepsilon(p_i, p_j) = \varepsilon(q_i', q_j')$ for $i < n$, i.e. that $\bar{p}$ and $\bar{q}'$ are in fact isomorphic as tuples (not just quasi-isomorphic). Every finite partial isomorphism of Cantor space extends to a isomorphism, so let $f$ be such an extension and let $q_n' = f(p_n)$. Then $(\bar{p}p_n, \bar{q}'q_n') \in \mathcal{Q}$. So $\mathcal{Q}$ satisfies [FORTH]. The argument for [BACK] is symmetric. Note that we did not need our construction of $\bar{q}'$ and $q_n'$ to be effective in our proof of the [FORTH] step, because all we are showing is that there is *some* extension in $\mathcal{Q}$ of $(\bar{p}, \bar{q})$ (essentially telling us that the space we are searching through to extend our path from length $n$ to length $n + 1$ is non-empty). The effectiveness comes from the fact that $\bar{Q}$ is effectively open, which can be seen by the fact that it is defined uniformly across $n$ as a finite union of preimages of effectively open sets in $[0, 1]$ under computable functions. ☐

You might be wondering about this trick involving partial isomorphisms extending to isomorphisms. In this case, it was that partial quasi isomorphisms can be "quasi-extended" to isomorphisms, meaning that (up to the finite error inherent in a partial quasi isomorphism), every partial quasi isomorphism was consistent (up to that error) with being the restriction of an isomorphism. You might expect that there is a general idea here, and in fact there is!

# Metric Scott Families

For this section, we assume we are working in the computable language $\mathcal{L}^c_\Omega(\sigma, \nu)$, where $\nu$ is a collection of connectives including all piecewise linear connectives, $\Omega$ is the universal 1-Lipschitz modulus, and the basic formulas are recursively enumerable. Note, these were the conditions we imposed to ensure the equivalence of *u.r.i.c.e.* and $\Sigma^c_1$ relations. There are other ways (such as using a different weak modulus) which can also ensure this, but we don't have an exact characterization yet. These conditions allow us to recursively enumerate the formulas in $\mathcal{L}^c_\Omega(\sigma, \nu)$.

**Definition 59.** An **effective quasi Scott family** for a $\sigma$-structure $\mathcal{M}$ is a computably enumerable family of $\Sigma^c_1$ formulas $\phi_{i,j}$ together with a computable function $a : \omega \times \omega \to \omega$ where

$$(\forall \bar{p} \in \mathcal{M}^{<\omega})(\forall j \in \omega)(\exists i \in \omega)(\phi^{\mathcal{M}}_{i,j}(\bar{p}) > 1 - 2^{-a(i,j)})$$

which satisfies that for any $\bar{p}, \bar{q} \in \mathcal{M}^n$, $j \in \omega$:

**IF** for some $i \in \omega$ we have $\phi^{\mathcal{M}}_{i,j}(\bar{p}) > 1 - 2^{-a(i,j)}$ and $\phi^{\mathcal{M}}_{ij}(\bar{q}) > 1 - 2^{-a(i,j)}$

**THEN** for some automorphism $f : \mathcal{M} \to \mathcal{M}$, $\varepsilon(p_k, f(q_k)) > 2^{-2^{-(j+1)}}$ for all $k < n$

**Intuition.** A quasi Scott family is essentially a collection of formulas which approximately isolate automorphism orbits of $\mathcal{M}$ by arbitrarily fine neighborhoods. In a structure like the real Hilbert space $\mathbb{R}^2$, there are uncountably many orbits (all concentric circles around the origin), so you could never hope to find countably many formulas exactly isolating all the automorphism orbits. But there is a natural sense in which some of the orbits are "close" to each other: in this case, they are close if they are both contained the same thin annulus. So it makes sense to say we can "approximately isolate" an orbit by a formula. The important thing here is that every point which "looks approximately like an element of the orbit $\mathcal{O}$" (according to the quasi Scott family) is close (in the sense of equalness) to the orbit $\mathcal{O}$.

**Motivation.** As indicated in the previous section, the motivation here is that we would like to figure out whether two tuples in different structures can be carried close to each other by an isomorphism of those structures, just by looking at their atomic diagrams. A quasi Scott family will let us figure out when we know enough about those tuples' atomic diagrams to determine whether there is such an isomorphism. In the case of Cantor space, there was a rather simple effective quasi Scott family: the formulas giving the approximate pairwise equalness between all the points in the tuple. Cantor space is an example of an **ultrahomogeneous** continuous structure: every partial isomorphism extends to a total isomorphism. But in continuous logic, it is perhaps more appropriate to think of it as an example of the following:

**Definition 60.** A continuous structure $\mathcal{M}$ over a finite relational language[11] is **ultra-quasi-homogeneous** if any partial $\tau$-quasi-isomorphism $f : \mathcal{M} \to \mathcal{M}$ (meaning from a finite substructure of $\mathcal{M}$ to a another finite substructure of $\mathcal{M}$) $\rho(\tau)$-quasi-extends to a total isomorphism $\hat{f} : \mathcal{M} \to \mathcal{M}$, where $\rho(\tau)$ is a non-decreasing function of $\tau$ with $\lim_{\tau\to 1}\rho(\tau) = 1$. Here a partial $\tau$-quasi-isomorphism preserves the truth values of atomic formulas up to $\tau$-equality, i.e. $\varepsilon(\varphi(f(\bar{p})), \varphi(\bar{p})) \geq \tau$ for any $\bar{p}$ from the domain of $f$ and any atomic formula $\phi(\bar{x})$ in our language. We say $g$ is a $\tau$-quasi-extension of $f$ if the domain of $g$ contains the domain of $f$, and $\varepsilon(f(p), g(p)) \geq \tau$ for every $p$ in the domain of $f$.

**Observation 19.** *Let $\mathcal{M}$ be a computable continuous structure over a 1-Lipschitz finite relational language, and suppose $\mathcal{M}$ is ultra-quasi-homogeneous with $\rho(\tau)$ lower computable. Then $\mathcal{M}$ has an effective quasi Scott family.*

*Proof Sketch.* For each tuple of points $\bar{q} = (q_{i_0}, \dots q_{i_{k-1}})$ from the countable dense set $D$ of $\mathcal{M}$, construct a formula $\varphi_{\langle i_0,\dots i_{k-1}\rangle, j}(\bar{x})$ isolating the orbit of $\bar{q}$ up to $\tau = 2^{-2^{-(j+1)}}$ equality by

$$\varphi(\bar{x}) := \bigwedge_{\psi(x_0,\dots,x_{k-1}) \text{ atomic}} \varepsilon\big(\psi(\bar{x}), [\psi^{\mathcal{M}}]([\bar{q}](j+3))(j+3)\big)$$

□

**Theorem 20** (originally proved by Ventsov c.1992 for classical first-order structures)**.** *Suppose $\mathcal{M}$ is a computable $\sigma$-structure. Then the following are equivalent:*

*(i) $\mathcal{M}$ is **effectively** $\Sigma^c_1$**-atomic**, i.e. there is an effective quasi Scott family $\{\phi_{i,j} : i, j \in \omega\}$ for $\mathcal{M}$.*

*(ii) $\mathcal{M}$ is **uniformly relatively computably categorical**. That is, there is uniform procedure which, given the diagram of a copy $\mathcal{N}$ of $\mathcal{M}$, computes an isomorphism $\mathcal{N} \to \mathcal{M}$.*

**Note.** The term "atomic" here refers to the fact that classically the Boolean algebra of definable subsets of $\mathcal{M}^n$ is an atomic Boolean algebra. In the continuous case, a computable structure being effectively $\Sigma^c_1$ atomic means that we have an effective procedure which takes a name for a tuple $\bar{p} \in \mathcal{M}^n$ and returns, for any degree of accuracy, a $\Sigma^c_1$ formula isolating the orbit of $\bar{p}$ to that degree of accuracy. If our language also includes the connective $\lim_n$, and our name for $\bar{p}$ is computable, then our procedure can also give us a $\Sigma^1_c$ formula isolating the orbit of $\bar{p}$ exactly. However, we cannot hope to have a $\Sigma^1_c$ formula exactly isolating *every* orbit, because there are only countably many such formulas but might be uncountably many orbits, hence the need for this somewhat more complicated definition of Scott family.

[11] Over a finite language, we include all atomic formulas in the atomic diagram of an $n$-tuple, not just the first $n$!. We restrict to finite languages here for simplicity. The restriction to relational languages is so that every subset of a structure is a substructure.

*Proof of (i) ⇒ (ii).* Let $\mathcal{N}$ be a copy of $\mathcal{M}$. Define a $\mathcal{D}(\mathcal{N})$-effectively open quasi back-and-forth set by

$$\mathcal{Q} := \{(\bar{p}, \bar{q}) \in \mathcal{M}^{<\omega} \times \mathcal{N}^{<\omega} : (\exists i, j \in \omega)\big(|\bar{p}| = |\bar{q}| = j \ \&\ \mathcal{M} \models_{a(i,j)} \phi_{i,j}(\bar{p}) \ \&\ \mathcal{N} \models_{a(i,j)} \phi_{i,j}(\bar{q})\big)\}$$

where "$\mathcal{M} \models_n \phi$" means $\phi^{\mathcal{M}} > 1 - 2^{-n}$, i.e. $\phi$ is at least $\tau$-true in $\mathcal{M}$ for some $\tau > 1 - 2^{-n}$.

It's clear this is $\mathcal{D}(\mathcal{N})$-effectively open (it's defined by a computable disjunction of open conditions), so suffices to show that $\mathcal{Q}$ is in fact a quasi back-and-forth set. Let $g : \mathcal{N} \to \mathcal{M}$ be an isomorphism. If $(\bar{p}, \bar{q}) \in \mathcal{Q}$ is witnessed by $\phi_{i,j}$, then $\mathcal{M} \models_{a(i,j)} \phi_{i,j}(g(\bar{q}))$ and $\mathcal{M} \models_{a(i,j)} \phi_{i,j}(\bar{p})$, so for some $f \in Aut(\mathcal{M})$, $\varepsilon(p_k, (f \circ g)(q_k)) > 2^{-2^{-(j+1)}}$ for all $k < j$. So by 1-Lipschitzness of formulas, $\varepsilon(\phi_k^{at,\mathcal{M}}(\bar{p}), \phi_k^{at,\mathcal{N}}(\bar{q})) = \varepsilon(\phi_k^{at,\mathcal{M}}(\bar{p}), \phi_k^{at,\mathcal{M}}((f \circ g)(\bar{q}))) > 2^{-2^{-(j+1)}}$ for all $k < j$! (in fact, for all $k$). Thus $\mathcal{Q}$ satisfies the property [QUASI-ISOMORPHIC DIAGRAMS].
If $(\bar{p}, \bar{q}) \in \mathcal{Q}$ is witnessed by $\phi_{i,j}$, again we have $\mathcal{M} \models_{a(i,j)} \phi_{i,j}(g(\bar{q}))$, (since $g$ is an isomorphism) so for some $f \in Aut(\mathcal{M})$, $\varepsilon(p_k, (f \circ g)(q_k)) > 2^{-2^{-(j+1)}}$ for all $k < j$. Suppose $c \in \mathcal{M}$. Then define $q'_k = (f \circ g)^{-1}(p_k)$ for $k < j$, and $d = (f \circ g)^{-1}(c)$. Notice that for $k < j$,

$$\varepsilon^{\mathcal{N}}(q'_k, q_k) = \varepsilon^{\mathcal{N}}((f \circ g)^{-1}(p_k), q_k) = \varepsilon^{\mathcal{M}}(p_k, (f \circ g)(q_k)) > 2^{-2^{-(j+1)}}$$

Most importantly, however, because an isomorphism $(f \circ g)^{-1}$ carries $\bar{p}c$ to $\bar{q}'d$, all formulas (including the ones in the Scott family) must take on the same values on $\bar{p}c$ and $\bar{q}'d$, and by definition of Scott family there must be some $i' \in \omega$ so that $\phi_{i',j+1}^{\mathcal{M}}(\bar{p}c) > 2^{-2^{-a(i',j+1)}}$, so we must have $(\bar{p}c, \bar{q}'d) \in \mathcal{Q}$. That is, $\mathcal{Q}$ satisfies [FORTH]. A similar argument shows $\mathcal{Q}$ satisfies [BACK].

We can thus obtain a $\big(\mathcal{D}(\mathcal{N}) \oplus \mathcal{D}(\mathcal{M})\big)$-effective ($= \mathcal{D}(\mathcal{N})$-effective) isomorphism by the theorem in the previous section. To see that this procedure is uniform, note that our definition of $\mathcal{Q}$ is uniform in $\mathcal{D}(\mathcal{N})$. Our procedure for taking our quasi back-and-forth set $\mathcal{Q}$ and producing a tail-dense branch is also uniform in $\mathcal{D}(\mathcal{N}) \oplus \mathcal{D}(\mathcal{M}) \equiv_T \mathcal{D}(\mathcal{N})$, as is our procedure for building an isomorphism from this branch. □

*Proof of (ii) ⇒ (i).* This is the harder direction. We need to take a uniform computational procedure, and produce a collection of syntactic objects. We glossed over the details of one direction of our argument in the proof of the equivalence between *u.r.i.c.e.* and $\Sigma_1^c$ relations in Section 1.4. We'll try now to include full details.

The idea is essentially to carefully track the convergence of the Turing functional $\Phi$ naming the isomorphism $f$, measured by the amount of information about the atomic diagram of $\mathcal{N}$ we give it, to figure out when we have approximately isolated an orbit. Recall that the atomic diagram of an $\omega$-presentation $\mathcal{N}$ only gives the values of the atomic formulas on tuples from $\mathcal{N}$'s dense subset, which is just $\omega$ (not just canonically identified with $\omega$, but *equal* to $\omega$). Also recall that as a mathematical object, our Turing functional with oracle for the diagram plugged in, $\Phi^{\mathcal{D}(\mathcal{N})}$, is a Turing functional which takes as oracle a name $[p]$ for $p$ and returns a function $\omega \to D$ which is name for $f(p)$ (where $D$ is the dense subset of $\mathcal{M}$). That is, $\Phi^{\mathcal{D}(\mathcal{N})}$ is a uniformly continuous partial function $\omega^\omega \to D^\omega$. But since $\Phi$ is a computational procedure, we can also think of $\Phi^{\mathcal{D}(\mathcal{N})}$ as a $\subseteq$-increasing sequence of partial functions $\Phi_s^{\mathcal{D}(\mathcal{N})} : \omega^s \to D^{<\omega}$, where $\Phi_s^{\mathcal{D}(\mathcal{N})}\big(([p](0), \dots, [p](s-1))\big)(i)$ is the accuracy $2^{-2^{-i}}$ approximation of $f(p)$ computed by $\Phi$ from $\mathcal{D}(\mathcal{N})$ in at most $s$ steps (if it converges).[12] For $\bar{n} \in \omega^k \subseteq \mathcal{N}^k$, we will by abuse of notation write $\Phi_s^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$ to stand for the tuple $\Big(\Phi_s^{\mathcal{D}(\mathcal{N})}\big((n_0, n_0, \dots, n_0)\big)(i), \dots, \Phi_s^{\mathcal{D}(\mathcal{N})}\big((n_{k-1}, n_{k-1}, \dots, n_{k-1})\big)(i)\Big)$, where these tuples $(n_i, n_i, \dots, n_i)$ are of length $s$. The important observation is that once we see $\Phi_s^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$ converge, we now know $f(\bar{n}) \in \mathcal{M}$ up to accuracy $\tau = 2^{-2^{-i}}$, and since $\Phi$ used only the first $s$ bits of information about $\mathcal{D}(\mathcal{N})$ in its computation, those $s$ bits of information are sufficient to determine the automorphism orbit of $\bar{n}$ up to $\tau$-equality. Why? Let's suppose to the contrary those $s$ bits were not enough to determine the automorphism orbit to this level of accuracy $\tau$. Then we could construct an alternate $\omega$-presentation of $\mathcal{M}$, $\mathcal{N}'$, for which those $s$ bits of the diagram of $\mathcal{N}$ used by $\Phi_s$ in this computation are equal to corresponding $s$ bits of the diagram of $\mathcal{N}'$, but for which $\bar{n}$ interpreted as a tuple from $\mathcal{N}'$ cannot be carried by an isomorphism to a $\tau$-neighborhood of $\bar{n}$ interpreted as a tuple from $\mathcal{N}$. But this implies that one of $f = \Phi^{\mathcal{D}(\mathcal{N})}$ or $g = \Phi^{\mathcal{D}(\mathcal{N}')}$ is not an isomorphism, for if they were both isomorphisms, then $f^{-1} \circ g$ would be an isomorphism carrying $\bar{n}$ interpreted as a tuple from $\mathcal{N}'$ to $\bar{n}$ interpreted as a tuple from $\mathcal{N}$.

We can also note that if $\bar{n} \in \omega^k$ is $\tau'$-equal to $\bar{p} \in \mathcal{N}^k$, then if $\Phi_s^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$ converges, we have also using those same $s$ bits of $\mathcal{D}(\mathcal{N})$ determined the automorphism orbit of $\bar{p}$ up to accuracy $\tau'\tau$, so this same analysis can be used to figure out how much of the atomic diagram of $\mathcal{N}$ is needed to isolate the automorphism orbit of an arbitrary $\bar{p} \in \mathcal{N}^k$.

Our idea now is to translate the information contained in those $s$-many bits of $\mathcal{D}(\mathcal{N})$ (the bits used to isolate the automorphism orbit of $\bar{n}$ to accuracy $\tau$) into a syntactic object (a formula) which expresses the same information. It's easy to do this for a fixed $\bar{n} \in \omega^k \subseteq \mathcal{N}^k$: those finitely many bits express finitely-many facts of the form $\tau_j^{lower} < \phi_j^{at,\mathcal{N}}(\bar{m}) < \tau_j^{upper}$, for some $\bar{m}$ extending $\bar{n}$.[13] Using unary 1-Lipschitz rational piecewise-linear connectives, we can can express each of these facts as $u_j(\phi_j^{at,\mathcal{N}}(\bar{m})) > 1 - 2^{-l}$ for some $l \in \omega$ common to all of the finitely many $\phi_j^{at}$[14]. Then the information in those $s$ many bits is expressed exactly

[12] We are using the convention that an oracle computation can only access the first $s$ bits of the oracle in its first $s$ steps of computation. This fits with the idea that the head of the oracle tape can only move one position left or right in a single step.

[13] The reason we need $\bar{m}$ extending $\bar{n}$ is that $\Phi_s$ might read more about diagram of $\mathcal{N}$ that just formulas involving elements of $\bar{n}$. But still, it can only read information about finitely many elements of $\mathcal{M}$, so $\bar{m}$ will be a finite tuple.

[14] The tighter the bound required for the values of the $\phi_j^{at}$, the larger $l$ must be. Note that if we allowed

by:

$$\psi(\bar{m}) := \bigwedge_j u_j(\phi_j^{at,\mathcal{N}}(\bar{m})) > 1 - 2^{-l}$$

where this conjunction (infimum) is finite. Now here's the important observation:

> If $\mathcal{M} \models_l \psi(\bar{q})$ for some $\bar{q} \in \mathcal{M}^{<\omega}$, $\bar{p} := \bar{q} \upharpoonright |\bar{n}|$ **must** be within $\tau = 2^{-2^{-i}}$ of the automorphism orbit of $f(\bar{n})$.

This is somewhat complicated, so we'll explain this observation step-by-step:

- By construction of $\psi(\bar{y})$, $\mathcal{M} \models_l \psi(\bar{q})$ **if and only if** for each $\phi_j^{at}$ mentioned in the $s$ bits of $\mathcal{D}(\mathcal{N})$ used by $\Phi_s^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$, we have $\tau_j^{upper} < \phi_j^{at,\mathcal{M}}(\bar{q}) < \tau_j^{lower}$.
- Build another $\omega$-presentation $\mathcal{N}'$ of $\mathcal{M}$ as follows. First, find a function $g : \omega \to \mathcal{M}$ with dense range such that $g(\bar{m}) = \bar{q}$. Next, define $\mathcal{D}(\mathcal{N}')$ to match $\mathcal{D}(\mathcal{N})$ **exactly** on the finite parts used in the computation of $\Phi^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$. We can do this because all of those bits are consistent with the values of $\phi_j^{at,\mathcal{M}}(\bar{q})$. Now define the rest of $\mathcal{D}(\mathcal{N}') = \bigoplus_n \mathcal{D}^{\mathcal{N}'}(0,1,...,n-1)$ not yet spoken for by setting

  $$\mathcal{D}^{\mathcal{N}'}(0,1,...,n-1)(j) := \text{ the least } \frac{r}{2^{n+2}} \text{ such that } \varepsilon\big(\phi_j^{at,\mathcal{M}}[x_m \mapsto g(m) : m < n], \frac{r}{2^{n+2}}\big) > 2^{-2^{-n+2}}$$

  .
- Note that $\Phi_s^{\mathcal{D}(\mathcal{N}')}(\bar{n})(i)$ outputs the exact same value that $\Phi_s^{\mathcal{D}(\mathcal{N})}(\bar{n})(i)$ does, since the oracles are identical on all the bits used in the computation. Thus $\Phi^{\mathcal{D}(\mathcal{N}')}(\bar{n})(i)$ is at least $\tau$-equal to $f(\bar{n}) \in \mathcal{M}^k$ (here $\bar{n}$ is thought of as an element of $\mathcal{N}$).
- But $\mathcal{N}'$ was constructed so that $\bar{p} := \bar{q} \upharpoonright |\bar{n}|$ is sent by an isomorphism $\mathcal{M} \to \mathcal{N}'$ (the inverse of the isomorphism induced by $g$) to $\bar{n}$ interpreted as an element of $\mathcal{N}'$. But then composing that isomorphism $\mathcal{M} \to \mathcal{N}'$ with the isomorphism $\mathcal{N}' \to \mathcal{M}$ given by $\Phi^{\mathcal{D}(\mathcal{N}')}$ gives an automorphism $\mathcal{M} \to \mathcal{M}$ sending $\bar{p} := \bar{q} \upharpoonright |\bar{n}|$ to something at least $\tau$-equal to $f(\bar{n})$. That is, $\bar{p} := \bar{q} \upharpoonright |\bar{n}|$ is $\tau$-equal to some point in the automorphism orbit of $f(\bar{n})$.

Amplifying this a bit: we can conclude that for any $\bar{p} \in \mathcal{M}^k$, if $\mathcal{M} \models_l \sup_{\bar{z}} \psi(\bar{p}^\frown \bar{z})$, then $\bar{p}$ is within $\tau$ of the automorphism orbit of $f(\bar{n}) \in \mathcal{M}^k$. Conversely, if $\bar{p}$ is *in* the automorphism orbit of $f(\bar{n})$, then since $\mathcal{M} \models_l \sup_{\bar{z}} \psi(\bar{n}^\frown \bar{z})$, and an isomorphism $\mathcal{N} \to \mathcal{M}$ carries $\bar{n}$ to $\bar{p}$, we also know $\mathcal{M} \models_l \sup_{\bar{z}} \psi(\bar{p}^\frown \bar{z})$. That is, the set $\{\bar{x} \in \mathcal{M}^k : \mathcal{M} \models_l \sup_{\bar{z}} \psi(\bar{x}^\frown \bar{z})\}$ contains the automorphism orbit of $f(\bar{n})$, and is contained with a $\tau$-neighborhood of the automorphism

ourselves formulas with n-Lipschitz modulus for any $n$, we could set $l = 1$, and we would have had no need for the function $\alpha(i, j)$ in our definition of effective Scott family.

orbit of $f(\bar{n})$.

