MODEL EXISTENCE AND SCOTT'S THEOREM FOR INFINITARY REAL-VALUED LOGIC: A TOPOLOGICAL PERSPECTIVE

OLIVIA AUBONE, EDUARDO DUEÑEZ, AND JOSÉ IOVINO

To Xavier Caicedo

ABSTRACT. In this paper we adopt a topological perspective to extend two theorems of classical infinitary logic to the infinitary logic of Keisler's general real-valued structures: the omitting types theorem and Scott's theorem. There the formulas need not be uniformly bounded, and no moduli of continuity are required, as they are in the metric extensions of the classical theorems; neither equality nor a metric is required. The omitting types theorem holds for every countable fragment, because the space of Henkin types of the fragment is Čech-complete. Scott's theorem characterizes every structure with countable sorts, among all structures, up to approximate equivalence, and the classes of this equivalence are Borel sets in the Polish space of structures on the natural numbers. With the monoid of surjections of the natural numbers in place of the group of permutations, we also prove López-Escobar's theorem for classical infinitary logic without equality.

1. INTRODUCTION

In this paper we extend two theorems of classical infinitary logic, the omitting types theorem and Scott's theorem, to the infinitary extension of Keisler's model theory for real-valued structures [Kei23], and we give each a topological form. In Keisler's general structures, function symbols are interpreted arbitrarily and predicate symbols by bounded real-valued functions with no moduli of continuity; the formulas need not have a common bound, but each carries a bound of its own, and neither logical equality nor a prescribed metric is required. In Part I, we characterize model existence through Henkin types (Theorem A), and the omitting types theorem becomes a Baire category argument in the Čech-complete space of Henkin types of a countable fragment (Theorem B). In Part II, Scott's theorem becomes a

2020 *Mathematics Subject Classification.* Primary 03C75; Secondary 03C66, 03C25, 03B50, 54E52.

Key words and phrases. infinitary logic, real-valued logic, continuous model theory, omitting types theorem, Henkin types, Čech-complete spaces, Baire category theorem, logic without equality, Scott's theorem, López-Escobar's theorem, Polish monoids.

This version: November 13, 2026. The most recent version of this paper is available at `https://utsa-quantum-math.github.io/resources/omitting-types-real-valued-logic/`; it may supersede the version deposited on arXiv.

statement about the Polish space of structures on the natural numbers, where surjections, acting on atomic diagrams, take the place of permutations.

For countable fragments of classical $L_{\omega_1\omega}$, the omitting types theorem is proved in Keisler's book by the model existence theorem [Kei71, Chapters 3 and 11], after Makkai had extended Smullyan's method of consistency properties to the infinitary logics on admissible sets [Mak69]. Keisler later proved a general omitting types theorem by forcing with finite sets of sentences [Kei73], and Morley derived the theorem for countable fragments from the Baire category theorem [Mor74], so that the omission of countably many types became the density of a countable intersection of open dense sets. All of these theorems presuppose equality. Keisler and Miller carried the omitting types theorem into first-order logic without equality, building the model from closed terms and testing the types to be omitted through tuples of terms [KM01, Section 4]; their setting is two-valued and finitary, and we follow them in building models from closed terms.

In first-order logic, the axioms of equality assert its reflexivity and the interchangeability of equal elements in every atomic formula, and thereby make equality a congruence for the function and predicate symbols. In real-valued logic, their counterparts are axioms of continuity.[1] The role of equality passed to a uniform structure with moduli prescribed by the signature, in Iovino's work on stable Banach space structures [Iov94, Iov99a, Iov99b], in Henson and Iovino's model theory of normed space structures [HI02, Sections 2 and 3], and in the continuous logic of Ben Yaacov and Usvyatsov [BU10, Section 2]. For a countable signature, the uniform structure has a countable base, and it is therefore induced by a single pseudometric (see [Eng89, Section 8.1]).

Within this metric setting Eagle, drawing in part on work of Caicedo and Iovino [CI14], applied Čech's form of the Baire category theorem to a space of countable metric structures [Eag14]. His theorem is formulated for structures whose signature prescribes a master metric, the real-valued counterpart of equality, together with a modulus of uniform continuity for every symbol, and whose predicates take values in $[0, 1]$ [Eag14, Definitions 1.4 and 1.5]; formulas that are not uniformly bounded, predicates without moduli of continuity, and structures without a prescribed master metric lie outside that setting. Keisler then obtained the pseudometric from the formulas themselves [Kei23, Section 3.3], and showed that general structures stand to first-order structures without equality as metric structures with a distance predicate stand to first-order structures with equality [Kei23, Section 1]; his model theory is first-order. Theorems A and B hold in a logic that none of these settings contains, infinitary, real-valued and free of each of the assumptions on which these settings rest.

The infinitary language $L_{\omega_1\omega}$ extends Keisler's first-order language by countable joins and infima, each sentence taking its values in a bounded interval of its own, so that the formal types, families of such values, are the points of a product of

[1] In Łukasiewicz predicate logic, the congruence axioms for a similarity, or graded equality, make every predicate 1-Lipschitz for the distance given by one minus the similarity (see [Háj98, Section 5.6] and the earlier work collected there).

compact intervals. We expand the vocabulary by countably many new constants of each sort and build models from closed terms, which themselves serve as elements; since the vocabulary has no equality symbol, no quotient is taken. The Henkin types, defined by the real-valued form of Henkin's requirement of witnesses, are exactly the types of such term structures. The key observation of Part I is that, for a countable vocabulary and a countable fragment, the space of Henkin types of the fragment is a G_δ subset of a compact space, hence Polish and Čech-complete, so that Morley's method operates in that space.

A type of the fragment in a finite tuple $\bar{x}$ of variables is a set of expressions $\varphi(\bar{x}) \geq r$, with r rational and with φ a formula of the fragment whose free variables are among $\bar{x}$. The type is omitted by a Henkin type of the fragment that satisfies none of the instances obtained by substituting closed terms of the expanded vocabulary for these variables.

Theorem A. Let a countable vocabulary be given, together with a formal type for a set of sentences of $L_{\omega_1\omega}$ for this vocabulary. The formal type is realized by a structure whose sorts are countable if and only if it extends to a Henkin type. If the given sentences belong to a countable fragment, then the formal type is realized by some structure if and only if it extends to a Henkin type of the fragment.

Theorem B. Let a countable fragment of $L_{\omega_1\omega}$ for a countable vocabulary be given, together with a theory on sentences of the fragment that has a model and a sequence of types of the fragment. Suppose that, for every instance of every type of the sequence, the Henkin types of the fragment that satisfy the theory and the instance form a set with empty interior in the space of Henkin types of the fragment that satisfy the theory. Then some Henkin type of the fragment that satisfies the theory omits every type of the sequence, and the theory has a model whose sorts are countable and in which no tuple of elements satisfies any type of the sequence.

The conclusion is generic in the Henkin space: under the hypothesis of Theorem B, the Henkin types satisfying the theory and omitting every type of the sequence form a dense G_δ subset of the space of Henkin types satisfying the theory. Omission can therefore be imposed together with any finite collection of strict rational conditions on sentences of the fragment whose corresponding basic open set in that space is nonempty. The countable model is obtained from a point of this dense set by the term construction. Theorem A is proved in Section 9 (Proposition 5), and Theorem B in Section 11 (Theorem 8), after the Čech completeness of the space of Henkin types is established in Section 10 (Corollary 7).

The hypothesis of Theorem B is the topological form of the hypothesis of Keisler's general omitting types theorem [Kei73, p. 108, General Omitting Types Theorem]. In classical two-valued logic, with Keisler's $\forall\exists$ formulas taken over the set of all formulas of the fragment, a model omits a type $\Gamma(\bar{x})$ exactly when it satisfies the sentence $\forall\bar{x} \bigvee_{\gamma\in\Gamma} \neg\gamma(\bar{x})$. For this sentence Keisler asks that every finite set of sentences satisfiable in the class remain satisfiable after the addition of $\bigvee_{\gamma\in\Gamma} \neg\gamma(\bar{c})$, for every tuple $\bar{c}$ of new constants. In topological terms, every instance has empty interior with respect to the basic sets given by finite sets of sentences. Since every $\forall\exists$ sentence $\forall\bar{x} \bigvee_i \theta_i(\bar{x})$ over the fragment holds in exactly

the models that omit the type $\{\neg\theta_i\}_i$, this reading covers every sentence to which Keisler's theorem applies.

Theorem B carries this case of the general omitting types theorem to Keisler's general structures with three refinements: its instances are taken at tuples of closed terms, as in the work of Keisler and Miller, rather than at tuples of constants; its types consist of closed conditions $\varphi \geq r$ in place of formulas; and its space is the space of Henkin types that satisfy the theory, with basic open sets given by finite theories, so that the hypothesis of Keisler's theorem becomes a statement about interiors in a Čech-complete space.

In the terms of Keisler and Miller, the hypothesis of Theorem B says that no type of the sequence is locally realized. They call a type locally realized when it is implied, through some tuple of terms, by a formula consistent with the theory [KM01, Definition 4.1]; they show that every type realized in every model of a consistent theory is locally realized [KM01, Theorem 4.2], and they state local realization as the existence of a term mapping between spaces of types under which the preimage of the type has nonempty interior [KM01, Section 4]. Caicedo and Iovino use tuples of terms in the same way in their ω-principal types [CI14, Definition 3.5], and Eagle does so in his principal types, crediting this use of terms to Keisler and Miller and to Caicedo and Iovino [Eag14, Definition 4.10].

In Part II we prove Scott's theorem for general structures (Theorem C): every structure with countable sorts is characterized, among all structures, by a single sentence up to approximate back-and-forth equivalence. This equivalence coincides with agreement on the full type, the values of all sentences of $L_{\omega_1\omega}$. We give Scott's theorem a topological form (Theorem D) and establish a version of López-Escobar's theorem that marks the reach of this form. These theorems hold in the same generality as those of Part I.

Scott proved that every countable structure for a countable vocabulary is characterized up to isomorphism, among countable structures, by a single sentence of $L_{\omega_1\omega}$ [Sco65]. The theorem has a topological reading: the structures whose universe is the set of natural numbers form a Polish space, the isomorphism classes are the orbits of the group of permutations of the natural numbers, and every orbit, as the set of models of a Scott sentence, is Borel. López-Escobar proved the converse, that every Borel set of such structures invariant under the group is the set of models of a sentence of $L_{\omega_1\omega}$ [LE65], and Vaught's transforms later gave the theorem a topological proof [Vau74]. Both theorems rest on equality: isomorphisms are bijections, and the group of permutations is the group of bijections of the universe. Ben Yaacov, Doucha, Nies and Tsankov extended Scott's theorem, together with a version of López-Escobar's theorem, to separable metric structures whose predicates are uniformly continuous, with moduli prescribed by the signature, with respect to a complete metric [BDNT17]; there the metric takes the part of equality, and structures without a metric lie outside the setting of their analysis. General structures require neither equality nor a metric, and their sentences register a structure only up to approximation (Example 19). In Theorem C, therefore, the place of isomorphism is taken by *approximate equivalence*: two structures are approximately equivalent if, for every real number $r > 0$, there is a back-and-forth

system between them that matches the values of the atomic formulas to within r times their a priori bounds.

Theorem C. Let a countable vocabulary and a structure $\mathfrak{M}$ whose sorts are countable be given. There is a sentence $\sigma_{\mathfrak{M}}$ of $L_{\omega_1\omega}$ such that, for every structure $\mathfrak{N}$, the following conditions are equivalent: $\mathfrak{N}$ satisfies the closed condition $\sigma_{\mathfrak{M}} \leq 0$; the structures $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent; and the structures $\mathfrak{M}$ and $\mathfrak{N}$ have the same full type. If, moreover, the vocabulary has a single sort and a predicate symbol for equality, and $\mathfrak{M}$ and $\mathfrak{N}$ are structures with countable sorts in which every atomic formula ψ other than 1 takes only the two values r_ψ and $-r_\psi$, the predicate symbol for equality taking the value 1 exactly on the diagonal, then $\mathfrak{N}$ satisfies $\sigma_{\mathfrak{M}} \leq 0$ exactly when $\mathfrak{N}$ is isomorphic to $\mathfrak{M}$.

The first assertion is Theorem 17, and the second is Corollary 18. The key observation is that the Scott formulas measure, by their values, the error with which a tuple of another structure matches a tuple of $\mathfrak{M}$ in the back-and-forth game, and that a strict inequality keeps this error from accumulating along the game. Approximate equivalence is intrinsic to this syntax. It can identify nonisomorphic countable structures even when both structures are complete, indeed discrete, for the pseudometrics induced by the formulas. Thus passing to these completions does not turn the equivalence characterized by the Scott sentence into isomorphism (Example 21).

Theorem C has a topological form, in which the logic enters only the proof. For a countable vocabulary with a single sort, let X_V be the Polish space of structures whose universe is the set $\mathbb{N}$ of natural numbers. The atomic diagram of a structure in X_V records the values of the atomic formulas at the tuples of natural numbers; surjections of $\mathbb{N}$ onto $\mathbb{N}$ act on these diagrams by reindexing them. For structures $\mathfrak{M}$ and $\mathfrak{N}$ in X_V, we compare their diagrams after reindexing each by a surjection. The number $\delta(\mathfrak{M}, \mathfrak{N})$ is the infimum, over all pairs of surjections, of the resulting uniform distances, normalized by the a priori bounds. Two structures are *exactly equivalent* if there is a back-and-forth system between them along which the values of the atomic formulas agree.

Theorem D. Let a countable vocabulary with a single sort be given, and let X_V and δ be as above.

(a) Two structures in X_V are approximately equivalent if and only if δ vanishes at them, and exactly equivalent if and only if their reindexed atomic diagrams coincide for some pair of surjections.
(b) The function δ is a pseudometric on X_V, and every class of approximate equivalence is the zero set of a bounded Borel function on X_V that is Lipschitz for δ; in particular, every class is a Borel set, and two structures in X_V are approximately equivalent if and only if no Borel function on X_V that is Lipschitz for δ separates them.
(c) A bounded function on X_V is Borel and takes the same value at exactly equivalent structures if and only if it is given by a sentence of the logic obtained from $L_{\omega_1\omega}$ by adding, as connectives, the continuous functions of finitely many real variables.

Part (a) is Lemma 22, part (b) is Theorem 24, and part (c) is Theorem 27. The monoid of surjections, rather than the group of permutations, is what the absence of equality calls for: a structure is exactly equivalent to the structure obtained by giving each of its elements a twin, yet in Example 23 the two atomic diagrams remain at distance 1 under every pair of permutations. The main tool for part (b) is Lemma 11, by which every sentence of $L_{\omega_1\omega}$ gives a function that is Lipschitz for δ, and the main tool for part (c) is the method of Vaught's transforms, carried out over the monoid of surjections.

Part (c) marks the reach of part (b). A sentence of the enlarged logic can distinguish the value $\frac{1}{2}$ of a predicate from values arbitrarily close to it, so some of the Borel functions invariant under exact equivalence separate approximately equivalent structures (Example 26); replacing the functions of part (b) by all such functions therefore changes the equivalence that they detect, from approximate equivalence to exact equivalence. When the values of the atomic formulas range over finite sets separated by a uniform gap relative to their a priori bounds, the two equivalences coincide, and the invariant Borel functions are given by sentences of $L_{\omega_1\omega}$ itself (Theorem 30); for two-valued structures, this is López-Escobar's theorem for classical $L_{\omega_1\omega}$ without equality, with the monoid of surjections in place of the group of permutations.

Organization of the paper. Part I consists of Sections 2–11. In Section 2 we recall Čech's form of the Baire category theorem, and in Sections 3–8 we set up the structures, languages, fragments, theories and formal types. In Section 9 we define Henkin types and prove Theorem A. In Section 10 we show that, in the countable case, the space of Henkin types of a fragment is a G_δ subset of the product of intervals in which it lies, and Theorem B, proved in Section 11, then follows from Čech's form of the Baire category theorem. Part II consists of Sections 12 and 13. In Section 12 we construct the Scott sentences and prove Theorem C, and in Section 13 we give the topological form of Theorem C and prove Theorem D.

Part I. Model Existence and the Omitting Types Theorem

2. Topological Preliminaries

The topological fact on which the paper rests is Čech's form of the Baire category theorem, which applies, in particular, to every countable intersection of open subsets of a compact Hausdorff space. The theorem is applied in the space of Henkin types of a fragment (Section 9), which, as we show in Section 10, is such an intersection when the vocabulary and the fragment are countable. We recall the definitions that the theorem involves, with Engelking's *General Topology* [Eng89] as our reference for topology.

A topological space is a **Baire space** if the intersection of every sequence of dense open subsets of the space is dense; Baire spaces were introduced by Bourbaki in 1948 (see [Eng89, notes to Section 3.9]). The spaces to which this property will be applied are Tychonov spaces, that is, completely regular T_1-spaces; every subspace of a compact Hausdorff space is one (see [Eng89, Theorem 3.2.6]). Let

X be a Tychonov space. A sequence $(\mathscr{U}_n)_{n<\omega}$ of open covers of X is **complete** if every family $(F_i)_{i\in I}$ of closed subsets of X that has the finite intersection property, and that contains, for every $n < \omega$, a member included in some element of $\mathscr{U}_n$, has nonempty intersection. The space X is **Čech-complete** if it has a complete sequence of open covers. Čech [Čec37] introduced these spaces in 1937, under the name of topologically complete spaces, as the Tychonov spaces that are G_δ subsets of their Stone–Čech compactifications. The characterization of Čech-complete spaces by sequences of open covers, adopted here as the definition, was established independently by Frolík in 1960 and by Arhangel'skiĭ in 1961 (see [Eng89, Theorem 3.9.2 and the notes to Section 3.9]). By Theorems 3.9.1 and 3.9.2 of Engelking [Eng89], a Tychonov space is Čech-complete if and only if it is a G_δ subset of some, or equivalently of every, compactification of it. Consequently, every G_δ subset X of a compact Hausdorff space K is Čech-complete. The ambient compact spaces used below are products $\prod_{i\in I}[a_i, b_i]$ of real intervals with $a_i \leq b_i$, endowed with the product topology; they are compact Hausdorff by Tychonov's theorem (see [Eng89, Theorem 3.2.4]). Čech also proved the following form of the Baire category theorem (see [Eng89, Corollary 3.9.4 and the notes to Section 3.9]).

Theorem 1 (Baire category theorem; Čech [Čec37])**.** Every Čech-complete space is a Baire space.

3. Vocabularies and Structures

The models whose existence the omitting types theorem asserts are real-valued structures. Such a structure consists of one or more nonempty sets, called its sorts, of functions from finite products of sorts to sorts, and of its predicates. The predicates of the structure are bounded real-valued functions on finite products of sorts. Classical first-order structures for vocabularies without equality correspond to the case in which every predicate takes only the values 1 and -1, read as true and false, respectively.

Formulas speak about structures through a vocabulary, which is a set of symbols for sorts, functions and predicates. A structure for the vocabulary interprets each symbol as an actual sort, function or predicate. Since symbols must be kept apart from their interpretations, the symbols of a vocabulary are printed in this color, and mathematical objects in black. For the same reason, the variables, the connectives and the formal expressions built from them are printed in this color.

A **sort descriptor** is a purely syntactic symbol M. Let $\mathbb{M}$ be a nonempty set of sort descriptors. A **function symbol for** $\mathbb{M}$ is a purely syntactic symbol f given together with

(i) a natural number $n = n_f$, called the **arity of** f, and
(ii) an $(n+1)$-tuple $(M_i)_{i\leq n}$ of sort descriptors in $\mathbb{M}$, called the **extended arity of** f.

We indicate the arity and the extended arity of f by writing

$$f : \prod_{i<n} M_i \to M_n.$$

A function symbol of arity 0 is called a **constant symbol of sort** M_0. A **predicate symbol for** $\mathbb{M}$ is a purely syntactic symbol p given together with

(i) a natural number $n = n_p$, called the **arity of** p,
(ii) an n-tuple $(M_i)_{i<n}$ of sort descriptors in $\mathbb{M}$, called the **extended arity of** p, and
(iii) a real number $r = r_p \geq 0$, called the **a priori bound of** p.

We indicate the arity, the extended arity and the a priori bound of p by writing

$$p\colon \prod_{i<n} M_i \to [-r, r].$$

A **vocabulary** is a triple $V = (\mathbb{M}, F, P)$ consisting of a nonempty set $\mathbb{M}$ of sort descriptors, a set F of function symbols for $\mathbb{M}$, and a set P of predicate symbols for $\mathbb{M}$. The set P is required to contain a nullary predicate symbol 1, called the **logical unity constant**, with a priori bound $r_1 = 1$. The vocabulary V is **countable** if the sets $\mathbb{M}$, F and P are countable. The a priori bounds provide the compactness on which the topology of the paper rests. Indeed, as we show in Section 4, the value of a formula in a structure lies in an interval $[-r, r]$ determined by the formula alone. Consequently, the types of all structures lie in a product of compact intervals, the compact Hausdorff space of formal types of Section 8.

Given a vocabulary $V = (\mathbb{M}, F, P)$, a V**-structure**, or real-valued structure with vocabulary V, is a triple

$$\mathfrak{M} = \langle (M^{\mathfrak{M}})_{M\in\mathbb{M}}, (f^{\mathfrak{M}})_{f\in F}, (p^{\mathfrak{M}})_{p\in P}\rangle$$

in which

- ♠ the interpretation $M^{\mathfrak{M}}$ of each sort descriptor $M \in \mathbb{M}$ is a nonempty set,
- ♠ the interpretation of each function symbol $f\colon \prod_{i<n} M_i \to M_n$ in F is a function
$$f^{\mathfrak{M}}\colon \prod_{i<n} M_i^{\mathfrak{M}} \to M_n^{\mathfrak{M}},$$
- ♠ the interpretation of each predicate symbol $p\colon \prod_{i<n} M_i \to [-r_p, r_p]$ in P is a function
$$p^{\mathfrak{M}}\colon \prod_{i<n} M_i^{\mathfrak{M}} \to [-r_p, r_p],$$
and
- ♠ the interpretation of the logical unity constant is $1^{\mathfrak{M}} = 1$.

Here a nullary symbol is interpreted on the one-point set $\prod_{i<0} M_i^{\mathfrak{M}}$, and its interpretation is identified with its value.

4. First-Order and Infinitary Languages

Types, and later Henkin types, are formed from the values that structures assign to formulas. We therefore build, from the symbols of a vocabulary, the terms and the formulas of a first-order language and of its infinitary extension, and interpret them in structures.

For the rest of the paper, we fix a vocabulary $V = (\mathbb{M}, F, P)$.

For each sort descriptor $M \in \mathbb{M}$, we fix a sequence $(x_i^M)_{i<\omega}$ of distinct symbols, not in V, chosen so that the sequences for distinct sort descriptors have no symbol in common. Each symbol x_i^M is called a **variable of sort** M.

Terms of sort M are defined recursively as follows:

- ♠ every variable of sort M is a term of sort M;
- ♠ for every function symbol $f : \prod_{i<n} M_i \to M$ in F and every tuple $\bar{\tau} = (\tau_i)_{i<n}$ of terms such that τ_i is of sort M_i for each $i < n$, the expression $f(\bar{\tau})$ is a term of sort M.

In particular, every constant symbol of sort M is a term of sort M. For a natural number n and a tuple $(M_i)_{i<n}$ of sort descriptors in $\mathbb{M}$, a tuple $(\tau_i)_{i<n}$ of terms such that τ_i is of sort M_i for each $i < n$ is called a **term of sort** $\prod_{i<n} M_i$; for $n = 0$, the only such term is the empty tuple. A term is **closed** if no variable occurs in it.

For a predicate symbol $p : \prod_{i<n} M_i \to [-r_p, r_p]$ in P and a term $\bar{\tau}$ of sort $\prod_{i<n} M_i$, the expression $p(\bar{\tau})$ is called an **atomic V-formula** with **a priori bound** $r_{p(\bar{\tau})} := r_p$. The other V-formulas are formed from the atomic ones by applying symbols called **logical connectives**:

- ♠ for each rational number s, the symbol $s\cdot$ is a logical connective;
- ♠ the symbols $+$ and $\vee$ are logical connectives;
- ♠ for each variable x, the symbol $\sup_x$ is a logical connective.

The V-formulas and their a priori bounds are defined recursively, starting from the clause that every atomic V-formula is a V-formula. If φ and ψ are V-formulas with a priori bounds r_φ and r_ψ, respectively, then the expressions $s\cdot\varphi$, $(\varphi + \psi)$, $(\varphi \vee \psi)$ and $\sup_x \varphi$ are V-formulas for every rational number s and every variable x. The a priori bounds of these four formulas are defined by

$$r_{s\cdot\varphi} := |s| r_\varphi, \qquad r_{\varphi+\psi} := r_\varphi + r_\psi,$$
$$r_{\varphi\vee\psi} := \max\{r_\varphi, r_\psi\}, \qquad r_{\sup_x \varphi} := r_\varphi,$$

respectively. Since the connectives $s\cdot$ and $\sup_x$ are written before the formula to which they apply, the parentheses in $(\varphi + \psi)$ and $(\varphi \vee \psi)$ ensure that every V-formula is formed in only one way. Hence the a priori bound of a V-formula, like the other notions defined below by recursion on formulas, is well defined; we omit the parentheses when no ambiguity arises. The set of all V-formulas is called the **first-order language for** V and is denoted by L^V.

The set $\mathrm{free}(\xi)$ of **variables free in a V-formula** ξ is defined recursively as follows:

- ♠ if ξ is atomic, of the form $p(\bar{\tau})$, then $\mathrm{free}(\xi)$ is the set of variables that occur in the terms of the tuple $\bar{\tau}$;
- ♠ if ξ is of the form $s\cdot\varphi$, then $\mathrm{free}(\xi) := \mathrm{free}(\varphi)$;
- ♠ if ξ is of the form $\varphi + \psi$ or $\varphi \vee \psi$, then $\mathrm{free}(\xi) := \mathrm{free}(\varphi) \cup \mathrm{free}(\psi)$;
- ♠ if ξ is of the form $\sup_x \varphi$, then $\mathrm{free}(\xi) := \mathrm{free}(\varphi) \setminus \{x\}$.

A **V-sentence** is a V-formula with no free variables, that is, a V-formula φ such that $\mathrm{free}(\varphi) = \emptyset$.

For a term τ and a tuple $\bar{x}$ of distinct variables, we write $\tau(\bar{x})$ to indicate that every variable occurring in τ is among $\bar{x}$. Similarly, for a formula φ, we write $\varphi(\bar{x})$

to indicate that every variable free in φ is among $\bar{x}$. If $\bar{x} = (x_i)_{i<n}$, where x_i is a variable of sort M_i for each $i < n$, we write

$$_{\bar{x}}M := \prod_{i<n} M_i$$

and say that $\bar{x}$ is a variable of sort $_{\bar{x}}M$. For a V-structure $\mathfrak{M}$, we write

$$_{\bar{x}}M^{\mathfrak{M}} := \prod_{i<n} M_i^{\mathfrak{M}}.$$

The infinitary language $L^V_{\omega_1\omega}$ is obtained by adding to the formation rules of L^V a new logical connective $\bigvee$, called the **infinitary connective of countable join**. More precisely, the language $L^V_{\omega_1\omega}$, together with the a priori bounds and the free variables of its members, is defined by recursion on the countable ordinals. Let $\mathcal{L}_0$ be the set of atomic V-formulas, and let $\mathcal{L}_\lambda := \bigcup_{\alpha<\lambda} \mathcal{L}_\alpha$ for every countable limit ordinal λ. For $\alpha < \omega_1$, the set $\mathcal{L}_{\alpha+1}$ is obtained by adding to $\mathcal{L}_\alpha$ two kinds of expressions formed from its members. The first are the expressions $s\cdot\varphi$, $\varphi + \psi$, $\varphi \vee \psi$ and $\sup_x \varphi$ with $\varphi, \psi \in \mathcal{L}_\alpha$, for all rational numbers s and all variables x; their a priori bounds and free variables are given by the clauses stated for L^V. The second concern the sequences $\Phi = (\varphi_n)_{n<\omega}$ of members of $\mathcal{L}_\alpha$ for which $\sup_{n<\omega} r_{\varphi_n}$ is finite and, for some finite tuple $\bar{x}$ of variables, every φ_n is of the form $\varphi_n(\bar{x})$. For every such sequence, the expression $\bigvee\Phi$ is added, with a priori bound $r_{\bigvee\Phi} := \sup_{n<\omega} r_{\varphi_n}$. The free variables of $\bigvee\Phi$ are those of the formulas φ_n, that is, $\mathrm{free}(\bigvee\Phi) := \bigcup_{n<\omega} \mathrm{free}(\varphi_n)$, a subset of the finite set of variables of $\bar{x}$. Finally, we set $L^V_{\omega_1\omega} := \bigcup_{\alpha<\omega_1} \mathcal{L}_\alpha$. Since ω_1 is regular, every sequence $(\varphi_n)_{n<\omega}$ of members of $L^V_{\omega_1\omega}$ consists of members of some $\mathcal{L}_\alpha$, and the set $L^V_{\omega_1\omega}$ is therefore closed under all these formation rules.

We also write $\bigvee_n \varphi_n$ for $\bigvee\Phi$, and we denote the set of sentences of $L^V_{\omega_1\omega}$ by $S^V_{\omega_1\omega}$. A formula of $L^V_{\omega_1\omega}$ may have infinite length, but only finitely many free variables. The two languages differ in size: if V is countable, the first-order language L^V is countable, whereas the infinitary language $L^V_{\omega_1\omega}$ has the cardinality of the continuum. For this reason, the Baire category argument of this paper takes place in countable fragments (see Section 6).

Let $\mathfrak{M}$ be a V-structure, and let $\tau(\bar{x})$ be a term of sort M, where $\bar{x} = (x_j)_{j<m}$. The **interpretation of $\tau(\bar{x})$ in $\mathfrak{M}$** is the function $\tau^{\mathfrak{M}} : {}_{\bar{x}}M^{\mathfrak{M}} \to M^{\mathfrak{M}}$ defined recursively, for $\bar{a} = (a_j)_{j<m}$ in $_{\bar{x}}M^{\mathfrak{M}}$, as follows:

- ♠ the interpretation of the variable x_j is the projection $x_j^{\mathfrak{M}}(\bar{a}) := a_j$;
- ♠ if $f : \prod_{i<n} M_i \to M$ is a function symbol and $\bar{\tau} = (\tau_i)_{i<n}$ is a term of sort $\prod_{i<n} M_i$, then the interpretation of the term $f(\bar{\tau})$ is given by $(f(\bar{\tau}))^{\mathfrak{M}}(\bar{a}) := f^{\mathfrak{M}}((\tau_i^{\mathfrak{M}}(\bar{a}))_{i<n})$.

The interpretation of a closed term τ is a function on the one-point set $_{\bar{x}}M^{\mathfrak{M}}$ with $\bar{x}$ empty, and is identified with its value, an element of $M^{\mathfrak{M}}$. The **interpretation of a formula $\xi(\bar{x})$ of $L^V_{\omega_1\omega}$ in $\mathfrak{M}$** is the function $\xi^{\mathfrak{M}}$ on $_{\bar{x}}M^{\mathfrak{M}}$ defined recursively, for $\bar{a}$ in $_{\bar{x}}M^{\mathfrak{M}}$, as follows:

- ♠ if ξ is atomic, of the form $p(\bar{\tau})$ with $\bar{\tau} = (\tau_i)_{i<n}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := p^{\mathfrak{M}}((\tau_i^{\mathfrak{M}}(\bar{a}))_{i<n});$$
- ♠ if ξ is of the form $s\cdot\varphi$, where s is a rational number and $\varphi \in L^V_{\omega_1\omega}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := s \cdot \varphi^{\mathfrak{M}}(\bar{a});$$
- ♠ if ξ is of the form $\varphi + \psi$, where $\varphi, \psi \in L^V_{\omega_1\omega}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := \varphi^{\mathfrak{M}}(\bar{a}) + \psi^{\mathfrak{M}}(\bar{a});$$
- ♠ if ξ is of the form $\varphi \vee \psi$, where $\varphi, \psi \in L^V_{\omega_1\omega}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := \max\{\varphi^{\mathfrak{M}}(\bar{a}), \psi^{\mathfrak{M}}(\bar{a})\};$$
- ♠ if ξ is of the form $\sup_y \varphi$, where $\varphi \in L^V_{\omega_1\omega}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := \sup_{b \in {}_yM^{\mathfrak{M}}} \varphi^{\mathfrak{M}}(\bar{a}[b/y]),$$
where, if y is not among $\bar{x}$, the formula φ is regarded as a formula $\varphi(\bar{x}y)$, and $\bar{a}[b/y]$ is the tuple $\bar{a}b$; if y is the variable x_j of $\bar{x}$, the formula φ is regarded as a formula $\varphi(\bar{x})$, and $\bar{a}[b/y]$ is the tuple obtained from $\bar{a}$ by replacing its entry a_j by b;
- ♠ if ξ is of the form $\bigvee \Phi$, where $\Phi = (\varphi_n(\bar{x}))_{n<\omega}$ is a sequence of formulas of $L^V_{\omega_1\omega}$, then
$$\xi^{\mathfrak{M}}(\bar{a}) := \sup_{n<\omega} \varphi_n^{\mathfrak{M}}(\bar{a}).$$

An induction on formulas shows that the suprema in this definition are finite and that $\xi^{\mathfrak{M}}$ takes its values in $[-r_\xi, r_\xi]$. When ξ is a sentence, the tuple $\bar{x}$ may be taken empty, and $\xi^{\mathfrak{M}}$ is identified with its value, a real number in $[-r_\xi, r_\xi]$.

The finite connectives suffice to approximate every continuous function of finitely many formula values on their prescribed intervals. Indeed, for a positive integer k and formulas $\varphi_1, \dots, \varphi_k$, write $K = \prod_{i=1}^k [-r_{\varphi_i}, r_{\varphi_i}]$. The coordinate functions on the compact box K, together with the constant function 1, generate under addition, rational scalar multiplication and maximum a rational vector lattice of continuous functions that separates points.[2] Its uniform closure is a real vector lattice containing the constants, so the lattice Stone–Weierstrass theorem makes it all of $C(K, \mathbb{R})$, the space of continuous real-valued functions on K. Thus, for every continuous function $f : K \to \mathbb{R}$ and every real number $\varepsilon > 0$, there is an expression p in these finite connectives such that $|f(t) - p(t)| < \varepsilon$ for every $t \in K$. In particular, the approximation is uniform over all structures when t is the tuple of values of $\varphi_1, \dots, \varphi_k$.

Several proofs below proceed by induction on formulas of infinite length, and for these we use the **rank** of a formula, defined recursively as follows. Atomic formulas have rank 0; the rank of $s\cdot\varphi$ and of $\sup_x \varphi$ is the rank of φ plus one; the

[2] Any fixed countable dense set of real scalar coefficients suffices for this approximation argument; the rational coefficients are a convenient choice. The logical unity constant already supplies every rational constant as $q \cdot 1$, with $q \in \mathbb{Q}$.

rank of $\varphi + \psi$ and of $\varphi \vee \psi$ is the larger of the ranks of φ and ψ plus one; and the rank of $\bigvee \Phi$ is the supremum of the ranks of the members of Φ plus one. Every rank is a countable ordinal, and the rank of a proper subformula of a formula is smaller than the rank of the formula. Since substituting terms for variables changes only the atomic subformulas, it preserves the rank. These definitions apply to every vocabulary, and in particular to the expanded vocabulary of Section 5.

5. Expansion by Constants

In the method introduced by Henkin in 1949, the elements of a structure are named by new constant symbols; the first step is to add these symbols to the fixed vocabulary. Let $C = \{c_n^M : n < \omega,\ M \in \mathbb{M}\}$ be a set of distinct constant symbols, not in V, such that c_n^M is of sort M for every $n < \omega$ and every $M \in \mathbb{M}$. The **expansion of V by C** is the vocabulary $V_C := (\mathbb{M}, F \cup C, P)$, which is countable if V is; the infinitary language $L_{\omega_1\omega}^{V_C}$ is denoted by $L_{\omega_1\omega}^{C}$, and its set of sentences by $S_{\omega_1\omega}^{C}$.

Given a V-structure $\mathfrak{M} = \langle (M^{\mathfrak{M}})_{M\in\mathbb{M}}, (f^{\mathfrak{M}})_{f\in F}, (p^{\mathfrak{M}})_{p\in P}\rangle$, an **expansion of $\mathfrak{M}$ to V_C** is a V_C-structure

$$\mathfrak{N} = \langle (M^{\mathfrak{N}})_{M\in\mathbb{M}}, (f^{\mathfrak{N}})_{f\in F\cup C}, (p^{\mathfrak{N}})_{p\in P}\rangle$$

that agrees with $\mathfrak{M}$ on the symbols of V, in the sense that $M^{\mathfrak{N}} = M^{\mathfrak{M}}$ for every $M \in \mathbb{M}$, that $f^{\mathfrak{N}} = f^{\mathfrak{M}}$ for every $f \in F$, and that $p^{\mathfrak{N}} = p^{\mathfrak{M}}$ for every $p \in P$. Each expansion $\mathfrak{N}$ of $\mathfrak{M}$ to V_C is determined by the elements $(c_n^M)^{\mathfrak{N}}$ of $M^{\mathfrak{M}}$ that interpret the new constant symbols. Conversely, every V_C-structure $\mathfrak{N}$ is an expansion of exactly one V-structure, its **V-reduct**, obtained by forgetting the interpretations of the symbols in C.

In the structures built by the Henkin construction, every element has a name. A V_C-structure $\mathfrak{N}$ is **canonical** if, for every sort descriptor $M \in \mathbb{M}$, every element of $M^{\mathfrak{N}}$ is the interpretation in $\mathfrak{N}$ of a closed V_C-term of sort M.

In a canonical structure, a supremum over a sort is a supremum over the values of closed terms, and it can be computed from sentences by substituting these terms for the variable. The following lemma, which justifies this computation, holds in every V_C-structure. For a formula $\psi(\bar{z}\bar{x})$ and a tuple $\bar{\tau}$ of closed terms, we write $\psi(\bar{z}\bar{\tau})$ for the formula obtained by substituting the terms of $\bar{\tau}$ for the free occurrences of the corresponding variables of $\bar{x}$; a closed term is free for every variable.

Lemma 2 (Substitution)**.** Let $\mathfrak{N}$ be a V_C-structure, let $\psi(\bar{z}\bar{x})$ be a formula of $L_{\omega_1\omega}^{C}$, where $\bar{x}$ is a tuple of distinct variables not among $\bar{z}$, and let $\bar{\tau}$ be a tuple of closed V_C-terms of sort ${}_{\bar{x}}M$. Then $\psi(\bar{z}\bar{\tau})^{\mathfrak{N}}(\bar{b}) = \psi^{\mathfrak{N}}(\bar{b}\,\bar{\tau}^{\mathfrak{N}})$ for every tuple $\bar{b} \in {}_{\bar{z}}M^{\mathfrak{N}}$.

Proof. The corresponding identity for terms is proved by induction on terms, and the lemma follows by induction on the rank of ψ. The parameters $\bar{z}$ are needed in the step for suprema, in which the variable of the supremum joins them. □

6. Fragments

Countability is indispensable to the Baire category argument, which writes the Henkin types of Section 9 as a countable intersection of open sets attached to sentences (see Section 10). The infinitary language, however, has the cardinality of the continuum even when the vocabulary is countable (see Section 4). The argument is therefore carried out in fragments: sets of formulas that contain the first-order language and are closed under the first-order connectives, under subformulas and under substitution of terms. A fragment generated by countably many formulas over a countable vocabulary is countable, as we show below.