One final modification before we continue is get a formula $\varphi(\bar{x})$ for which $\{\bar{x} : \mathcal{M} \models_l \varphi(\bar{x})\}$ is ensured to contain not only the automorphism orbit of $f(\bar{n})$, but an open $\tau$-neighborhood of the automorphism orbit. It will become clear exactly why we want this, but the basic idea is that we want to ensure that for any degree of accuracy, the open sets determined by the formulas in our Scott family for that degree of accuracy actually cover all of $\mathcal{M}^{<\omega}$. If we don't perform this modification, it could be possible that we cover only a dense set of orbits without covering all of $\mathcal{M}^{<\omega}$.[15] We define our formula $\varphi(\bar{x})$ to be

$$\varphi(\bar{x}) := \sup_{\bar{y}} \sup_{\bar{z}} \Big( \psi(\bar{y}^\frown \bar{z}) \wedge u(\varepsilon(\bar{x}, \bar{y})) \Big)$$

Where $u : [0,1] \to [0,1]$ is a unary 1-Lipschitz continuous logical connective defined so that $u(t) > 1 - 2^{-l}$ if and only if $t > \tau$.[16] We now have, with this modified formula, that:

- $\{\bar{x} \in \mathcal{M}^k : \mathcal{M} \models_l \varphi(\bar{x})\}$ is contained in a $\tau^2$ neighborhood of the orbit of $f(\bar{n})$.
- $\{\bar{x} \in \mathcal{M}^k : \mathcal{M} \models_l \varphi(\bar{x})\}$ contains a $\tau$-neighborhood of the orbit of $f(\bar{n})$.

We have one more trick up our sleeves. We have so far a quantifier-free formula $\varphi(\bar{x})$ and $l \in \omega$ *for a particular* $\bar{n} \in \mathcal{N}^k$ and *a particular* $\tau < 1$. But we still need to find a *uniform* way of doing this. The trick now is to notice that we made **no** assumptions about $\mathcal{N}$, except that it was isomorphic to $\mathcal{M}$. So everything we just proved is still true if we take $\mathcal{N} = \mathcal{M}$. Given that the diagram of $\mathcal{M}$ is computable, we can now, uniformly in $i, j \in \omega$, define $\varphi_{i,j}$ and $\alpha(i,j)$ to be the $\varphi(\bar{x})$ and $l \in \omega$ arising from our construction for $\bar{n} = \bar{p}_j$ and $\tau = 1-2^{-(i+3)}$.[17] We can see each because of these formulas is $\Sigma_1^c$ (in fact, $\Sigma_1$), and this family is recursively enumerable because we can recursively enumerate triples $(j, s, i)$ for which $\Phi^{\mathcal{D}(\mathcal{M})\restriction s}(p_{jk})(i) \downarrow$ for $k < |\bar{p}_j|$ (which is all that is needed for this construction). Let's verify that this does indeed form a quasi Scott family:

- For any $\bar{p} \in \mathcal{M}^{<\omega}$ and $i \in \omega$, for some $j \in \omega$, $\bar{p}$ will belong to an open $(1 - 2^{-(i+1)})$-neighborhood of $f(\bar{p}_j)$, since the the $\bar{p}_j$ are dense. But then by construction, since $\{\bar{x} : \mathcal{M} \models_{\alpha(i,j)} \varphi_{i,j}(\bar{x})\}$ contains a $(1 - 2^{-(i+1)})$-neighborhood of the orbit of $f(\bar{p}_j)$, we have $\phi_{i,j}^{\mathcal{M}}(\bar{p}) > 1 - 2^{-\alpha(i,j)}$.
- On the other hand, if $\phi_{i,j}^{\mathcal{M}}(\bar{p}) > 1 - 2^{-\alpha(i,j)}$, then $\bar{p}$ is contained within an open $\tau^2$-neighborhood of the automorphism orbit of $f(\bar{p}_j)$, so for some automorphism $g$ of $\mathcal{M}$, $\varepsilon(\bar{p}, g(f(\bar{p}_j))) > \tau^2 = (1 - 2^{-(i+1)})^2 = 1 - 2 * 2^{-(i+1)} + (2^{-(i+1)})^2 > 1 - 2^{-i}$.

[15] In a similar way that it is possible for an open cover of $\mathbb{Q}$ to not cover $\mathbb{R}$.

[16] We can explicitly define $u(t) := max\{1, t + (1 - 2^{-l}) - \tau\}$.

[17] We need to choose $\tau = 1-2^{-(i+3)}$ to account for the fact that we only guaranteed $\{\bar{x} \in \mathcal{M}^k : \mathcal{M} \models_l \phi(\bar{x})\}$ contains a $\tau^2$ neighborhood of the automorphism orbit of $f(\bar{n})$, and for the fact that if two tuples $\bar{p}_1$ and $\bar{p}_2$ are guaranteed to be within $\tau'$ of the automorphism orbit of $f(\bar{n})$, they are only guaranteed to be withing $\tau'^2$ of each other. So we chose $\tau$ so that $\tau^4 > 1 - 2^{-(i+1)}$.

□

Please forgive the length of this proof! In a trade-off between brevity and explanation, we chose to err on the side of too much explanation. To reiterate the basic idea: whenever this Turing functional $\Phi^{\mathcal{D}(\mathcal{M})}$ has computed for enough time to find a $\tau$-approximation to $f(\bar{p})$, the finite amount of information about $\mathcal{D}(\mathcal{M})$ used in that computation is sufficient to isolate the orbit up to $\tau$-equality, and that finite amount of information can be expressed as a formula, which is more-or-less just the conjunction of all the facts about the atomic formulas used in the computation. The worry we had was that perhaps $\Phi$ is using some special facts about the presentation of $\mathcal{M}$ to compute an approximation to $f(\bar{p})$ without actually reading all the bits of information needed to isolate the orbit to that degree of accuracy. But we proved this could not happen, because $\Phi$ has to work for *any* any copy of $\mathcal{M}$: if $\Phi$ did not actually read all the bits necessary to isolate the orbit of $f(\bar{p})$ to appropriate degree of accuracy, then $\Phi$ would not compute an isomorphism when presented with a copy of $\mathcal{N}'$ of $\mathcal{M}$ specifically constructed to fool it.

# Degree Spectra

Recall that we have computable continuous structures, but we also have non-computable $\omega$-presentations (i.e. "copies") of continuous structures. Some continuous structures have no computable copies, for example $(\mathbb{N}, 0, 1, +, *, H)$, where $H$ is a unary relation for the halting set.[18] If we want to measure the intrinsic complexity of a structure, the naïve idea would be to measure it by the least Turing degree which computes a copy of it. In this case, the complexity would be $0'$.

However, Slaman [16] and Wehner [15] independently showed that there is a structure which has no computable copy, but every non-recursive degree computes a copy of it. This tells us our naïve way of measuring the complexity of a structure does not work. Instead, we need to use the following:

**Definition 61.** Let $\mathcal{M}$ be a complete, separable continuous structure. We then define the **degree spectrum** of $\mathcal{M}$ to be the set of Turing degrees which compute a copy of $\mathcal{M}$, i.e. $\{d \in \mathbb{T} : \exists \mathcal{A} \cong \mathcal{M}, \mathcal{D}(\mathcal{A}) \leq_T d\}$.[19]

[18] This is a discrete structure, but thus also a continuous structure with the discrete equalness relation $\varepsilon(x, y) = 1$ if $x = y$ and 0 otherwise.

[19] Here $\mathbb{T}$ is the set of Turing degrees, equivalence classes of elements of $2^\omega$ under the equivalence relation of Turing equivalence. $\leq_T$ is the relation of being Turing reducible to, i.e. $x \leq_T y$ if there is an oracle Turing machine which computes $x$ given $y$ as an oracle. Turing equivalence is defined by $x \equiv_T y$ iff $x \leq_T y$ and $y \leq_T x$. We are, of course, blurring the distinction between the relation of Turing reduction on elements of $2^\omega$ and elements of $\mathbb{T}$, but this doesn't really matter since this relation is invariant under $\equiv_T$.

One thing we might wonder is: is the degree spectrum of a structure the same as the set of Turing degrees of copies of that structure? The problem is whether we can encode arbitrary information into a copy of the structure. Classically, there are trivial structures (e.g., a countable set with only equality) which only have computable copies. The degree spectrum of such a structure would be the entire set of Turing degrees $\mathbb{T}$, but in this case there is a mismatch because a copy of such of a structure cannot really encode arbitrary information.

For continuous structures, this question is a bit more subtle, because we have a lot of free choice in terms of our approximations. Take a countable set $\{p_0, p_1, p_2, \dots\}$ with just the equalness relation $\varepsilon(p_i, p_j) = 1$ if $i = j$ and $\frac{1}{\sqrt{2}}$ otherwise. In making a copy of this, even if we use atomic diagrams instead of named atomic diagrams, as we are doing, we still have a free choice about whether the $k$-th approximation to $\varepsilon(p_i, p_j)$ for $i \neq j$ will be equal to the greatest rational of the form $\frac{m}{2^{k+2}}$ less than $\frac{1}{\sqrt{2}}$, or the least such rational greater than $\frac{1}{\sqrt{2}}$. By making these choices systematically, we can code infinitely many arbitrary bits into the atomic diagram of a copy of this structure.

**Note.** If we used the discrete equalness relation, we couldn't do this, as our definition of the "atomic diagram" implies that if $\varepsilon(a, b) = 0$ (respectively 1) we must choose as our $k$-th approximation $\frac{0}{2^{k+2}}$ (respectively $\frac{2^{k+2}}{2^{k+2}}$). The same is true if we had a more complicated discrete structure with all relations being $\{0, 1\}$-valued. We have set up our definition of "atomic diagram" so that the atomic diagram of a copy of a classical first-order structure interpreted as a continuous structure is Turing equivalent to its classical atomic diagram, and uniformly so. Our worry here is not about whether we are faithfully generalizing the classical case, but whether our particular faithful generalization is the right one. In particular, it seems strange that by simply modifying the equalness relation of a structure by applying the computable linear bijection $\tau \mapsto (1 - \frac{1}{\sqrt{2}})\tau + \frac{1}{\sqrt{2}}$ we completely change whether we can encode arbitrary information into the atomic diagram of a copy of that structure.

Our concern relates to the following theorem about classical structures, due to Julia Knight:

**Theorem 21** (Knight [8]). *Suppose a Turing degree $d$ computes a copy of a non-trivial structure $\mathcal{M}$. Then there is a copy of $\mathcal{M}$ whose atomic diagram is of Turing degree $d$.*

Here a structure is **trivial** if we can find finitely many points such that every permutation fixing those points is an automorphism. (An example would be a countable set just with equality and finitely many named constants). It turns out this theorem is a precise characterization of trivial structures: a structure is trivial if and only if for some Turing degree $d$ which computes a copy of $\mathcal{M}$, $\mathcal{M}$ has no copy of Turing degree $d$.

Our previous example with the equalness relation taking values in $\{\frac{1}{\sqrt{2}}, 1\}$ seems like it ought to be trivial, but this equivalence does not seem to hold. This might seem unconcerning, because an analogue of Knight's theorem *does* hold for continuous structures (for a somewhat trivial reason), it just no longer characterizes trivial structures. But this is an indication that we might we using the wrong definition somewhere. The point of Knight's theorem is lost if our definition of $\omega$-presentation allows us include auxiliary information totally unrelated to the structure itself.

The ideal solution, which we will not pursue now, would be to define a continuous model of oracle computation. Consider our named atomic diagrams. To a certain extent, modifying the approximations by some tiny amounts still within bounds shouldn't matter, since these are just approximations to the values of formulas anyway. You might imagine two people, with the same "copy" of a continuous structure, giving different answers for their approximations to the equality of points in the dense set, even if they are given the dense set in the same order. Our problem only arises when we have knowledge that the person making the measurements is hiding information in a systematic way by the way they choose to make their approximations. But this hidden information shouldn't really be considered part of the data: it should be considered measurement noise. In the case of classical structures, there is no way to hide information in measurement noise, so this problem doesn't arise. By moving from named atomic diagrams to atomic diagrams, we cut down on the amount of information that can be hidden in measurement noise from arbitrarily large finite amounts of information per step of the approximation to to a single bit per step of the approximation. But it would be better if we could just use our simpler definition of named atomic diagrams, and force computations not to take into account things that shouldn't matter, like which primes divide the denominator of the rational approximation.

One potential way to ensure this would be to require that the computation give the exact same result for all oracles which are the same up to measurement noise. The problem with this, however, is that "being the same up to measurement noise" is not a transitive relation. Instead, we want our computations to depend *uniformly continuously* on the input. We'd like to think of an oracle as a sequence $(a_n : n \in \omega) \in [0, 1]^\omega$, and an oracle computation as giving a continuous function which depends continuously on this oracle sequence. An ideal model of such computations would involve only continuous operations on the oracle and the input. In fact there is a candidate model of universal analogue computation based on Claude Shannon's "General Purpose Analog Computer". A paper by Bournez, Campagnolo, Graça, and Hainry have reinterpreted Shannon's model in [3] to give a Turing complete model of analog computation (in the sense that it can represent all Turing computable functions in $\mathbb{N}$ using analog approximations to step functions).

It might help to make a physical analogy here, in the spirit of the Church-Turing thesis. Imagine you are given a sequence of metal rods as an oracle, and you have for every $k \in \omega$ a tool which can measure length to within accuracy $2^{-k}$ (assuming classical physics, so there are no obstructions to arbitrarily precise measurements). Which continuous functions can you now compute to arbitrary precision using this oracle, and which computable structures can you now compute a copy of? Towards the first question, Shannon essentially argued that differential analyzers built from a small number of units, including “multipliers”, “adders”, and “integrators”, can represent a large class of such functions, and in fact these were successful in solving ODEs in practice before digital computers were capable of doing so.[20]

One way of capturing this without moving to an analog model of computation might be to model the “randomness” of measurement noise. We could use a model like classical register machines, augmented with a non-deterministic measurement operation which takes values $k$ and $j$ from two registers, finds the length of the $j$-th rod up to error $2^{-k}$, and outputs that approximation $\frac{m}{2^k}$ into a third register. We might imagine flipping a coin to decide between $\frac{m}{2^k}$ and $\frac{m+1}{2^k}$ if the end of the rod seems to fall partway between two marks on a ruler. But it's important that we think of this operation as being non-deterministic, rather than random. It might seem reasonable to define a continuous function $f$ to be computable from this oracle of metal rods if there is an augmented register machine $M$ and a 1-random (or sufficiently random) sequence of choices of under/over for the error of the approximations, such that $M$ computes a name for $f$ when the measurement operation under/over approximates according to that random sequence. If the measurements over/under approximate according to a random sequence, surely there is no information to be gleaned from this, right? Wrong! If we use this definition, then from an oracle all of whose rod-lengths are $\frac{1}{\sqrt{2}}$ we would be able to compute every 1-random, and there are 1-randoms of arbitrary high Turing degree.

OK, perhaps you think we can fix this by changing the existential quantifier over 1-random sequences to a universal quantifier: we want the function to be computable by a fixed augmented register machine for every 1-random choice of errors. This will work fine for capturing whether a fixed function is computable or not.[21] But then we run into the question of whether it gives the right answer for which continuous structures have computable copies relative to a given oracle. This is a more subtle problem, and it has to do with the uniformity of Theodore Slaman's construction in [16]: there is in fact a *uniform* procedure for taking a non-computable binary sequence and obtaining a copy of the Slaman-Wehner structure (which is not computable). Using this fact, we can see that there is a fixed augmented register machine with the oracle consisting of rods all of whose lengths are $\frac{1}{\sqrt{2}}$ such that for *any*

[20]They were used to do ballistics calculations on battleships, taking into account the movement of the ship. One of the many examples showing that pure mathematics can have ethical implications downstream.

[21]This actually needs proof, but it should be a straightforward generalization of the classical theorem that every relation $R$ on $\omega$ computable from each of a positive measure set of oracles is computable. The idea behind the classical proof is this: there are only countably many computational procedures, and a countable union of measure zero sets of measure zero, so some fixed machine must compute the relation for a positive measure set of oracles. Now pick some basic open set of oracles where measure $> \frac{3}{4}$ of them compute $R$ via that machine. To figure out whether $n \in R$, use majority rule: compute using all oracles in that basic open set in parallel, and as soon as you see $> \frac{1}{2}$ agree on whether $n \in R$ or $n \notin R$, go with the majority answer. Since $> \frac{1}{2}$ of the oracles compute $R$, $> \frac{1}{2}$ must give the same answer about whether $n \in R$, and it must be the correct answer, since $< \frac{1}{2}$ of them compute the wrong answer. The modification for computing a continuous function would be to replace the questions about whether $n \in R$ with questions about whether $q_1 < R(c) < q_2$ for pairs of rational numbers $q_1 < q_2$.

1-random sequence of choices for over/under approximation, the register machine computes a copy of the Slaman-Wehner structure. But we shouldn't be able to do this, as this oracle is computationally trivial, while the Slaman-Wehner structure is not.

All this is to say that randomness is wrong notion to use here. We instead want to say that for *any* choice of under/over approximation at all (i.e., for any possible errors in measurement), a fixed machine computes the function or computes a copy of the structure. More specifically, for a continuous structure or function to be computable from an oracle is for there to be a fixed augmented register machine such that for any possible choice of error, that register machine computes some copy of that structure or some name for that function. We could think of the choice of error as a choice of path in the tree of possible execution traces of a non-deterministic program, so this could be rephrased as saying that the augmented register machine with a given oracle computes a structure (or function) if for every execution trace of the program given that oracle it computes a copy of that structure (or name for that function).

The only worry about this definition is that it involves a universal quantifier over $2^\omega$. Classically, the ternary relation of "$R \subseteq \omega$ is computed by Turing machine $M$ relative to oracle $X$" is arithmetic (it's not hard to write down a $\Pi^0_2$ definition). But we have here what looks to be a $\Pi^1_1$ definition of the corresponding ternary relation in the continuous setting "$R \subseteq \omega$ is computed by Turing machine $M$ relative to continuous oracle $X$", since we need to quantify over an uncountable set of possible errors. It's not as bad as it looks, however. If we are clever, we can give it an arithmetic definition by noting that it is equivalent to something like "for any input $x$ and error bound $2^{-(k+1)}$, and any two finite strings of errors of sufficient length, the computation of $f(x)$ up to error $2^{-(k+1)}$ via machine $M$ with the given oracle computes values which are within $2^{-k}$ of each other along the two execution paths given by those two different strings of errors". The reason we can do this is because any particular halting computation can only make finitely many requests from the oracle.

It's worth noting that while this trick works for continuous functions/relations, it does not work for continuous structures. The intuitive reason we should expect you really need a quantifier over $2^\omega$ to capture the relation "A copy of the structure $\mathcal{M}$ is computed by Turing machine $M$ relative to oracle X" is that to check whether what we compute is really a copy of $\mathcal{M}$, we need to know there exists an isomorphism between $\mathcal{M}$ and the structure whose diagram we are computing, but even classically this cannot be done in a $\Sigma^0_\alpha$ way for any $\alpha < \omega_1^{CK}$. This follows from work in a paper by Fokina, Friedman, Harizanov, Knight, McCoy, and Montalbán *Isomorphism Relations on Computable Structures*[6]. Here they construct a family of computable trees $T_\alpha$ for each $\alpha < \omega_1^{CK}$ such that the equivalence relation of a being isomorphic to $T_\alpha$ restricted to computable trees is not computable relative to the $\alpha$-th iterate of the halting set. If our ternary relation had a $\Sigma^0_\alpha$ definition, then by quantifying over Turing machines $M$ and setting $X = 0$, we would have a $\Sigma^0_\alpha$ definition of being a computable structure isomorphic to $\mathcal{T}_\alpha$, a contradiction. So it's no worry to us that we end up with a $\Sigma^1_1$ definition of our ternary relation for continuous structures too. Our ternary relation can just say: "Turing machine M computes a continuous structure $\mathcal{M}'$ relative to continuous oracle $X$ and there exists a (classical) oracle $Z$ and a Turing functional $\Phi$ relative to $Z$ such that $\Phi$ computes an isomorphism $\mathcal{M}' \to \mathcal{M}$". The relation of a given function being an isomorphism if $\Pi^0_1$, and the relation of a Turing machine $M$ computing some continuous structure or other from a continuous oracle $X$ is $\Pi^0_2$, by the argument above, so this gives a $\Sigma^1_1$ definition (the only second order quantifier is an existential quantifier over classical oracles $Z$).

We haven't made much progress here besides laying out definitions, but hopefully others can work upon this basis. We expect that the "no two cones" theorem (the degree spectrum of a structure cannot be the union of two cones with neither one contained in the other) should also hold for continuous structures. The basic difficulty is that there is a difference between what we might call representations and presentations of a continuous structure.

**Definition 62.** An $\omega$-**representation** of a continuous structure $\mathcal{M}$ as an equalness relation $\varepsilon$ on $\omega$ together with a uniformly continuous function $\bar{\omega}^n \to \bar{\omega}$ for each $n$-ary function symbol and a uniformly continuous function $\bar{\omega} \to [0, 1]$ for each $n$-ary relation symbol (where $\bar{\omega}$ is the completion of $\omega$ under this equalness relation) such that $\bar{\omega}$ with these functions/relations is isomorphic to $\mathcal{M}$.

Another way of thinking about it is that an $\omega$-representation of $\mathcal{M}$ is a structure obtained by pulling back the continuous structure on $\mathcal{M}$ to $\omega$ via a map $\omega \to \mathcal{M}$ with dense range. While a continuous structure can be presented as a discrete structure by its atomic diagram, there are usually many such presentations corresponding to each representation of a continuous structure, corresponding to the many ways to present a function on $\bar{\omega}$. We don't see a straightforward way to apply the classical analysis of degree spectra to continuous structures by considering their representations or presentations. Thinking about representations for a moment, to represent a uniformly continuous function $\bar{\omega}^n \to \bar{\omega}$, we can naturally replace this with a uniformly continuous function $\omega^n \to \bar{\omega}$ (since the extension from $\omega^n$ to $\bar{\omega}^n$ is determined uniquely). $\bar{\omega}$ can be naturally identified with a sort of "quotient" of a $\Pi^0_1$ subset of $\omega^\omega$ by the $\Pi^0_1$ similarity relation $x \sim y$ iff $\forall i \varepsilon(x(i), y(i)) \geq 2^{-2^{-i}}$. So there may be some idea to apply classical results about degree spectra by thinking of a function $f : \omega \to \bar{\omega}$ as an "similarity class" (if we can make sense of such things) of functions $g : \omega \times \omega \to \omega$ under the similarity relation $g_1 \sim g_2$ iff $\forall i \forall j \varepsilon(g_1(i, j), g_2(i, j)) \geq 2^{-2^{-i}}$. We probably should not replace this universal quantifier with a limit to get an equivalence relation, because this seems to wash away too much of the computational structure we want to observe. But perhaps thinking about the degree spectra of classical quotient structures can help here.