Let ξ be a formula of $L^V_{\omega_1\omega}$. The set $\mathrm{sub}(\xi)$ of **subformulas of** ξ is defined recursively as follows:

- ♠ if ξ is atomic, we set $\mathrm{sub}(\xi) := \{\xi\}$;
- ♠ if ξ is of the form $s\cdot\varphi$, we set $\mathrm{sub}(\xi) := \{\xi\} \cup \mathrm{sub}(\varphi)$;
- ♠ if ξ is of the form $\varphi + \psi$ or $\varphi \vee \psi$, we set $\mathrm{sub}(\xi) := \{\xi\} \cup \mathrm{sub}(\varphi) \cup \mathrm{sub}(\psi)$;
- ♠ if ξ is of the form $\sup_x \varphi$, we set $\mathrm{sub}(\xi) := \{\xi\} \cup \mathrm{sub}(\varphi)$;
- ♠ if ξ is of the form $\bigvee_n \varphi_n$, we set $\mathrm{sub}(\xi) := \{\xi\} \cup \bigcup_{n<\omega} \mathrm{sub}(\varphi_n)$.

A **fragment** of $L^V_{\omega_1\omega}$ is a set $\mathcal{F} \subseteq L^V_{\omega_1\omega}$ with the following properties:

- (F1) it contains L^V;
- (F2) it is closed under subformulas, that is, $\mathrm{sub}(\varphi) \subseteq \mathcal{F}$ for every $\varphi \in \mathcal{F}$;
- (F3) it is closed under the first-order connectives, that is, if $\varphi, \psi \in \mathcal{F}$, then $s\cdot\varphi \in \mathcal{F}$ for every rational number s, $\varphi + \psi \in \mathcal{F}$, $\varphi \vee \psi \in \mathcal{F}$, and $\sup_x \varphi \in \mathcal{F}$ for every variable x;
- (F4) it is closed under substitution of terms, that is, if $\varphi(\bar{x}\, y) \in \mathcal{F}$ and τ is a term of sort ${}_yM$ that is free for y in φ (no free occurrence of y in φ lies within a subformula $\sup_z \psi$ of φ such that z occurs in τ), then the formula $\varphi(\bar{x}\tau)$, obtained from φ by substituting τ for the free occurrences of y, belongs to $\mathcal{F}$.

The intersection of a nonempty family of fragments of $L^V_{\omega_1\omega}$ is a fragment, and the language $L^V_{\omega_1\omega}$ is itself a fragment. Hence, for every set $\Phi \subseteq L^V_{\omega_1\omega}$, the intersection of all fragments containing Φ is the smallest fragment containing Φ, which is called the **fragment generated by** Φ. If V and Φ are countable, then the fragment generated by Φ is countable. Indeed, let $\mathcal{F}_0 := L^V \cup \bigcup_{\varphi\in\Phi} \mathrm{sub}(\varphi)$; for each $k < \omega$, let $\mathcal{F}_{k+1}$ be obtained by adding to $\mathcal{F}_k$ the formulas that (F2)–(F4) require for members of $\mathcal{F}_k$. Since every formula has countably many subformulas, and since there are countably many rational numbers, variables and terms, an induction on k shows that every set $\mathcal{F}_k$ is countable. The union of the increasing sequence $(\mathcal{F}_k)_{k<\omega}$ is a fragment that contains Φ and is contained in every fragment containing Φ. Consequently, the fragment generated by Φ is this union, which is countable.

Let $\mathcal{F}$ be a fragment of $L^V_{\omega_1\omega}$. Since $L^V_{\omega_1\omega}$ is contained in the infinitary language $L^C_{\omega_1\omega}$ of the expanded vocabulary V_C of Section 5, we may consider the fragment of $L^C_{\omega_1\omega}$ generated by $\mathcal{F}$, which we denote by $\mathcal{F}_C$. If V and $\mathcal{F}$ are countable, then

so is $\mathcal{F}_C$, by the preceding paragraph applied to the countable vocabulary V_C. We denote the set of sentences of $\mathcal{F}$ by $S^{\mathcal{F}}$ and the set of sentences of $\mathcal{F}_C$ by $S_C^{\mathcal{F}}$.

7. Conditions and Theories

In real-valued logic, the role of the statements of classical logic is played by conditions, formal strict inequalities between a sentence and a rational number. Sets of conditions, called theories, specify the models to be constructed, whereas the types to be omitted are expressed by closed conditions. Strict conditions are taken as the basic ones because each of them defines an open set in the topology of Section 10, as the Baire category argument requires.

A **lower-bound condition** is a formal expression $\varphi > r$, and an **upper-bound condition** is a formal expression $\varphi < r$. Each of these expressions is formed from a sentence φ of the infinitary language $L^C_{\omega_1\omega}$ of the expanded vocabulary V_C (see Section 5) and a rational number r. A **condition** is a lower-bound condition or an upper-bound condition, and a **theory** is a set of conditions. For a set Σ of sentences, a **Σ-theory** is a theory whose conditions are on sentences in Σ.

For a rational number r, we write

- ♠ $\varphi \geq r$ for the countable theory $\{\varphi > r - \frac{1}{n} : n \in \mathbb{Z}_+\}$, and
- ♠ $\varphi \leq r$ for the countable theory $\{\varphi < r + \frac{1}{n} : n \in \mathbb{Z}_+\}$.

The expressions $\varphi \geq r$ and $\varphi \leq r$ are called **closed conditions**.

Let $\mathfrak{N}$ be a V_C-structure. We say that $\mathfrak{N}$ **satisfies** the condition $\varphi > r$, and write $\mathfrak{N} \vDash \varphi > r$, if $\varphi^{\mathfrak{N}} > r$; it satisfies the condition $\varphi < r$ if $\varphi^{\mathfrak{N}} < r$. A V_C-structure that satisfies every condition of a theory is called a **model** of the theory; the same term is used for V-structures and theories whose conditions are on V-sentences. Thus a structure satisfies the closed condition $\varphi \geq r$ exactly when the value of φ in it is at least r.

Under the reading of 1 as true and -1 as false, the connectives generalize those of classical first-order logic, as the following equivalences show. Each of these equivalences holds for every V_C-structure $\mathfrak{N}$ and every rational number r. Multiplication by -1 plays the role of negation: for a sentence φ, if $-\varphi$ denotes the sentence obtained by applying the connective $s\cdot$ with $s = -1$, then

$$\begin{aligned} \mathfrak{N} \vDash -\varphi > r \quad &\text{iff} \quad (-\varphi)^{\mathfrak{N}} > r \\ &\text{iff} \quad -(\varphi^{\mathfrak{N}}) > r \\ &\text{iff} \quad \varphi^{\mathfrak{N}} < -r. \end{aligned}$$

In particular, for $r = 0$, a V_C-structure satisfies $-\varphi > 0$ exactly when it satisfies $\varphi < 0$. The binary connective $\vee$ is the counterpart of disjunction: for sentences φ and ψ, we have

$$\begin{aligned} \mathfrak{N} \vDash \varphi \vee \psi > r \quad &\text{iff} \quad \max\{\varphi^{\mathfrak{N}}, \psi^{\mathfrak{N}}\} > r \\ &\text{iff} \quad \mathfrak{N} \vDash \varphi > r \text{ or } \mathfrak{N} \vDash \psi > r. \end{aligned}$$

The supremum $\sup_x$ takes the place of the existential quantifier: for a formula $\varphi(x)$ of $L^C_{\omega_1\omega}$, the condition $\sup_x \varphi > r$ is characterized as follows:

$$\begin{aligned} \mathfrak{N} \vDash \sup_x \varphi > r \quad &\text{iff} \quad (\sup_x \varphi)^{\mathfrak{N}} > r \\ &\text{iff} \quad \sup\left\{\varphi^{\mathfrak{N}}(a) : a \in {}_xM^{\mathfrak{N}}\right\} > r \\ &\text{iff} \quad \varphi^{\mathfrak{N}}(a) > r \text{ for some } a \in {}_xM^{\mathfrak{N}}. \end{aligned}$$

Finally, the join $\bigvee$ generalizes countable disjunction: for a sequence Φ of sentences of $L^C_{\omega_1\omega}$ whose join is a sentence of $L^C_{\omega_1\omega}$, the condition $\bigvee \Phi > r$ is characterized in the same way:

$$\begin{aligned} \mathfrak{N} \vDash \bigvee \Phi > r \quad &\text{iff} \quad \left(\bigvee \Phi\right)^{\mathfrak{N}} > r \\ &\text{iff} \quad \sup\left\{\varphi^{\mathfrak{N}} : \varphi \in \Phi\right\} > r \\ &\text{iff} \quad \varphi^{\mathfrak{N}} > r \text{ for some } \varphi \in \Phi. \end{aligned}$$

8. Types of Structures and Formal Types

The type of a structure records the values that the structure assigns to a set of sentences; a formal type is a family of values in the prescribed intervals, whether or not these values come from a structure. The Henkin conditions of Section 9 select formal types that do come from structures. These Henkin types form the subspace of the compact space of formal types in which the Baire category argument takes place.

Let Φ be a set of sentences of the infinitary language $L^C_{\omega_1\omega}$ of the expanded vocabulary V_C. We write $\prod \Phi$ for the product $\prod_{\varphi\in\Phi}[-r_\varphi, r_\varphi]$ endowed with the product topology, where r_φ is the a priori bound of φ. In this topology, the product $\prod \Phi$ is a compact Hausdorff space (see Section 2); its elements are called **formal Φ-types**. For a formal Φ-type $\mathfrak{m}$ and a sentence $\varphi \in \Phi$, we write $\varphi^{\mathfrak{m}}$ for the φ-coordinate of $\mathfrak{m}$ and say that $\mathfrak{m}$ satisfies a condition $\varphi > r$ or $\varphi < r$ if $\varphi^{\mathfrak{m}} > r$ or $\varphi^{\mathfrak{m}} < r$, respectively. If $\mathfrak{N}$ is a V_C-structure, the **Φ-type of $\mathfrak{N}$** is the formal Φ-type

$$\Phi\text{-tp}(\mathfrak{N}) \coloneqq (\varphi^{\mathfrak{N}})_{\varphi\in\Phi}.$$

When Φ is contained in the set $S^V_{\omega_1\omega}$ of sentences of $L^V_{\omega_1\omega}$, the same definition applies to a V-structure $\mathfrak{M}$. The $S^V_{\omega_1\omega}$-type $S^V_{\omega_1\omega}\text{-tp}(\mathfrak{M})$ is called the **full type** of $\mathfrak{M}$; similarly, the full type of a V_C-structure is its type for the set $S^C_{\omega_1\omega}$ of sentences of $L^C_{\omega_1\omega}$.

9. Henkin Types

In his proof of the completeness theorem, Henkin [Hen49] built a model of a consistent theory out of new individual constants, chosen to provide every existential sentence of the theory with a witness among them. The Henkin conditions below demand of a formal type that its values respect the connectives and, in the real-valued form of Henkin's requirement, that the value of a supremum over a

variable be the supremum of the values obtained by substituting closed terms for that variable. In this section we show that the formal types satisfying the Henkin conditions are exactly the types of canonical structures, structures in which every element is the value of a closed term. With the help of a downward Löwenheim–Skolem theorem for countable fragments, we then determine which formal types are realized by structures.

Let $\mathcal{G}$ be a fragment of the infinitary language $L^C_{\omega_1\omega}$ of the expanded vocabulary V_C, and let $S^{\mathcal{G}}$ be the set of its sentences. Let x be a variable, let φ be a formula whose free variables are among x, and let τ be a closed V_C-term of sort ${}_xM$. We write $\varphi(\tau)$ for the sentence obtained by substituting τ for the free occurrences of x in φ, as in condition (F4) of the definition of a fragment (see Section 6). A formal type $\mathfrak{m} \in \prod S^{\mathcal{G}}$ is an $S^{\mathcal{G}}$**-Henkin type** if it satisfies the following **Henkin conditions**:

(H1) $1^{\mathfrak{m}} = 1$;
(H2) $(s\cdot\varphi)^{\mathfrak{m}} = s\,\varphi^{\mathfrak{m}}$ for every $\varphi \in S^{\mathcal{G}}$ and every $s \in \mathbb{Q}$;
(H3) $(\varphi+\psi)^{\mathfrak{m}} = \varphi^{\mathfrak{m}} + \psi^{\mathfrak{m}}$ for all $\varphi, \psi \in S^{\mathcal{G}}$;
(H4) $(\varphi\vee\psi)^{\mathfrak{m}} = \max\{\varphi^{\mathfrak{m}}, \psi^{\mathfrak{m}}\}$ for all $\varphi, \psi \in S^{\mathcal{G}}$;
(H5) $(\sup_x \varphi)^{\mathfrak{m}} = \sup\{\varphi(\tau)^{\mathfrak{m}} : \tau \text{ is a closed } V_C\text{-term of sort } {}_xM\}$ for every variable x and every formula $\varphi \in \mathcal{G}$ whose free variables are among x;
(H6) $\big(\bigvee\Phi\big)^{\mathfrak{m}} = \sup\{\varphi^{\mathfrak{m}} : \varphi \in \Phi\}$ for every sentence $\bigvee\Phi \in S^{\mathcal{G}}$.

Every sentence that occurs in these conditions belongs to $S^{\mathcal{G}}$: the logical unity constant 1 by (F1), and the sentences $s\cdot\varphi$, $\varphi+\psi$, $\varphi\vee\psi$ and $\sup_x\varphi$ by (F3). The members of Φ belong to $S^{\mathcal{G}}$ by (F2), since the free variables of $\bigvee\Phi$ are those of its members; the sentences $\varphi(\tau)$ belong to $S^{\mathcal{G}}$ by (F4). The language $L^C_{\omega_1\omega}$ is a fragment of itself, whose set of sentences is $S^C_{\omega_1\omega}$; its $S^C_{\omega_1\omega}$-Henkin types lie in $\prod S^C_{\omega_1\omega}$ and are simply called **Henkin types**.

To pass from formal types back to structures, we follow Henkin and take as elements the closed terms themselves; the structure so obtained depends only on the values of the atomic sentences. Let $\mathcal{A}$ be the set of atomic V_C-sentences; by (F1), it is contained in $S^{\mathcal{G}}$ for every fragment $\mathcal{G}$ of $L^C_{\omega_1\omega}$. An **atomic type** is a formal type $\mathfrak{a} \in \prod\mathcal{A}$ such that $1^{\mathfrak{a}} = 1$. Every atomic type $\mathfrak{a}$ determines a V_C-structure $\mathfrak{N}_{\mathfrak{a}}$, the **term structure of** $\mathfrak{a}$, as follows. For each $M \in \mathbb{M}$, the set $M^{\mathfrak{N}_{\mathfrak{a}}}$ consists of the closed V_C-terms of sort M; it is nonempty, since it contains the constants c^M_n. Each function symbol f of V_C is interpreted syntactically, by $f^{\mathfrak{N}_{\mathfrak{a}}}(\bar\tau) := f(\bar\tau)$, and each predicate symbol p by $p^{\mathfrak{N}_{\mathfrak{a}}}(\bar\tau) := p(\bar\tau)^{\mathfrak{a}}$. The interpretation $p^{\mathfrak{N}_{\mathfrak{a}}}$ is a function into $[-r_p, r_p]$, because the a priori bound of the atomic sentence $p(\bar\tau)$ is r_p. Since, moreover, the logical unity constant satisfies $1^{\mathfrak{N}_{\mathfrak{a}}} = 1^{\mathfrak{a}} = 1$, the term structure $\mathfrak{N}_{\mathfrak{a}}$ is a V_C-structure. No quotient of the set of terms is needed, since the vocabulary has no equality symbol. An induction on terms shows that $\tau^{\mathfrak{N}_{\mathfrak{a}}} = \tau$ for every closed V_C-term τ; hence the term structure $\mathfrak{N}_{\mathfrak{a}}$ is canonical and satisfies $\mathcal{A}\text{-tp}(\mathfrak{N}_{\mathfrak{a}}) = \mathfrak{a}$.

Lemma 3. Let $\mathcal{G}$ be a fragment of $L^C_{\omega_1\omega}$.

(a) If $\mathfrak{N}$ is a canonical V_C-structure, then $S^{\mathcal{G}}\text{-tp}(\mathfrak{N})$ is an $S^{\mathcal{G}}$-Henkin type.

(b) Two $S^{\mathcal{G}}$-Henkin types that agree on $\mathcal{A}$ are equal.
(c) A formal type $\mathfrak{m} \in \prod S^{\mathcal{G}}$ is the $S^{\mathcal{G}}$-type of a canonical V_C-structure if and only if it is an $S^{\mathcal{G}}$-Henkin type; in that case, it is the $S^{\mathcal{G}}$-type of the term structure of its restriction to $\mathcal{A}$.

Proof. The idea is that the Henkin conditions repeat the recursive definition of the values of sentences, except that a supremum over a sort becomes a supremum over closed terms. The Henkin conditions therefore determine the value of every sentence from the values of the atomic ones, which is part (b). Part (a) asserts that the same conditions hold in a canonical structure, in which every element is the value of a closed term and the two suprema coincide.

(a) Conditions (H1), (H2), (H3), (H4) and (H6) hold for the values $\sigma^{\mathfrak{N}}$ of the sentences $\sigma \in S^{\mathcal{G}}$, by the definition of a V_C-structure and the recursive definition of the interpretation of formulas. For (H5), let $\varphi \in \mathcal{G}$ be a formula whose free variables are among x. By the substitution lemma (Lemma 2), applied with $\bar{z}$ empty, we have $\varphi(\tau)^{\mathfrak{N}} = \varphi^{\mathfrak{N}}(\tau^{\mathfrak{N}})$ for every closed V_C-term τ of sort ${}_x\boldsymbol{M}$. Since $\mathfrak{N}$ is canonical, every element of ${}_x\boldsymbol{M}^{\mathfrak{N}}$ is of the form $\tau^{\mathfrak{N}}$ for a closed V_C-term τ of sort ${}_x\boldsymbol{M}$; this gives

$$(\sup_x \varphi)^{\mathfrak{N}} = \sup_{b \in {}_x\boldsymbol{M}^{\mathfrak{N}}} \varphi^{\mathfrak{N}}(b) = \sup_\tau \varphi^{\mathfrak{N}}(\tau^{\mathfrak{N}}) = \sup_\tau \varphi(\tau)^{\mathfrak{N}}.$$

(b) Let $\mathfrak{m}$ and $\mathfrak{m}'$ be $S^{\mathcal{G}}$-Henkin types that agree on $\mathcal{A}$. We show by induction on the rank of σ that $\sigma^{\mathfrak{m}} = \sigma^{\mathfrak{m}'}$ for every $\sigma \in S^{\mathcal{G}}$, the atomic case being the hypothesis. Suppose that σ is of the form $s \cdot \varphi$, $\varphi + \psi$, $\varphi \vee \psi$ or $\bigvee \Phi$. In that case, the formulas φ, ψ and the members of Φ are sentences in $S^{\mathcal{G}}$ of smaller rank, on which the types $\mathfrak{m}$ and $\mathfrak{m}'$ agree by the induction hypothesis. Conditions (H2), (H3), (H4) and (H6), respectively, compute $\sigma^{\mathfrak{m}}$ and $\sigma^{\mathfrak{m}'}$ alike from the values of these sentences. Suppose, finally, that σ is of the form $\sup_x \varphi$. For every closed V_C-term τ of sort ${}_x\boldsymbol{M}$, the sentence $\varphi(\tau)$ then belongs to $S^{\mathcal{G}}$ and has smaller rank. The types $\mathfrak{m}$ and $\mathfrak{m}'$ agree on each of these sentences for the same reason, and condition (H5) therefore gives $\sigma^{\mathfrak{m}} = \sigma^{\mathfrak{m}'}$.

(c) By (a), the $S^{\mathcal{G}}$-type of a canonical V_C-structure is an $S^{\mathcal{G}}$-Henkin type. Conversely, let $\mathfrak{m}$ be an $S^{\mathcal{G}}$-Henkin type, and let $\mathfrak{a}$ be its restriction to $\mathcal{A}$, an atomic type by (H1). The term structure $\mathfrak{N}_{\mathfrak{a}}$ is canonical, with $\mathcal{A}\text{-tp}(\mathfrak{N}_{\mathfrak{a}}) = \mathfrak{a}$. By (a), the $S^{\mathcal{G}}$-type of $\mathfrak{N}_{\mathfrak{a}}$ is an $S^{\mathcal{G}}$-Henkin type that agrees with $\mathfrak{m}$ on $\mathcal{A}$; by (b), the two types are equal. □

Let $\mathfrak{a}$ be an atomic type. By Lemma 3(a), applied to $L^C_{\omega_1\omega}$, the full type $S^C_{\omega_1\omega}\text{-tp}(\mathfrak{N}_{\mathfrak{a}})$ is a Henkin type that extends $\mathfrak{a}$. Part (b) of that lemma shows that no other Henkin type extends $\mathfrak{a}$. We call this full type the **Henkin type of** $\mathfrak{a}$ and denote it by $\hat{\mathfrak{a}}$. For every fragment $\mathcal{G}$ of $L^C_{\omega_1\omega}$, the same lemma shows that the restriction of $\hat{\mathfrak{a}}$ to $S^{\mathcal{G}}$ is the only $S^{\mathcal{G}}$-Henkin type that extends $\mathfrak{a}$.

Let $\mathfrak{N}$ be a V_C-structure, and let $\mathfrak{a} \coloneqq \mathcal{A}\text{-tp}(\mathfrak{N})$. If $\mathfrak{N}$ is canonical, then Lemma 3(a) shows that the full type $S^C_{\omega_1\omega}\text{-tp}(\mathfrak{N})$ is a Henkin type that extends $\mathfrak{a}$, and therefore equals $\hat{\mathfrak{a}}$. In other words, we have $\varphi^{\mathfrak{N}} = \varphi^{\hat{\mathfrak{a}}}$ for every $\varphi \in S^C_{\omega_1\omega}$.

If $\mathfrak{N}$ is not canonical, the two values may differ, and in either direction. For example, suppose that V has a single sort descriptor M and no function symbols, and that the only predicate symbol of V besides 1 is $p\colon M \to [-1,1]$. Let $\mathfrak{N}$ be the V_C-structure with $M^{\mathfrak{N}} = \{0,1\}$, $p^{\mathfrak{N}}(0) = -1$ and $p^{\mathfrak{N}}(1) = 1$, in which every constant of C is interpreted as 0. For a variable x of sort M, the sentence $\sup_x p(x)$ takes the value 1 in $\mathfrak{N}$, whereas $(\sup_x p(x))^{\hat{\mathfrak{a}}} = \sup_n p(c_n^M)^{\mathfrak{a}} = -1$, the constants c_n^M being the only closed V_C-terms. Applying the connective $s\cdot$ with $s = -1$ to this sentence reverses the inequality.

Let Φ be a subset of the set $S^V_{\omega_1\omega}$ of sentences of $L^V_{\omega_1\omega}$, and let $\mathfrak{m} \in \prod \Phi$ be given. We say that the formal type $\mathfrak{m}$ is **realized** if there is a V-structure $\mathfrak{M}$ such that $\Phi\text{-tp}(\mathfrak{M}) = \mathfrak{m}$; in that case, we say that $\mathfrak{M}$ **realizes** $\mathfrak{m}$. When V is countable, there are countably many closed V_C-terms. In that case, a V-structure can therefore have a canonical expansion to V_C, to which Lemma 3 applies, only if its sorts are countable. The following form of the downward Löwenheim–Skolem theorem replaces every V-structure by one whose sorts are countable and which assigns the same values to the sentences of a given countable fragment.

Lemma 4. Assume that the vocabulary V is countable, and let $\mathcal{F}$ be a countable fragment of $L^V_{\omega_1\omega}$. For every V-structure $\mathfrak{M}$ there is a V-structure $\mathfrak{M}_0$ whose sorts are countable such that $\varphi^{\mathfrak{M}_0} = \varphi^{\mathfrak{M}}$ for every φ in the set $S^{\mathcal{F}}$ of sentences of $\mathcal{F}$.

Proof. The structure $\mathfrak{M}_0$ is obtained by closing countably many elements of $\mathfrak{M}$ under the functions of $\mathfrak{M}$ and under the choice of approximate witnesses for the suprema of formulas of $\mathcal{F}$. Let $\psi \in \mathcal{F}$, let y be a variable, and let $\bar{x}$ be a finite tuple of distinct variables among which are the free variables of $\sup_y \psi$. For every tuple $\bar{a} \in {}_{\bar{x}}M^{\mathfrak{M}}$ and every integer $k \geq 1$, choose an element $b_{\psi,y,\bar{x},\bar{a},k} \in {}_yM^{\mathfrak{M}}$ such that $\psi^{\mathfrak{M}}(\bar{a}[b_{\psi,y,\bar{x},\bar{a},k}/y]) > (\sup_y \psi)^{\mathfrak{M}}(\bar{a}) - 1/k$, with the notation of the clause for suprema in Section 4. For each $M \in \mathbb{M}$, let A_0^M consist of one element of $M^{\mathfrak{M}}$. For each $j < \omega$ and each $M \in \mathbb{M}$, let A_{j+1}^M consist of the elements of A_j^M, the values in $M^{\mathfrak{M}}$ of the functions $f^{\mathfrak{M}}$ at tuples from A_j, and the witnesses $b_{\psi,y,\bar{x},\bar{a},k}$ in $M^{\mathfrak{M}}$ with $\bar{a}$ from A_j. Here a tuple is *from* A_j if each of its entries lies in the set A_j^N of its sort N. Since V and $\mathcal{F}$ are countable, and since there are countably many variables and finite tuples of variables, an induction on j shows that every set A_j^M is countable; hence so is $A^M \coloneqq \bigcup_j A_j^M$. Moreover, the sets A_j^M increase with j, whence the family $(A^M)_{M\in\mathbb{M}}$ is closed under the functions $f^{\mathfrak{M}}$. Let $\mathfrak{M}_0$ be the V-structure whose sorts are the nonempty sets A^M and whose functions and predicates are the restrictions of those of $\mathfrak{M}$. By induction on the rank of $\varphi \in \mathcal{F}$ (see Section 4), we show that $\varphi^{\mathfrak{M}_0}(\bar{a}) = \varphi^{\mathfrak{M}}(\bar{a})$ for every tuple $\bar{a}$ from $\mathfrak{M}_0$ of the appropriate sort. For atomic φ this holds because terms take the same values in $\mathfrak{M}_0$ and in $\mathfrak{M}$. The steps for $s\cdot$, $+$, $\vee$ and $\bigvee$ follow from the definitions, the subformulas of members of $\mathcal{F}$ being members of $\mathcal{F}$ by (F2). Let φ be of the form $\sup_y \psi$, and let $\bar{a}$ be a tuple from $\mathfrak{M}_0$ of sort ${}_{\bar{x}}M$, where $\bar{x}$ is a finite tuple of distinct variables among which are the free variables of φ. The induction hypothesis gives $\varphi^{\mathfrak{M}_0}(\bar{a}) \leq \varphi^{\mathfrak{M}}(\bar{a})$, since the supremum defining $\varphi^{\mathfrak{M}_0}(\bar{a})$ ranges over a subset of

${}_yM^{\mathfrak{M}}$. Conversely, the tuple $\bar{a}$ is from A_j for some j. For every $k \geq 1$, the witness $b \coloneqq b_{\psi,y,\bar{x},\bar{a},k}$ is therefore an element of $\mathfrak{M}_0$, and the induction hypothesis applied to ψ gives $\varphi^{\mathfrak{M}_0}(\bar{a}) \geq \psi^{\mathfrak{M}_0}(\bar{a}[b/y]) = \psi^{\mathfrak{M}}(\bar{a}[b/y]) > \varphi^{\mathfrak{M}}(\bar{a}) - 1/k$. The case of sentences is the lemma. □

Recall from Section 6 that, for a fragment $\mathcal{F}$ of $L^V_{\omega_1\omega}$, we write $\mathcal{F}_C$ for the fragment of $L^C_{\omega_1\omega}$ generated by $\mathcal{F}$, and $S^{\mathcal{F}}_C$ for its set of sentences.

Proposition 5 (Henkin Realizability)**.** Assume that the vocabulary V is countable, and let $\Phi \subseteq S^V_{\omega_1\omega}$ and $\mathfrak{m} \in \prod \Phi$ be given.

(a) The formal type $\mathfrak{m}$ is realized by a V-structure whose sorts are countable if and only if it extends to a Henkin type in $\prod S^C_{\omega_1\omega}$.

(b) If $\Phi \subseteq S^{\mathcal{F}}$ for a countable fragment $\mathcal{F}$ of $L^V_{\omega_1\omega}$, then $\mathfrak{m}$ is realized if and only if it extends to an $S^{\mathcal{F}}_C$-Henkin type.

Proof. The idea is that, since V is countable, a V-structure has countable sorts exactly when it has a canonical expansion to V_C. By Lemma 3, moreover, the types of canonical V_C-structures are exactly the Henkin types. In part (b), Lemma 4 first replaces the realizing structure by one whose sorts are countable.

(a) Suppose first that a V-structure $\mathfrak{M}$ whose sorts are countable realizes $\mathfrak{m}$. Since every sort of $\mathfrak{M}$ is countable and nonempty, there is an expansion $\mathfrak{M}[C]$ of $\mathfrak{M}$ to V_C in which, for each $M \in \mathbb{M}$, the map $n \mapsto (c^M_n)^{\mathfrak{M}[C]}$ carries ω onto $M^{\mathfrak{M}}$, and every such expansion is canonical. By Lemma 3(a), applied to $L^C_{\omega_1\omega}$, the full type $S^C_{\omega_1\omega}$-tp$(\mathfrak{M}[C])$ is a Henkin type. This Henkin type extends $\mathfrak{m}$, because the value of a V-sentence in $\mathfrak{M}[C]$ is its value in $\mathfrak{M}$. Conversely, let $\mathfrak{m}' \in \prod S^C_{\omega_1\omega}$ be a Henkin type that extends $\mathfrak{m}$. By Lemma 3(c), applied to $L^C_{\omega_1\omega}$, the type $\mathfrak{m}'$ is the full type of a canonical V_C-structure $\mathfrak{N}$, whose V-reduct therefore realizes $\mathfrak{m}$. Since V_C is countable together with V, there are only countably many closed V_C-terms. The sorts of the V-reduct of $\mathfrak{N}$ are therefore countable, every element of $\mathfrak{N}$ being the value of such a term.

(b) If a V-structure $\mathfrak{M}$ realizes $\mathfrak{m}$, then so does a V-structure $\mathfrak{M}_0$ whose sorts are countable, by Lemma 4 and the inclusion $\Phi \subseteq S^{\mathcal{F}}$. By Lemma 3(a), applied to $\mathcal{F}_C$, the $S^{\mathcal{F}}_C$-type of the canonical expansion $\mathfrak{M}_0[C]$ constructed in (a) is an $S^{\mathcal{F}}_C$-Henkin type, and it extends $\mathfrak{m}$. Conversely, let $\mathfrak{m}'$ be an $S^{\mathcal{F}}_C$-Henkin type that extends $\mathfrak{m}$. By Lemma 3(c), applied to $\mathcal{F}_C$, the type $\mathfrak{m}'$ is the $S^{\mathcal{F}}_C$-type of a canonical V_C-structure, whose V-reduct realizes $\mathfrak{m}$. □

10. Topology on the Space of Henkin Types

Čech's form of the Baire category theorem needs a Čech-complete space, and the Henkin types supply one. When the vocabulary and the fragment are countable, the Henkin types for the sentences of the expanded fragment form a G_δ subset of the compact Hausdorff space of all formal types for these sentences, and therefore a Čech-complete space and a Baire space.

Let $\mathcal{F}$ be a fragment of $L^V_{\omega_1\omega}$, and recall from Section 6 that $S^{\mathcal{F}}_C$ is the set of sentences of the fragment $\mathcal{F}_C$ of $L^C_{\omega_1\omega}$ generated by $\mathcal{F}$. We denote the set of $S^{\mathcal{F}}_C$-Henkin types by $\mathcal{H}^{\mathcal{F}}_C$. As recalled in Section 2, the product $\prod S^{\mathcal{F}}_C = \prod_{\varphi\in S^{\mathcal{F}}_C}[-r_\varphi, r_\varphi]$ of Section 8, where r_φ is the a priori bound of φ, is a compact Hausdorff space when endowed with the product topology. The product topology is generated by the projections (see [Eng89, Section 2.3 and Proposition 1.4.8]), and the open rays with rational endpoints form a subbasis of the topology of each interval $[-r_\varphi, r_\varphi]$. Together, these two facts show that the sets

$$\langle\!\langle \varphi > r \rangle\!\rangle := \left\{ \mathfrak{m} \in \prod S^{\mathcal{F}}_C : \varphi^{\mathfrak{m}} > r \right\}$$
$$\langle\!\langle \varphi < r \rangle\!\rangle := \left\{ \mathfrak{m} \in \prod S^{\mathcal{F}}_C : \varphi^{\mathfrak{m}} < r \right\}$$

where $\varphi \in S^{\mathcal{F}}_C$ and $r \in \mathbb{Q}$, form a subbasis of the product topology. For an $S^{\mathcal{F}}_C$-theory T, we write $\langle\!\langle T \rangle\!\rangle$ for the set of formal types $\mathfrak{m} \in \prod S^{\mathcal{F}}_C$ that satisfy every condition of T, in the sense of Section 8. We say that $\mathfrak{m}$ **satisfies** T if $\mathfrak{m} \in \langle\!\langle T \rangle\!\rangle$; thus each $\langle\!\langle T \rangle\!\rangle$ is the intersection of the subbasic sets given by the conditions of T, and the set $\langle\!\langle \emptyset \rangle\!\rangle$ is all of $\prod S^{\mathcal{F}}_C$. For each $\varphi \in S^{\mathcal{F}}_C$, the projection $\mathfrak{m} \mapsto \varphi^{\mathfrak{m}}$ is continuous on $\prod S^{\mathcal{F}}_C$ (see [Eng89, Proposition 1.4.8]). Hence, for a closed condition $\varphi \geq r$ with $\varphi \in S^{\mathcal{F}}_C$, the set $\langle\!\langle \varphi \geq r \rangle\!\rangle = \{\mathfrak{m} \in \prod S^{\mathcal{F}}_C : \varphi^{\mathfrak{m}} \geq r\}$ is closed, being the preimage of the closed ray $[r, \infty)$ under the projection $\mathfrak{m} \mapsto \varphi^{\mathfrak{m}}$. The sets in the proof of Proposition 6 below are likewise built from sets defined by equations and inequalities between finitely many coordinates: if $\sigma_1, \ldots, \sigma_k \in S^{\mathcal{F}}_C$ and $g : \mathbb{R}^k \to \mathbb{R}$ is continuous, then the set $\{\mathfrak{m} : g(\sigma_1^{\mathfrak{m}}, \ldots, \sigma_k^{\mathfrak{m}}) > 0\}$ is open. Under the same hypotheses, the set $\{\mathfrak{m} : g(\sigma_1^{\mathfrak{m}}, \ldots, \sigma_k^{\mathfrak{m}}) = 0\}$, being the intersection of the open sets $\{\mathfrak{m} : |g(\sigma_1^{\mathfrak{m}}, \ldots, \sigma_k^{\mathfrak{m}})| < 1/n\}$ with $n \in \mathbb{Z}_+$, is a G_δ set.

Proposition 6. If the vocabulary V and the fragment $\mathcal{F}$ are countable, then the set $\mathcal{H}^{\mathcal{F}}_C$ of $S^{\mathcal{F}}_C$-Henkin types is a G_δ subset of $\prod S^{\mathcal{F}}_C$.

Proof. The idea is that the Henkin conditions consist of countably many equations, each either between finitely many coordinates or between one coordinate and the supremum of countably many others. Each such equation is, moreover, equivalent to countably many requirements defining G_δ sets. All the index sets below are countable: since V is countable, so are V_C, the set of variables and the set of closed V_C-terms. The fragment $\mathcal{F}_C$ is countable as well, by Section 6.

By the continuity of the projections, every set in the following list is a G_δ subset of $\prod S^{\mathcal{F}}_C$. In addition, for each $i \in \{1, \ldots, 6\}$, a formal type belongs to all the sets listed under (Hi) if and only if it satisfies the Henkin condition (Hi):

(H1) the set $\{\mathfrak{m} : 1^{\mathfrak{m}} = 1\}$;
(H2) for $\varphi \in S^{\mathcal{F}}_C$ and $s \in \mathbb{Q}$, the set $\{\mathfrak{m} : (s\cdot\varphi)^{\mathfrak{m}} = s\,\varphi^{\mathfrak{m}}\}$;
(H3) for $\varphi, \psi \in S^{\mathcal{F}}_C$, the set $\{\mathfrak{m} : (\varphi + \psi)^{\mathfrak{m}} = \varphi^{\mathfrak{m}} + \psi^{\mathfrak{m}}\}$;
(H4) for $\varphi, \psi \in S^{\mathcal{F}}_C$, the set $\{\mathfrak{m} : (\varphi \vee \psi)^{\mathfrak{m}} = \max\{\varphi^{\mathfrak{m}}, \psi^{\mathfrak{m}}\}\}$;

(H5) for each variable x and each formula $\varphi \in \mathcal{F}_C$ whose free variables are among x, the closed sets $\{\mathfrak{m} : (\sup_x \varphi)^{\mathfrak{m}} \geq \varphi(\tau)^{\mathfrak{m}}\}$ with τ a closed V_C-term of sort ${}_xM$, and the set

$$\bigcap_{n\in\mathbb{Z}_+} \bigcup_{\tau} \left\{ \mathfrak{m} : (\sup_x \varphi)^{\mathfrak{m}} < \varphi(\tau)^{\mathfrak{m}} + \tfrac{1}{n} \right\},$$

which is a G_δ set because each union is open;

(H6) for each sentence $\bigvee \Phi \in S_C^{\mathcal{F}}$, the closed sets $\{\mathfrak{m} : (\bigvee \Phi)^{\mathfrak{m}} \geq \varphi^{\mathfrak{m}}\}$ with $\varphi \in \Phi$, and the set

$$\bigcap_{n\in\mathbb{Z}_+} \bigcup_{\varphi\in\Phi} \left\{ \mathfrak{m} : (\bigvee \Phi)^{\mathfrak{m}} < \varphi^{\mathfrak{m}} + \tfrac{1}{n} \right\}.$$

Each set listed under (H1), (H2), (H3) and (H4), and each closed set listed under (H5) and (H6), is the zero set of a continuous function of finitely many coordinates (such as $\mathfrak{m} \mapsto \max\{\varphi^{\mathfrak{m}} - (\bigvee \Phi)^{\mathfrak{m}}, 0\}$). Each union listed is a union of sets of the form $\{\mathfrak{m} : g(\sigma_1^{\mathfrak{m}}, \dots, \sigma_k^{\mathfrak{m}}) > 0\}$; in both cases, the remark preceding the proposition therefore applies. Being indexed by countable sets, the family of all these sets is countable, and its intersection is $\mathcal{H}_C^{\mathcal{F}}$. Since a countable intersection of G_δ sets is a G_δ set, the proposition follows. □

Corollary 7. If the vocabulary V and the fragment $\mathcal{F}$ are countable, then $\mathcal{H}_C^{\mathcal{F}}$ is Čech-complete, Polish, and a Baire space.

Proof. By Proposition 6, the space $\mathcal{H}_C^{\mathcal{F}}$ is a G_δ subset of the compact Hausdorff space $\prod S_C^{\mathcal{F}}$. As recalled in Section 2, every such subset is Čech-complete, and the Baire category theorem for Čech-complete spaces (see Theorem 1) shows that it is a Baire space. Since $S_C^{\mathcal{F}}$ is countable, the product $\prod S_C^{\mathcal{F}}$ is compact and metrizable, and a G_δ subset of a Polish space is Polish [Kec95, Theorem 3.11]. □

In Section 11 we express in terms of conditions that a set of Henkin types has empty interior. For this purpose we use a basis of the topology of $\mathcal{H}_C^{\mathcal{F}}$ described by finite theories. Since $\langle\langle S \rangle\rangle$ is a finite intersection of subbasic open sets whenever S is a finite $S_C^{\mathcal{F}}$-theory, the sets $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle S \rangle\rangle$ with S finite form a basis of the topology of $\mathcal{H}_C^{\mathcal{F}}$ (see [Eng89, Sections 1.1 and 2.1]).