# Chapter 3

# Hilbert Spaces and Observations

It's worth at least discussing an alternative method of rigorously introducing continuity into our underlying logic, as mentioned in the introduction. There are no new mathematical results in this chapter, just philosophical discussion of an expository nature. Instead of thinking of expanding the set of truth values from $\{0, 1\}$ to $[0, 1]$, and continuing to think of properties as deterministic functions on a space of objects given by formulas, we can instead think of properties as inherently non-deterministic, properly being valued in something like a space of probability distributions. Let's go over a line of thinking which might lead you to this approach.

Our modern conception of logic can largely be traced back to Gottlob Frege's *Begriffsschrift* and *Grundgesetze der Arithmetik*. One of his principles is what we might think of as a duality between sets and properties. Given any property $P$, we have a corresponding set $\{x : P(x)\}$[1], and conversely any set determines the property of being a member of that set. Frege thought of properties (or in a more direct English translation of his terminology, "concepts") as functions from objects to truth values.

Let's focus just on unary properties for a moment (rather than more general properties which relate multiple objects). If $X$ is the collection of all objects, then the set of functions $X \to \{0, 1\}$ is the collection of all unary properties. These properties form an atomic Boolean algebra under the operations $P \vee Q := \max(P, Q)$, $P \wedge Q := \min(P, Q)$, and $\neg P := 1 - P$. The atoms of this Boolean algebra, which are characterstic functions of singletons, are in one-to-one correspondance with objects.

However, if we reflect on how we actually apply logic to real problems, we notice that what we count as an "object" varies from context to contex. We might think of a cities as objects, or tables and chairs as objects, or atoms as objects, or electrons as objects, depending on context. There are also paradoxes such as the ship of Theseus which challenge our naive

[1]This idea turns out to have some problems, namely Russell's Paradox. But this problem can be worked around by weaking the duality to apply only to subsets of a set in the von Neunmann heirarchy.

conception of identity of objects. Given this, we might have some skepticism about whether there really is a canonical way to break up the world into a collection of objects, as we seemed to be assuming when making this correspondance between the collection of unary properties and this atomic Boolean algebra. At the very least, we should probably think about what an object is, exactly.

One way to represent the non-canonicalness of breaking up the world into objects is to replace the Boolean algebra of properties with a vector space over $\mathbb{F}_2$. If we start with, say, a finite collection of atoms, we can form a vector space over $\mathbb{F}_2$ by applying the functor from sets to $\mathbb{F}_2$ vector spaces which takes the set $X$ to the vector space $\mathbb{F}_2^X$ of functions $X \to \mathbb{F}_2$ with pointwise addition. We can think of $X$ as a basis for $\mathbb{F}_2^X$, and define a nondegenerate "Hermitian form" on $\mathbb{F}_2^X$ by specifying $\langle x, y\rangle = \delta_{xy}$ for all $x, y \in X$. Subspaces given as the span of a subset of $X$ are in one-to-one correspodance with properties in the sense of Frege, and conjunction corresponds to intersection of subspaces, disjunction to sum of subspaces, and negation to othrogonal complement. The classical bottom element $\bot = \emptyset$ of the Boolean algebra corresponds to the trivial subspace $\{0\}$, and the the top element $\top = X$ corresponds to the entire space $\mathbb{F}_2^X$.

To model breaking up the world in two different ways, we might choose a different basis $Y$ for $\mathbb{F}_2^X$ which is orthonormal with respect to $\langle\cdot,\cdot\rangle$. For example, if $X = \{a, b, c, d\}$, then $Y = \{b+c+d, a+c+d, a+b+d, a+b+c\}$ is an orthonormal basis as well.[2] Subspaces given by spans of subsets of $Y$ correspond to new properties that weren't in our original atomic Boolean algebra, which arise by breaking up the world in a different way. In general, we might want to admit any (closed) subspace of $\mathbb{F}_2^X$ as a valid property. Another idea might be instead of breaking up $\mathcal{F}_2^X$ in different ways by choosing different orthonormal bases, we can quotient by a linear map to lower-dimensional space (corresponding to grouping the atoms into larger objects, which we consider to be our new atoms). We can get quite far with this, but a 1995 paper by Maria Pia Solèr *Characterization of hilbert spaces by orthomodular spaces*[14] gives a limitative result on what we can do over a field like $\mathbb{F}_2$: the only division $*$-rings[3] over which there exists a vector space $V$ with an infinite collection of vectors orthonormal with respect to an *orthomodular* Hermitian form $\langle\cdot,\cdot\rangle$ are $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ (the quaternions). Solèr defines a Hermitian inner product to be "othomodular" if for any subspace $U$ of $V$ such that $(U^\perp)^\perp = U$, $U^\perp + U = V$, where $\perp$ and $+$ are orthogonal complement and sum of subspaces defined in the usual way. Failure of orthomodularity is quite serious from a logical point of view, because it corresponds to a failure of the classical law of excluded middle, $x \vee \neg x = \top$. Intuitionist logic allows failures of excluded middle, but quantum logic, developed by Birkhoff and von Neumann to formalize the logic of quantum mechanics, has excluded middle as an axiom (it is rather distributivity of conjunction over

[2] This trick depends crucially on the dimension. If $|X| = 3$, for example, there is only one orthonormal basis up to permutation, a side effect of the quirky nature of considering "Hermitian forms" over a finite field.

[3] $\mathbb{F}_2$ is a division-$*$-ring with the only possible conjugation action $x^* = x$.

disjunction which fails).

If we believe excluded middle should hold, and we want to work with infinite spaces of objects (like points on a continuum), then Solèr's theorem more or less necessitates that if we want to continue in this direction, we need to work over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. Let's start this discussion again, but this time working over $\mathbb{C}$.

Given a finite set $X$, we can form a Hilbert space $H(X)$ with orthonormal basis $X$. The underlying set of $H(X)$ can be considered to consist of elements of the form $\sum_{x\in X}\alpha_x x$ with $\alpha \in \mathbb{C}$, and the inner product is given by $\langle \sum_{x\in X}\alpha_x x, \sum_{x\in X}\beta_x x\rangle = \sum_{x\in X}\bar{\alpha}\beta$. Every subset of $X$ corresponds to a closed subspace of $H(X)$. Intersections of subsets correspond to intersections of the corresponding subspaces, and unions of subsets correspond to sums of subspaces. In this way, the Boolean algebra of subsets of $X$ (think, unary relations on $X$) can be represented as a subalgebra of the algebra of subspaces of $H(X)$.

To properly generalize this idea, we should think of $H(X)$ as a space of complex measures on $X$. Given $f = \sum_{x\in X}\alpha_x x \in H(X)$, we can obtain a real measure on $X$ by $\mu_f(Y) = \sum_{y\in Y}\bar{\alpha}_y\alpha_y = \sum_{y\in Y}|\alpha_y|^2 = ||Proj_{Span(Y)}f||^2$. The interesting observation here is that given another orthonormal basis $\mathcal{B}$ for $H(X)$, we can also interpret $f \in H(X)$ as a complex measure on $\mathcal{B}$. Such a change of basis corresponds to a symmetry of this space of complex measures (just ike a permutation of $X$ would correspond to a symmetry of the Boolean algebra of its subsets). In physics, this symmetry is useful to describe phenomenona like the polarization of photons. Fixing an orientation of your polarizing filter corresponds to choosing an orthonormal basis. Once you've done this, you can obtain a real probability distribution on a set of orthogonal directions of polarization. But you can just as well rotate your polarizing filter, obtaining a real probability distribution on a different set of orthogonal directions. The same element $f \in H(X)$ (thought of a the state of the photon) gives a probability distribution over outcomes for any orientation of the polarizing filter. It has been experimentally verified that photons with a given polarization (corresponding to a particular choice of complex measure) will either be blocked or not blocked by the polarizing filter with probabilities determined by the real measure obtained from that complex measure and the orthonormal basis given by the orientatation of the polarizing filter.

For general $X$, we ought to require $X$ be a topological space with a group action, and a chosen invariant measure $\mu$. For finite $X$, this group action can just be the permutation group on $X$ and $\mu$ the uniform probability measure on $X$. For something like $X = \mathbb{R}^n$ (thought of as a space of positions), the group action might be the affine group, and the invariant measure Lebesgue measure. In this setting, we can define the underlying set of $H(X)$ to be complex measures of the form $\psi_f(Y) = \int_X \mathbf{1}_Y f d\mu$, where $f \in \mathbb{L}^2(X,\mu,\mathbb{C})$[4] , the space of $\mu$-square-integrable complex-valued functions on $X$, and $\mathbf{1}_Y$ is the characteristic function of

[4]It is also useful to generalize to distributions, e.g. to allow atomic measures.

$Y$. This defines what we'll call a **complex measure** on the $\sigma$-algebra of $\mu$-measurable sets which is absolutely continuous with respect to $\mu$. The space of complex measures of this form can be made into a Hilbert space using the inner product $\langle \psi_f, \psi_g \rangle = \int_X \bar{f} g d\mu$. Note we could have equivalently taken the objects in $H(X)$ to be the functions themselves rather than the corresponding complex measures.

Given an orthonormal basis $\mathcal{B}$ of $H(X)$, each $\psi_f \in H(X)$ determines a real measure $\mu_f$ on $\mathcal{B}$ by $\mu_f(Y) = ||Proj_{Span(Y)}\mu_f||^2$. We can also think of this as giving a kind of measure on subspaces of $H(X)$. For a collection of orthogonal subspaces, the measure of their sum is the sum of their measures. It is often convenient to normalize this real measure to get a probability measure, by means of: $\hat{\mu_f}(Y) = \frac{||Proj_{Span(Y)}\mu_f||^2}{||\mu_f||^2}$.

We might ask: what is the motivation for this construction? The idea is that there are a few principles in the underlying philosophy of classical physics which seem to be at odds. The state space of the universe is presumably a continuum (in particular, connected). We also think that the result of an observation should depend continuously on the state of the universe (i.e., be a continuous function of the state of the universe). Finally, we think that there are some observations which are binary: they always have an answer, and it's either "yes" or "no" (one example is whether or not a photon passes through a polarizing grid). But if all of these are so, such a binary observation must always have the same answer: a continuous function from a continuum to a discrete set must be constant. The simplest solution to this is to broaden our notion of what an observable quantity is, so that there is a continuous path between a definite "yes" result and a definite "no" result. We can do this by interpreting a definite "yes" result or "no" result to be an atomic probability distribution over outcomes, and to situate these into a broader space of probability distributions over outcomes. We then see ex post facto that for polarization of photons (and many other physical experiments), for reasons of symmetry we should work in a space of complex measures with the 2-norm, rather than a space of real measures with the 1-norm. It's easiest to see with polarization of photons: if we used classical probability distributions, these would form a simplex whose vertices are the classical outcomes. But this space of probability distributions mathematically favors the classical outcomes we started with, while we know from experiment that there is no favored set of polarizations: any rotation of the polarizing grid will give a different set of classical outcomes. This argument about the symmetries of the 2-norm versus the 1-norm giving an a priori reason to use the form comes from Scott Aaronson's book *Quantum Computing Since Democritus*[1]. All this together seems to give good reason for the foundations of quantum mechanics to be developed over Hilbert spaces. An it is important that we have some kind of a priori justification for our foundations if we wish to have an understanding of physics which goes beyond a mere predictive model. It would be unsattisfying if the axioms of quantum mechanics were justified only by their agreement with experiment. The existence of a priori justifications for certain foundational choices also gives us an excuse to do the kind of "armchair physics" we are currently doing.

It is at this point in any discussion of the foundations of physics that we reach some serious philosophical issues having to do with how this mathematical formalism is to be interpreted physically. We have, apparently, given up on the idea that the results of observations are *determined* by the state of the universe. This runs counter to centuries of physical theory. But on closer inspection, the situation is not as absurd as it seems. While the results of observations are not determined by the state of the universe, it can still be said that observable quantities are determined by the state of the universe. The idea is that an observable quantity *is* a sort of probability distribution over a set of observation outcomes. We cannot, in a single experiment, see the entire probability distribution, but if we repeat the same experiment many times, we can deduce the probability distrubtion over outcomes for various physical situations. We also have deterministic laws which govern how this probability distribution evolves over time. So we have not lost the principle that observable quantities are deterministic. Rather we have just reinterpreted what an observable quantity is.

We should say a bit about how what we've said so far relates to the most common "practical" interpretion of quantum mechanics: the Copenhagen interpretation. In the Copenhagen interpretation, an observable quantity corresponds to a self-adjoint operator on $H(X)$. Self-adjoint operators on a Hilbert space are in bijective correspondance with spanning collections of closed orthogonal subspaces each with an associated real number. The subspaces are the eigensapces of the operator, and the real number associated to a subspace is the corresponding eigenvalue. It's important to note that, from our discussion so far, there is no need to suppose that these real numbers (the eigenvalues) have any physical meaning. They could be replaced by an abstract set of observation outcomes. The reason why physicists usually suppose outcomes of observations are real numbers is because many physical quantities like energy and momentum in a certain direction seem to naturally lie on a linear scale. But we need not assume this, and we won't. So instead of an observable quantity corresponding to a self-adjoint operator on $H(X)$, we will think of an observable quantity simply as an orthogonal collection of closed subspaces of $H(X)$ together with a one-to-one function from this collection of subspaces to a set of observation outcomes. In this context, the Copenhagen interpretation is based on the following claim (stated in rough first-approximation for a finite set of outcomes): when you perform an observation with associated subspaces $\{W_\alpha : \alpha \in I\}$ corresponding to distinct outcomes $\{o_\alpha : \alpha \in I\}$, and the state of the system is $\psi_f$, then the probability of outcome $o_\alpha$ is $\frac{||Proj_{W_\alpha}\psi_f||^2}{||\psi_f||^2}$, and if the observation has outcome $o_\alpha$, then the state of the system after the observation is $Proj_{W_\alpha}\psi_f$. In particular, if you perform the same measurement twice in a row, it will not change the state of the system as at all, because with probability 1, you will just end up applying the same projection operator onto $W_\alpha$ as the last measurement, and projections are idempotent. In general, the claim is as follows: for a subset $J \subseteq I$, the probability of an outcome in $\{o_\alpha : \alpha \in J\}$ is $\frac{1}{||\psi_f||^2}||Proj_{Span(\{W_\alpha : \alpha \in J\})}\psi_f||^2$, and if we learn precisely that our outcome is in $\{o_\alpha : \alpha \in J\}$, then after our observation the state of the system is the projection of $\psi_f$ onto $Span(\{W_\alpha : \alpha \in J\})$.[5] The "precisely" is important: learning precisely that a particle has $x$-position in the interval $(-1/2, 1/2)$ means we know for sure it is in that interval, and cannot know from the result of the experiment that it is in some strict subinterval. Although we can conclude logically that the particle must have $x$-position in $(-1, 1)$, the resulting state after the observation depends on the fact that we are projecting $\psi_f$ onto the particular subspace corresponding to $(-1/2, 1/2)$. This is important to keep track of to make the correct preductions for experiments like the delayed choice quantum erasure experiment: the correct prediction results only when you are careful to keep track of exactly what you are learning from each observation. We can see this even in the classical double slit experiments: if you know which slit the photon travelled through, the state of that photon is different (as witnessed by a different interference pattern) than if you only know it passed through one or the other of the slits. One way of enforcing this in our theory itself is to make the set of outcomes for each particular observation finite (even if the Hilbert space is infinite dimensional), e.g. instead of a single $x$-position observation, we have many different $x$-position observations with various levels of precision. For example, when we detect a photon has travelled through one of the slits, we do not know the exact position as it pases through, only the position up to the width of the slit. This would correspond to a different observable quantity than if we had used a more precise means of measuring the position of the photon.

This phenomenon of the state being projected to a particular subspace based on what you learn from the observation is known as "wavefunction collapse". The general idea is that when a system is not being observed, it evolves continuously according to Schrödinger's equation. When it is observed, it undergoes collapse, which is a discontinuous process (a projection). It's interesting to note that the Copenhagen interpretation essentially integrates epsitomology into the rules governing the behavior of physical systems: what you can theoretically learn from an experiment has a lot of effect on its distributions of outcomes. This leads to a very strong objection to the Copenhagen interpretation: it depends on having a system with external observers. An isolated system (such as the universe itself) does not have external observers, and so ought to (according to the Copenhagen interpretation) never undergo collapse. This does not mean that the Copenhagen interpretation is not valid: in fact it is incredibly good at making practical predictions. It's just that it cannot form the foundations of a complete physical theory. Physicists who care about foundations have tried to get around this by deriving apparent wavefunction collapse as a feature of coupled systems consisting of an environment and an observed subsystem. This is known as "quantum decoherence", and in this picture, rather than an immediate and discontinuous wavefunction collapse, there is rather a very rapid yet continuous evolution of the system from one in which the state of the observed subsystem is not correlated to the state of the environment,

[5] Since we are working with Hilbert spaces, when we use the word"span", we mean orthogonal span (so infinite linear combinations are allowed, not just finite linear combinations). This is necessary to ensure that the span of a collection of closed subspaces is again a closed subspace.

to one in which the state of the observed subsystem is heavily correlated to the state of the environment. Roughly, before a photon passes through a polarizing grid, the content of your lab notebook doesn't tell you anything about the polarization of the photon, but after your experiment, it does. According to this picture, your lab notebook will in fact end up in a superposition of states, one in which you've written "photon passed through" and another in which you've written "photon blocked": this is a necessary consequence of the linearity of Schrödinger's equation if we suppose (as these physicists do) that it describes the dynamics completely. The idea is that the new correlation between the state of the environment (e.g. your lab notebook) and the polarization of the photon gives the appearance of wavefunction collapse, because each of the classical states which compose the superposition which the entire system is in after the experiment will be states in which experiments on the photon give a deterministic outcome: if your lab notebook says the photon passed through, and you repeat the experiment, the photon will again pass through (and vice versa). This was not true before the experiment: the future observed polarization of the yet-to-be-observed photon will not be correlated with whatever you chose to write beforehand. We will explore this idea later, because conditional statements when describing quantum systems have to be interpreted very carefully. The main point of this part of the discussion is only to point out that the Copenhagen interpretation has problems, and that there is a proposed solution. There is also something unsatisfactory about this proposed solution, namely that we never seem to consciously experience such superpositions. One solution to this, known as the "many-worlds hypothesis", is that each classical component of a superposition corresponds to one of many equally real worlds in a massive branching tree, and our conscious mind happens to be experiencing just one branch along that tree, but other versions of our mind are travelling along all the other branches just as well. This may sound like an aburdly strong metaphysical claim to a philosopher, but it is the most popular interpretation amongst physicsts beside the Copenhagen interpretation.

All of this talk about interpretations has just been to put what we are going to do in context with the main lines of thought in the foundations of physics. We do not wish to resolve what the "correct" interpretation is, or even make claims about physical reality. Rather, we want to witness some interesting mathematical phenomena which logicians who are interested in the interplay between the discrete and the continuous may find interesting. We are using physics just as an inspiration.

In this thesis, our general theme is that discrete structures often are at home in a broader continuous context. In this case we have an analogue of Stone's representation theorem, where instead of representing a Boolean algebra a colleciton of subsets of some set, we represent a Boolean algebra as a collection of closed subspaces of a Hilbert space. Elements of a Boolean algebra are propositions, and conjunction, disjunction, and negation (when represented as an algebra of sets) correspond to intersection, union, and complement. This idea can be traced back to Frege, and is ubiquitious in mathematical logic. When represented as an algebra of closed subspaces of a Hilbert space, conjunction, disjunction, and negation correspond to intersection, sum[6], and orthogonal complement of subspaces.

The phenomenon of interest to logicians here is that if we start with a Boolean algebra of propositions, and represent it as an algebra of closed subspaces of a Hilbert space, we see in this representation many new propositions (many new closed subspaces). The phenomenon we want to look at now is the role of orthogonality. The way we chose to represent Boolean algebras, given any two propositions, $p$ and $q$, the corresponding subspaces $W_p$ and $W_q$ have the property that $W_p \cap W_q$, $W_p \cap W_q^\perp$, $W_p^\perp \cap W_q$, and $W_p^\perp \cap W_q^\perp$ are mutually orthogonal and direct sum to the entire space. We can think of a proposition as corresponding to a pair of orthogonal subspaces which direct sum to the entire space, one of which is associated with the proposition being true, and the other with the proposition being false, or in the language of quantum mechanics: each proposition corresponds to a self-adjoint operator on a Hilbert space with two eigenvalues, 1 (corresponding to true) and 0 (corresponding to false). It is a standard theorem that a collection of self-adjoint operators commute if and only if they share a common othogonal eigenbasis. Physicists interpret commuting self-adjoint operators to be simultaneously observable quantities. An example of such quantities are the x-position, y-position, and z-position of a photon. We can see that our way of representing Boolean algebras as algebras of subspaces of Hilbert spaces assigns the elements of the Boolean algebra to simultaneously observable quantities. This makes complete sense, as classically for any set of propositions, we can evaluate the truth of any number of them that we like, in any order we like, and then take Boolean combinations afterwards without fear that evaluating the truth of of one proposition might change the truth of another proposition. But this does not hold in reality. Hiesenberg's uncertainty principle says that x-position and x-momentum are not simultaneously observable quantities. They correspond to self-adjoint operators which do not commute.

One way of witnessing this is a failure the law of conditional probability. This says if we have two events $A$ and $B$ (think, sets of outcomes for different experiments), then the probability $P(A|B)$ of $A$ conditional on event $B$ happening is $\frac{P(A \cap B)}{P(B)}$. This has been experimentally demonstrated to not hold for position and momentum of a particle in the same direction (as well as many other observable quantities), and is even part of the standard labwork of undergraduate physics students.

The question we wish to pose to the reader is: should we take this as a serious challenge to the foundations of mathematics? On the one hand, because of the representation theorem we just gave of Boolean algebas inside algebras of closed subspaces of Hilbert spaces, classical logic is at least consistent, and applicable to a wide range of circumstances. On the other hand, it seems to challenge the notion that classical logic has any kind of undisputed a priori status as a basis for all mathematical thought. In particular, it excludes as logical

[6]We define the sum of an infinite collection of subspaces to be the closure of their span, rather than just the span, so our operations always give closed subspaces when applied to closed subspaces

impossibilities actual situations which seem to have perfectly fine mathematical descriptions.

One defense of classical logic here is to claim that there is some yet-unknown but completely classical explanation for the observations of physicists. Bohm's "Pilot Wave" theory, for example, attempts to bring quantum mechanics closer to a classical explanation: this essentially states that that there are hidden variables which guide the evolution of the universe classically. But this explanation has the unfortunate feature of violating one of main tennets of physics, namely that all interactions happen locally. It could also be considered to violate Occam's razor, by proposing a theory that is strictly more complicated (and observably indistinguishable) from other interpretations of quantum mechanics. The situation is that for around a century, physicists have tried to find a clear, coherent, and simple classical explanation of their observations, and failed. Although this is in no way a proof that our universe operates on a kind of non-classical logic, it means we cannot be so quick to dismiss it.

Another defense would be to point at that quantum mechanics and its strange logic is, in fact, interpreted inside classical mathematics. So although what scientists may call "propositions" turn out not to really be propositions in a classical mathematical/philosophical sense, this is just a problem with scientists using confusing and wrong terminology, not with mathematics itself.