11. Omitting Types Theorem

The Henkin types satisfying a theory are the types of the canonical structures (those in which every element is the value of a closed term) whose reducts are models of the theory. To omit countably many types in a model of the theory, it therefore suffices to find such a Henkin type that realizes none of them. In this section we show that, when the vocabulary and the fragment are countable and the theory has a model, the Henkin types satisfying the theory form a nonempty Baire space. We also show that the required Henkin type exists if, for each instance of each of the given types, the members of this space satisfying the instance form a nowhere dense set. The members of this space that realize one of the given types

then form a meager subset, and the Baire category theorem provides a Henkin type outside it.

Henceforth, assume that the vocabulary V and the fragment $\mathcal{F}$ of Section 10 are countable. We also fix an $S^{\mathcal{F}}$-theory B (a theory whose conditions are on sentences of $\mathcal{F}$) that has a model. Recall from Section 10 that $\mathcal{H}_C^{\mathcal{F}}$ is the set of $S_C^{\mathcal{F}}$-Henkin types, where $S_C^{\mathcal{F}}$ is the set of sentences of the fragment $\mathcal{F}_C$ of $L^C_{\omega_1\omega}$ generated by $\mathcal{F}$. The set $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ of $S_C^{\mathcal{F}}$-Henkin types that satisfy B (see Section 10) is nonempty. Indeed, by Lemma 4, the theory B has a model $\mathfrak{M}$ whose sorts are countable. The canonical expansion $\mathfrak{M}[C]$ of $\mathfrak{M}$ constructed in the proof of Proposition 5(a) satisfies B, since the values of V-sentences in $\mathfrak{M}[C]$ are their values in $\mathfrak{M}$. By Lemma 3(a), the $S_C^{\mathcal{F}}$-type of $\mathfrak{M}[C]$ belongs to $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$. Conversely, by Lemma 3(c), every member of $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ is the $S_C^{\mathcal{F}}$-type of a canonical V_C-structure, whose V-reduct is a model of B. Thus the members of $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ are exactly the $S_C^{\mathcal{F}}$-types of the canonical V_C-structures whose V-reducts are models of B.

Since $S^{\mathcal{F}}$ is countable, so is the theory B, whence the set $\langle\langle B \rangle\rangle$ is a countable intersection of open sets. By Proposition 6, the set $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ is therefore a G_δ subset of the compact Hausdorff space $\prod S_C^{\mathcal{F}}$. As recalled in Section 2, a G_δ subset of a compact Hausdorff space is Čech-complete; by Theorem 1, the space $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ is a nonempty Baire space.

Let $\bar{x}$ be a finite tuple of distinct variables. An **$\mathcal{F}$-type in** $\bar{x}$ is a set $\Gamma(\bar{x})$ of expressions $\varphi(\bar{x}) \geq r$, where $r \in \mathbb{Q}$ and $\varphi \in \mathcal{F}$ is a formula whose free variables are among $\bar{x}$. Substituting closed terms for $\bar{x}$ turns each of these expressions into a closed condition in the sense of Section 7. For a tuple $\bar{\tau}$ of closed V_C-terms of sort ${}_{\bar{x}}M$, the **instance** $\Gamma(\bar{\tau})$ of $\Gamma(\bar{x})$ is the union of the closed conditions $\varphi(\bar{\tau}) \geq r$ for the members $\varphi(\bar{x}) \geq r$ of $\Gamma(\bar{x})$. Each of these conditions is a countable theory; the instance $\Gamma(\bar{\tau})$ is an $S_C^{\mathcal{F}}$-theory because the fragment $\mathcal{F}_C$ is closed under substitution of terms (see (F4) in Section 6). An $S_C^{\mathcal{F}}$-Henkin type $\mathfrak{m} \in \mathcal{H}_C^{\mathcal{F}}$ **realizes** $\Gamma(\bar{x})$ if it satisfies some instance $\Gamma(\bar{\tau})$, that is, if $\mathfrak{m} \in \langle\langle \Gamma(\bar{\tau}) \rangle\rangle$ for some tuple $\bar{\tau}$; otherwise it **omits** $\Gamma(\bar{x})$. Similarly, if $\mathfrak{M}$ is a V-structure, a tuple $\bar{a} \in {}_{\bar{x}}M^{\mathfrak{M}}$ **satisfies** $\Gamma(\bar{x})$ in $\mathfrak{M}$ if $\varphi^{\mathfrak{M}}(\bar{a}) \geq r$ for every member $\varphi(\bar{x}) \geq r$ of $\Gamma(\bar{x})$. We say that $\Gamma(\bar{x})$ **has nonempty interior in** $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ if there exist a tuple $\bar{\tau}$ of closed V_C-terms of sort ${}_{\bar{x}}M$ and a finite $S_C^{\mathcal{F}}$-theory S such that

$$\emptyset \neq \mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle \cap \langle\langle S \rangle\rangle \subseteq \langle\langle \Gamma(\bar{\tau}) \rangle\rangle .$$

Otherwise, we say that $\Gamma(\bar{x})$ **has empty interior in** $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$. The sets $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle \cap \langle\langle S \rangle\rangle$, with S a finite $S_C^{\mathcal{F}}$-theory, form a basis of the topology of $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$ (see Section 10). Consequently, the type $\Gamma(\bar{x})$ has nonempty interior in this space exactly when, for some tuple $\bar{\tau}$, the closed subset $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle \cap \langle\langle \Gamma(\bar{\tau}) \rangle\rangle$ of this space has nonempty interior. An $\mathcal{F}$-type is expressed by closed conditions so that this subset, being an intersection of the closed sets $\langle\langle \varphi(\bar{\tau}) \geq r \rangle\rangle$, is closed (see Section 10). When this subset has empty interior, its complement in $\mathcal{H}_C^{\mathcal{F}} \cap \langle\langle B \rangle\rangle$

is therefore a dense open set. The Baire category theorem applies to countable intersections of such dense open sets, as the proof of Theorem 8 shows.

Theorem 8 (Omitting Types Theorem)**.** Assume that the vocabulary V and the fragment $\mathcal{F}$ are countable, and let B be an $S^{\mathcal{F}}$-theory that has a model. Let $(\Gamma_n(\bar{x}_n))_{n<\omega}$ be a sequence of $\mathcal{F}$-types each of which has empty interior in $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$. Then there is an $S^{\mathcal{F}}_C$-Henkin type $\mathfrak{m} \in \mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$ that omits every $\Gamma_n(\bar{x}_n)$, and the theory B has a model whose sorts are countable and in which, for every $n < \omega$, no tuple of elements satisfies $\Gamma_n(\bar{x}_n)$.

Proof. The idea is to write the set of Henkin types that satisfy B and omit every $\Gamma_n(\bar{x}_n)$ as a countable intersection of open dense subsets of the nonempty Baire space $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$. Fix $n < \omega$, let $\bar{x} \coloneqq \bar{x}_n$ and $\Gamma \coloneqq \Gamma_n$, and let $\bar{\tau}$ be a tuple of closed V_C-terms of sort ${}_{\bar{x}}M$. Since the set $\langle\!\langle \varphi(\bar{\tau}) \geq r \rangle\!\rangle$ is closed for every member $\varphi(\bar{x}) \geq r$ of $\Gamma(\bar{x})$ (see Section 10), the set $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle \cap \langle\!\langle \Gamma(\bar{\tau}) \rangle\!\rangle$ is closed in the space $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$. By the definition of empty interior, this closed set has empty interior in $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$, and the complement

$$O_{\Gamma,\bar{\tau}} \coloneqq \{\mathfrak{m} \in \mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle \,:\, \mathfrak{m} \text{ does not satisfy } \Gamma(\bar{\tau})\}$$

is therefore an open dense subset of $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$. Since V is countable, there are only countably many tuples $\bar{\tau}$; consequently, the set $O_\Gamma \coloneqq \bigcap_{\bar{\tau}} O_{\Gamma,\bar{\tau}}$ of members of $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$ that omit $\Gamma(\bar{x})$ is a countable intersection of open dense subsets of $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$. As n was arbitrary, the same holds for each set O_{Γ_n}, and hence for the set $O \coloneqq \bigcap_{n<\omega} O_{\Gamma_n}$ of members of $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$ that omit $\Gamma_n(\bar{x}_n)$ for every $n < \omega$. Since $\mathcal{H}^{\mathcal{F}}_C \cap \langle\!\langle B \rangle\!\rangle$ is a nonempty Baire space, the set O is dense in it, and in particular nonempty.

Let $\mathfrak{m}$ be a member of O. By Lemma 3(c), the type $\mathfrak{m}$ is the $S^{\mathcal{F}}_C$-type of a canonical V_C-structure $\mathfrak{N}$, whose V-reduct $\mathfrak{M}$ is a model of B. The sorts of $\mathfrak{M}$ are countable, since every element of $\mathfrak{N}$ is the value of one of the countably many closed V_C-terms. Now fix $n < \omega$ and a tuple $\bar{a} \in {}_{\bar{x}_n}M^{\mathfrak{M}}$. Since $\mathfrak{N}$ is canonical, we may write $\bar{a} = \bar{\tau}^{\mathfrak{N}}$ for a tuple $\bar{\tau}$ of closed V_C-terms. By the substitution lemma (Lemma 2), with $\bar{z}$ empty, we have $\varphi^{\mathfrak{M}}(\bar{a}) = \varphi^{\mathfrak{N}}(\bar{\tau}^{\mathfrak{N}}) = \varphi(\bar{\tau})^{\mathfrak{m}}$ for every formula $\varphi \in \mathcal{F}$ whose free variables are among $\bar{x}_n$. Since $\mathfrak{m}$ omits $\Gamma_n(\bar{x}_n)$, it does not satisfy $\Gamma_n(\bar{\tau})$, and the tuple $\bar{a}$ therefore does not satisfy $\Gamma_n(\bar{x}_n)$ in $\mathfrak{M}$. □

Part II. Scott's Theorem and Its Topological Form

12. Scott Sentences

In 1965, Scott proved that every countable structure for a countable vocabulary is characterized, among countable structures and up to isomorphism, by a single sentence of the infinitary logic of two-valued structures [Sco65]. Ben Yaacov, Doucha, Nies and Tsankov extended the theorem to separable metric structures whose predicates are uniformly continuous with respect to a complete metric, with moduli prescribed by the signature [BDNT17]. In this section we extend Scott's theorem to the general structures of this paper, which require no prescribed metric

and no prescribed continuity. In $L_{\omega_1\omega}$, the value of a sentence is invariant under approximate back-and-forth equivalence, so that a sentence registers a structure only up to approximation. Accordingly, the theorem characterizes every structure whose sorts are countable, among all structures, up to approximate back-and-forth equivalence, and, for structures with a predicate for equality in which every atomic formula ψ takes only the two values r_ψ and $-r_\psi$, the theorem yields the content of Scott's theorem.

Throughout the section, the vocabulary $V = (\mathbb{M}, F, P)$ is countable. We use the following abbreviations for formulas φ and ψ of $L^V_{\omega_1\omega}$, a variable x and a sequence $\Phi = (\varphi_n)_{n<\omega}$ whose join is a formula: $-\varphi$ is $s\cdot\varphi$ with $s = -1$; $\varphi \wedge \psi$ is $-(-\varphi \vee -\psi)$; $\inf_x \varphi$ is $-\sup_x -\varphi$; $\bigwedge \Phi$ is $-\bigvee_n -\varphi_n$; $|\varphi|$ is $\varphi \vee -\varphi$; and $\varphi \dot{-} \psi$ is $(\varphi + -\psi) \vee (0\cdot 1)$. Apart from $-\varphi$, their values are $\min\{\varphi^{\mathfrak{M}}, \psi^{\mathfrak{M}}\}$, the infimum of the values of φ over the sort of x, the infimum of the values of the members of Φ, $|\varphi^{\mathfrak{M}}|$ and $\max\{\varphi^{\mathfrak{M}} - \psi^{\mathfrak{M}}, 0\}$, and their a priori bounds are $\max\{r_\varphi, r_\psi\}$, r_φ, $\sup_n r_{\varphi_n}$, r_φ and $r_\varphi + r_\psi$, respectively.

Real numbers enter the language as countable meets of rational multiples of the logical unity constant.

Lemma 9. Let real numbers t and $\epsilon > 0$ be given. Then there is a V-sentence $\underline{t}$ of $L^V_{\omega_1\omega}$ that takes the value t in every V-structure and has a priori bound at most $|t| + \epsilon$.

Proof. Choose rational numbers q_n with $t \leq q_n \leq t + \epsilon$ and $q_n \to t$, and let $\underline{t}$ be the meet of the sentences $q_n\cdot 1$. These sentences have the common a priori bound $|t| + \epsilon$, and the value of their meet is $\inf_n q_n = t$. □

For each finite sequence $(\boldsymbol{M}_i)_{i<k}$ of sort descriptors, we use the tuple $\bar{x} = (x_i^{\boldsymbol{M}_i})_{i<k}$ of variables, whose entries are distinct because their indices are; a tuple of elements of a structure of sort $_{\bar{x}}\boldsymbol{M}$ is said to be of **sort** $(\boldsymbol{M}_i)_{i<k}$. The variable that extends such a tuple by an entry of sort $\boldsymbol{M}$ is $x_k^{\boldsymbol{M}}$, distinct from the entries of $\bar{x}$. An **atomic formula in** $\bar{x}$ is an atomic V-formula whose variables are among $\bar{x}$; since V is countable, there are countably many of them.

Definition 10. Let V-structures $\mathfrak{M}$ and $\mathfrak{N}$ and a real number $r > 0$ be given. An **r-back-and-forth system** between $\mathfrak{M}$ and $\mathfrak{N}$ is a set I of pairs $(\bar{a}, \bar{b})$ of tuples of the same sort, $\bar{a}$ of elements of $\mathfrak{M}$ and $\bar{b}$ of elements of $\mathfrak{N}$, such that the pair of empty tuples belongs to I and, for every pair $(\bar{a}, \bar{b})$ in I of sort $(\boldsymbol{M}_i)_{i<k}$, with $\bar{x} = (x_i^{\boldsymbol{M}_i})_{i<k}$:

(i) $|\psi^{\mathfrak{M}}(\bar{a}) - \psi^{\mathfrak{N}}(\bar{b})| \leq r\, r_\psi$ for every atomic formula ψ in $\bar{x}$;
(ii) for every sort descriptor $\boldsymbol{M} \in \mathbb{M}$ and every $c \in \boldsymbol{M}^{\mathfrak{M}}$ there is $d \in \boldsymbol{M}^{\mathfrak{N}}$ with $(\bar{a}c, \bar{b}d) \in I$;
(iii) for every sort descriptor $\boldsymbol{M} \in \mathbb{M}$ and every $d \in \boldsymbol{M}^{\mathfrak{N}}$ there is $c \in \boldsymbol{M}^{\mathfrak{M}}$ with $(\bar{a}c, \bar{b}d) \in I$.

The structures $\mathfrak{M}$ and $\mathfrak{N}$ are **approximately equivalent** if there is an r-back-and-forth system between them for every real number $r > 0$.

Lemma 11. Let V-structures $\mathfrak{M}$ and $\mathfrak{N}$, a real number $r > 0$ and an r-back-and-forth system I between them be given. Then $|\varphi^{\mathfrak{M}}(\bar{a}) - \varphi^{\mathfrak{N}}(\bar{b})| \leq r\, r_\varphi$ for every pair $(\bar{a}, \bar{b})$ in I, of sort $(M_i)_{i<k}$ say, and every formula $\varphi(\bar{x})$ of $L^V_{\omega_1\omega}$, where $\bar{x} = (x_i^{M_i})_{i<k}$.

Proof. We argue by induction on the rank of φ (see Section 4), for all pairs in I at once. For atomic formulas the inequality is condition (i). For $s\cdot\varphi$, $\varphi + \psi$ and $\varphi \vee \psi$, the inequality follows from the inequalities for φ and ψ, since $|\max\{u, v\} - \max\{u', v'\}| \leq \max\{|u-u'|, |v-v'|\}$ for real numbers u, v, u' and v'; for $\bigvee \Phi$, the inequality follows from $|\sup_n u_n - \sup_n u'_n| \leq \sup_n |u_n - u'_n|$ for bounded sequences of real numbers. Let $\varphi(\bar{x})$ be of the form $\sup_y \psi$, with y of sort M. Rename the bound variables of ψ away from the variable x_k^M, and let $\psi'(\bar{x}x_k^M)$ be obtained by substituting x_k^M for the free occurrences of y. The substitution preserves the rank, the a priori bound and the admissibility of every join, and the formulas $\sup_y \psi$ and $\sup_{x_k^M} \psi'$ take the same values; the induction hypothesis therefore applies to ψ'. Let $c \in M^{\mathfrak{M}}$ be given; condition (ii) gives $d \in M^{\mathfrak{N}}$ with $(\bar{a}c, \bar{b}d) \in I$, and the induction hypothesis gives

$$\psi'^{\mathfrak{M}}(\bar{a}c) \leq \psi'^{\mathfrak{N}}(\bar{b}d) + r\, r_\psi \leq \varphi^{\mathfrak{N}}(\bar{b}) + r\, r_\varphi.$$

Taking the supremum over c gives $\varphi^{\mathfrak{M}}(\bar{a}) \leq \varphi^{\mathfrak{N}}(\bar{b}) + r\, r_\varphi$, and condition (iii) gives the reverse inequality in the same way. □

Corollary 12. Approximately equivalent V-structures have the same full type.

Proof. Applied to the pair of empty tuples, Lemma 11 gives $|\varphi^{\mathfrak{M}} - \varphi^{\mathfrak{N}}| \leq r\, r_\varphi$ for every sentence φ of $L^V_{\omega_1\omega}$ and every $r > 0$. □

Let a V-structure $\mathfrak{M}$ whose sorts are countable be given. The Scott formulas of $\mathfrak{M}$, defined below, measure the error with which a tuple of another structure matches a tuple of $\mathfrak{M}$ in the back-and-forth game. For an atomic formula ψ in $\bar{x} = (x_i^{M_i})_{i<k}$ with $r_\psi > 0$, choose a rational number c_ψ with $1/(6r_\psi) \leq c_\psi \leq 1/(3r_\psi)$ and a rational number ϵ_ψ with $0 < \epsilon_\psi \leq r_\psi$. For a tuple $\bar{a}$ of elements of $\mathfrak{M}$ of sort $(M_i)_{i<k}$, let $t = \psi^{\mathfrak{M}}(\bar{a})$, and let

$$\delta_{\psi,\bar{a}}(\bar{x}) \coloneqq c_\psi \cdot |\psi + -\underline{t}\,|,$$

where $\underline{t}$ is the sentence of Lemma 9 for t and ϵ_ψ. The a priori bound of $\delta_{\psi,\bar{a}}$ is at most $c_\psi(r_\psi + |t| + \epsilon_\psi) \leq 3c_\psi r_\psi \leq 1$, and, for every V-structure $\mathfrak{N}$ and every tuple $\bar{b}$ of elements of $\mathfrak{N}$ of sort $(M_i)_{i<k}$,

$$|\psi^{\mathfrak{N}}(\bar{b}) - \psi^{\mathfrak{M}}(\bar{a})| \leq 6r_\psi\, \delta^{\mathfrak{N}}_{\psi,\bar{a}}(\bar{b}). \tag{1}$$

Definition 13. For every countable ordinal α and every tuple $\bar{a}$ of elements of $\mathfrak{M}$ of sort $(M_i)_{i<k}$, the **Scott formula** $\varphi^\alpha_{\bar{a}}(\bar{x})$, with $\bar{x} = (x_i^{M_i})_{i<k}$, is defined by recursion on α. The formula $\varphi^0_{\bar{a}}$ is the join of the formula $0\cdot 1$ and of the formulas $\delta_{\psi,\bar{a}}$ for the atomic formulas ψ in $\bar{x}$ with $r_\psi > 0$. For every countable ordinal α, the formula

$\varphi^{\alpha+1}_{\bar{a}}$ is the join of the formula $\varphi^{\alpha}_{\bar{a}}$ and, for every sort descriptor $M \in \mathbb{M}$, of the formulas

$$\inf_{x^M_k} \varphi^{\alpha}_{\bar{a}c}(\bar{x}x^M_k) \quad \text{for } c \in M^{\mathfrak{M}}, \qquad \sup_{x^M_k} \bigwedge_{c \in M^{\mathfrak{M}}} \varphi^{\alpha}_{\bar{a}c}(\bar{x}x^M_k).$$

For every countable limit ordinal λ, the formula $\varphi^{\lambda}_{\bar{a}}$ is the join of the formulas $\varphi^{\alpha}_{\bar{a}}$ with $\alpha < \lambda$.