However, we could point out that the development of continuous logic proceeded by first giving an interpretation of continuous logic inside classcal mathematics. This will be so with almost any new mathematics: one of the very first things you try to do is interpret your new mathematical structures inside the existing framework of mathematics. This is the easiest way of establishing the consistency of a new theory, and it's important to ensure that mathematics as a whole has a coherent ontology. This does not mean that we think that classical logic should be held to be more fundamental than continuous logic. In fact, we think the opposite! It would certainly be a challenge to try to "boot-strap" continuous logic, working only on first-principles which are easily explained to someone who does not know any classical logic or mathematics. But this can be explained by the fact that humans have been refining classical logic and classical mathematics for thousands of years, and have developed a relatively useful pedagody for it. In practice, a professional geometer will rarely think of themselves as dealing with Zermelo-Fraenkel sets, and may not even know what the axioms are. Most of the body of mathematics has never been proved from the ZFC axioms. All this is to say that, at the very least, the meaning of mathematical concepts don't seem to depend upon them being interpreted in classical set theory. So it would probably be premature to claim that just because quantum mechanics has been developed using largely classical mathematical structures which have standard interpretations as sets, quantum mechanics is ultimately a classical theory.

As an area of future study, it would be interesting if there were a common generalization of the logic of quantum mechanics and continuous logic. It's hard to see, at the moment, how supremums and infimums would play a role, since these do not have analogues in the complex numbers. At the very least, we encourage the reader to keep an open mind about the foundations of mathematics, and take serious the question of whether it would benefit mathematics to found it on a continuous base, or a quantum logic base.

# Chapter 4

# Bridging Continuous Structures and Descriptive Set Theory?

The theory of Polish spaces, some might consider, is squarely in the realm of descriptive set theory. A classical theorem says that, topologically, every Polish space is a quotient of a subset $\omega^\omega$, Baire space. The construction is what you might expect given some of our definitions in Chapter 1: take a countable dense set $D$ of your Polish space $X$, identify $D$ with $\omega$, then identify sequences from $D$ with elements of $\omega^\omega$. Take the subset of $\omega^\omega$ corresponding to fast Cauchy sequences, and quotient by the relation of converging to the same point in $X$. A refinement of this argument shows that every non-empty Polish space is a quotient of $\omega^\omega$ itself, meaning Baire space can be considered a kind of universal Polish space to do descriptive set theory with. See Moschovakis' *Descriptive Set Theory* [11] for more background.

We might wonder, now that we can think of effective Polish spaces as computable continuous structures, how much of effective descriptive set theory can we develop in this framework? Usually descriptive set theory is done in a second-order language, with quantifiers both over reals[1] and over the naturals. Descriptive set theory is quite powerful, allowing us for example to quantify of models of fragments of set theory. We can immediately see that our continuous first-order theory of $\omega^\omega$, even if we add additional functions and relations, is going to be weaker, in the sense that we can define less. One problem is that because all of our formulas are uniformly continuous, we can only ever define closed sets (according to our definition where a set is a definable if the equalness relation to that set is definable). Actually there is a subtlety here, which is that our language cannot distinguish between a set and its closure. In a sense the "definable sets" in a continuous structure are equivalence classes of subsets under the relation $X_1 \equiv X_2$ iff $\overline{X_1} = \overline{X_2}$. But in any case, it makes sense to say we will only ever be able to define sets that are $\mathbf{\Pi}^0_1$ in the (boldface) Borel hierarchy. However, there are still complicated closed sets when measured with the lightface Borel hierarchy, where unions and intersections are required to be effective.

[1] The term "reals" is sometimes used to describe elements of $\omega^\omega$, since it is homeomorphic to $\mathbb{R} \setminus \mathbb{Q}$

This is brings up an interesting consideration for the philosophy of science. Some questions, like whether the distance from Earth to the center of the galaxy is rational or irrational at a given moment, are neither falsifiable nor verifiable. Even someone with access to arbitrarily precise instruments, all mathematical knowledge, and access to all $< \omega_1^{CK}$ iterates of the halting problem as an oracle could find no physical evidence either for or against this distance being rational.[2] But we might think that this is a silly example of something neither falsifiable nor verifiable: it doesn't really correspond to any kind of natural physical property. After all, nature abhors the discontinuous. The complication here doesn't really have to do with the descriptive complexity of the property, but more with the fact that the corresponding relation is nowhere continuous. But there can be uniformly continuous properties which are not falsifiable or verifiable, because the process of converting measurement data into a decision about to whether or not (or rather, to what extent) the property holds is computationally non-trivial. In other words, an all-knowing being might be able to use the data we gather to figure out to what extent the property holds, but we are unable to, because we don't know how to put the pieces together. An example along these lines might be the "spectral gap problem", which asks whether there is a gap between the energy of the ground state of a physical system and its other states. A paper by Cubitt et. al. "Undecidability of the spectral gap" [5] showed there is a family of physical systems for which it is undecidable which have a spectral gap. Meaning even with a complete description of the physical system and its governing laws, there are properties of the system we can't determine. In general, we might wonder about the role of uniformly continuous relations at high levels of the effective Borel hierarchy. These are relations which *would* be falsifiable or verifiable if we had access to an oracle. If the Church-Turing thesis turns out to be false (e.g., if there is some physical procedure to solve the $n$-th iterate of the halting problem), then some of these relations which we initially thought were neither falsifiable nor verifiable might turn out to be both verifiable and falsifiable.

One other question to consider when thinking about definability in continuous logic is "what if infinite mathematical objects don't really exist?". We've already seen that some well-behaved uncountable continuous structures can be described completely using only a countable amount of information, including many of the standard uncountable spaces used by practicing mathematicians[3], such as Cantor space, $\mathbb{R}$, separable Hilbert spaces, etc. Furthermore for any point $p$ in one of these nice uncountable spaces, and any degree of accuracy

[2]There is of course the measure-theoretic argument that rational distances are very rare, but how do we know the universe was not set up so the Earth is in a perfectly circular orbit around the center of galaxy? Or perhaps suppose this is being asked by an omnipotent demon who is trying to trick you by asking you this question about a moment of their choosing with an unlikely answer? If these hypotheticals bother you, you could replace this by breaking up $\mathbb{R}$ into two disjoint pieces $A$ and $B$ which are both dense in $\mathbb{R}$ and both have measure $r$ inside any interval $[-r, r]$, but then there is the question of how to communicate the question to someone. Let's not dwell on this, though.

[3]Logicians, of course, prefer impractical mathematics.

we like, we can find a point $\tilde{p}$ which is indistinguishable from $p$ (to that degree of accuracy) which can be described by only a finite amount of information. Going further, a function between such spaces, which in general would require an uncountable amount of information to describe as a set-theoretic function, will only need a countable amount of information to describe if it is uniformly continuous (only its values on the points which have finite descriptions are needed); and if the function is computable, we would actually only need a finite amount of information to describe it! Given all of this, it's interesting to consider whether we really need infinite objects at all. If we are satisfied with only having finite approximations, how much mathematics would we really be missing? Recall, a finitist might say something like:

> I don't really believe in Cantor space, the uncountable set. Nor do I believe in the completed infinity of $2^{<\omega}$, the set of all finite binary strings. I am sure, however, that the finite binary strings I can write down exist, and that for any two finite binary strings $\sigma$ and $\tau$ that exist, their concatenation $\sigma^\frown\tau$ exists. I also believe that, given a nice finite description of a well-defined computational procedure, I can apply it to any binary string that exists, and the result will be another existing binary string.

This might seem hopelessly limiting at first, but there is a way of interpreting statements with many alternations of quantifiers over infinite sets as iterated modal statements about finite objects. For example, consider the $\Pi_2$ statement "$(\forall x \in \mathbb{N})(\exists y \in \mathbb{N})(x + x = y + x)$". There are $\Pi_2$ statements which are independent of ZFC, so just by looking at the syntactic form, for all we know this could be an impossible problem to solve. But we can translate it into a kind of modal statement which seems more palatable: "If you give me a natural number $x$, I can produce for you a natural number $y$ so that $x + x = y + x$". There is no longer an explicit reference to the set of all natural numbers. Rather, there is a claim about your ability to fulfill requests of a given form. In this case, you can fulfill the request trivially by handing back the number $x$ you are given: the identity function is a Skolem function for this $\Pi_2$ statement. We don't need to believe in infinite sets to make sense of this statement.

In the context of continuous logic, consider a statement like "$\inf_x \sup_y \phi(x, y) > \tau$", where $\phi(x, y)$ is computable and 1-Lipschitz, and $\tau \in [0, 1]$ is computable, evaluated in a computable continuous structure based on the continuous space $2^\omega$ with additional relations/functions. While $\inf_x$ and $\sup_y$ could be thought of as quantifiers over an uncountable set, the result is the same if they are interpreted as quantifiers over a dense subset of $2^\omega$, which we can identify with $2^{<\omega}$. So translating, this becomes a statement with quantifiers over only countable sets:

$$(\exists m \in \omega)(\forall k \in \omega)(\forall \sigma_1 \in 2^{<\omega})(\exists \sigma_2 \in 2^{<\omega})\big([\phi](\sigma_1, \sigma_2)(m + k + 2) > [\tau](m + k + 2) + 2^{-m}\big)$$

Because our function is 1-Lipschitz, however, we can actually create an a priori bound on $|\sigma_1|$ and $|\sigma_2|$ depending on $m$. So this turns out to be equivalent to something like:

$$(\exists m \in \omega)(\forall k \in \omega)(\forall \sigma_1 \in 2^{m+3})(\exists \sigma_2 \in 2^{m+3})\big([\phi](\sigma_1, \sigma_2)(m+k+3) > [\tau](m+k+3) + 2^{-m}\big)$$

Optimizing further, we don't need to know the value of $\phi$ or $\tau$ to arbitrary precision, only to sufficient precision to see the value of one is bounded strictly above the other, meaning we don't need to quantify over $k$. In fact, $k = 0$ suffices here:

$$(\exists m \in \omega)(\forall \sigma_1 \in 2^{m+3})(\exists \sigma_2 \in 2^{m+3})\big([\phi](\sigma_1, \sigma_2)(m+3) > [\tau](m+3) + 2^{-m}\big)$$

In other words, this statement about given by a $\Pi_2$ continuous formula turned out to be equivalent to a classical $\Sigma_1$ statement! And our argument can be generalized: for any 1-Lipschitz $\Pi_n$ sentence for any $n \in \omega$, evaluating whether $\psi > \tau$ in $2^\omega$ for $\tau$ computable will be classically $\Sigma_1$. The reason for this, essentially, is that $2^\omega$ is compact. For any given level of precision, we can evaluate any formula to that given level of precision by considering only a finite collection of points in $2^\omega$ to which every point in $2^\omega$ is sufficiently close. The only reason this turns out to be $\Sigma_1$, and not recursive, is because in general deciding for two computable reals whether $r_1 > r_2$ is a $\Sigma_1$ problem. If we only wanted to be able to compute $\varepsilon(\phi(\bar{x}), \tau)$ to arbitrary precision, this would turn out to be computable for any $\Sigma_n$ or $\Sigma_n$ 1-Lipschitz formula $\phi(\bar{x})$ for any $n$: for any fixed $\bar{x} \in (2^\omega)^n$, simply search in parallel for each rational $\tau_1 < \tau_2$ for a witness to $\tau_1 < \psi(\bar{x}) < \tau_2$. It's also worth noting this can be done uniformly across $\omega$-presentations of $2^\omega$, because to find an analogue of $2^n \subseteq 2^\omega$ in any copy of $2^\omega$ given its atomic diagram, we can just search for $2^n$ points with pairwise equalness $2^{-2^{-n}}$, so in fact $\Sigma_k$ relations for every $k \in \omega$ are both u.r.i.c.e. an co-u.r.i.c.e. But this means, by our previous result in chapter 1, that they are all $\Delta_1^c$, which gives us a form of effective quantifier elimination for Cantor space.[4]

**Note.** A similar argument would *not* work for a non-compact space like $\mathbb{R}$ or $\omega^\omega$, because no finite collection of points can approximate every point to within distance, say $\frac{1}{3}$. What's important, philosophically, is that we have found a beautiful expression of the intuition that compact structures are "basically finite". It's also worth noting that the condition our formulas be 1-Lipschitz can be weakened, as long as we can compute a modulus for them.

We'll go into more detail about compact continuous structures in a later section, the point of this discussion was to show that the hierarchy of $\Sigma_n$ and $\Pi_n$ continuous relations on $2^\omega$ (in our finitary continuous language) bears very little resemblance to the hierarchy of classically $\Sigma_n^0$ and $\Pi_n^0$ relations on $2^\omega$ as classical structure that we look at in effective descriptive set theory. This will be true even if we additional 1-Lipschitz functions and relations to our continuous language, like coordinate-wise addition (mod 2). This is all to say that continuous logic at least at first glance, doesn't give us a new intrinsic way to study to study the effective Borel hierarchy on $2^\omega$. The way it handles uncountable spaces is quite different

[4]We can replace $\Delta_1^c$ with $\Delta_1$ if we have the connective $\lim_n$ in our language.

from effective descriptive set theory.

It's also worth trying to understand what goes wrong if we try to, say, add some function and relations in our language which allow us to interpret $\mathbb{N}$, giving us back some of the expressive power we take advantage of in effective descriptive set theory. For example, perhaps we use a prefix-free coding of natural numbers to identify it with a subset of $2^{<\omega}$, and represent natural numbers as equivalence classes under a nice equivalence relation on strings in $2^\omega$. The problem will be that we will end up with our representations of natural numbers having accumulation points in $2^\omega$. If we, say, we identify 0 with the equivalence class of strings starting with 00 and 1 with the equivalence class of string starting with 01, then any uniformly continuous binary relation representing the graph of a function $\mathbb{N} \to \{0, 1\}$ will need to need to relate all (representatives of) natural numbers sufficiently close to an accumulation point to (a representative of) the same value. If we have a single accumulation point, this means our function must be eventually constant.

However, we might still have a chance on $\omega^\omega$, since it is not compact. Indeed, we can identify natural numbers with equivalence classes under the computable equivalence relation $x \equiv y$ iff $x(0) = y(0)$. We can add functions corresponding to addition and multiplication to our language (pointwise addition and multiplication would work). The problem here is that the evaluation map $(x, i) \mapsto x(i)$ (represented using our interpretation of natural numbers in $\omega^\omega$) will not be uniformly continuous. We can see this because for each $i \in \omega$, the map $x \mapsto x(i)$ is $2^i$-Lipschitz, but no better. So our binary modulus would need to be bounded from below by the function $(\tau_1, \tau_2) \mapsto \tau_2^{2^i}$ for each $i$, inconsistent with the modulus axioms. The fact that we can't define the graph of the evaluation map seems to suggest that we won't have much luck defining general closed sets. We can also see that in any compact subset of $\omega^\omega$, we will have the same limitation on definability of sets as we did in the case of $2^\omega$.

# Chapter 5

# Notions of Genericity

This chapter is not really related to continuous logic, but we think it fits with the general theme of studying uncountable structures from a logical point of view. It might have something to say about extending the notion of genericity to points in arbitrary continuous spaces, but we have not done that here. We want to study a more intrinsic notion of genericty inside closed perfect subsets of Cantor space. It turns out after the fact that this replicates some work of Bernard Anderson in [2], which is a good lesson to properly review the literature before starting work on your own. We include this mostly because it might serve some expository purpose. This work was done to try to replicate some of what Jan Reimann and Theodore Slaman did for measure in *Measures and the Random Reals*[13].

**Lemma 22** (Fixed-Point Lemma)**.** *Let $G \in 2^\omega$ be 1-generic relative to $z$, and $\Phi$ be a truth-table functional $2^\omega \to 2^\omega$ recursive in $z$. Then if $\Phi(G) = G$, there is basic clopen $[\sigma] \ni G$ with $\Phi|_{[\sigma]} = Id_{[\sigma]}$.*

*Proof.* Consider the $z$-effective open set $V := \{x \in 2^\omega : (\exists n \in \omega)\Phi(x)(n) \downarrow\neq x(n)\}$. $G \notin V$ (since $\Phi(G) = G$), so by 1-genericity of G relative to z, there is some condition forcing $G \notin V$, i.e., there is some $\sigma \subset G$ such that $[\sigma] \cap V = \emptyset$. So if $x \in [\sigma]$, then $(\forall n \in \omega)\Phi(x)(n) \downarrow= x(n)$, i.e, $\Phi(x) = x$. $\square$

**Conventions.** If $X \subseteq 2^\omega$, then $W \subseteq X$ is **open in the relative topology of X** (abbreviated $\mathbf{open_X}$) if $W = U \cap X$ for some open $U \subseteq 2^\omega$. Likewise we use the abbreviations $\partial_\mathbf{X}$, $\mathbf{Cl_X}$, and $\mathbf{Int_X}$ to refer to the boundary, closure, and interior operations relative to $X$, and use the same notation for other topological properties. We abide by the convention that if we apply a typed adjective to a new name, the name is of the corresponding type (e.g., if we say "$Z$ is open$_X$", we are implicitly stating that $Z \subseteq X$).

**Definition 63.** Let $X \subseteq 2^\omega$ be a perfect subset of $2^\omega$, and $z \in 2^\omega$. Then $x \in X$ is **weakly 1-generic in X relative to z** (or **weakly z-1-generic in X**) if x belongs to every z-effectively open set $U$ such that $X \cap U$ is dense in $X$.[1] (Equivalently, $x$ belongs to every $z$-effectively open$_X$ dense$_X$ set)[2]

**Definition 64.** Let $X \subseteq 2^\omega$ be a perfect subset of $2^\omega$, and $z \in 2^\omega$. Then $x \in X$ is **1-generic in X relative to z** (or **z-1-generic in X**) if x is not in $\partial_X U = Cl_X(U) \setminus Int_X(U)$ for any z-effectively open$_X$ set $U$.[3]

The intuition behind the property of 1-genericity is that if $x$ is 1-generic, any $\Sigma^0_1$ property $x$ might have or not have is either forced to hold or forced to not hold by some finite initial segment of $x$. The following observation shows this intuition extends to 1-genericity in $X$:

**Observation 23.** *Let $X = [T] \subseteq 2^\omega$, where $T$ is a perfect tree, and $x \in X$[4]. The following are equivalent:*

*(i) $x$ is z-1-generic in $X$.*

*(ii) For any z-effectively open set $U$, if $X \cap U$ is dense along $x$[5], then $x \in X \cap U$.*

*(iii) For every $\Sigma^0_1(z)$ set of finite strings $S$, there is an initial segment $\sigma \subset x$ such that either $\sigma \in S \cap T$ or $(\forall \tau \supseteq \sigma)(\tau \notin S \cap T)$.*

*(iv) For every Turing functional relative to $z$, say $\Phi$, there is an initial segment $\sigma \subset x$ such that either $\Phi^\sigma(0)\downarrow$ or $(\forall \tau \supseteq \sigma)(\tau \in T \to \Phi^\tau(0)\uparrow)$*

**Definition 65.** A **z-uniformly perfect** subset of $2^\omega$ is a perfect set $X \subseteq 2^\omega$ such that the function $f_X(\sigma) := \begin{cases} 1 & [\sigma] \cap X \neq \emptyset \\ 0 & [\sigma] \cap X = \emptyset \end{cases}$ is recursive in $z$.[6]

[1]We say $x \in X$, but we could have equivalently said $x \in 2^\omega$, since if $x \notin X$, there is a basic clopen set $[\sigma] \ni x$ disjoint from $X$, and thus given any $z$-effectively open set $U$ with $X \cap U$ dense in $X$, $U \setminus [\sigma]$ is a z-effectively open set whose intersection with $X$ is dense in $X$ not containing $x$.

[2]We didn't define what a $z$-effectively open$_X$ set is, but you could imagine the definition: we can consider set of the form $[\sigma] \cap X$ as basic open sets in $X$, and then a $z$-effective open$_X$ set is a set of the form $\bigcup_{i\in\omega}[\sigma_i] \cap X$, where the map $i \mapsto \sigma_i$ is recursive.

[3]It is worth saying why 1-generic in $X$ is not just the same as 1-generic which happens to belong to $X$. It is true since $X$ is closed that for $Z \subseteq X$ (no topological assumptions on $Z$), $Cl_X(Z) = X \cap Cl(Z)$. Moreover, if $U \subseteq X$ is open$_X$ (say $U = X \cap V$ with $V$ open), then $Int_X(U) = U = X \cap V = X \cap Int(V)$, so $\partial_X(U) = (X \cap Cl(U)) \setminus (X \cap Int(V)) = X \cap (Cl(U) \setminus Int(V))$. This looks very close to $X \cap \partial V$, and if we are not careful this might make us suspect that the boundary points of an open$_X$ set $U = X \cap V$, where $V$ is open, are just the boundary points of $V$ which happen to lie in $X$. This is in fact true if $X$ is clopen, which makes 1-genericity in $X$ trivial for $X$ clopen. However, for an arbitrary perfect $X$, it need not be true for all open $V \subseteq 2^\omega$ that $Cl(X \cap V) = X \cap Cl(V)$. For example, if $X$ is the perfect set of infinite binary sequences with all even digits being "0", and $V$ is the open set of binary sequences which have at least one even digit being "1", then $V$ is dense in $2^\omega$, so $X \cap Cl(V) = X \cap 2^\omega = X$, but $Cl_X(X \cap V) = Cl_X(\emptyset) = \emptyset$. In other words, this fails spectacularly.

[4]Here we actually have to assume $x \in X$, because condition (iii) is trivially satisfied if $x \notin X$

[5]i.e., $(\forall n \in \omega)(\exists y \supset x \restriction n)(y \in X \cap U)$

[6]Note that for a $z$-uniformly perfect set X, $x \in X$ if and only if $\forall n f_X(X \restriction n) = 1$, so such a set is $\Pi^0_1(z)$.

**Observation 24.** *Let $X \subseteq 2^\omega$ be non-empty. The following are equivalent:*

*(i) $X$ is $z$-uniformly perfect.*

*(ii) $X = [T]$ for some $z$-recursive perfect tree $T \subset 2^\omega$.*

*(iii) $X = \Phi(2^\omega)$ for some injective, order-preserving, 1-Lipschitz[7] functional $\Phi$ relative to $z$.*

*(iv) $X = \Phi(2^\omega)$ for some injective truth-table functional $\Phi$ relative to $z$.*

*Proof.* $(i) \implies (ii)$: Suppose that $X$ is a non-empty $z$-uniformly perfect subset of $2^\omega$. Then $T := \{\sigma : f_X(\sigma) = 1\} \subseteq 2^{<\omega}$ is a z-recursive perfect tree such that $X = [T]$.

$(ii) \implies (iii)$: Define $\Phi$ by recursion as follows. Given $\Phi(\sigma)$, let $\tau$ be the longest string $\tau \supseteq \sigma$ in $T$ such that there is no string $\tau' \supseteq \sigma$ in $T$ of the same length incompatible with $\tau$. Then define $\Phi(\sigma^\frown 0) = \tau^\frown 0$ and $\Phi(\sigma^\frown 1) = \tau^\frown 1$.

$(iii) \implies (iv)$: An injective 1-Lipschitz functional is an injective truth-table functional.