By induction on α, every Scott formula has a priori bound at most 1, so that every join and every meet in the recursion is a formula of $L^V_{\omega_1\omega}$, the sorts of $\mathfrak{M}$, the set $\mathbb{M}$, the set of atomic formulas in $\bar{x}$ and the set of ordinals below a countable ordinal being countable. By the same induction, every Scott formula takes values in $[0, 1]$, and $\varphi^{\alpha}_{\bar{a}}$ is at most $\varphi^{\beta}_{\bar{a}}$ in every V-structure for $\alpha \leq \beta$. In the recursion, the infima measure the moves from $\mathfrak{M}$ to the other structure and the suprema the moves back.

Lemma 14. The value of every Scott formula $\varphi^{\alpha}_{\bar{a}}$ at the tuple $\bar{a}$ in $\mathfrak{M}$ is 0.

Proof. For $\alpha = 0$, every formula $\delta_{\psi,\bar{a}}$ vanishes at $\bar{a}$. For a successor ordinal $\alpha+1$, let $M \in \mathbb{M}$ be given. The induction hypothesis gives $\inf_b \varphi^{\alpha,\mathfrak{M}}_{\bar{a}c}(\bar{a}b) \leq \varphi^{\alpha,\mathfrak{M}}_{\bar{a}c}(\bar{a}c) = 0$ for every $c \in M^{\mathfrak{M}}$, and, for every $e \in M^{\mathfrak{M}}$, the infimum of the values $\varphi^{\alpha,\mathfrak{M}}_{\bar{a}c}(\bar{a}e)$ over $c \in M^{\mathfrak{M}}$ is at most $\varphi^{\alpha,\mathfrak{M}}_{\bar{a}e}(\bar{a}e) = 0$. The limit case is immediate. □

Lemma 15. There is a countable ordinal $\alpha_{\mathfrak{M}}$ such that $\varphi^{\alpha_{\mathfrak{M}}+1,\mathfrak{M}}_{\bar{a}}(\bar{b}) = \varphi^{\alpha_{\mathfrak{M}},\mathfrak{M}}_{\bar{a}}(\bar{b})$ for all tuples $\bar{a}$ and $\bar{b}$ of elements of $\mathfrak{M}$ of the same sort.

Proof. For fixed tuples $\bar{a}$ and $\bar{b}$, the values $\varphi^{\alpha,\mathfrak{M}}_{\bar{a}}(\bar{b})$, for $\alpha < \omega_1$, form a nondecreasing family of real numbers. The open intervals between the values at α and at $\alpha + 1$, for the ordinals α at which the family increases, are pairwise disjoint, so these ordinals are countably many; since ω_1 is regular, there is a countable ordinal from which the value at every ordinal equals the value at its successor. At a countable limit ordinal λ the value is the supremum of the earlier values, since $\varphi^{\lambda}_{\bar{a}}$ is the join of the earlier formulas; by induction, the family is constant from that ordinal on. There are countably many pairs $(\bar{a}, \bar{b})$, and the supremum of countably many countable ordinals is countable. □

Definition 16. The **Scott sentence** $\sigma_{\mathfrak{M}}$ of $\mathfrak{M}$ is the join of the sentence $\varphi^{\alpha_{\mathfrak{M}}+1}_{\varnothing}$, for the empty tuple $\varnothing$, and of the sentences

$$\sup_{x^{M_0}_0} \cdots \sup_{x^{M_{k-1}}_{k-1}} \left(\varphi^{\alpha_{\mathfrak{M}}+1}_{\bar{a}} \dot{-} \varphi^{\alpha_{\mathfrak{M}}}_{\bar{a}}\right)$$

for the tuples $\bar{a}$ of elements of $\mathfrak{M}$, of every sort $(M_i)_{i<k}$, where the ordinal $\alpha_{\mathfrak{M}}$ is given by Lemma 15.

The members of this join have a priori bound at most 2, so $\sigma_{\mathfrak{M}}$ is a sentence of $L^V_{\omega_1\omega}$; the values of $\sigma_{\mathfrak{M}}$ are nonnegative, and the value of $\sigma_{\mathfrak{M}}$ in $\mathfrak{M}$ is 0, by Lemmas 14 and 15.

Theorem 17 (Scott's theorem for infinitary real-valued logic)**.** Assume that the vocabulary V is countable, let $\mathfrak{M}$ be a V-structure whose sorts are countable, and let $\sigma_{\mathfrak{M}}$ be the Scott sentence of $\mathfrak{M}$. For every V-structure $\mathfrak{N}$, the following conditions are equivalent:

(a) $\mathfrak{N}$ satisfies the closed condition $\sigma_{\mathfrak{M}} \le 0$;
(b) $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent;
(c) $\mathfrak{M}$ and $\mathfrak{N}$ have the same full type.

Proof. Condition (b) implies (c) by Corollary 12, and (c) implies (a) because $\sigma_{\mathfrak{M}}^{\mathfrak{M}} = 0$.

Suppose (a); since $\sigma_{\mathfrak{M}}$ takes nonnegative values, $\sigma_{\mathfrak{M}}^{\mathfrak{N}} = 0$. Let a real number $r > 0$ be given, and let I be the set of pairs $(\bar{a}, \bar{b})$ of tuples of the same sort, $\bar{a}$ of elements of $\mathfrak{M}$ and $\bar{b}$ of elements of $\mathfrak{N}$, such that $\varphi_{\bar{a}}^{\alpha_{\mathfrak{M}},\mathfrak{N}}(\bar{b}) < r/6$. Since $\sigma_{\mathfrak{M}}^{\mathfrak{N}} = 0$, the value in $\mathfrak{N}$ of $\varphi_{\varnothing}^{\alpha_{\mathfrak{M}}+1}$ is 0, and so is the smaller value of $\varphi_{\varnothing}^{\alpha_{\mathfrak{M}}}$; the pair of empty tuples belongs to I. Moreover, $\varphi_{\bar{a}}^{\alpha_{\mathfrak{M}}+1,\mathfrak{N}}(\bar{b}) \le \varphi_{\bar{a}}^{\alpha_{\mathfrak{M}},\mathfrak{N}}(\bar{b})$ for all tuples $\bar{a}$ of $\mathfrak{M}$ and $\bar{b}$ of $\mathfrak{N}$ of the same sort. Let $(\bar{a}, \bar{b})$ be a pair in I, of sort $(M_i)_{i<k}$. For condition (i) of Definition 10, let ψ be an atomic formula in $\bar{x} = (x_i^{M_i})_{i<k}$; if $r_\psi = 0$, both values of ψ vanish, and otherwise $\varphi_{\bar{a}}^{\alpha_{\mathfrak{M}}} \ge \varphi_{\bar{a}}^{0} \ge \delta_{\psi,\bar{a}}$ and (1) give $|\psi^{\mathfrak{N}}(\bar{b}) - \psi^{\mathfrak{M}}(\bar{a})| < 6r_\psi \cdot r/6 = r\, r_\psi$. For condition (ii), let $M \in \mathbb{M}$ and $c \in M^{\mathfrak{M}}$ be given. Then $\inf_{d \in M^{\mathfrak{N}}} \varphi_{\bar{a}c}^{\alpha_{\mathfrak{M}},\mathfrak{N}}(\bar{b}d) \le \varphi_{\bar{a}}^{\alpha_{\mathfrak{M}}+1,\mathfrak{N}}(\bar{b}) \le \varphi_{\bar{a}}^{\alpha_{\mathfrak{M}},\mathfrak{N}}(\bar{b}) < r/6$, so there is $d \in M^{\mathfrak{N}}$ with $(\bar{a}c, \bar{b}d) \in I$. For condition (iii), let $M \in \mathbb{M}$ and $d \in M^{\mathfrak{N}}$ be given. Then the infimum of the values $\varphi_{\bar{a}c}^{\alpha_{\mathfrak{M}},\mathfrak{N}}(\bar{b}d)$ over $c \in M^{\mathfrak{M}}$ is at most $\varphi_{\bar{a}}^{\alpha_{\mathfrak{M}}+1,\mathfrak{N}}(\bar{b}) < r/6$, so there is $c \in M^{\mathfrak{M}}$ with $(\bar{a}c, \bar{b}d) \in I$. Thus I is an r-back-and-forth system, and (b) holds. □

The proof rests on a strict inequality: a match with error less than $r/6$ extends, one element at a time, to matches with error less than $r/6$, so that the error does not accumulate along the back-and-forth game. The structure $\mathfrak{N}$ is arbitrary, and Theorem 17 is the real-valued counterpart of Scott's theorem together with the characterization of infinitary equivalence by back-and-forth systems. For structures with a predicate for equality in which every atomic formula ψ takes only the two values r_ψ and $-r_\psi$, approximate equivalence among countable structures is isomorphism (Corollary 18).

Corollary 18 (The content of Scott's theorem)**.** Assume that V has a single sort descriptor M and a predicate symbol $E : M \times M \to [-1, 1]$. Let $\mathfrak{M}$ and $\mathfrak{N}$ be V-structures whose sorts are countable, in which every atomic formula ψ other than 1 takes only the two values r_ψ and $-r_\psi$ and E takes the value 1 exactly on the diagonal. Then $\mathfrak{N}$ satisfies $\sigma_{\mathfrak{M}} \le 0$ if and only if $\mathfrak{M}$ and $\mathfrak{N}$ are isomorphic.

Proof. Isomorphic structures have the same full type, and $\sigma_{\mathfrak{M}}^{\mathfrak{M}} = 0$; so isomorphism gives $\sigma_{\mathfrak{M}}^{\mathfrak{N}} = 0$. Conversely, suppose that $\sigma_{\mathfrak{M}}^{\mathfrak{N}} = 0$; by Theorem 17 there is an r-back-and-forth system I between $\mathfrak{M}$ and $\mathfrak{N}$ for $r = 1$. The two values r_ψ and $-r_\psi$ of an atomic formula ψ differ by $2r_\psi > r\, r_\psi$, so condition (i) of Definition 10

gives $\psi^{\mathfrak{M}}(\bar{a}) = \psi^{\mathfrak{N}}(\bar{b})$ for every pair $(\bar{a}, \bar{b})$ in I and every atomic formula ψ in the corresponding variables. Applied to the atomic formulas $E(x_i, x_j)$ and $E(f(\bar{\tau}), x_j)$, for function symbols f and tuples $\bar{\tau}$ of variables, this shows that the map $a_i \mapsto b_i$ is well defined, injective, and compatible with the predicates and with the functions on the entries of the tuples. Enumerate the sorts of $\mathfrak{M}$ and $\mathfrak{N}$, and choose pairs in I alternately by conditions (ii) and (iii), each extending the previous one, so that every element of $\mathfrak{M}$ appears in the first tuples and every element of $\mathfrak{N}$ in the second; the union of the corresponding maps is an isomorphism. □

Approximation is a feature of the logic itself, as the following example shows.

Example 19. Assume that V has a single sort descriptor M, no function symbols, and, besides 1, a single predicate symbol $p\colon M \to [-1, 1]$. Let $\mathfrak{M}$ and $\mathfrak{N}$ be the V-structures with $M^{\mathfrak{M}} = \{0\} \cup \{1/n\colon n \geq 1\}$ and $M^{\mathfrak{N}} = \{1/n\colon n \geq 1\}$, in which p is interpreted as the identity. For $r > 0$, let I consist of the pairs $(\bar{a}, \bar{b})$ of tuples of the same length with $|a_i - b_i| \leq r$ for every i. The atomic formulas in $\bar{x}$ other than 1 are the formulas $p(x_i)$, so condition (i) of Definition 10 holds; condition (ii) holds because every element of $\mathfrak{M}$ lies within r of an element of $\mathfrak{N}$, and condition (iii) because every element of $\mathfrak{N}$ is an element of $\mathfrak{M}$. Thus $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent, and, by Corollary 12, they have the same full type. They are not isomorphic, since p vanishes at an element of $\mathfrak{M}$ and at no element of $\mathfrak{N}$. The formula $\bigvee_n \big(n{\cdot}p(x) \wedge 1\big)$, which would separate the element 0 of $\mathfrak{M}$ from the elements of $\mathfrak{N}$, is not a formula of $L^V_{\omega_1\omega}$, since the a priori bounds n of the members of the join have no common bound.

Remark 20. The formulas of $L^V_{\omega_1\omega}$ induce on each V-structure $\mathfrak{N}$ a pseudometric on each sort: for elements a and a' of $M^{\mathfrak{N}}$, the number $d_{\mathfrak{N}}(a, a')$ is the supremum of $|\varphi^{\mathfrak{N}}(\bar{b}a) - \varphi^{\mathfrak{N}}(\bar{b}a')|/r_\varphi$ over the formulas φ of $L^V_{\omega_1\omega}$ with $r_\varphi > 0$ and the tuples $\bar{b}$ of elements of $\mathfrak{N}$ of the appropriate sort. By definition, every predicate of $\mathfrak{N}$ is Lipschitz for these pseudometrics, with constant given by the a priori bound of the predicate symbol in each argument. In Example 19, $d_{\mathfrak{N}}(a, a') = |a - a'|$ for distinct elements a and a': the formula $p(x)$ gives one inequality, and Lemma 11, applied to the s-back-and-forth system between $\mathfrak{N}$ and itself consisting of the pairs $(\bar{c}, \bar{e})$ with $|c_i - e_i| \leq s$ for every i, where $s = |a - a'|$, gives the other. The completion of $\mathfrak{N}$ for $d_{\mathfrak{N}}$, with p extended by continuity, is therefore a copy of $\mathfrak{M}$. Keisler obtained a distance predicate from first-order formulas for every countable vocabulary [Kei23, Section 3.3]. Approximately equivalent structures need not, however, have isomorphic completions for these pseudometrics, even when both structures are already complete, as the next example shows.

Example 21. Assume that V has a single sort descriptor M, no function symbols, and, besides 1, predicate symbols $p\colon M \to [-1, 1]$ and $E\colon M \times M \to [-1, 1]$. Let $\mathfrak{M}$ and $\mathfrak{N}$ be the V-structures with $M^{\mathfrak{M}} = \mathbb{Q} \cap [0, 1]$ and $M^{\mathfrak{N}} = \mathbb{Q} \cap (0, 1]$, in which p is interpreted as the identity and E takes the value 1 on the diagonal and -1 elsewhere. For $r > 0$, the pairs $(\bar{a}, \bar{b})$ of tuples of the same length with $a_i = a_j$ exactly when $b_i = b_j$ and with $|a_i - b_i| < r$ for every i form an r-back-and-forth

system, since every interval of length $2r$ centered at a point of $[0,1]$ contains infinitely many elements of each structure; so $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent. For each of the two structures and distinct elements a and a' of it, the formula $E(y,x)$ with the parameter a for y gives $d(a,a') \geq 2$, where d is the pseudometric of that structure (Remark 20), and the a priori bounds give $d(a,a') \leq 2$; so both pseudometrics are discrete and both structures are complete. The structures are not isomorphic, since p takes the value 0 in $\mathfrak{M}$ and not in $\mathfrak{N}$.

13. Scott's Theorem in the Space of Countable Structures

In classical logic, Scott's theorem has a topological form. The structures whose universe is the set of natural numbers form a Polish space, the isomorphism classes are the orbits of the group of permutations of the natural numbers, and Scott's theorem implies that every orbit is a Borel set. López-Escobar's theorem identifies the Borel sets that are invariant under the group with the sets defined by sentences of the infinitary logic [LE65, Vau74].

We give Theorem 17 a corresponding topological form. Since general structures carry no equality, we replace the group of permutations by the monoid of surjections of the natural numbers, acting on atomic diagrams. For structures on the natural numbers, exact equivalence means that the orbits of their diagrams meet, and approximate equivalence means that these orbits come arbitrarily close for a uniform distance between diagrams. The resulting theorem (Theorem 24) involves only a Polish space, the action of a Polish monoid and a distance, and the logic enters only the proof, through the Scott sentences of Section 12.

A version of López-Escobar's theorem then marks the reach of this topological form: the bounded Borel functions invariant under exact equivalence are the values of the sentences of a richer logic, and some of them separate approximately equivalent structures, so that replacing the functions of Theorem 24 by all such functions changes the equivalence they detect, from approximate equivalence to exact equivalence.

Throughout the section, the vocabulary $V = (\mathbb{M}, F, P)$ is countable and has a single sort descriptor M; for a countable set of sort descriptors, the same arguments apply with one copy of the set of natural numbers for each sort descriptor. We write x_j for the variable x_j^M, and, for an atomic formula ψ in $(x_j)_{j<m}$ and a tuple $\bar{k} = (k_j)_{j<m}$ of natural numbers, we write $\psi(x_{\bar{k}})$ for the atomic formula obtained from ψ by substituting x_{k_j} for x_j, for every $j < m$; the two formulas have the same predicate symbol, and hence the same a priori bound.

The space of structures on the natural numbers. Let X_V be the set of V-structures $\mathfrak{M}$ with $M^{\mathfrak{M}} = \mathbb{N}$. A structure in X_V is determined by the interpretations of the symbols of V, and X_V is identified with the product

$$\prod_p [-r_p, r_p]^{\mathbb{N}^{n_p}} \times \prod_f \mathbb{N}^{\mathbb{N}^{n_f}},$$

over the predicate symbols p in P other than 1 and the function symbols f in F, with the product topology and the discrete topology on $\mathbb{N}$. As a countable product

of Polish spaces, X_V is a Polish space. For every formula $\varphi(x_0, \dots, x_{m-1})$ of $L^V_{\omega_1\omega}$ and every tuple $\bar{k} \in \mathbb{N}^m$, the map $\mathfrak{M} \mapsto \varphi^{\mathfrak{M}}(\bar{k})$ is Borel on X_V: when φ is atomic, the map is continuous, since, by induction on terms, the value of a term at $\bar{k}$ is locally constant on X_V, the set $\mathbb{N}$ carrying the discrete topology, so that near each structure the value of φ at $\bar{k}$ depends on finitely many coordinates; the connectives $s\cdot$, $+$ and $\vee$ are continuous; and the values of $\sup_x \varphi$ and $\bigvee \Phi$ are suprema of countably many Borel functions.

Let A be the set of pairs $(\psi, \bar{k})$ in which, for some natural number m, the formula ψ is an atomic formula in $(x_j)_{j<m}$ with $r_\psi > 0$ and $\bar{k}$ belongs to $\mathbb{N}^m$. The set A is countable, and the product

$$Y \coloneqq \prod_{(\psi,\bar{k})\in A} [-r_\psi, r_\psi]$$

is a compact metrizable space. The **atomic diagram** of a structure $\mathfrak{M} \in X_V$ is the point $D(\mathfrak{M})$ of Y with coordinates $D(\mathfrak{M})(\psi, \bar{k}) \coloneqq \psi^{\mathfrak{M}}(\bar{k})$, and the map $D\colon X_V \to Y$ is continuous. For points y and z of Y, let

$$\rho(y, z) \coloneqq \sup_{(\psi,\bar{k})\in A} \frac{|y(\psi, \bar{k}) - z(\psi, \bar{k})|}{r_\psi},$$

the uniform distance normalized by the a priori bounds; as a supremum of continuous functions, ρ is lower semicontinuous on $Y \times Y$.