$(iv) \implies (i)$: Suppose that $\Phi$ is an injective truth-table functional relative to $z$ with $\Phi(2^\omega) = X$. Then there is a $z$-recursive function $u : \omega \to \omega$ such that $(\forall \rho \in 2^{u(n)})(\forall i < n)(\Phi(\rho)(i)\downarrow)$. Then define $f_X(\sigma) = 1$ if and only if $(\exists \tau \in 2^{u(|\sigma|)})(\forall i < |\sigma|)(\Phi(\tau)(i) = \sigma(i))$. □

**Note.** The function u in the previous proof is a uniform upper bound on the "use" function of $\Phi$. The use function tells us how many bits of $x \in 2^\omega$ are needed to compute the first $k$ digits of $\Phi(x)$. Why can we get a uniform bound (one that works for *all* $x \in 2^\omega$ at once)? Well, for any k, the function: $x \mapsto$ "number of bits needed to compute $\Phi(x)(0), ..., \Phi(x)(k-1)$" is continuous, and $2^\omega$ is compact, so its image must be a compact subset of $\omega$, thus bounded.

**Observation 25.** *The image of a $z$-effective open set $U \subseteq 2^\omega$ under an injective 1-Lipschitz functional $\Phi$ is a $z$-effective open$_{\Phi(2^\omega)}$ set.*

*Proof.* Fix $x \in U$. First we'll prove that for every basic clopen $[\tau] \subseteq U$ with $x \in [\tau]$, there is a basic clopen of the form $[\Phi(x) \restriction n]$ whose preimage is contained in $[\tau]$. Consider the continuous function $f : [\tau]^c \to \mathbb{R}$ defined by $f(y) = d(\Phi(x), \Phi(y))$[8]. $f$ is continuous, and $[\tau]^c$ is compact, so $f$ achieves a minimum. Since $y \in [\tau]^c \implies y \neq x \implies \Phi(x) \neq \Phi(y) \implies d(\Phi(x), \Phi(y)) > 0$, this minimum cannot be 0, so must be $\frac{1}{2^m}$ for some $m \in \omega$. But then $y \in \Phi^{-1}([\Phi(x) \restriction m+1]) \implies \Phi(y) \in [\Phi(x) \restriction m+1] \implies d(\Phi(x), \Phi(y)) < \frac{1}{2^m} \implies y \notin [\tau]^c \implies y \in [\tau]$.

[7] A **1-Lipschitz functional relative to z** is a Turing functional relative to $z$ such that for all $\sigma \in 2^{<\omega}$, $\Phi^\sigma(i)\downarrow$ for all $i < |\sigma|$.

[8] The metric $d(x, y) := 2^{-k}$ where $k \in \omega$ is the greatest such that $x \restriction k = y \restriction k$ (or 0 if $x = y$). This turns $2^\omega$ into a complete metric space inducing the usual topology.

To prove the observation, define $V \subseteq 2^\omega$ by $V = \bigcup\{[\sigma] : \Phi^{-1}([\sigma]) \subseteq U\}$. Note that $V$ is a z-effective open subset of $2^\omega$, since $z$ can compute a finite cover of $\Phi^{-1}([\sigma])$ by basic clopen sets uniformly in $\sigma$: $\Phi^{-1}([\sigma]) = \bigcup\{[\tau] : \tau \in 2^{|\sigma|} \wedge (\forall i < |\sigma|)(\Phi(\tau)(i) = \sigma(i))\}$. We can check $\Phi(U) = \Phi(2^\omega) \cap V$, since

$$\Phi(x) \in \Phi(U) \iff x \in U \iff (\exists \tau \in 2^{<\omega})(x \in [\tau] \wedge [\tau] \subseteq U)$$

$$\iff (\exists \sigma \in 2^{<\omega})(x \in \Phi^{-1}([\sigma]) \wedge \Phi^{-1}([\sigma]) \subseteq U)$$

$$\iff (\exists \sigma \in 2^{<\omega})(\Phi(x) \in [\sigma] \wedge \Phi^{-1}([\sigma]) \subseteq U) \iff \Phi(x) \in V$$

Thus $\Phi(U)$ is a z-effective open$_{\Phi(2^\omega)}$ set. □

**Observation 26.** *If $\Phi$ is an injective truth-table functional, then $\Phi$ has a truth-table left inverse (i.e., a truth-table functional $\Psi$ such that $\Psi \circ \Phi = Id_{2^\omega}$).*

*Proof.* Let $u$ be a recursive uniform bound on the use of $\Phi$, as defined in the proof of Observation 24. Define a z-recursive labeling function $l : 2^{<\omega} \to 2^{<\omega}$ as follows. If $\Phi(\tau) \upharpoonright |\sigma| = \sigma$ for some $\tau \in 2^{u(|\sigma|)}$, let $l(\sigma) = \rho$, where $\rho$ is the longest common initial segment of the strings $\{\tau \in 2^{u(|\sigma|)} : \Phi(\tau) \upharpoonright |\sigma| = \sigma\}$. If $\Phi(\tau) \upharpoonright |\sigma| \neq \sigma$ for any $\tau \in 2^{u(|\sigma|)}$, let $l(\sigma) = l(\sigma \upharpoonright (|\sigma| - 1))^\frown 0$. Notice that if $\sigma \subseteq \tau$, then $l(\sigma) \subseteq l(\tau)$. We'll now show that for any $z \in 2^\omega$, $l(z \upharpoonright (n+1)) \subsetneq l(z \upharpoonright n)$ infinitely often. For $z \notin \Phi(2^\omega)$ this is clear, so suppose $z = \Phi(x)$, and suppose to the contrary that $l(\Phi(x) \upharpoonright (n+1)) = l(\Phi(x) \upharpoonright n)$ for all $n \geq N$. Then for all $n \in \omega$, there exists some $y_n \in 2^\omega$ with $d(\Phi(x), \Phi(y_n)) \leq 2^{-n}$ but $d(x, y_n) \geq 2^{-|l(x \upharpoonright N)|-1}$. Since $C := \{y \in 2^\omega : d(x, y) \geq 2^{-|l(x \upharpoonright N)|-1}\}$ is compact, the continuous function $y \mapsto d(\Phi(x), \Phi(y))$ restricted to $C$ must achieve its minimum in $C$, which by our previous sentence must be 0. But then there is some $y \in C$ with $\Phi(x) = \Phi(y)$. This is impossible, however, since $x \notin C$ and $\Phi$ is injective. Now simply define $\Psi(\sigma)(i) = l(\sigma)(i)$ for $i < |l(\sigma)|$. We just verified that $\Psi$ is total and $z$-recursive, and we can also see that $\Psi(\Phi(x)) = x$ for any $x \in 2^\omega$. □

**Theorem 27.** *The following are equivalent:*

*(i) $x$ is z-1-generic in $X$ for some z-uniformly perfect set $X \subseteq 2^\omega$.*

*(ii) $x$ is truth-table equivalent relative to $z$ to a z-1-generic $G$.*

*Proof.* (i) $\Rightarrow$ (ii): Fix an injective 1-Lipschitz functional $\Phi$ relative to $z$ with $\Phi(2^\omega) = X$. We claim $G := \Phi^{-1}(x)$ is 1-generic relative to $z$. Let $U \subseteq 2^\omega$ be a $z$-effective open set dense along $G$. Then by Observation 25, $\Phi(U)$ is $z$-effective open$_X$. Let's check that $\Phi(U)$ is dense along $x$ in $X$. Let $\sigma \subset x$. Suppose for the sake of contradiction that $(X \cap [\sigma]) \cap \Phi(U) = \emptyset$, then $[\sigma] \cap \Phi(U) = \emptyset$ (since $\Phi(U) \subseteq X$). But then $\emptyset = \Phi^{-1}([\sigma] \cap \Phi(U)) = \Phi^{-1}([\sigma]) \cap \Phi^{-1}(\Phi(U)) = \Phi^{-1}([\sigma]) \cap U$. But then $\Phi^{-1}([\sigma])$ is an open neighborhood of $G = \Phi^{-1}(x)$ disjoint from $U$, contradicting that $U$ was dense along $G$. So $\Phi(U)$ is a $z$-effective open$_X$ set dense along $x$. Thus $x \in \Phi(U)$, so $G \in \Phi^{-1}(\Phi(U)) = U$. This shows $G$ is $z$-1-generic. And $G$ and $x$ are truth-table equivalent relative to $z$, because by assumption $x = \Phi(G)$ and $\Phi$ is truth-table relative to $z$, and by (the relativization of) Observation 26, there is a $\Psi$ truth-table relative to $z$ with $\Psi(G) = \Psi(\Phi(x)) = x$.

*(ii)* $\Rightarrow$ *(i)*: Let $\Phi$ and $\Psi$ be truth-table functionals relative to $z$ such that $\Psi(x) = G$ and $\Phi(G) = x$. Then $\Psi \circ \Phi$ is a truth-table functional relative to $z$ with $(\Psi \circ \Phi)(G) = G$. By the fixed-point lemma, there is a basic clopen set $[\sigma] \ni G$ such that $(\Psi \circ \Phi)|_{[\sigma]} = Id_{[\sigma]}$. Let $X := Range(\Phi|_{[\sigma]})$. Then $\Psi|_X$ gives a homeomorphism $X \to [\sigma]$, with inverse $\Phi|_{[\sigma]}$. Note that $X$ is $z$-uniformly perfect, since $\Phi'(\tau) := \Phi(\sigma^\frown\tau)$ is a one-to-one truth-table functional relative to $z$ with $\Phi'(2^\omega) = X$. We can check $x$ is 1-generic in $X$ relative to $z$: Suppose $U$ is a $z$-effective open$_X$ set dense along $x$. Then $\Phi^{-1}(U)$ is a $z$-effective open set in $2^\omega$. But $\Phi^{-1}(U) \supseteq \Psi|_X(U)$, which is dense along $G = \Psi|_X(x)$, because all topological properties are preserved by homeomorphisms, and $\Psi|_X$ is a homeomorphism $X \to [\sigma]$). Hence $\Phi^{-1}(U)$ is a $z$-effective open set dense along $G$, so $G \in \Phi^{-1}(U)$. But then $x = \Phi(G) \in \Phi(\Phi^{-1}(U)) = U$. Thus $x$ is $z$-1-generic in the $z$-uniformly perfect set $X := Range(\Phi|_{[\sigma]})$. □

The philosophical interpretation of this is that 1-genericity in $2^\omega$ is *universal* for 1-genericity in perfect spaces.

# Chapter 6

# Model Theoretic Constructions

We'll give an overview of some model-theoretic constructions over continuous logic, and discuss possible effectivizations of these.

## 6.1 Effective Type Omitting

**Definition 66.** Fix a language $\mathcal{L}$. A **condition** $E(x_1, \ldots, x_n)$ is a formal expression of the form $\phi(x_1, \ldots, x_n) \in A$, where $\phi(x_1, \ldots, x_n)$ is an $\mathcal{L}$-formula, and $A \subseteq [0, 1]$. We say that a tuple $(a_1, \ldots, a_n) \in \mathcal{M}^n$ **satisfies** E, or $\mathcal{M} \models E[a_1, \ldots, a_n]$, if $\phi^{\mathcal{M}}(a_1, \ldots, a_n) \in A$.

**Definition 67.** A condition $E$ is **closed** (resp. **open**) if $A \subseteq [0, 1]$ is **closed** (resp. **open**).

**Note.** If in some presentation of $\mathcal{M}$ we can compute $\phi^M(x_1, \ldots, x_n)$ to arbitrary precision, and $A$ is an effective closed (resp. open) set, then we can effectively falsify (resp. verify) whether a given $(a_1, \ldots, a_n) \in \mathcal{M}^n$ satisfies the condition $\phi(x_1, ..., x_n) \in A$.

**Proposition 28.** *Assuming we use the complete space of logical connectives, every closed condition is equivalent to one for which* $A = \{1\}$*.*

*Proof:* Define a unary logical connective $u : \mathbb{R} \to \mathbb{R}$ by $u(x) = \sup_{y \in C} \varepsilon(x, y)$. Then $\phi \in C$ if and only if $u(\phi) = 1$. It's easy to see $u$ is uniformly continuous (in fact, 1-Lipschitz).

**Definition 68.** An $\mathcal{L}$-theory $T$ is a set of a closed conditions with no free variables, where all conditions have $A = \{1\}$

**Question for the Reader.** Why do we use closed conditions for a theory?

**Answer:** We want a natural proof theory. In Gentzen's sequent calculus, we interpret $A_1, \ldots, A_k \vdash B_1, \ldots, B_j$ to mean what we might abbreviate in our usual terminology as $A_1 \wedge \cdots \wedge A_k \longrightarrow B_1 \vee \cdots \vee B_j$. This is logically equivalent in classical first order logic to what we might express in English as: "Either one of the $A_i$ fails to hold, or one of the $B_i$ holds, or both". If we abide by the convention that $A_i$ should be closed conditions, and $B_i$ should be open conditions, then this entailment relation turns out to itself be an open condition (as it is a disjunction of open conditions). As you only need a finite amount of information to verify an open condition, this allows semantic validities to have finite witnesses, i.e. proofs. Take a look at Ben Yaacov & Pedersen's "A Proof of Completeness for Continuous First-order Logic"[20] for a detailed account of a sound and complete proof system. Another reason to have a theory be formed by closed conditions is a worry about compactness: the set of open conditions $\{1 - 2^n < \varepsilon(a, b) < 1 : n \in \omega\}$ is finitely satisfiable, but not satisfiable. This is not a worry for closed conditions.

**Definition 69.** Suppose $A \subseteq \mathcal{M} \models T$. Let $T_A$ be the theory of $(\mathcal{M}, (a : a \in A))$. An **n-type over A** is a set $p$ of $\mathcal{L}(A)$-conditions of the form $\phi(\bar{x}) = 1$ with free variables $x_1, ..., x_n$ such that $T_A \cup p$ is consistent.

**Definition 70.** An $n$-type is **principal** if there exist formulas $\phi_k(x_1, ..., x_n)$ such that

$$\Big(\forall \bar{x} \in \mathcal{M}^n\Big)\varepsilon\Big(\phi_k(\bar{x}), \sup_{\bar{y} \in p(\mathcal{M})} \varepsilon(\bar{x}, \bar{y})\Big) \geq 2^{-2^{-k}}\Big)$$

**Definition 71.** For an $n$-type $p$, we define $p(\mathcal{M}) := \{\bar{x} \in \mathcal{M}^n : \mathcal{M} \models p(\bar{x})\}$. This is the set of tuples in $\mathcal{M}$ **realizing** the type $p$.

In what follows, we assume our continuous language $\mathcal{L}$ is relational. Recall $\mathcal{D}^{\mathcal{M}}(\bar{a})$ is the atomic diagram of $\bar{a}$.

**Definition 72.** Given a class $\mathbb{K}$ of $\mathcal{L}$-structures, we define $\mathbb{K}^{fin} = \{\mathcal{D}^{\mathcal{M}}(\bar{a}) : \mathcal{M} \in \mathbb{K}, \bar{a} \in \mathcal{M}^{<\omega}\}$.

**Theorem 29.** *Let $\mathcal{K}$ be a non-empty $\Pi_2^c$ class for which $\mathcal{K}^{fin}$ is c.e.. Then there is at least one computable structure in $\mathbb{K}$.*

**Proof Sketch:** We'll build a computable path through $K^{fin}$, giving an ascending chain of quasi-substructures $\mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq \ldots$, meeting finitely many requirements at each stage. The "limit" of this chain, $\mathcal{M}$, will be in $\mathbb{K}$.

We are trying to ensure in the limit that $\mathcal{M}$ satisfies a condition of the form:

$$\bigwedge_{i \in \omega} \inf_{\bar{y}_i} \bigvee_{j \in \omega} \sup_{\bar{x}_{ij}} \phi_{ij}(\bar{y}_i, \bar{x}_{ij}) = 1$$

To do this, it suffices to ensure that for any tuple $\bar{b}$ from $\mathcal{M}_s$ of length $|y_s|$, and any $i < s$, we have the following at the next stage $s + 1$:

$$\mathcal{M}_{s+1} \models \bigvee_{j \in \omega} \sup_{\bar{x}_{ij}} \phi_{ij}(\bar{b}, \bar{x}_{ij}) > 1 - 2^{-s}$$

But this is simply an search problem: we can enumerate $\mathbb{K}^{fin}$, and look for a structure $\mathcal{M}_{s+1}$ quasi-extending $\mathcal{M}_s$ (to the appropriate degree of accuracy), a $j \in \omega$, and a tuple $\bar{c} \in \mathcal{M}_{s+1}^{|\bar{x}_{ij}|}$ such that

$$\mathcal{M}_{s+1} \models \phi_{ij}(\bar{b}, \bar{c}) > 1 - 2^{-s}$$

Here we are glossing over some details dealing with error bounds (since we only know the values of atomic formulas in $\mathcal{M}_{s+1}$ up to some error).

Checking whether one atomic diagram "extends" another in an appropriate sense can be done effectively (since these are finite objects). And we are ultimately checking a finite number of open conditions, so once we have sufficiently tight error bounds on the values of the atomic formulas, we can verify positive instances of $\phi_{ij}(\bar{b}, \bar{c}) > 1 - 2^{-s}$.

Now just notice that the truth values of $\Sigma_c^1$ formulas can't decrease by more than an appropriate error bound when quasi-extending a structure. By choosing the degree of accuracy of the quasi-extension from $\mathcal{M}_s$ to $\mathcal{M}_{s+1}$ to be at least $1 - 2^{-(s+2)}$, we can conclude that for any $\bar{b} \in \mathcal{M}^{|\bar{y}_i|}$, any $i \in \omega$, and any $s \in \omega$, we have:

$$\mathcal{M} \models \bigwedge_{j \in \omega} \sup_{\bar{x}_{ij}} \phi_{ij}(\bar{b}, \bar{x}_{ij}) > 1 - 2^{-(s-1)}$$

since $\mathcal{M}$ is an accuracy $1 - 2^{-(s+1)}$ quasi-extension of $\mathcal{M}_s$, for all $s \in \omega$.

But this implies:

$$\mathcal{M} \models \bigwedge_{i \in \omega} \inf_{\bar{y}_i} \bigvee_{j \in \omega} \sup_{\bar{x}_{ij}} \phi_{ij}(\bar{y}_i, \bar{x}_{ij}) = 1$$

## Ultraproducts, Compactness, and Inspiration from Nonstandard Analysis

One way to prove the compactness theorem in classical first-order logic is through ultraproducts. If you have a countably-infinite theory $T = \{\phi_i : i \in \omega\}$ which is finitely consistent, then choose $\mathcal{M}_n \models \{\phi_i : i < n\}$ for each $n \in \omega$, take a non-principle ultrafilter $\mathcal{U}$ on $\omega$, and consider $\mathcal{M} := (\prod_n \mathcal{M}_n)/\mathcal{U}$, whose elements are equivalence classes of sequences of the form $\bar{a} = (a_n)_{n \in \omega}$ with $a_n \in \mathcal{M}_n$. Two elements $\bar{a}$ and $\bar{b}$ are equal in the ultraproduct (or equivalent in the product) if $\{n : a_n = b_n\} \in \mathcal{U}$. We say $\mathcal{M} \models R(\bar{a}^1, \ldots, \bar{a}^k)$ if $\{n : \mathcal{M}_n \models R(a_n^1, \ldots a_n^k)\} \in \mathcal{U}$. By induction on formula complexity, one can then prove Łoś's theorem that $\mathcal{M} \models \varphi$ if and only if $\{n : \mathcal{M}_n \models \varphi\} \in \mathcal{U}$. Since for each $i \in \omega$, $\{n : \mathcal{M} \not\models \varphi_i\}$ is not in $\mathcal{U}$ (since $\mathcal{U}$ is non-principle), we must have $\mathcal{M} \models \varphi_i$ for each $i \in \omega$. This can be generalized to theories of arbitrary cardinality by replacing $\omega$ with the set of finite subsets of $T$.

If we try to naively port ultraproducts to continuous logic, we run into some problems. Recall that a closed condition is something of the form $[\varphi \in A]$ for $\varphi$ a formula in our language and $A \subseteq [0,1]$ a closed set. The idea might be to define $\mathcal{M} \models [R(\bar{a}^1, \ldots \bar{a}^k) \in A]$ if and only if $\{n : \mathcal{M}_n \models [\varphi(a_n^1, \ldots, a_n^k) \in A]\} \in \mathcal{U}$. But unfortunately this doesn't quite work. Consider $A = \{1\}$, and continuous structures $\mathcal{M}_n$ over the signature with two constant symbols "a" and "b" such that $\mathcal{M}_n \models [\varepsilon(a,b) = 1 - 2^{-n}]$. Then $\{n : \mathcal{M}_n \models [\varepsilon(a,b) = 1]\}$ is the empty set, so certainly not in $\mathcal{U}$. But $\{n : \mathcal{M}_n \models [\varepsilon(a,b) \geq \tau]\}$ for any $\tau < 1$ is cofinite, so certainly in $\mathcal{U}$. This implies in the ultraproduct $\mathcal{M} \models [\varepsilon(a,b) \geq \tau]$ for every $\tau < 1$, but that $\mathcal{M} \not\models [\varepsilon(a,b) = 1]$. In other words, $\mathcal{M}$ must assign to $\varepsilon(a,b)$ a truth value strictly between $\tau$ and 1 for every $\tau < 1$. The difference between this truth value and 1 is thus a non-zero infinitesimal. We seem to be *forced* into using non-standard truth values.

However, there is a nice observation we can use from non-standard analysis. In non-standard analysis, a function $f : X \to Y$ between two metric spaces is uniformly continuous if any only if the non-standard extension of that function $f^* : X^* \to Y^*$ between their ultrapowers satisfies the property $\forall \bar{x}, \bar{y} \in X^*$, $\bar{x} \sim \bar{y} \Rightarrow f^*(\bar{x}) \sim f^*(\bar{y})$. To put it in other words, a function/relation is uniformly continuous **if and only if** it respects the equivalence relation of being infinitesimally close. This fact can be extended to ultraproducts of continuous spaces over the same language. In our case, for every $k$-ary relation symbol $R$, we have a relation $R^* : (\prod_n (\mathcal{M}_n)/\mathcal{U})^k \to (\prod_n [0,1])/\mathcal{U} = [0,1]^*$. To obtain a function to the standard interval $[0,1]$, we simply quotient both $(\prod_n \mathcal{M}_n)/\mathcal{U}$ and $(\prod_n [0,1])/\mathcal{U}$ by the equivalence relation $\sim$ defined by $\bar{x} \sim \bar{y} \Leftrightarrow (\forall j \in \omega)(\{n : \varepsilon(x_n, y_n) > 1 - 2^{-j}\} \in \mathcal{U})$. Since $\mathcal{R}$ is uniformly continuous with the same modulus in each of the $\mathcal{M}_n$, we can see that if $\bar{x} \sim \bar{y}$, then $R^*(\bar{x}) \sim R^*(\bar{y})$, and so quotienting both the domain and codomain of $R^*$ by $\sim$ yields a well-defined function $(((\prod_n \mathcal{M}_n)/\mathcal{U})/\sim)^k \to [0,1]^*/\sim$. Note $[0,1]^*/\sim$ is isomorphic to the standard interval $[0,1]$, so we end up with a standard predicate on $((\prod_n \mathcal{M}_n)/\mathcal{U})/\sim$.