The set $\mathrm{Surj}(\mathbb{N})$ of surjections of $\mathbb{N}$ onto $\mathbb{N}$ is the intersection, over $n \in \mathbb{N}$, of the open subsets $\{g : n \in g(\mathbb{N})\}$ of the Baire space $\mathbb{N}^{\mathbb{N}}$, and is therefore a Polish space [Kec95, Theorem 3.11]; under composition, $\mathrm{Surj}(\mathbb{N})$ is a monoid. For $g \in \mathrm{Surj}(\mathbb{N})$ and $y \in Y$, let g^*y be the point of Y with coordinates $(g^*y)(\psi, \bar{k}) \coloneqq y(\psi, g\bar{k})$, where $g\bar{k} = (g(k_j))_{j<m}$. The map $(g, y) \mapsto g^*y$ is continuous, since each coordinate of g^*y depends on finitely many values of g and on one coordinate of y; moreover $(g\circ h)^* = h^*\circ g^*$ for $g, h \in \mathrm{Surj}(\mathbb{N})$, and $\rho(g^*y, g^*z) \le \rho(y, z)$.

A surjection onto the universe of a structure also determines a structure on the natural numbers. Let $\mathfrak{M}$ be a V-structure whose universe $M^{\mathfrak{M}}$ is countable, and let $g : \mathbb{N} \to M^{\mathfrak{M}}$ be a surjection. The **pullback** $g^*\mathfrak{M}$ is the member of X_V in which, for every predicate symbol p and every function symbol f of arity n and every tuple $\bar{k} \in \mathbb{N}^n$,

$$p^{g^*\mathfrak{M}}(\bar{k}) \coloneqq p^{\mathfrak{M}}(g\bar{k}), \qquad f^{g^*\mathfrak{M}}(\bar{k}) \coloneqq \min g^{-1}\big(f^{\mathfrak{M}}(g\bar{k})\big),$$

the least preimage under g of the value of $f^{\mathfrak{M}}$. By induction on terms, $g(\tau^{g^*\mathfrak{M}}(\bar{k})) = \tau^{\mathfrak{M}}(g\bar{k})$ for every term $\tau(x_0, \dots, x_{m-1})$ and every $\bar{k} \in \mathbb{N}^m$; consequently, $\psi^{g^*\mathfrak{M}}(\bar{k}) = \psi^{\mathfrak{M}}(g\bar{k})$ for every atomic formula $\psi(x_0, \dots, x_{m-1})$, and $D(g^*\mathfrak{M}) = g^*D(\mathfrak{M})$ when $\mathfrak{M}$ belongs to X_V. When V has function symbols, the pullbacks chosen by least preimages need not compose as an action of the monoid $\mathrm{Surj}(\mathbb{N})$ on X_V; the action is the reindexing of atomic diagrams, and the proofs below use only the identity $D(g^*\mathfrak{M}) = g^*D(\mathfrak{M})$ and the exact equivalence of $g^*\mathfrak{M}$ with $\mathfrak{M}$, established below.

The comparison of structures uses the distance given by back-and-forth systems. For V-structures $\mathfrak{M}$ and $\mathfrak{N}$, let $\delta(\mathfrak{M}, \mathfrak{N})$ be the infimum of the real numbers $r > 0$

for which there is an r-back-and-forth system between $\mathfrak{M}$ and $\mathfrak{N}$ (Definition 10). The values of an atomic formula ψ lie in $[-r_\psi, r_\psi]$, so the set of all pairs of tuples of the same length is a 2-back-and-forth system, and $\delta(\mathfrak{M}, \mathfrak{N}) \leq 2$; the structures $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent exactly when $\delta(\mathfrak{M}, \mathfrak{N}) = 0$. An **exact back-and-forth system** between $\mathfrak{M}$ and $\mathfrak{N}$ is a set of pairs of tuples that contains the pair of empty tuples and satisfies conditions (ii) and (iii) of Definition 10 and condition (i) with $r = 0$, that is, with equality of the values of every atomic formula; the structures are **exactly equivalent** if there is such a system between them. Exactly equivalent structures are approximately equivalent. For a surjection $g : \mathbb{N} \to M^{\mathfrak{M}}$, the pairs $(\bar{k}, g\bar{k})$, for the tuples $\bar{k}$ of natural numbers, form an exact back-and-forth system between $g^*\mathfrak{M}$ and $\mathfrak{M}$: condition (ii) holds with the element $g(k)$ for $k \in \mathbb{N}$, and condition (iii) with a preimage under g of the given element. Every V-structure whose universe is countable is therefore exactly equivalent to a member of X_V, and the space X_V represents all such structures.

Lemma 22. For all structures $\mathfrak{M}$ and $\mathfrak{N}$ in X_V,

$$\delta(\mathfrak{M}, \mathfrak{N}) = \inf_{p,q \in \mathrm{Surj}(\mathbb{N})} \rho\big(p^*D(\mathfrak{M}), q^*D(\mathfrak{N})\big),$$

and $\mathfrak{M}$ and $\mathfrak{N}$ are exactly equivalent if and only if $p^*D(\mathfrak{M}) = q^*D(\mathfrak{N})$ for some surjections p and q.

Proof. Let surjections p and q and a real number $r \geq \rho(p^*D(\mathfrak{M}), q^*D(\mathfrak{N}))$ be given, and let I be the set of pairs $(p\bar{k}, q\bar{k})$ for the tuples $\bar{k}$ of natural numbers. The set I contains the pair of empty tuples. For condition (i) of Definition 10, let ψ be an atomic formula in $(x_j)_{j<m}$ and let $\bar{k} \in \mathbb{N}^m$; if $r_\psi = 0$, both values of ψ vanish, and otherwise $|\psi^{\mathfrak{M}}(p\bar{k}) - \psi^{\mathfrak{N}}(q\bar{k})|$ is the absolute value of the coordinate $(\psi, \bar{k})$ of $p^*D(\mathfrak{M}) - q^*D(\mathfrak{N})$, and is at most $r\, r_\psi$. For condition (ii), an element c of $\mathfrak{M}$ is $p(j)$ for some $j \in \mathbb{N}$, since p is surjective, and the pair $(p\bar{k}\, p(j), q\bar{k}\, q(j))$ belongs to I; condition (iii) holds in the same way. Hence I is an r-back-and-forth system for every such $r > 0$, and an exact one when $p^*D(\mathfrak{M}) = q^*D(\mathfrak{N})$; in particular, $\delta(\mathfrak{M}, \mathfrak{N})$ is at most the infimum in the statement.

Conversely, let I be an r-back-and-forth system between $\mathfrak{M}$ and $\mathfrak{N}$, for a real number $r > 0$, or an exact one, for $r = 0$. Choosing pairs alternately by conditions (ii) and (iii), each extending the previous one, we obtain sequences $(a_j)_{j\in\mathbb{N}}$ and $(b_j)_{j\in\mathbb{N}}$ of natural numbers such that the pair $((a_j)_{j<m}, (b_j)_{j<m})$ belongs to I for every m, and every natural number occurs among the numbers a_j and among the numbers b_j. The maps $p : j \mapsto a_j$ and $q : j \mapsto b_j$ are surjections. Let $(\psi, \bar{k}) \in A$, with ψ in $(x_j)_{j<m}$, and let m' be a natural number larger than every entry of $\bar{k}$. The formula $\psi(x_{\bar{k}})$ is an atomic formula in $(x_j)_{j<m'}$ with a priori bound r_ψ, and the values of $\psi(x_{\bar{k}})$ at the tuples $(a_j)_{j<m'}$ and $(b_j)_{j<m'}$ are $\psi^{\mathfrak{M}}(p\bar{k})$ and $\psi^{\mathfrak{N}}(q\bar{k})$; condition (i) gives $|\psi^{\mathfrak{M}}(p\bar{k}) - \psi^{\mathfrak{N}}(q\bar{k})| \leq r\, r_\psi$. Hence $\rho(p^*D(\mathfrak{M}), q^*D(\mathfrak{N})) \leq r$, so the infimum in the statement is at most $\delta(\mathfrak{M}, \mathfrak{N})$, and $p^*D(\mathfrak{M}) = q^*D(\mathfrak{N})$ when I is exact. □

Two structures in X_V are therefore approximately equivalent when, for every real number $r > 0$, the orbits of their atomic diagrams under the monoid $\mathrm{Surj}(\mathbb{N})$

come within distance r for ρ, and exactly equivalent when the two orbits have a common point. In classical logic, the group of permutations of $\mathbb{N}$ plays this part. Without equality the group does not suffice, since a structure is exactly equivalent to the structure obtained by giving each element a twin.

Example 23. Assume that the only predicate symbol of V besides 1 is $p\colon M \times M \to [-1,1]$, and that V has no function symbols. Let $\mathfrak{M}_1$ and $\mathfrak{M}_2$ be the members of X_V in which $p^{\mathfrak{M}_1}(a,b)$ is 1 for $a = b$ and 0 otherwise, and $p^{\mathfrak{M}_2}(a,b)$ is 1 for $\lfloor a/2 \rfloor = \lfloor b/2 \rfloor$ and 0 otherwise. The surjection $p\colon a \mapsto \lfloor a/2 \rfloor$ satisfies $p^* D(\mathfrak{M}_1) = D(\mathfrak{M}_2)$, so that $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are exactly equivalent by Lemma 22. For permutations p and q of $\mathbb{N}$, however, let a and b be the natural numbers with $q(a) = 0$ and $q(b) = 1$; the coordinate $(p(x_0, x_1), (a,b))$ of $q^* D(\mathfrak{M}_2)$ is 1, and that of $p^* D(\mathfrak{M}_1)$ is 0, since $p(a) \neq p(b)$. Thus $\rho(p^* D(\mathfrak{M}_1), q^* D(\mathfrak{M}_2)) = 1$ for all permutations p and q.

Scott's theorem. The following theorem is stated in the language of the space X_V alone.

Theorem 24 (Scott's theorem in topological form)**.** Assume that the vocabulary V is countable and has a single sort descriptor, and let X_V, Y, D, ρ and the action of $\mathrm{Surj}(\mathbb{N})$ on Y be as above. For $\mathfrak{M}$ and $\mathfrak{N}$ in X_V, let

$$\delta(\mathfrak{M}, \mathfrak{N}) \coloneqq \inf_{p,q \in \mathrm{Surj}(\mathbb{N})} \rho\big(p^* D(\mathfrak{M}), q^* D(\mathfrak{N})\big).$$

Then δ is a pseudometric on X_V, and, for every $\mathfrak{M} \in X_V$, there is a Borel function $f_{\mathfrak{M}}\colon X_V \to [0,2]$, Lipschitz for δ, whose zero set is $\{\mathfrak{N} \in X_V : \delta(\mathfrak{M}, \mathfrak{N}) = 0\}$. In particular, every class of the equivalence relation $\delta(\mathfrak{M}, \mathfrak{N}) = 0$ is a Borel subset of X_V, and two members of X_V lie in the same class if and only if no Borel function on X_V that is Lipschitz for δ separates them.

Proof. By Lemma 22, the function δ is the distance given by back-and-forth systems. The pairs $(\bar{k}, \bar{k})$ form an exact back-and-forth system between a structure and itself, the inverse of an r-back-and-forth system is an r-back-and-forth system, and, for an r-back-and-forth system I between $\mathfrak{M}$ and $\mathfrak{N}$ and an s-back-and-forth system J between $\mathfrak{N}$ and $\mathfrak{P}$, the pairs $(\bar{a}, \bar{c})$ for which some tuple $\bar{b}$ satisfies $(\bar{a}, \bar{b}) \in I$ and $(\bar{b}, \bar{c}) \in J$ form an $(r+s)$-back-and-forth system between $\mathfrak{M}$ and $\mathfrak{P}$; hence δ is a pseudometric.

Let $\sigma_{\mathfrak{M}}$ be the Scott sentence of $\mathfrak{M}$ (Definition 16), and let $f_{\mathfrak{M}}(\mathfrak{N}) \coloneqq \sigma_{\mathfrak{M}}^{\mathfrak{N}}$. The values of $\sigma_{\mathfrak{M}}$ are nonnegative and the a priori bound of $\sigma_{\mathfrak{M}}$ is at most 2, so $f_{\mathfrak{M}}$ takes values in $[0,2]$, and $f_{\mathfrak{M}}$ is Borel, as the value of a sentence. For $\mathfrak{N}$ and $\mathfrak{N}'$ in X_V and an r-back-and-forth system between them, Lemma 11, applied to the pair of empty tuples, gives $|\sigma_{\mathfrak{M}}^{\mathfrak{N}} - \sigma_{\mathfrak{M}}^{\mathfrak{N}'}| \leq 2r$; hence $|f_{\mathfrak{M}}(\mathfrak{N}) - f_{\mathfrak{M}}(\mathfrak{N}')| \leq 2\,\delta(\mathfrak{N}, \mathfrak{N}')$. By Theorem 17, $f_{\mathfrak{M}}(\mathfrak{N}) = 0$ exactly when $\mathfrak{M}$ and $\mathfrak{N}$ are approximately equivalent, that is, when $\delta(\mathfrak{M}, \mathfrak{N}) = 0$. The class of $\mathfrak{M}$ is the zero set of the Borel function $f_{\mathfrak{M}}$, and is therefore Borel. A function that is Lipschitz for δ is constant on every class, and $f_{\mathfrak{M}}$ separates the class of $\mathfrak{M}$ from every other class. □

In classical logic, the theorem gives back the fact that the isomorphism classes of countable structures are Borel.

Corollary 25. Assume that V is countable and has a single sort descriptor M and a predicate symbol $E : M \times M \to [-1, 1]$. Let X_E be the set of members of X_V in which every atomic formula ψ other than 1 takes only the two values r_ψ and $-r_\psi$ and E takes the value 1 exactly on the diagonal. For $\mathfrak{M}$ and $\mathfrak{N}$ in X_E, the equality $\delta(\mathfrak{M}, \mathfrak{N}) = 0$ holds if and only if $\mathfrak{N} = g^*\mathfrak{M}$ for some permutation g of $\mathbb{N}$, and every orbit of the group of permutations of $\mathbb{N}$ on X_E is a Borel subset of X_V.

Proof. For a permutation g, the pullback $g^*\mathfrak{M}$ is the structure transported from $\mathfrak{M}$ along g, and g is an isomorphism from $g^*\mathfrak{M}$ onto $\mathfrak{M}$; conversely, an isomorphism g from $\mathfrak{N}$ onto $\mathfrak{M}$ is a permutation with $\mathfrak{N} = g^*\mathfrak{M}$. By Theorem 17 and Corollary 18, $\delta(\mathfrak{M}, \mathfrak{N}) = 0$ exactly when $\mathfrak{N}$ is isomorphic to $\mathfrak{M}$. The set X_E is closed in X_V, and the orbit of $\mathfrak{M}$ is the intersection of X_E with the class of $\mathfrak{M}$, a Borel set by Theorem 24. □

López-Escobar's theorem and the limits of the topological form. Theorem 24 separates the classes of approximate equivalence by Borel functions that are Lipschitz for δ. Theorem 27 below represents all bounded Borel functions invariant under exact equivalence by the sentences of a logic with continuous connectives, and Example 26 shows that some of these functions separate approximately equivalent structures, so that the two classes of functions differ.

Let $L^{c,V}_{\omega_1\omega}$ be the logic obtained from $L^V_{\omega_1\omega}$ by adding, for every integer $\ell \geq 1$ and every continuous function $u : \mathbb{R}^\ell \to \mathbb{R}$, a connective that forms from formulas $\varphi_1, \dots, \varphi_\ell$ the formula $u(\varphi_1, \dots, \varphi_\ell)$, with value $u(\varphi_1^{\mathfrak{M}}(\bar{a}), \dots, \varphi_\ell^{\mathfrak{M}}(\bar{a}))$ and with a priori bound the maximum of $|u|$ on the product of the intervals $[-r_{\varphi_i}, r_{\varphi_i}]$; joins are formed, as in $L^V_{\omega_1\omega}$, from sequences of formulas with a common a priori bound and finitely many free variables. By induction on formulas, the value of every formula of $L^{c,V}_{\omega_1\omega}$ lies within the a priori bound of the formula, the map $\mathfrak{M} \mapsto \varphi^{\mathfrak{M}}(\bar{k})$ is Borel on X_V for every formula φ and every tuple $\bar{k}$, and exactly equivalent structures give every sentence the same value; for the last assertion, the proof of Lemma 11 applies with equalities in place of the inequalities.

Example 26. Assume that the only predicate symbol of V besides 1 is $p : M \to [-1, 1]$, and that V has no function symbols. Let $\mathfrak{M}$ be a member of X_V in which p takes each of the values $\frac{1}{2}$ and $\frac{1}{2} + \frac{1}{k}$, for integers $k \geq 2$, at infinitely many elements and takes no other value, and let $\mathfrak{N}$ be a member of X_V in which p takes each of the values $\frac{1}{2} + \frac{1}{k}$, for integers $k \geq 2$, at infinitely many elements and takes no other value. For a real number $r > 0$, the pairs $(\bar{a}, \bar{b})$ of tuples of the same length with $|p^{\mathfrak{M}}(a_i) - p^{\mathfrak{N}}(b_i)| \leq r$ for every i form an r-back-and-forth system, since every value of p in either structure lies within r of the values of p at infinitely many elements of the other; thus $\delta(\mathfrak{M}, \mathfrak{N}) = 0$. For integers $k \geq 1$, let $u_k(t) := \max\{0, 1 - k|t - \frac{1}{2}|\}$, and let θ be the sentence $\sup_{x_0} \bigwedge_k u_k(p(x_0))$ of $L^{c,V}_{\omega_1\omega}$, with a priori bound 1. The value of θ is 1 in $\mathfrak{M}$ and 0 in $\mathfrak{N}$. The function $\mathfrak{N}' \mapsto \theta^{\mathfrak{N}'}$ is therefore Borel on X_V and invariant under exact equivalence, but not Lipschitz for δ; by Lemma 11, no sentence of $L^V_{\omega_1\omega}$ has the same values. As in Example 19, the obstruction is

the absence of a common a priori bound: θ distinguishes the value $\frac{1}{2}$ from values arbitrarily close to $\frac{1}{2}$.

López-Escobar's theorem for $L^{c,V}_{\omega_1\omega}$ states that the sentences of $L^{c,V}_{\omega_1\omega}$ give all the invariant Borel functions. Ben Yaacov, Doucha, Nies and Tsankov proved a version of López-Escobar's theorem for separable metric structures [BDNT17]; the version below requires no metric.

Theorem 27 (López-Escobar's theorem for exact equivalence)**.** Assume that the vocabulary V is countable and has a single sort descriptor, and let $f : X_V \to \mathbb{R}$ be a bounded function. Then f is Borel and takes the same value at exactly equivalent structures if and only if there is a sentence φ of $L^{c,V}_{\omega_1\omega}$ with $f(\mathfrak{M}) = \varphi^{\mathfrak{M}}$ for every $\mathfrak{M} \in X_V$.

The proof rests on Vaught's transforms [Vau74], [Kec95, Section 16.B], taken here over the monoid Surj(ℕ). For a finite sequence $s = (s_j)_{j<n}$ of natural numbers, let U_s be the set of surjections g with $g(j) = s_j$ for every $j < n$. The sets U_s are nonempty and open, they form a basis of the topology of Surj(ℕ), and every nonempty open subset of U_s includes a set U_t for some finite sequence t that extends s. Let $F : Y \to \mathbb{R}$ be a bounded Borel function, let $y \in Y$, and let U be a nonempty open subset of Surj(ℕ). The map $g \mapsto F(g^* y)$ is Borel on U, and therefore has the Baire property [Kec95, Proposition 8.22]. The **Vaught transforms** of F are

$$F^{*U}(y) \coloneqq \inf\big\{t \in \mathbb{R} : \{g \in U : F(g^* y) > t\} \text{ is meager}\big\}$$
$$F_{*U}(y) \coloneqq -(-F)^{*U}(y),$$

the essential supremum and the essential infimum, in the sense of category, of the function $g \mapsto F(g^* y)$ on U.