So for continuous structures, we define the ultraproduct to be not $(\mathcal{M}_n)/\mathcal{U}$, but the further quotient $((\mathcal{M}_n)/\mathcal{U})/\sim$. This completely generalizes ultraproducts in classical logic, because in classical structures $\sim$ is trivial (roughly because $\{0,1\}^* = \{0,1\}$. A more detailed exposition of this, with a proof of Łoś's theorem can be found in Appendix D. This has of course been proven before for continuous logic based on metrics, but we figured it is worthwhile to show it still holds when we use equalness relations, and our proof which factors through the classical ultraproduct may be illuminating.

This relation of "being infinitesimally close" from non-standard analysis playing an important motivating role in defining ultraproducts of continuous structures raises the question of whether continuous logic might conversely be useful in non-standard analysis. The immediate stumbling block is that multiplication in the reals is not uniformly continuous. So while $\mathbb{R}$ as a 1-dimensional vector spaces over the reals has a relatively natural presentation as a continuous structure matching its usual metric, $\mathbb{R}$ as a ring does not. We showed how previously that we can create a theory of Hilbert spaces with a modified version of the inner product, of which multiplication on the reals is a special case, but we need to be careful because this modified multiplication does not necessarily behave as expected in elementary extensions.

There is a work-around, developed by Ben Yaacov, Henson, et. al. The idea would be to break up the reals into a countable collection of sorts corresponding to the intervals $[-n, n]$. Then multiplication would be broken up into a countable collection of operations, e.g. multiplication on the sort corresponding to $[-n, n]$ could map to sort corresponding to $[-k, k]$ for any $k \geq n^2$. There is no obstacle to treating arbitrary continuous functions like $e^x$, $\sin(x)$, etc. likewise, because their restrictions to any of these bounded intervals are bounded. The sacrifice we make by pulling this trick is that we can now only quantify over bounded intervals of reals. Because we have natural embeddings of "smaller" sorts into "larger" sorts, we have an atlas for $\mathbb{R}$ (in the sense of differential geometry). We have definitions of multiplication (or whatever other function we might like to add) on each map in the atlas, which are compatible with the transition maps of the atlas. From an outside perspective, we now have multiplication on all of $\mathbb{R}$, essentially by taking the union of the sorts and quotienting by the transition maps, as one does when one constructs a manifold from an atlas. But from an internal perspective, we only have the atlas, as our language lacks precisely the expressive capability needed to turn an atlas whose maps are given by sorts into a manifold. In particular, we cannot interpret the presentation of the reals as a discrete continuous structure into this presentation of the reals.

We should think of this work-around using sorts as being necessary if we want to keep our usual ring operation and also keep compactness. The following observation is instructive:

**Exercise for the Reader:** The ring of real numbers, presented as a continuous structure in the way suggested above with sorts, is isomorphic to all of its ultrapowers. Hint: for every $\tau < 1$, and any sequence $(a_i; i \in I) \subseteq [-n, n]^{\mathcal{I}}$, there is some $r_\tau \in [-n, n]$ with $\{i \in \mathcal{I} : \varepsilon(a_i, r_\tau) > \tau\} \in \mathcal{U}$.

In fact, a more general statement is true. If all the sorts of a structure are compact, then it is isomorphic to all of its ultrapowers. This is a generalization of the classical fact that any finite structure is isomorphic to all of its ultrapowers. So indeed it is a fairly substantial trade-off to accommodate multiplication via this work-around, for if we present the one-dimensional vector space of the reals as a continuous structure without sorts, its ultrapowers will have points an infinite distance apart (or in our terminology, entirely unequal points), but as a ring with sorts it will not. By using the work-around we are losing the ability to "take advantage" of the compactness theorem, because every consistent type will already be realized. Let's actually state this as a theorem:

**Theorem:** If $\mathcal{M}$ is a continuous structure all of whose sorts are compact, then $\mathcal{M}$ realizes every type over $\mathcal{M}$.

*Proof:* Take an ultrapower of $\mathcal{M}$ with a $|\mathcal{M}|^+$-complete ultrafilter. This will be isomorphic to $\mathcal{M}$, and realize every type over $\mathcal{M}$, so $\mathcal{M}$ realizes every type over $\mathcal{M}$.

All of this suggests that continuous logic doesn't have much to say about non-standard analysis. We can't add infinitesimals by an elementary extension, which is entire point of non-standard analysis. However, there is a positive spin on this. The fact that an ultraproduct of a compact structure is always canonically isomorphic to that structure in continuous logic is important for the philosophical thesis that compactness is for continuous structures what finiteness is for discrete structures. Logicians are no doubt familiar with the connection between topological compactness and the compactness theorem given by the fact that spaces of $n$-types are compact with a topology generated by formulas with $n$ free variables (essentially the compactness of Stone spaces). But we have here now a connection with the topology of the structure itself, not some abstract logical space of ideal points: if a continuous structure is compact, all its types are realized in that structure. In continuous logic, you can study a wide class of relatively rich structures *without* needing to move to saturated elementary extensions.

An example here that might be illustrative is the p-adic integers, which are the completion of the integers under a different norm. We gave a presentation of this structure as a computable continuous structure in Chapter 1. We now give another definition of the p-adics as a limit of $\mathbb{Z}/p^k\mathbb{Z}$. Our approach to thinking about the p-adic integers is not necessarily unique, because it's widely understood that the p-adic integers are a kind of limit of $\mathbb{Z}/p^k\mathbb{Z}$. But we think continuous logic can present the p-adic integers in a beautiful way which illustrates the interplay between logic and topology. It may be a good idea to re-read the presentation in Chapter 1 to recall the continuous structure on the $p$-adics (the ring operations are 1-Lipschitz).

Our idea is to think of $\mathbb{Z}_p$ as a as a limit of successively-better quasi-isomorphic structures $\mathbb{Z}/p^k\mathbb{Z}$. In [18], Ben Yaacov and others generalized the notion of Gromov-Hausdorff distance between metric spaces to metric structures by defining a "back-and-forth distance" between structures (which also has an equivalent syntactic definition). Using this idea, it's possible to define a notion of convergence of continuous structures. Their definition is a bit technical (although you have seen some of the main ideas in the section of quasi back-and-forth sets), but for our purposes now we don't need to go into it, because we are in a special case where this distance is easy to compute. On $\mathbb{Z}/p^k\mathbb{Z}$, we are using the equality predicate obtained from the metric $d(a,b) = p^{-\max\{i:p^i|(a-b)\}}$ if $a \neq b$, and 0 otherwise, which is basically the same definition we gave for the metric on $\mathbb{Z}_p$, but using divisibility in $\mathbb{Z}/p^k\mathbb{Z}$ instead. We can define $\tau_k = 2^{-p^{-k}}$ quasi-isomorphisms $f_k : \mathbb{Z}/p^k\mathbb{Z} \to \mathbb{Z}_p$ by sending $a \in \mathbb{Z}/p^k\mathbb{Z}$ identified with a number in $\{0, 1, 2, \ldots, p^k - 1\}$ to $1 + 1 + 1 + \cdots + 1$ ($a$ times) thought of as an element of $\mathbb{Z}_p$. We should say what we mean by a$\tau$-quasi-isomorphism. We'll give a not completely general definition here, because a general definition is a bit involved:

**Definition:** A $\tau$-quasi-isomorphism $f : \mathcal{M} \to \mathcal{M}$ between two continuous structures is a total function such that:

- For every $y \in \mathcal{N}$, exists $x \in \mathcal{M}$ with $\varepsilon(f(x), y) \geq \tau$
- For every atomic formula $\varphi(\bar{x})$, and any $\bar{a}$ from $\mathcal{M}$, $\varepsilon\big(\varphi(f(\bar{a})), \varphi(\bar{a})\big) \geq \tau$

In this case, every atomic formula is equivalent to one of the form $\varepsilon(q(x_0, \ldots, x_{j-1}), 0)$, where $q(x_0, \ldots, x_{j-1})$ is a multivariate polynomial with integer coefficients. We can see that the second condition for a $\tau_k$-quasi-isomorphism is satisfied by $f_k$, because there exists a map $g_k : \mathbb{Z}_p \mapsto \mathbb{Z}/p^k\mathbb{Z}$ which sends $x \in \mathbb{Z}_p$ to $x \bmod p^k \in \mathbb{Z}/p^k\mathbb{Z}$. $g_k \circ f_k$ is the identity map, and $g_k$ is a ring homomorphism. This means that $g_k(q(x_0, \ldots, x_{j-1})) = q(g_k(x_0), \ldots, g_k(x_{j-1}))$. $g_k$ also satisfies $\varepsilon(\varepsilon(g_k(x), g_k(y)), \varepsilon(x, y)) \geq 2^{-p^{-k}}$. This last fact is a long calculation, which we've put in Appendix E Using these facts, we can then see:

$$\varepsilon\left(\varepsilon\left(g_k\big(q_k(f_k(a_0), \ldots, f_k(a_{j-1}))\big), g_k(0)\right), \varepsilon\big(q_k(f_k(a_0), \ldots, f_k(a_{j-1})), 0\big)\right) \geq 2^{-p^{-k}}$$

$$\varepsilon\left(\varepsilon\left(q_k\big((g_k \circ f_k)(a_0), \ldots, (g_k \circ g_k)(a_{j-1})\big)\big), 0\right), \varepsilon\left(q_k\big(f_k(a_0), \ldots, f_k(a_{j-1})\big), 0\right)\right) \geq 2^{-p^{-k}}$$

$$\varepsilon\left(\varepsilon\left(q_k(a_0, \ldots, a_{j-1}), 0\right), \varepsilon\left(q_k(f_k(a_0), \ldots, f_k(a_{j-1})), 0\right)\right) \geq 2^{-p^{-k}}$$

$$\varepsilon\left(\varphi(a_0, \ldots a_{j-1}), \varphi(f_k(a_0), \ldots, f_k(a_{j-1}))\right) \geq 2^{-p^{-k}}$$

We can also verify the first property $\tau_k$ quasi-isomorphism for $f_k$: given $x \in \mathbb{Z}_p$, $f_k(g_k(x))$ is within $\tau_k$ of $x$, so every point inf $\mathbb{Z}_p$ is at least $\tau$-equal to an element of the range of $f_k$. $f_k$ is thus a $\tau_k$-quasi-isomorphism. It's not hard to see that $g_k$ is also a $\tau_k$-quasi-isomorphism (and in a sense, it is a $\tau_k$ inverse of $f_k$).

It's a standard argument by induction on the complexity of formulas that if we work in a finitary 1-Lipschitz language (a language with weak modulus $\Omega$ the universal 1-Lipschitz modulus), then if two continuous structures over that language are $\tau$-quasi-isomorphic, then the truth value of the interpretations in those two structures of any formula $\varphi$ of that language are at least $\tau^{\frac{1}{d}}$-equal, where $d$ is the maximum depth of nesting of quantifiers (sups and infs) in $\varphi$. When $\tau = 1$, this is simply saying that isomorphic structures have the same theory. This more general claim can be summarized by saying that quasi-isomorphic structures have quasi-equal theories. In our case, because the $\tau_k$ converge to 1, this means is that the theory of $\mathbb{Z}_p$ is equal to the limit of the theories of $\mathbb{Z}/p^k\mathbb{Z}$. And we have an explicit bound on the rate of convergence of the truth value of any given sentence, meaning given *any* sentence $\varphi$ in the 1-Lipschitz continuous language of rings with a maximum quantifier nesting depth of $d$, we have a uniform-in-$k$ procedure for evaluating its truth in $\mathbb{Z}_p$ to accuracy $\tau_k^{1/d}$: simply evaluate that formula in $\mathbb{Z}/p^k\mathbb{Z}$. Since $\mathbb{Z}/p^k\mathbb{Z}$ are finite structures, we can evaluate any formula there in time exponential in the length of that formula by a brute force evaluation: just evaluate $\sup_x \psi(x)$ or $\inf_x \psi(x)$ recursively by checking all $p^k$ possibilities for $x$. Each nested quantifier corresponds to another nested search over a space of size $p^k$. In particular the continuous theory of $\mathbb{Z}_p$ is decidable, in the sense that we have a uniform procedure to compute the truth values of sentences to arbitrary accuracy.

This decidability result is not unique to the p-adic integers. In fact it holds for any "effectively compact" computable continuous structure. We define a computable, continuous structure $\mathcal{M}$ to be **effectively compact** if we can compute, uniformly in $\tau < 1$, a finite $\tau$-covering of $\mathcal{M}$. A **$\tau$-covering** $X \subseteq \mathcal{M}$ is a set such that every $p \in \mathcal{M}$ is $\tau$-equal to some point in $X$. The idea is that if $\mathcal{M}$ is effectively compact, and $\psi(x)$ a 1-Lipschitz formula, we can use the $\tau$-covering $X$ to approximate the truth value of $\inf_x \psi(x)$ to within $\tau$ by the finite conjunction $\bigwedge_{x\in X} \psi(x)$ (likewise with $\sup_x \psi(x)$). One way of viewing this is model-theoretic terms is that in a 1-Lipschitz continuous language (or generally any continuous language on which we impose a weak modulus), if a type is realized in a structure by some point, then nearby points approximately realize that type (or rather, realize an approximately equal type). This means even if we are working in an infinitary language (with a weak modulus, as we will always assume), we always have the $\tau$-equivalence of $\inf_x \psi(x)$ and $\wedge_{x\in X}\psi(x)$ for *any* formula in that language whenever $X$ is a $\tau'$-covering of $\mathcal{M}$, where $\tau'$ depends only on $\tau$ and the weak modulus $\Omega$ (it is equal to $\tau$ if $\Omega$ is the universal 1-Lipschitz modulus). In the case that $\mathcal{M}$ is effectively compact, we have a uniform procedure for finding this $X$. However, in computable infinitary languages, this doesn't allow us to compute every formula, because we can only simplify quantifiers, not infinite conjunctions or disjunctions. But if we *are* working in a finitary language (as we have been with the p-adic integers), and our continuous structure is computable and effectively compact, then that structure has a decidable theory. This is analogous to the fact that finite computable structures in classical first-order logic have decidable theories, but we now have the ability to apply it to a much wider range of structures, like the p-adic integers.

Going back to our p-adic integers example, we've explained why the continuous theory of $\mathbb{Z}/p^k\mathbb{Z}$ converges to the continuous theory of $\mathbb{Z}_p$, but there is also a sense in which the structures themselves converge. One way to see this is to note that $h_{jk} := g_j \circ f_k$ will also quasi-isomorphism $\mathbb{Z}/p^k\mathbb{Z} \to \mathbb{Z}/p^j\mathbb{Z}$, in fact, a $\tau_{jk}$ quasi-isomorphism with $\tau_{jk} = \min(2^{-p^{-k}}, 2^{-p^{-j}})$. If we consider the family of these maps with $j \leq k$, we get a directed system whose inverse limit is $\mathbb{Z}_p$. This is the standard algebraic approach to constructing the p-adic integers, and the reason algebraists prefer it is that the maps $h_{jk}$ preserve addition and multiplication exactly instead of just approximately (they are classical ring homomorphisms). However viewing them as continuous structures, we also see it is equally valid to consider the directed system of these maps $h_{jk}$ for $j \geq k$. If we consider this family of these maps with $j \geq k$, and take the direct limit of this directed system, we also get $\mathbb{Z}_p$. This fits more in line with the model-theoretic approach to constructing a structure by a chain of embeddings (rather than quotients). Continuous logic allows us to generalize this classical limit of a chain of embeddings to allow converging chains of quasi-embeddings.

We think that this is a useful way to think about $\mathbb{Z}_p$ because it explains better than the inverse limit definition why the theory of $\mathbb{Z}_p$ is connected to that of $\mathbb{Z}/p^k\mathbb{Z}$ better than the inverse limit definition. It is useful to consider Hensel's lifting lemma in this context. This tells us that if $p(a) = 0$ and $p'(a) \neq 0$ in $\mathbb{Z}/p^k\mathbb{Z}$, then we have some $a' \in \mathbb{Z}/p^{k+1}\mathbb{Z}$ with $g_k(a') = a$ and $p(a') = 0$, $p'(a) \neq 0$. This works more generally for polynomials with coefficients in $\mathbb{Z}_p$ (we can add these polynomials to our language without changing our previous results). This allows us, in certain special cases, to compute the value of a formula exactly, not just compute its value to arbitrary precision. And it indicates a general technique to study infinite structures: construct that infinite structure as a direct limit of quasi-isomorphic substructures for which you have an easy finite combinatorial relation between the truths of certain formulas in one of those finite substructures to another.

# Chapter 7

# Why limit truth values to $[0, 1]$?

## 7.1 $\mathcal{T} = [0, 1]$ as a continuous structure

Our set of truth values, $\mathcal{T} = [0, 1]$ is playing double-duty: it is both a "multiplicative value semigroup" for continuous equality predicates on continuous structures, and a space of values for relations. But we may want to think of it as a sort of canonical continuous structure in its own right, analogous to the way that $\{0, 1\}$ can be thought of as a the simplest Boolean algebra (it is the initial object in the category of Boolean algebras with $0 \neq 1$ and morphisms Boolean homomorphism). This is somewhat tricky, though. We are, in a sense, thinking of $\mathcal{T} = [0, 1]$ as a one-point compactification $\mathbb{R}_{\geq 0} \cup \{\infty\}$ of the non-negative reals, with $0 \in \mathbb{R}_{\geq 0}$ corresponding to $1 \in \mathcal{T}$, and $\infty$ corresponding to $0 \in \mathcal{T}$.

Our first question is: what is the appropriate continuous equality predicate on $\mathcal{T}$? One proposal would be to use $\varepsilon^{\mathcal{T}}(\tau_1, \tau_2) = 2^{-|log_2(\tau_1)/log_2(\tau_2)|} = \min(\frac{\tau_1}{\tau_2}, \frac{\tau_2}{\tau_1})$. The idea here is that these truth-values $\tau$ usually arise from distances, and these distances obey a multiplicative version of the triangle inequality: $\varepsilon(x, z) \geq \varepsilon(x, y) * \varepsilon(y, z)$, or in terms of distances $2^{-d(x,z)} \geq 2^{-d(x,y)} * 2^{-d(y,z)}$. Unfortunately, negation $\tau \mapsto 1 - \tau$ is not continuous with respect to this continuous equality predicate, nor are alternative negations like $\tau \mapsto 2^{1/log_2(\tau)}$ which one might consider to be more natural. The basic problem is that the topology on $\mathbb{R}_{\geq 0} \cup \infty$ does not arise from a metric for which $\mathbb{R}_{\geq 0}$ is a metric subspace with its usual metric. A solution to this is simply to use the metric $d(\tau_1, \tau_2) = |\tau_1 - \tau_2|$, and corresponding continuous equality predicate $2^{-|\tau_1 - \tau_2|}$. An unsatisfactory feature of this, however, is that iterated equality does not agree with classical equality, in the sense that

$$((0 =^{\mathcal{T}} 1) =^{\mathcal{T}} 1) = (0 =^{\mathcal{T}} 1)^{\mathcal{T}} = 0 = (1 =^{\mathcal{T}} 0)$$

for $\mathcal{T} = \{0, 1\}$, while

$$\varepsilon^{\mathcal{T}}(\varepsilon^{\mathcal{T}}(0, 1), 1)) = \varepsilon^{\mathcal{T}}(\frac{1}{2}, 1) = \frac{1}{\sqrt{2}} \neq \frac{1}{2} = \varepsilon^{\mathcal{T}}(0, 1)$$

for $\mathcal{T} = [0,1]$. But an answer to this is that it only makes sense to consider iterated equality in a few special circumstances (e.g. comparing equality in two quasi-isomorphic structures, as in Appendix **??**). What matters more for us a specific enumeration of a base for the topology of $\mathcal{T}$ (up to computable re-ordering), for which Turing functional computable total functions $\mathcal{T}^n \to \mathcal{T}$ induce computable functions from unions of recursively enumerable open sets in the codomain to unions of recursively enumerable families basic open sets in the domain. We should think of uniform continuity as a computational notion, rather than a purely topological notion: its purpose is to give us a means to estimate the values of a function in small open balls around a point given the value at the center of that ball, and it depends on the specific way that we enumerate a base for the topology of our space. The precise equality predicate we use is not as important as the uniform structure that it imposes.

The classical set of truth values, $\{0,1\}$, is used because it is the smallest set which is sufficient for separating elements in an arbitrary set by functions to it. It can also be considered the initial object in the category of non-degenerate[1] Boolean algebras. The set of all homomorphisms from a particular Boolean algebra to $\{0,1\}$ naturally corresponds with the set of all ultrafilters on that Boolean algebra, which can be thought of as consistent assignments of truth to propositions, given the relations between the propositions imposed by that Boolean algebra. In the realm of first-order logic, we can think of n-ary predicates on a first-order structure $\mathcal{M}$ as functions $\mathcal{M}^n \to \{0,1\}$, and complex formulas can be thought of as being built from simple operations on these functions spaces. For example, existential quantification is just projection along one of the coordinates.

We can make an analogy to the role of $\{0,1\}$ in classical logic to various other "classifiers" in mathematics. In classical homotopy theory, one considers maps from $S^n$ into a manifold to study its topology. We classify manifolds by looking at the structure of the space of all continuous functions from a simple manifold to the space under consideration, up to homotopy equivalence. In Morse theory, one considers differentiable maps from a differentiable manifold to $\mathbb{R}$, and classifies manifolds by invariants computed from the critical points of those functions. In analysis, the underlying field, usually $\mathbb{R}$ or $\mathbb{C}$, serves a role as a classifier. Here the appropriate functions are bounded linear functionals, and one often uses the weakest topology for which these linear functionals are continuous. One might ask the same question here: why should we limit ourselves to only maps from $S^n$, or only maps to $\mathbb{R}$? Ultimately, the real explanation is that they are sufficient to develop a general theory which has rich enough consequences. The set of truth values $\{0,1\}$ has more than proven its sufficiency in classical logic.

[1]Some people allow Boolean Algebras to have $0 = 1$ (or $\bot = \top$ in another terminology), because it gives Boolean algebras a universal theory. But there are reasons to exclude this, similar to the reasons why algebraists usually exclude $0 = 1$ in the definition of fields. We use the term "non-degenerate" to clarify we are requiring $0 \neq 1$.

Continuous logic can perhaps be motivated by these examples in other areas of mathematics. The three examples given all make crucial use of topology. From a high-level perspective, a topology restricts the the ways in which you are allowed to categorize the objects of the domain of a structure. If we only use constructions which obey this restriction (e.g., using only continuous functions), we can prove more powerful results about the structures so constructed, because we have excluded pathological situations. We gave the example in a previous chapter of the pathological property of two orbiting bodies being rational distance apart. This leads to a motivation for topology arising purely from considerations in the philosophy of science. Roughly, the axioms of a topology can be thought of as axioms of verifiable properties:

- The conjunction of two verifiable properties is verifiable
- Arbitrary disjunctions of verifiable properties are verifiable
- Trivially necessary and trivially impossible properties are verifiable

These correspond, respectively, to the following axioms of a topology $\tau$ on a set $X$:

- $U, V \in \tau \implies U \cap V \in \tau$
- $U_\alpha \in \tau$ for $\alpha \in I \implies \bigcup_{\alpha \in I} U_\alpha \in \tau$
- $\emptyset \in \tau$ and $X \in \tau$

In this sense, topological spaces can be seen as an alternative to Boolean algebras as a means of putting a structure on a space of properties. The absence of closure under complementation (negation) can be considered a feature in philosophy of science, because it corresponds to the fact that the negation of a verifiable property may not be verifiable. Falsifiable properties are dual to verifiable properties: closed sets in a topological space can be thought of as corresponding to falsifiable properties. In the special case where the topology is induced by a metric, verifiability of a property corresponds to the fact that if an object has the given property, there is some finite degree of error (measured according to the metric) for which the result of a measurement with that degree of error can guarantee the property holds of that object.

The relevance of this to a theory of computability is that computations (classically) can only make use of a finite amount of information. The way computers are physically instantiated, even for discrete computation, requires a careful analysis of error bounds. A mental picture of the abstraction of Turing machines can illustrate some general technique for reducing these errors, e.g. using only two symbols, using a tape with uniform spacing, using a head that moves only at unit speed, having a finite set of clearly defined state transitions for the head, etc. The purpose of this model was to illustrate one direction of the strong Church-Turing thesis: that every recursive function can (in an ideal world) be computed by a physical machine. In fact in the historical development of digital computers, there were serious concerns over the ability to compensate for such errors, so much so that John von Neumann was forced by public criticism to write a paper *Probabilistic logics and synthesis of reliable organisms from unreliable components*[12] proving a fault-tolerance theorem. These concerns recently re-emerged with the burgeoning development of quantum computers, and a corresponding theorem has been proven here called the "threshhold theorem", which gives a bound on the error rate below which it is possible to perform computations which give the correct answer with arbitrarily high probability. The connection between computability theory and topology is quite deep. The class of functions $2^\omega \to 2^\omega$ which can be computed by a Turing functional relative to some oracle or other is equal to the class of all continuous functions $2^\omega \to 2^\omega$.

This reflection seems to indicate that perhaps computability theory should not be isolated to the realm of the discrete. The explanation for why so much of computability theory has been developed for discrete structures over a classical logic base can be explained more by our human tendency to want binary answers to questions for the sake of clarity. By moving to continuous logic, we start to see interesting phenomena like the distinction between definability and codefinability, corresponding to falsifiability and verifiability, respectively. Note that falsifiability corresponds to definability, which might be the opposite of what we might expect. The reason for this is that definability is a closed condition. The way definability is usually used is in the antecedent of a conditional: when you quantify over a definable set X, the condition $x \in X$ is in the antecedent, and if we want the entire statement to be an open condition, then the antecedent should be a closed condition. In philosophy of science, this would correspond to the fact that scientific theories tend to be falsifiable, and only verifiable in the much weaker sense that we can find inductive evidence for them by doing many experiments which do not falsify the theory. So this provides one motivation to not use a continuous space for our set of truth values rather than a discrete set of truth values.

But perhaps more can be said about why we choose $[0, 1]$ instead of some other continuous space as our set of truth values. Supposing we keep the rest of our logic the same (requiring some form of uniform continuity, having function symbols generating term algebras, etc.), but change predicates to take values in a more general space $\mathcal{T}$ of truth values, what might we stand to gain or lose by changing $\mathcal{T}$?

The first thing we can notice is that $[0, 1]$ is compact. What happens if we allow a non-compact $\mathcal{T}$? Then it seems we will lose the compactness theorem: pick an open cover $\{U_i : i \in I\}$ of $\mathcal{T}$ that doesn't have a finite subcover, and consider the closed conditions $[R(a) \in U_i^c]$. These are finitely consistent, because the condition $[R(a) \in \bigcap_{j<n} U_{i_j}^c]$ is realized by setting $R(a) = \tau$ for some $\tau \in \mathcal{T} \setminus \bigcup_{j<n} U_{i_j}$ which is non-empty because $\bigcup_{j<n} U_{i_j}$ does not cover $\mathcal{T}$.[2] But its inconsistent because $\bigcap_i U_i^c = \emptyset$. Also recall that our construction

[2]It's worth noting that Ben Yaacov et. al in [19] allow predicates to take values in $\mathbb{R}$, which is not compact, but they build into the language constraints for each predicate to take values in a compact interval

of ultraproducts of continuous structures depended on the fact that $\big((\prod_i \mathcal{T})/\mathcal{U})\big)/\sim$ is isomorphic to $\mathcal{T}$, where $\sim$ is the "infinitesimally close" relation, and this depended on $\mathcal{T}$ being compact.

Another thing to point out about $[0, 1]$ is that by using a continuous equality predicate (equalness relation) rather than a metric, we could represent in an easy way metric spaces which have points arbitrarily far apart (even infinitely far apart), so using $[0, 1]$ rather than some larger interval did not require us to sacrifice the ability to represented unbounded metric spaces.

Finally, $[0, 1]$ is totally ordered and has all sups and infs, which allows us to make sense of quantifiers in our language, in a way that an arbitrary compact continuous space used as our set of truth values would not, and it has an algebraic structure which allows us to express a multiplicative analog of the triangle inequality, which enables us to express a weak kind of transitivity for equalness.

There are in fact other reasonable choices here, however. In the early 80s Kopperman developed a theory of "value semigroups", which allow a substantial generalization of metric spaces where every topological space can be represented. A good overview of this idea (as well as the related idea of value quantales) is in Ittay Weiss's *Value semigroups, value quantales, and positivity domains*[17]. The general theory of value semigroups and value quantales is a bit technical, however, so we have chosen to use $[0, 1]$ to not obscure our presentation, with the trade-off being less generality in our results.

(possibly different for different predicates), which allows the proof of the compactness theorem to still go through. This is equivalent to having a multi-sorted language each of whose sorts has only predicates taking values in a fixed compact interval depending on the sort.

# Appendix A

# Miscellaneous Facts about Moduli

## Moduli and Limits

**Fact:** Suppose $(f_i : i \in \omega)$ and $f$ are functions $M \to N$ between two continuous spaces. If $(f_i : i \in \omega)$ converges to $f$ pointwise, and each $f_i$ obeys modulus $\Lambda$, then $f$ obeys modulus $\Lambda$.

*Proof:* Fix $x, y \in M$. We can compute

$$\varepsilon(f(x), f(y)) \geq \varepsilon(f(x), f_i(x))\varepsilon(f_i(x), f_i(y))\varepsilon(f_i(y), f(y))$$

$\varepsilon(f(x), f_i(x))$ converges to 1 in $i$, because $(f_i : i \in \omega)$ converges to $f$ pointwise. Likewise, $\varepsilon(f_i(y), f(y))$ converges to 1 in $i$. But also we have $\varepsilon(f_i(x), f_i(y)) \geq \Lambda(\varepsilon(x, y))$ for each $i$, since $f_i$ obeys modulus $\Lambda$. Taking a liminf over $i$, we find

$$\varepsilon(f(x), f(y)) \geq \liminf_i \varepsilon(f(x), f_i(x))\varepsilon(f_i(x), f_i(y))\varepsilon(f_i(y), f(y)) \geq \liminf_i \varepsilon(f_i(x), f_i(y)) \geq \Lambda(\varepsilon(x, y))$$

.

**Fact:** If $f : M \to N$ is a continuous function between two continuous spaces, and $f$ obeys modulus $\Lambda$ on a dense set of points $D \subseteq M$, it obeys the modulus everywhere.

*Proof:* Fix $x, y \in M$. We can find $(x_i : i \in \omega)$ and $(y_i : i \in \omega)$ in $D^\omega$ converging to $x$ and $y$, respectively. Then

$$\varepsilon(f(x), f(y)) \geq \varepsilon(f(x), f(x_i))\varepsilon(f(x_i), f(y_j))\varepsilon(f(y_j), f(y)) \geq \varepsilon(f(x), f(x_i))\Lambda(\varepsilon(x_i, y_j))\varepsilon(f(y_j), f(y))$$

Since $f$ is continuous, $\varepsilon(f(x), f(x_i))$ converges to 1 in $i$, and $\varepsilon(f(y_j), f(y))$ converges to 1 in $j$, so taking a liminf over $i$ and $j$, and using the fact that $\Lambda$ is monotonic and continuous, we find

$$\varepsilon(x, y) \geq \liminf_{i,j} \Lambda(\varepsilon(x_i, y_j)) = \Lambda(\liminf_{i,j} \varepsilon(x_i, y_j)) = \Lambda(\varepsilon(x, y))$$

**Note:** neither of these proofs depended on the continuous spaces being complete or separable.

## Moduli for Infimums and Supremums

Our definitions of $\Lambda_{sup_x \phi}$ and $\Lambda_{\inf_x \phi}$ might seem strange, because they suggest that taking an infimum or supremum over a variable should make a function better behaved, in the sense that it now obeys a (non-strictly) tighter modulus. This is counter-intuitive, but true:

**Fact:** Let $\phi(\bar{x})$ be a function $\prod_{i<n} \mathcal{M}_i \to [0,1]$ obeying modulus $\Lambda_\phi$. Then for $i < n$, $sup_{x_i}\phi$ obeys $\Lambda_{sup_{x_i}\phi}$ defined by $\Lambda_{sup_{x_i}\phi}(\tau_0, ..., \tau_{n-2}) = \Lambda_\phi(\tau_0, ...., \tau_{i-1}, 1, \tau_i, ..., \tau_{n-2})$.

*Proof:* Fix $i < n$. Fix $a_j, b_j \in M_j$ for $j \neq i$, $j < n$. Without loss of generality, assume $\sup_{x_i} \phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}) \geq \phi(b_0, ..., b_{i-1}, b, b_{i+1}, ..., b_{n-1})$ for all $b \in M$. Choose a sequence $(p_k : k \in \omega)$ from $M_i$ (not necessarily convergent) so that $\sup_{x_i} \phi(b_0, ..., b_{i-1}, x_i, b_{i+1}, ..., b_{n-1}) = \lim_k \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})$ Then we can compute:

$$\begin{aligned}
&\varepsilon\big(\sup_{x_i} \phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}), \sup_{x_i} \phi(b_0, ..., b_{i-1}, x_i, b_{i+1}, ..., b_{n-1})\big) \\
&= \varepsilon\big(\sup_{x_i} \phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}), \lim_k \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})\big) \\
&= \lim_k \varepsilon\big(\sup_{x_i} \phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}), \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})\big) \\
&= \liminf_k \varepsilon\big(\sup_{x_i} \phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}), \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})\big) \\
&\geq \liminf_k \sup_a \varepsilon\big(\phi(a_0, ..., a_{i-1}, a, a_{i+1}, ..., a_{n-1}), \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})\big)
\end{aligned}$$

$$\geq \liminf_k \sup_a \Bigg(\varepsilon\big(\phi(a_0, ..., a_{i-1}, a, a_{i+1}, ..., a_{n-1}), \phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1})\big)$$

$$*\varepsilon\big(\phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1}), \phi(b_0, ..., b_{i-1}, p_k, b_{i+1}, ..., b_{n-1})\big)\Bigg)$$

$$\geq \liminf_k \sup_a \Bigg(\varepsilon\big(\phi(a_0, ..., a_{i-1}, a, a_{i+1}, ..., a_{n-1}), \phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1})\big)$$

$$*\Lambda_\phi\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), \varepsilon(p_k, p_k), \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)\Bigg)$$

$$= \liminf_k \sup_a \Bigg( \varepsilon\big(\phi(a_0, ..., a_{i-1}, a, a_{i+1}, ..., a_{n-1}), \phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1})\big)$$

$$*\Lambda_\phi\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), 1, \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)\Bigg)$$

$$= \liminf_k \sup_a \Bigg( \varepsilon\big(\phi(a_0, ..., a_{i-1}, a, a_{i+1}, ..., a_{n-1}), \phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1})\big)\Bigg)$$

$$*\Lambda_\phi\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), 1, \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)$$

$$\geq \liminf_k \Bigg( \varepsilon\big(\phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1}), \phi(a_0, ..., a_{i-1}, p_k, a_{i+1}, ..., a_{n-1})\big)\Bigg)$$

$$*\Lambda_\phi\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), 1, \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)$$

$$= \Lambda_\phi\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), 1, \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)$$

$$= \Lambda_{\inf_{x_i}\phi}\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)$$

Thus $\sup_{x_i}\phi$ obeys the modulus $\Lambda_{\sup_{x_i}\phi}$. Note the conversion of limits to liminfs is just because we always know liminfs exist (so we don't need to ensure the terms inside when we break things up converge individually).

By a symmetric argument, we can also show:

$$\varepsilon\big(\inf_{x_i}\phi(a_0, ..., a_{i-1}, x_i, a_{i+1}, ..., a_{n-1}), \inf_{x_i}\phi(b_0, ..., b_{i-1}, x_i, b_{i+1}, ..., b_{n-1})\big)$$
$$\geq \Lambda_{\inf_{x_i}\phi}\big(\varepsilon(a_0, b_0), ..., \varepsilon(a_{i-1}, b_{i-1}), \varepsilon(a_{i+1}, b_{i+1}), ..., \varepsilon(a_{n-1}, b_{n-1})\big)$$

so we also have that $\inf_{x_i}\phi$ obeys the n-ary modulus $\Lambda_{\inf_{x_i}\phi}$.

**Note:** $\inf_{x_i}\phi$ also obeys $\Lambda_{sup_x\phi}$, and vice versa: in fact they are the same modulus.

# Appendix B

# Decidability

The appropriate analog of decidability for a relation $R$ on a computable continuous structure $\mathcal{M}$ is that $R$, thought of as a function $\mathcal{M}^n \to [0,1]$, has a computable name. $R$ is semidecidable if $R$ has a computable lower name, cosemidecidable if $R$ has a computable upper name (equivalently, if $\neg R$ has a computable lower name), and decidable if $R$ is both semidecidable and cosemidecidable.

The equalness relation on any computable continuous space is decidable, and this can be used to show that every computable continuous structure has a computably isomorphic presentation with an injective enumeration of its countable dense subset:

**Theorem:** Let $\mathcal{M} = (M, \varepsilon), D, (q_i)_{i\in\omega}, [\hat{\varepsilon}])$ be a computable continuous space, where $(q_i)_{i\in\omega}$ is a not-necessarily injective enumeration of $D$, and $[\hat{\varepsilon}]$ is a computable function $(\omega^2) \times \omega \to \mathbb{Q} \cap [0,1]$ such that $|[\hat{\varepsilon}(i,j)(k) - \varepsilon(q_i, q_j)| \geq 2^{-k}$ for all $i, j, k \in \omega$. Then there is a computable continuous space $\mathcal{M}' = (M, \varepsilon), D', (q'_i)_{i\in\omega}, [\hat{\varepsilon'}])$ which is computably isomorphic to $\mathcal{M}$, uniformly in $\mathcal{M}$.

*Proof:* We define $\iota : \omega \to \omega$ by induction by setting $\iota(0) = 0$, and defining $\iota(i+1) = k$, where $\langle k, l\rangle$ is the least element of the recursive set of codes[1] for pairs

$$\{\langle k, l\rangle \in \omega : [\hat{\varepsilon}](\iota(j), k)(l) < 2^{-2^{-(l-1)}} \text{ for all } j \leq i\}$$

If we define $q'_i = q_{\iota(i)}$, then $(q'_i)_{i\in\omega}$ is injective $\omega \to D$, because we defined $\iota$ so that we always have a witness to the unequalness of $q_{\iota(i+1)}$ with $q_{\iota(0)}, q_{\iota(1)}, \ldots, q_{\iota(i)}$. On the other hand, the sequence $(q'_i)_{i\in\omega}$ is dense in $D$. Suppose to the contrary that it is not dense. Let $k_0$ be the least such that $q_{k_0}$ is not in the closure of $\{q'_i : i \in \omega\}$. Choose $l_0$ to be the least such that $[\varepsilon](\iota(i), k_0)(l_0) < 2^{-2^{-(l_0-1)}}$ for every $i \in \omega$. Pick $i_0$ so that for all $k < k_0$, for some $i \leq i_0$, $[\varepsilon](\iota(i), k)(l_0) \geq 2^{-2^{-(l_0-1)}}$. Then at stage $i_0 + 1$ of our induction, we have that $\langle k_0, l_0\rangle$ is the least element of $\{\langle k, l\rangle \in \omega : [\hat{\varepsilon}](\iota(j), k)(l) < 2^{-2^{-(l-1)}}$ for all $j \leq i\}$, so $k_0$ should have been

[1] We are using the coding $\langle x, y\rangle = 2^{(i+1)}3^{(j+1)}$.

$\iota(i_0+1)$, a contradiction. So $(q'_i)_{i\in\omega}$ is a one-to-one enumeration of a dense subset of $D$. Now set $D' = \{q'_i : i \in \omega\}$, and $[\hat{\varepsilon}'](i,j) = [\hat{\varepsilon}](\iota(i), \iota(j))$. Since $D'$ is dense in $D$, and $D$ is dense on $(M, \varepsilon)$, $D'$ is dense in $M$. A computable isomorphism $\mathcal{M}' \to \mathcal{M}$ can be obtained from the embedding of $D'$ into $D$ induced by $\iota$. Because our construction was uniform in $\mathcal{M}$, this isomorphism is uniform in $\mathcal{M}$.

**Comments:** The basic idea here was to delay enumerating a point into $D'$ (possibly indefinitely) until we had computed the distance from that point to the elements we have already enumerated into $D'$. There are two cases: either we enumerate it, in which case that point will be in $D'$; or else we must enumerate points arbitrarily equal to it, in which case that point is in the closure of $D'$. This theorem tells us that whether or not we require the enumeration of the dense subset of $\mathcal{M}$ to be one-to-one or not, we get essentially the same class of computable continuous spaces. To strengthen this result to continuous structures, just note that the computable isomorphism $\phi : \mathcal{M}' \to \mathcal{M}$ we construct is the identity on $(M, \varepsilon)$ (it only differs in its presentation), so given any relation $R$ or function $f$ computable on $\mathcal{M}$, we can pull it back to a computable relation/function $R'$ or $f'$ on $\mathcal{M}'$ by defining $R'(\bar{x}) = R(\phi(\bar{x}))$ or $f(\bar{x}) = \phi^{-1}(f(\phi(\bar{x})))$, and $\phi$.

# Appendix C

# Computing Moduli and Functions

**Theorem:** Suppose $\Lambda : [0,1]^n \to [0,1]$ is an $n$-ary modulus, and we have a computable lower name $[\hat{\Lambda}] : \mathbb{Q}^n \cap [0,1]^n \to [0,1]$ such that $\lim_k [\hat{\Lambda}](\bar{q})(k) = \Lambda(\bar{q})$ for every $\bar{q} \in \mathbb{Q}^n \cap [0,1]^n$. Then $\Lambda$ has a lower name computable by a Turing functional.

*Proof:* We'll give a uniform procedure to compute $\Lambda$ from below. Suppose that $(\bar{q}_i)_{i\in\omega}$ is a coordinatewise-non-decreasing sequence converging to $\bar{\tau} \in [0,1]^n$ from below. Consider $q'_j := \max_{k<j} [\hat{\Lambda}](\bar{q}_k)(j)$. Since $[\hat{\Lambda}]$ is a non-decreasing function, $q'_j$ is also non-decreasing. We can also compute $q'_j$ uniformly from $(\bar{q}_i)_{i\in\omega}$. Note that since $\Lambda$ is continuous, for any $\epsilon > 0$, and the $\bar{q}_k$ converge to $\tau$ coordinatewise, for some $k_0$, $\Lambda(\bar{\tau}) - \Lambda(\bar{q}_{k_0}) < \epsilon/2$. But also eventually in $j$, $\Lambda(\bar{q}_{k_0}) - [\hat{\Lambda}](\bar{q}_{k_0})(j) < \epsilon/2$. Chaining these together, we get that eventually in $k$, eventually in $j$, $\Lambda(\bar{\tau}) - [\hat{\Lambda}(\bar{q}_k)(j)] < \epsilon$. So $(q'_j)_{j\in\omega}$ converges to $\Lambda(\tau)$ from below.

**Theorem:** Suppose that $f : \mathcal{M} \to \mathcal{N}$ obeys the lower computable modulus $\Lambda_f$, $D$ is a dense subset of $\mathcal{M}$, and $[\hat{f}]$ is a computable name for $f|_D : D \to \mathcal{N}$. Then $f$ has a name computable by a Turing functional.

*Proof:* Suppose $[x] = (q_i)_{i\in\omega}$ is a fast Cauchy sequence from $D$ converging to $x \in \mathcal{M}$. To compute $f(x)$ to within accuracy $2^{-2^{-k}}$, lower compute in parallel for all $j$ the modulus $\Lambda_f(2^{-2^{-j}})$ until you find a $j(k)$ for which $\Lambda_f(2^{-2^{-j(k)}}) > 2^{-2^{-(k+1)}}$, and define

$$[f]([x])(k) := [\hat{f}](q_{j(k)})(k+1)$$

.
Then we can see

$$\varepsilon([f][x](k), f(x)) \geq \varepsilon([\hat{f}](q_{j(k)})(k+1), f(q_{j(k)}))\varepsilon(f(q_{j(k)}), f(x))$$

$$\geq 2^{-2^{-(k+1)}} \Lambda_f(\varepsilon(q_{j(k)}, x)) \geq 2^{-2^{-(k+1)}} 2^{-2^{-(k+1)}} = 2^{-2^{-k}}$$

# Appendix D

# Compactness and Łoś's Theorem

The basic idea of how to prove the compactness theorem for continuous logic: take a classical ultraproduct of structures satisfying finite subtheories of a given theory, and then quotient by an "infinitesimally close" relation. In order for the quotient to be a well-defined continuous structure, we will use that the functions and relations are all uniformly continuous with moduli defined in the language (so the same for every structure over that language).

Recall that a theory consists of a set of closed conditions $[\varphi \in A]$ for $A \subseteq [0,1]$ a closed set, and if we have a rich enough set of logical connectives, we can assume without loss of generality that $A = \{1\}$ for each condition. We will work with single-sorted structures, although this can be generalized to multi-sorted structures by a similar, but more verbose, argument.