Lemma 28. Let $F : Y \to \mathbb{R}$ be a bounded Borel function, let $y \in Y$, and let s be a finite sequence of natural numbers. Then $F_{*U_s}(y)$ is the infimum of the numbers $F^{*U_t}(y)$ over the finite sequences t that extend s.

Proof. Let $h(g) \coloneqq F(g^* y)$ and $e \coloneqq F_{*U_s}(y)$. For a finite sequence t that extends s and a real number $\epsilon > 0$, the set $\{g \in U_s : h(g) > e - \epsilon\}$ is comeager in U_s, hence in the open subset U_t, and is therefore not meager in U_t; hence $F^{*U_t}(y) \geq e - \epsilon$, and so $F^{*U_t}(y) \geq e$. Conversely, let $q > e$ be a real number. The set $\{g \in U_s : h(g) < q\}$ is not meager and has the Baire property, and is therefore comeager in some nonempty open subset of U_s [Kec95, Proposition 8.26], and hence in a set U_t for a finite sequence t that extends s. Then $F^{*U_t}(y) \leq q$. □

A function $F : Y \to \mathbb{R}$ is **basic** if, for some integer $\ell \geq 1$, there are points $(\psi_i, \bar{k}_i)$ of A, for $i = 1, \dots, \ell$, and a continuous function $u : \mathbb{R}^\ell \to \mathbb{R}$ such that

$$F(y) = u\big(y(\psi_1, \bar{k}_1), \dots, y(\psi_\ell, \bar{k}_\ell)\big)$$

for every $y \in Y$.

Lemma 29. Let $c \geq 0$ be a real number and let $F\colon Y \to [-c, c]$ be a Borel function. For every natural number n, there are formulas $\varphi^+_{F,n}$ and $\varphi^-_{F,n}$ of $L^{c,V}_{\omega_1\omega}$, with free variables among $x_0, \dots, x_{n-1}$ and with a priori bounds at most $c+2$, such that

$$F^{*U_s}(D(\mathfrak{M})) = \big(\varphi^+_{F,n}\big)^{\mathfrak{M}}(s) \qquad \text{and} \qquad F_{*U_s}(D(\mathfrak{M})) = \big(\varphi^-_{F,n}\big)^{\mathfrak{M}}(s)$$

for every $\mathfrak{M} \in X_V$ and every $s \in \mathbb{N}^n$.

Proof. Let $\mathcal{C}$ be the class of Borel functions $G\colon Y \to [-c-2, c+2]$ for which such formulas exist for every n, with a priori bounds at most $c+2$.

Let G be a basic function with values in $[-c-2, c+2]$, given by points $(\psi_i, \bar{k}_i)$ of A, for $i = 1, \dots, \ell$, and a continuous function u, and let J be the finite set of the entries of the tuples $\bar{k}_i$. For $\mathfrak{M} \in X_V$ and $g \in \mathrm{Surj}(\mathbb{N})$, the value $G(g^*D(\mathfrak{M})) = u\big(\psi_1^{\mathfrak{M}}(g\bar{k}_1), \dots, \psi_\ell^{\mathfrak{M}}(g\bar{k}_\ell)\big)$ depends only on the restriction of g to J. For $s \in \mathbb{N}^n$, the surjections in U_s take the prescribed values s_j at the points $j < n$ of J, and the surjections in U_s that extend a given assignment of natural numbers to the points $j \geq n$ of J form a nonempty open set, on which the function $g \mapsto G(g^*D(\mathfrak{M}))$ is constant. The essential supremum on U_s is therefore the supremum over all such assignments, repeated values included, and is the value at s of the formula

$$\varphi^+_{G,n} \coloneqq \sup_{x_{j_1}} \cdots \sup_{x_{j_e}} u\big(\psi_1(x_{\bar{k}_1}), \dots, \psi_\ell(x_{\bar{k}_\ell})\big),$$

where $j_1, \dots, j_e$ are the points of J that are at least n; the formula $\varphi^-_{G,n}$ is the same with infima in place of suprema. The points $(\psi_i, \bar{k}_i)$ need not be distinct, so the maximum of $|u|$ on the product of the intervals $[-r_{\psi_i}, r_{\psi_i}]$ may exceed that of $|G|$; composing u with the retraction of $\mathbb{R}$ onto $[-c-2, c+2]$, which changes no value of G, we may assume that $|u| \leq c+2$ everywhere. The formula under the quantifiers is the connective u applied to atomic formulas, and the a priori bound of $\varphi^\pm_{G,n}$, the maximum of $|u|$ on the product of the intervals $[-r_{\psi_i}, r_{\psi_i}]$, is at most $c+2$; so $G \in \mathcal{C}$. No condition of distinctness enters the formula; at this step surjections, unlike permutations, dispense with equality.

If $G \in \mathcal{C}$, then $-G \in \mathcal{C}$, with $\varphi^\pm_{-G,n} \coloneqq -\varphi^\mp_{G,n}$. Let $G = \sup_k G_k$ for a sequence $(G_k)_{k\in\mathbb{N}}$ in $\mathcal{C}$. A countable union of meager sets is meager, so $G^{*U} = \sup_k G_k^{*U}$ for every nonempty open set U, and we let $\varphi^+_{G,n} \coloneqq \bigvee_k \varphi^+_{G_k,n}$. By Lemma 28, $G_{*U_s}(D(\mathfrak{M}))$ is the infimum, over the integers $m \geq n$ and the tuples $t \in \mathbb{N}^m$ that extend s, of the values $G^{*U_t}(D(\mathfrak{M}))$, and we let

$$\varphi^-_{G,n} \coloneqq \bigwedge_{m\geq n} \inf_{x_n} \cdots \inf_{x_{m-1}} \varphi^+_{G,m}.$$

Hence $\mathcal{C}$ is closed under countable suprema and, with negation, under countable infima, and therefore under pointwise limits of sequences, since $\lim_k G_k = \inf_j \sup_{k\geq j} G_k$; every a priori bound in these formulas is at most $c+2$.

The basic functions form an algebra of continuous functions on the compact metrizable space Y that contains the constants and separates points, and the Stone–Weierstrass theorem gives, for every continuous function $G\colon Y \to [-c, c]$, basic

functions G_k with $|G - G_k| \leq 2^{-k}$ for $k \in \mathbb{N}$; the functions $G_k + 2^{-k}$ are basic, with values in $[-c, c+2]$, and G is their infimum. Hence $\mathcal{C}$ contains every continuous function from Y to $[-c, c]$. Composing with the retraction of $\mathbb{R}$ onto $[-c, c]$, a continuous function, we see that every Borel function from Y to $[-c, c]$ is obtained from continuous functions from Y to $[-c, c]$ by iterated pointwise limits of sequences [Kec95, Theorem 24.3], and therefore belongs to $\mathcal{C}$. □

Proof of Theorem 27. The values of a sentence of $L^{c,V}_{\omega_1\omega}$ form a bounded Borel function on X_V that takes the same value at exactly equivalent structures.

Conversely, let f be such a function, with values in $[-c, c]$. Two members of X_V with the same atomic diagram are exactly equivalent, by Lemma 22, so $f = F \circ D$ for a function F on the set $D(X_V)$, which is analytic, as the continuous image of a Polish space. For a rational number q, the sets $F^{-1}((q, \infty)) = D(f^{-1}((q, \infty)))$ and $D(X_V) \setminus F^{-1}((q, \infty)) = D(f^{-1}((-\infty, q]))$ are disjoint analytic sets, and Lusin's separation theorem [Kec95, Theorem 14.7] gives a Borel subset B_q of Y with $F^{-1}((q, \infty)) = B_q \cap D(X_V)$. Let B'_q be the union of the sets $B_{q'}$ for the rational numbers $q' \geq q$, and, for $y \in Y$, let $G(y)$ be the supremum of the rational numbers $q \in [-c, c]$ with $y \in B'_q$, or $-c$ if there are none. For $-c \leq t < c$, the set $\{y : G(y) > t\}$ is the union of the Borel sets B'_q for the rational numbers $q \in (t, c]$; for $t < -c$ it is Y, and for $t \geq c$ it is empty. Thus $G : Y \to [-c, c]$ is Borel. For $y \in D(X_V)$, the point y belongs to B'_q exactly when $F(y) > q$, so $G(y) = F(y)$. Let $\mathfrak{M} \in X_V$. For every $g \in \mathrm{Surj}(\mathbb{N})$, the structure $g^*\mathfrak{M}$ is exactly equivalent to $\mathfrak{M}$, and

$$G(g^* D(\mathfrak{M})) = G(D(g^*\mathfrak{M})) = f(g^*\mathfrak{M}) = f(\mathfrak{M}).$$

The function $g \mapsto G(g^* D(\mathfrak{M}))$ is constant on $\mathrm{Surj}(\mathbb{N}) = U_\varnothing$, for the empty sequence $\varnothing$, and Lemma 29 gives

$$f(\mathfrak{M}) = G^{*U_\varnothing}(D(\mathfrak{M})) = \left(\varphi^+_{G,0}\right)^{\mathfrak{M}}.$$

The sentence $\varphi^+_{G,0}$ of $L^{c,V}_{\omega_1\omega}$ is the required one. □

Theorems 24 and 27 together mark the reach of the topological form. The functions of Theorem 24 are the values of sentences of $L^V_{\omega_1\omega}$, which are Lipschitz for δ by Lemma 11, and hence constant on the classes of approximate equivalence. By Theorem 27, the Borel functions invariant under exact equivalence are the values of the sentences of $L^{c,V}_{\omega_1\omega}$, and Example 26 shows that some of them separate approximately equivalent structures. To replace the separating functions by all bounded Borel functions invariant under exact equivalence is therefore to change the equivalence that they detect, from approximate equivalence to exact equivalence. Whether the sentences of $L^V_{\omega_1\omega}$ give all the bounded Borel functions that are Lipschitz for δ is the remaining question; when the atomic values range over finite sets with a uniform gap, they do.

Theorem 30 (Uniformly separated finite values: representation in the original syntax)**.** Assume that the vocabulary V is countable and has a single sort descriptor. For every predicate symbol p of V other than 1, let Λ_p be a finite subset of $[-r_p, r_p]$,

and assume that there is a real number $\gamma > 0$ with $|\lambda - \lambda'| > \gamma\, r_p$ for every predicate symbol p other than 1 and all distinct $\lambda, \lambda' \in \Lambda_p$. Let X_Λ be the set of members of X_V in which every predicate symbol p other than 1 takes values in Λ_p. Then a bounded function $f : X_\Lambda \to \mathbb{R}$ is Borel and takes the same value at approximately equivalent structures if and only if there is a sentence φ of $L^V_{\omega_1\omega}$ with $f(\mathfrak{M}) = \varphi^{\mathfrak{M}}$ for every $\mathfrak{M} \in X_\Lambda$.

Proof. Sentences of $L^V_{\omega_1\omega}$ give Borel functions that take the same value at approximately equivalent structures, by Corollary 12. For the converse, we may assume that X_Λ is nonempty, since otherwise every sentence represents f. For an atomic formula ψ with predicate symbol p other than 1, let $\Lambda_\psi := \Lambda_p$, and let $\Lambda_1 := \{1\}$. For $r < \gamma$, condition (i) of Definition 10 between members of X_Λ forces equality of the values of every atomic formula, so an r-back-and-forth system between members of X_Λ is exact, and f takes the same value at exactly equivalent members of X_Λ. The closed subset $Y_\Lambda := \prod_{(\psi,\bar{k})\in A} \Lambda_\psi$ of Y is compact metrizable, invariant under the action of $\mathrm{Surj}(\mathbb{N})$, and includes $D(X_\Lambda)$; the set X_Λ is closed in X_V and invariant under pullbacks by surjections. The proofs of Lemmas 28 and 29 and of Theorem 27 apply with X_Λ and Y_Λ in place of X_V and Y, and yield formulas of $L^V_{\omega_1\omega}$, once the basic functions are given by formulas of $L^V_{\omega_1\omega}$ with a priori bounds that depend only on c and γ.

Let $u : \mathbb{R}^\ell \to \mathbb{R}$ be a continuous function, and let $\psi_1, \dots, \psi_\ell$ be atomic formulas with $r_{\psi_i} > 0$. On Y_Λ the arguments of u range over the finite set $\Lambda_u := \prod_{i=1}^{\ell} \Lambda_{\psi_i}$. Choose a rational number C with $\max_{\Lambda_u} |u| \le C \le \max_{\Lambda_u} |u| + 1$ and, for each i, a rational number c_i with $1/(\gamma r_{\psi_i}) \le c_i \le 2/(\gamma r_{\psi_i})$, and, for $\lambda = (\lambda_i)_{i=1}^{\ell} \in \Lambda_u$, let

$$\chi_\lambda := 1 \dot{-} \bigvee_{i=1}^{\ell} c_i \cdot |\psi_i + -\underline{\lambda_i}\,|, \qquad \eta := \bigvee_{\lambda\in\Lambda_u} \Big(\underline{u(\lambda)} \wedge C \cdot \big(2 \cdot \chi_\lambda + -1 \big) \Big),$$

where the finite joins are iterations of $\vee$ and the constants are given by Lemma 9 with $\epsilon = r_{\psi_i}$ for $\underline{\lambda_i}$ and $\epsilon = 1$ for $\underline{u(\lambda)}$. At a point of Λ_u, the formula χ_λ takes the value 1 if the point is λ, and the value 0 otherwise, since some coordinate then differs from λ_i by more than $\gamma\, r_{\psi_i}$; so the member of the join indexed by λ takes the value $u(\lambda)$ at λ and the value $-C \le u$ elsewhere, and η takes the value of u at every point of Λ_u. The a priori bound of χ_λ is at most $1 + \max_i c_i \cdot 3 r_{\psi_i} \le 1 + 6/\gamma$, and that of η is at most $C(3 + 12/\gamma) + 1$. In the proof of Lemma 29, after the retraction of u onto $[-c-2, c+2]$, the maximum of $|u|$ is at most $c+2$, so that $C \le c+3$ for the basic functions that occur, and the a priori bounds of all the formulas in the proof are therefore at most a number that depends only on c and γ. □

For structures in which every atomic formula ψ other than 1 takes only the values r_ψ and $-r_\psi$, Theorem 30 is López-Escobar's theorem for classical $L_{\omega_1\omega}$ without equality, with the monoid of surjections in place of the group of permutations. In general, the question is the following.

Question 31. Assume that the vocabulary V is countable and has a single sort descriptor. Is every bounded Borel function on X_V that is Lipschitz for δ given by a sentence of $L^V_{\omega_1\omega}$?

A positive answer to Question 31 would complement the Scott theorems with a representation theorem in the original syntax, as López-Escobar's theorem complements Scott's theorem in classical logic. Theorem 30 is the positive case: when the atomic values range over finite sets with a uniform gap γ, distinct classes of approximate equivalence lie at distance at least γ for δ, so that every bounded function constant on the classes is Lipschitz on X_Λ; when such a function is also Borel, Theorem 30 supplies a sentence, with an a priori bound that depends only on the range of the function and on γ. When the atomic values accumulate, the problem is to represent these Lipschitz functions while keeping the a priori bounds uniform through the construction. Example 26 shows why invariance under exact equivalence alone does not suffice; its separating function is not Lipschitz for δ, so it leaves the representation problem open.

REFERENCES

[BDNT17] I. Ben Yaacov, M. Doucha, A. Nies and T. Tsankov, *Metric Scott analysis*, Adv. Math. **318** (2017), 46–87.

[BU10] I. Ben Yaacov and A. Usvyatsov, *Continuous first order logic and local stability*, Trans. Amer. Math. Soc. **362** (2010), no. 10, 5213–5259; arXiv:0801.4303.

[CI14] X. Caicedo and J. Iovino, *Omitting uncountable types, and the strength of* $[0, 1]$*-valued logics*, Ann. Pure Appl. Logic **165** (2014), no. 6, 1169–1200; arXiv:1202.5981.

[Čec37] E. Čech, *On bicompact spaces*, Ann. of Math. (2) **38** (1937), no. 4, 823–844.

[Eag14] C. J. Eagle, *Omitting types for infinitary* $[0, 1]$*-valued logic*, Ann. Pure Appl. Logic **165** (2014), no. 3, 913–932; arXiv:1304.5208.

[Eng89] R. Engelking, *General topology*, revised and completed ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989.

[Háj98] P. Hájek, *Metamathematics of fuzzy logic*, Trends in Logic—Studia Logica Library, vol. 4, Kluwer Academic Publishers, Dordrecht, 1998.

[Hen49] L. Henkin, *The completeness of the first-order functional calculus*, J. Symbolic Logic **14** (1949), no. 3, 159–166.

[HI02] C. W. Henson and J. Iovino, *Ultraproducts in analysis*, Analysis and logic (Mons, 1997), London Math. Soc. Lecture Note Ser., vol. 262, Cambridge Univ. Press, Cambridge, 2002, pp. 1–110.

[Iov94] J. Iovino, *Stable theories in functional analysis*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1994.

[Iov99a] J. Iovino, *Stable Banach spaces and Banach space structures. I. Fundamentals*, Models, algebras, and proofs (Bogotá, 1995), Lecture Notes in Pure and Appl. Math., vol. 203, Dekker, New York, 1999, pp. 77–95.

[Iov99b] J. Iovino, *Stable Banach spaces and Banach space structures. II. Forking and compact topologies*, Models, algebras, and proofs (Bogotá, 1995), Lecture Notes in Pure and Appl. Math., vol. 203, Dekker, New York, 1999, pp. 97–117.

[Kec95] A. S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.

[Kei71] H. J. Keisler, *Model theory for infinitary logic: Logic with countable conjunctions and finite quantifiers*, Studies in Logic and the Foundations of Mathematics, vol. 62, North-Holland, Amsterdam–London, 1971.

[Kei73] H. J. Keisler, *Forcing and the omitting types theorem*, Studies in model theory (M. D. Morley, ed.), MAA Stud. Math., vol. 8, Math. Assoc. America, Buffalo, NY, 1973, pp. 96–133.

[Kei23] H. J. Keisler, *Model theory for real-valued structures*, Beyond First Order Model Theory, Vol. II (J. Iovino, ed.), Chapman & Hall/CRC, Boca Raton, FL, 2023, Chapter 2; arXiv:2005.11851.

[KM01] H. J. Keisler and A. W. Miller, *Categoricity without equality*, Fund. Math. **170** (2001), no. 1, 87–106.

[LE65] E. G. K. López-Escobar, *An interpolation theorem for denumerably long formulas*, Fund. Math. **57** (1965), 253–272.

[Mak69] M. Makkai, *An application of a method of Smullyan to logics on admissible sets*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **17** (1969), 341–346.

[Mor74] M. Morley, *Applications of topology to* $L_{\omega_1\omega}$, Proceedings of the Tarski Symposium (L. Henkin et al., eds.), Proc. Sympos. Pure Math., vol. 25, Amer. Math. Soc., Providence, RI, 1974, pp. 233–240.

[Sco65] D. Scott, *Logic with denumerably long formulas and finite strings of quantifiers*, The Theory of Models (J. W. Addison, L. Henkin and A. Tarski, eds.), North-Holland, Amsterdam, 1965, pp. 329–341.

[Vau74] R. L. Vaught, *Invariant sets in topology and logic*, Fund. Math. **82** (1974/75), 269–294.

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT SAN ANTONIO, SAN ANTONIO, TX 78249, USA

Email address: olivia.aubone@utsa.edu

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT SAN ANTONIO, SAN ANTONIO, TX 78249, USA

Email address: eduardo.duenez@utsa.edu

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF TEXAS AT SAN ANTONIO, SAN ANTONIO, TX 78249, USA

Email address: jose.iovino@utsa.edu