Let $\mathcal{I}$ be an index set, and $\mathcal{U}$ an ultrafilter on $\mathcal{P}(\mathcal{I})$. Let $(\mathcal{M}_i : i \in \mathcal{I})$ be an $\mathcal{I}$-indexed family of continuous structures over the same continuous language $\mathcal{L}(\sigma,\nu)$. We first consider $(\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U}$, whose elements equivalence classes of sequences $\bar{a} = (a_i : i \in \mathcal{I})$ in $\prod_{i\in\mathcal{I}} \mathcal{M}_i$ under the equivalence $\bar{a} \equiv \bar{b}$ iff $\{i \in \mathcal{I} : a_i = b_i\} \in \mathcal{U}$. We then define a further equivalence relation, the "infinitesimally close" relation, by

$$\bar{a} \sim \bar{b} \iff (\forall \tau < 1)\{i \in \mathcal{I} : \varepsilon(a_i, b_i) > \tau\} \in \mathcal{U}$$

For an $n$-ary relation symbol $R$ in our language, we define

$$R^* : \big((\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U}\big)^n \to (\prod_{i\in\mathcal{I}} [0,1])/\mathcal{U}$$

by

$$R^*([\bar{a}^0], \ldots, [\bar{a}^{n-1}]) = [(R^{\mathcal{M}_i}(a_i^0, \ldots, a_i^{n-1}) : i \in \mathcal{I})]$$

where $[\bar{a}]$ denotes the equivalence class of $\bar{a}$ modulo $\equiv$. We likewise define for $f$ an $n$-ary function symbol in our language a function

$$f^* : \big((\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U}\big)^n \to (\prod_{i\in\mathcal{I}} \mathcal{M})/\mathcal{U}$$

by

$$f^*([\bar{a}^0], \ldots, [\bar{a}^{n-1}]) = [(f^{\mathcal{M}_i}(a_i^0, \ldots, a_i^{n-1}) : i \in \mathcal{I})]$$

Our task now is to show that these functions $R^*$ and $f^*$ respect the equivalence relation on $\sim$ in the sense that the following commutative diagrams can be completed with $f$ and $R$ respectively, where $\pi$ is the quotient map for $\sim$ on $(\prod_{i\in\mathcal{I}} \mathcal{M})/\mathcal{U}$, and $\pi'$ is the quotient map for $\sim$ on $(\prod_{i\in\mathcal{I}}[0,1])/\mathcal{U}$:

$$\begin{array}{ccc} (\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U} & \xrightarrow{f^*} & (\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U} \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ ((\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U})/\sim & \overset{f}{\dashrightarrow} & ((\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U})/\sim \end{array} \qquad \begin{array}{ccc} (\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U} & \xrightarrow{R^*} & (\prod_{i\in\mathcal{I}}[0,1]/\mathcal{U} \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi'} \\ ((\prod_{i\in\mathcal{I}} \mathcal{M}_i)/\mathcal{U})/\sim & \overset{R}{\dashrightarrow} & ((\prod_{i\in\mathcal{I}}[0,1]/\mathcal{U})/\sim \end{array}$$

Let's look only at functions, since the argument for relations is the same. Suppose that $\bar{a}^0 \sim \bar{b}^0$, $\bar{a}^1 \sim \bar{b}^1$, ..., and $\bar{a}^{n-1} \sim \bar{b}^{n-1}$. Fix $\tau < 1$, and recall that we have for each $i \in \mathcal{I}$

$$\varepsilon(f^{\mathcal{M}_i}(a_i^0, \ldots, a_i^{n-1}), f^{\mathcal{M}_i}(b_i^0, \ldots, b_i^{n-1})) \geq \Lambda_f(\varepsilon^{\mathcal{M}_i}(a_i^0, b_i^0), \ldots, \varepsilon^{\mathcal{M}_i}(a_i^{n-1}, b_i^{n-1}))$$

But then we have for any $\tau < 1$ and any $\tau_0, \ldots \tau_{n-1} < 1$ such that $\Lambda_f(\tau_0, \ldots, \tau_{n-1}) > \tau$:

$$\{i \in \mathcal{I} : \varepsilon(f(a_i^0, \ldots a_i^{n-1}), f(b_i^0, \ldots b_i^{n-1})) > \tau\} \supseteq \bigcap_{j<n} \{i \in \mathcal{I} : \varepsilon(a_i^j, b_i^j) > \tau_j\} \in \mathcal{U}$$

because $\mathcal{U}$ is closed under finite intersections. Since $\lim_{\tau_0,\ldots,\tau_{n-1}} \Lambda_f(\tau_0, \ldots, \tau_{n-1}) = 1$, we can conclude $f^*([\bar{a}^0], \ldots, [\bar{a}^{n-1}]) \sim f^*([\bar{b}^0], \ldots, [\bar{b}^{n-1}])$. This means that quotienting both the domain and codomain of $f^*$ by $\sim$ yields a well-defined function. This gives an an interpretation of function symbols in the continuous ultraproduct, and a similar argument holds for relations. In fact, the same proof yields a well-defined interpretation for terms in the term algebra and formulas in our language in the continuous ultraproduct.

One important question is whether these interpretations actually turn the continuous ultraproduct into a continuous $\mathcal{L}(\sigma, \nu)$ structure. The follow observation tells us that are formulas in the continuous ultraproduct are in fact $[0,1]$-valued:

**Observation:** $\big(\prod_{i\in\mathcal{I}}[0,1]\big)/\mathcal{U})/\sim$ is naturally isomorphic to $[0,1]$.

*Proof:* Let $\bar{\tau} \in \prod_{i\in\mathcal{I}}[0,1]$. For $\sigma \in 2^{<\omega}$, define

$$I_\sigma = \big[\sum_{j<|\sigma|} 2^{-\sigma(i)}, 2^{-|\sigma|} + \sum_{j<|\sigma|} 2^{-\sigma(i)}\big] \subseteq [0,1]$$

For example $I_0 = [0, \frac{1}{2}]$, $I_1 = [\frac{1}{2}, 1]$, $I_{10} = [\frac{2}{4} \cdot \frac{3}{4}]$, $I_{11} = [\frac{3}{4}, 1]$. If $\sigma \subseteq \sigma'$, then $I_\sigma \supseteq I_{\sigma'}$, and for any length $l \in \omega$, $\{I_\sigma : |\sigma| = l\}$ cover $[0,1]$. We now define a sequence of strings by induction: $\sigma_0 = \langle\rangle$, the empty string, and $\sigma_{k+1} \supseteq \sigma_k$ is defined to be $\sigma_k{}^\frown 0$ if $\{i : \tau_i \in I_{\sigma_k{}^\frown 0}\} \in \mathcal{U}$, and

$\sigma_k^\frown 1$ otherwise. Because $\mathcal{U}$ is an ultrafilter, we can prove by induction on $k$ that for all $k \in \omega$, $\{i : \tau_i \in I_{\sigma_k}\} \in \mathcal{U}$. Because $[0,1]$ is compact, $\bigcap_{k\in\omega} I_{\sigma_k} \neq \emptyset$, and since their diameters go to 0, it is a singleton. Let $\tau_\infty \in [0,1]$ be this singleton. Now consider the constant sequence $\bar{\tau}_\infty$, considered as an element of the ultraproduct. We can see by construction that $\bar{\tau} \sim \bar{\tau}_\infty$, so they correspond to equal elements in the continuous ultraproduct (after quotienting by $\sim$). But this means that the diagonal embedding $[0,1] \to \big((\prod_{i\in\mathcal{I}}[0,1])/\mathcal{U}\big)/\sim$ is actually a surjection. It's not hard to verify it is also one-to-one, and preserves the equalness relation, so $[0,1]$ and $\big((\prod_{i\in\mathcal{I}}[0,1])/\mathcal{U}\big)/\sim$ are isomorphic.

A natural question, now, is: are formulas in the ultraproduct assigned truth values consistent with the semantics of continuous logic? The classical analog of this question is answered in the affirmative by Łoś's Theorem. Let's prove the continuous analog of this. Note that we *allow* infinitary logical connectives (like lim), as long as they are uniformly continuous, so we have a new phenomenon here! Continuous logic can remain compact even when adding non-trivial infinitary connectives. We do need to restrict to languages with a weak modulus, however (e.g., the universal 1-Lipschitz modulus), however.

**Continuous Łoś's Theorem:** Let $\psi(\bar{x})$ be an $\mathcal{L}_\Omega(\sigma,\nu)$ formula. Let $\mathcal{M}$ be the ultraproduct of $(\mathcal{M}_i : i \in \mathcal{I})$. We'll prove that $\psi^{\mathcal{M}}(\bar{x})$ (the interpretation of $\psi(\bar{x})$ in $\mathcal{M}$ defined inductively by the semantics of continuous logic) is equal to $\psi*(\bar{x})/\sim$, the function obtained more directly from $\psi^*(\bar{x})$ by quotienting both the domain and codomain by $\sim$). The semantics for continuous logic is compositional, and formulas are defined inductively by closure under certain logical operations, so it suffices to show that each of these operations respects the semantics clauses of continuous logic. Since our space of logical connectives is quite large, we cannot just do a proof by cases, so we'll use the following lemma:

**Lemma**: Suppose $\mathcal{M} = \big((\prod_{i\in\mathcal{I}}\mathcal{M}_i)/\mathcal{U}\big)/\sim$, $\mathcal{N} = \big((\prod_{i\in\mathcal{I}}\mathcal{N}_i)/\mathcal{U}\big)/\sim$, and $\mathcal{K} = \big((\prod_{i\in\mathcal{I}}\mathcal{K}_i)/\mathcal{U}\big)/\sim$, where $\mathcal{M}_i$,$\mathcal{N}_i$, and $\mathcal{K}_i$ are continuous spaces. Suppose $f_i : \mathcal{M}_i \to \mathcal{N}_i$ each obey modulus $\Lambda_f$, and $g_i : \mathcal{N}_i \to \mathcal{K}_i$ each obey modulus $\Lambda_g$. Then for $f := \big((\prod_{i\in\mathcal{I}} f_i/\mathcal{U}\big)/\sim$, and $g := \big((\prod_{i\in\mathcal{I}} f_i/\mathcal{U}\big)/\sim$, $f \circ g : \mathcal{M} \to \mathcal{K}$ obeys modulus $\Lambda_g \circ \Lambda_f$, and $f \circ g = \big((\prod_{i\in\mathcal{I}} f_i/\mathcal{U}\big)/\sim$.

In fact we need a more general version of this lemma for $n$-ary (and even infinitary) functions, but this can be done by replacing a higher arity function on $\mathcal{M}$ by a unary function on a "power" of $\mathcal{M}$. We have not said what an infinite product of continuous spaces should be, but we will do this after we finish the proof of this theorem. There is some subtlety, because we want the topology to agree with the product topology, but it is well-known the the category of metric spaces with 1-Lipschitz maps as morphism does not have arbitrary products.

Continuing our proof of the theorem, let's start with terms. If $t(\bar{x}) = f(t_0(\bar{x}),\ldots,t_{n-1}(\bar{x}))$, then our lemma tells us that $(t^*/\sim)(\bar{x}) = (f^*/\sim)\big((t_0^*/\sim)(\bar{x}),\ldots,(t_{n-1}^*/\sim)(\bar{x})\big)$ and it obeys the composition their moduli. For atomic formulas, if $\psi(\bar{x}) = R(t_0(\bar{x}),\ldots,t_{n-1}(\bar{x}))$, then our lemma tells us that $(\psi^*/\sim)(\bar{x}) = (R^*/\sim)\big((t_0^*/\sim)(\bar{x}), \ldots, (t_{n-1}^*/\sim)(\bar{x})\big)$ and it obeys the composition their moduli. For the outermost construction in $\psi$ being a continuous logical connective, say $\psi(\bar{x}) = u(\varphi_0(\bar{x}), \varphi_1(\bar{x}), \varphi_2(\bar{x}), \ldots)$ for $u$ a continuous (possibly infinitary) connective, then (the generalization of) our lemma says that

$$\psi^*(\bar{x})/\sim = (u*/\sim)((\varphi_0^*/\sim)(\bar{x}), (\varphi_1^*/\sim)\big(\bar{x}), (\varphi_2^*/\sim)(\bar{x}), \ldots\big)$$

It remains only to check the cases $\psi(\bar{x}) = \sup_y \varphi(x^0, \ldots, x^{k-1}, y)$ and $\psi(\bar{x}) = \sup_y \varphi(x^0, \ldots, x^{k-1}, y)$. We'll check only sup, since the argument for inf is symmetric:

$$((\sup_y \varphi)^*/\sim)(\bar{x}^0, \ldots, \bar{x}^{k-1}) \leq \tau \iff \{i \in \mathcal{I} : \sup_{y_i \in \mathcal{M}_i} \varphi(x_i^0, \ldots, x_i^{k-1}, y_i) \leq \tau\} \in \mathcal{U}$$

$$\iff \{i \in \mathcal{I} : (\forall y_i \in \mathcal{M}_i)(\varphi(x_i^0, \ldots, x_i^{k-1}, y_i) \leq \tau)\} \in \mathcal{U}$$

$$\iff \left(\forall (y_i : i \in I) \in \prod_{i \in I} \mathcal{M}_i\right)\{i \in \mathcal{I} : \varphi(x_i^0, \ldots, x_i^{k-1}, y_i) \leq \tau)\} \in \mathcal{U}$$

$$\iff \left(\forall \bar{y} \in \mathcal{M}\right)(\varphi^*/\sim)(\bar{x}^0, \ldots, \bar{x}^{k-1}, \bar{y}) \leq \tau \iff \sup_{\bar{y} \in \mathcal{M}} \mathcal{M}_i(\varphi^*/\sim)(\bar{x}^0, \ldots, \bar{x}^{k-1}, \bar{y}) \leq \tau$$

This is essentially the standard idea from the classical proof of Łoś's Theorem. □

*Proof of Lemma:* Suppose $g_\alpha^i : \prod_j \mathcal{M}_\alpha^{ij} \to \mathcal{N}_\alpha^i$ obey moduli $\Lambda_{ij}$ and $f_\alpha : \prod_i \mathcal{N}_\alpha^i \to \mathcal{K}_\alpha$ obey modulus $\Lambda_j$. Consider for each $\alpha \in \mathcal{I}$ their composition $h_\alpha := f_\alpha \circ \bar{g}_\alpha$. Let's first look at $(g^i)^* : \prod_j \mathcal{M}^{ij*} \to \prod_i \mathcal{N}^{i*}$, $f^* : \prod_i \mathcal{N}^{i*} \to \mathcal{K}^*$, and $h^* : \prod_{ij} \mathcal{M}^{ij*} \to \mathcal{K}^*$ (i.e. as functions between the classical ultraproducts, before quotienting by $\sim$.[1] We want first to show that $h^* = f^* \circ \bar{g}^*$. Let $x^i j_k \in \prod_j \mathcal{M})k^{ij}$. We use for notational simplification that $\bar{x}^i$ refers to the double array of points $x_\alpha^{ij}$ for $i$ fixed. Then:

$$(f^* \circ \bar{g}^*)([\bar{x}]) = f^*((g^0)^*([\bar{x}^i]), (g^1)^*([\bar{x}^1]), (g^2)^*([\bar{x}^2]), \ldots)$$

$$= f^*\big([(g_\alpha^0(\bar{x}_\alpha^0) : \alpha \in \mathcal{I})], [(g_\alpha^1(\bar{x}_\alpha^1) : \alpha \in \mathcal{I})], [(g_\alpha^2(\bar{x}_\alpha^2) : \alpha \in \mathcal{I})], \ldots\big)$$

$$= \left[\left(f_\alpha\big(g_\alpha^0(\bar{x}_\alpha^0), g_\alpha^1(\bar{x}_\alpha^1), g_\alpha^2(\bar{x}_\alpha^2), \ldots\big) : \alpha \in \mathcal{I}\right)\right]$$

$$= \left[\left((f_\alpha \circ \bar{g}_\alpha)(\bar{x}_\alpha) : \alpha \in \mathcal{I}\right)\right] = \left[\big(h_\alpha(\bar{x}_\alpha) : \alpha \in \mathcal{I}\big)\right] = h^*([\bar{x}])$$

It follows that $(h^*/\sim) = (f^*/\sim) \circ (\bar{g}^*/\sim)$. Note that $g_\alpha^i$ obeys $\Lambda_j$ for each $j$, so $(g^i)^*/\sim$ also obeys $\Lambda_j$ (since closed conditions holding in each factor in an ultraproduct must also hold in the ultraproduct). Likewise $f^*/\sim$ must obey $\Lambda$. Since $(h^*/\sim) = (f^*/\sim) \circ (\bar{g}^*/\sim)$, and the composition of functions obeys the composition of their moduli, $h^*$ obeys modulus $\Lambda(\Lambda_0, \Lambda_1, \Lambda_2, \ldots)$.

[1] The classical ultraproduct of a product space is just the product of the ultraproduct of the factors.

# Appendix E

# Notes on Equalness

Sometimes we want to know the equalness of the equalness of two pairs of elements in two different structures. This ends up being somewhat complicated to compute in continuous logic, so in the following we describe some useful techniques to make these calculations.

**Fact 1:** For any $d \geq 0$, $1 - 2^{-d} \leq d$.

*Proof:* Let $f(d) = d - (1 = 2^{-d})$. We can see $f(0) = 0$, and compute for $d \geq 0$:

$$f'(d) = 1 - \ln(2)2^{-d} \geq 1 - \ln(2)2^{-0} = 1 - 0.693 \cdots > 0$$

So $f(d) \geq 0$ for all $d \geq 0$. But then $d \geq 1 - 2^{-d}$ for all $d \geq 0$.

**Fact 2:** Suppose our equalness predicates in two structures arise from respective metrics by $\varepsilon(x, y) = 2^{-|d(x,y)|}$ Then $\varepsilon(\varepsilon(a, b), \varepsilon(x, y)) \geq 2^{-|d(a,b)-d(x,y)|}$.

*Proof:* Assume without loss of generality $\varepsilon(a, b) \geq \varepsilon(x, y)$. Then

$$\varepsilon(a, b) - \varepsilon(x, y) = 2^{-d(a,b)} - 2^{-d(x,y)} = 2^{-d(a,b)}(1 - 2^{-(d(x,y)-d(a,b))})$$

$$\leq 1 - 2^{-(d(x,y)-d(a,b))} \leq d(x, y) - d(a, b)$$

where this last inequality is due to fact 1. By symmetry, then, we always have:

$$|\varepsilon(a, b) - \varepsilon(x, y)| \leq |d(x, y) - d(a, b)|$$

Taking 2 to the negative of both sides:

$$\varepsilon(\varepsilon(a, b), \varepsilon(x, y)) = 2^{-|\varepsilon(a,b)-\varepsilon(x,y)|} \geq 2^{-|d(x,y)-d(a,b)|}$$

# Bibliography


[1] Scott Aaronson. *Quantum Computing since Democritus*. Cambridge University Press, 2013. DOI: 10.1017/CBO9780511979309.

[2] Bernard A. Anderson. "Reals n-generic relative to some perfect tree". In: *Journal of Symbolic Logic* 73.2 (2008), pp. 401–411. DOI: 10.2178/jsl/1208359051.

[3] Olivier Bournez et al. "The General Purpose Analog Computer and Computable Analysis are Two Equivalent Paradigms of Analog Computation". In: *Theory and Applications of Models of Computation*. Ed. by Jin-Yi Cai, S. Barry Cooper, and Angsheng Li. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 631–643. ISBN: 978-3-540-34022-5.

[4] C.C. Chang et al. *Continuous Model Theory*. Annals of mathematics studies. Princeton University Press, 1966. ISBN: 9780691079295. URL: https://books.google.com/books?id=uTGdPSI5rI4C.

[5] Toby S. Cubitt, David Perez-Garcia, and Michael M. Wolf. "Undecidability of the spectral gap". In: *Nature* 528 (2015/12/09/online), 207–?? DOI: 10.1038/nature16059. URL: https://doi.org/10.1038/nature16059.

[6] EKATERINA B. FOKINA et al. "ISOMORPHISM RELATIONS ON COMPUTABLE STRUCTURES". In: *The Journal of Symbolic Logic* 77.1 (2012), pp. 122–132. ISSN: 00224812. URL: http://www.jstor.org/stable/23208238.

[7] Dominic Hyde and Diana Raffman. "Sorites Paradox". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2018. Metaphysics Research Lab, Stanford University, 2018.

[8] J.F. Knight. "Chapter 7 Degrees of models". In: *Handbook of Recursive Mathematics*. Ed. by Yu. L. Ershov et al. Vol. 138. Studies in Logic and the Foundations of Mathematics. Elsevier, 1998, pp. 289–309. DOI: https://doi.org/10.1016/S0049-237X(98)80008-6. URL: http://www.sciencedirect.com/science/article/pii/S0049237X98800086.

[9] Alexander G. Melnikov. "Computably isometric spaces". In: *J. Symbolic Logic* 78.4 (Dec. 2013), pp. 1055–1085. DOI: 10.2178/jsl.7804030. URL: https://doi.org/10.2178/jsl.7804030.

[10] Antonio Montalbán. *Computable Structure Theory*. Berkeley, California: (draft), 2018.

[11] Y.N. Moschovakis. *Descriptive Set Theory*. Studies in Logic and the Foundations of Mathematics. Elsevier Science, 1987. ISBN: 9780080963198. URL: `https://books.google.com/books?id=c7FQWlkA3KIC`.

[12] J. von Neumann. "Probabilistic logics and synthesis of reliable organisms from unreliable components". In: *Automata Studies*. Ed. by C. Shannon and J. McCarthy. Princeton University Press, 1956, pp. 43–98.

[13] Jan Reimann and Theodore A. Slaman. "Measures and their random reals". In: *Transactions of the American Mathematical Society* 367 (Feb. 2008). DOI: `10.1090/S0002-9947-2015-06184-4`.

[14] M. P. Solèr. "Characterization of hilbert spaces by orthomodular spaces". In: *Communications in Algebra* 23.1 (1995), pp. 219–243. DOI: `10.1080/00927879508825218`. eprint: `https://doi.org/10.1080/00927879508825218`. URL: `https://doi.org/10.1080/00927879508825218`.

[15] Wehner Stephan. "Enumerations, Countable Structures and Turing Degrees." In: *Proceedings of the American Mathematical Society* 126.7 (1998), p. 2131. ISSN: 00029939. URL: `https://libproxy.berkeley.edu/login?qurl=http%3a%2f%2fsearch.ebscohost.com%2flogin.aspx%3fdirect%3dtrue%26db%3dedsjsr%26AN%3dedsjsr.118896%26site%3deds-live`.

[16] Slaman Theodore A. "Relative to Any Nonrecursive Set." In: *Proceedings of the American Mathematical Society* 126.7 (1998), p. 2117. ISSN: 00029939. URL: `https://libproxy.berkeley.edu/login?qurl=http%3a%2f%2fsearch.ebscohost.com%2flogin.aspx%3fdirect%3dtrue%26db%3dedsjsr%26AN%3dedsjsr.118894%26site%3deds-live`.

[17] Ittay Weiss. "Value semigroups, value quantales, and positivity domains". In: *Journal of Pure and Applied Algebra* 223.2 (2019), pp. 844–866. ISSN: 0022-4049. DOI: `https://doi.org/10.1016/j.jpaa.2018.05.002`. URL: `http://www.sciencedirect.com/science/article/pii/S0022404918301257`.

[18] I. B. Yaacov et al. "Metric Scott analysis". In: *Advances in Mathematics* 318 (2017), pp. 46–87. DOI: `doi:10.1016/j.aim.2017.07.021`.

[19] I. B. Yaacov et al. "Model theory for metric structures". In: *Model Theoy with Applications to Algebra and Analysis* (1000), pp. 315–427. DOI: `doi:10.1017/cbo9780511735219.011`.

[20] Itaï Ben Yaacov and Arthur Paul Pedersen. "A proof of completeness for continuous first-order logic". In: *The Journal of Symbolic Logic* 75.1 (2010), pp. 168–190. DOI: `10.2178/jsl/1264433914`.

[21] L.A. Zadeh. "Fuzzy sets". In: *Information and Control* 8.3 (1965), pp. 338–353. ISSN: 0019-9958. DOI: `https://doi.org/10.1016/S0019-9958(65)90241-X`. URL: `http://www.sciencedirect.com/science/article/pii/S001999586590241X`.