METRIC SCOTT ANALYSIS

ITAÏ BEN YAACOV, MICHAL DOUCHA, ANDRÉ NIES, AND TODOR TSANKOV

Abstract. We develop an analogue of the classical Scott analysis for metric structures and infinitary continuous logic. Among our results are the existence of Scott sentences for metric structures and a version of the López-Escobar theorem. We also derive some descriptive set theoretic consequences: most notably, that isomorphism on a class of separable structures is a Borel equivalence relation iff their Scott rank is uniformly bounded below ω_1. Finally, we apply our methods to study the Gromov–Hausdorff distance between metric spaces and the Kadets distance between Banach spaces, showing that the set of spaces with distance 0 to a fixed space is a Borel set.

Contents

1. Introduction 1
2. Continuous infinitary logic 4
3. The back-and-forth hierarchy and Scott ranks 8
4. The space of Polish structures 13
5. The universal weak modulus 15
6. A López-Escobar theorem 17
7. Bounded rank and Borelness of isomorphism 22
8. Two examples for the 1-Lipschitz modulus 26
9. A characterization of CLI Polish groups 28
10. Connections with classical logic 30
References 31

1. Introduction

Two of the foundational results in classical infinitary logic are Scott's theorem [S], producing for every countable structure a canonical sentence that uniquely describes it up to isomorphism, and the theorem of López-Escobar [LE], characterizing the isomorphism-invariant Borel sets of models. These results and the techniques developed around them, *Scott analysis* (based on the back-and-forth method of Ehrenfeucht and Fraïssé) and the *Vaught transforms* (introduced by Vaught [V] to give a new proof of the López-Escobar theorem) have become a cornerstone of infinitary model theory as well as of the descriptive set theoretic study of the complexity of isomorphism of countable models. See, e.g., Gao [G2] for an exposition of the general theory, and Hjorth–Kechris [HK] and Hjorth–Kechris–Louveau [HKL] for some more detailed results. The notion of *Scott rank*, an ordinal that measures the model-theoretic complexity of structures, is also an indispensable tool.

Date: July 19, 2017.
2010 *Mathematics Subject Classification.* Primary 03C75, 03E15.
Key words and phrases. continuous logic, infinitary logic, Scott rank, Scott sentence, López-Escobar theorem, Borel equivalence relations, Gromov–Hausdorff distance, Kadets distance, CLI groups.

The goal of this paper is to develop a parallel theory for infinitary continuous logic. In recent years, there has been a lot of activity in first order continuous logic (see [BBHU] for a gentle introduction) and it turns out that much of the classical first order model theory extends to this setting, often with interesting twists. Perhaps more importantly, it seems that continuous logic is the "correct" setting for applying model-theoretic ideas to functional analysis and operator algebras, areas that have been hitherto far removed from model theory (see, for example, [FHS] and the references therein).

Some progress has been made towards the study of metric structures using classical infinitary logic (see, for example, [D1]). However, it seems that classical logic is too expressive in this setting and continuous logic is more appropriate for descriptive set theoretic applications. We discuss the connections of our approach with classical logic in Section 10.

An extended form of continuous logic, called infinitary here, allows connectives to act on certain infinite collections of formulas. It was introduced in [BI], where the authors obtain some applications to Banach space theory. Two other papers that focus on various versions of the omitting types theorem for infinitary continuous logic (that we also use here) are Caicedo and Iovino [CI] and Eagle [E]. However, their framework is somewhat more general (for example, they do not always require that structures be complete), while we keep the setting from [BI] as it seems to be the most relevant to our purposes.

Our results are inspired by their classical counterparts but, as is usual in continuous logic, new difficulties and interesting phenomena appear that have no analogue in the classical setting. On a philosophical level, this is perhaps best explained by descriptive set theory: the isomorphism equivalence relation of classical countable models (which, in view of Scott's results, is one of the main objects of study) is strictly less complicated than isomorphism of separable metric structures. More precisely, the former is a universal orbit equivalence relation of an action of S_∞ and the latter is (bi-reducible with) the universal orbit equivalence relation for Polish group actions [EFP$^+$]; by Hjorth's results [H] on turbulence, the latter is strictly more complicated.

The basis for classical Scott analysis is given by the back-and-forth equivalence relations $\equiv_\alpha$ (originally defined by Fraïssé) indexed by the countable ordinals α. These can be considered as Borel approximations of the analytic equivalence relation of isomorphism. The first novelty in the metric setting is that these equivalence relations are replaced by pseudo-distances r_α (i.e., distinct points can have distance 0) that measure how different two tuples, of the same length and possibly coming from different structures, are. These pseudo-distances naturally give rise to equivalence relations E_α: $\bar{a}\, E_\alpha\, \bar{b} \iff r_\alpha(\bar{a}, \bar{b}) = 0$, and $E_\infty = \bigcap_\alpha E_\alpha$. The inductive definition is mostly uneventful, apart from the base case. For classical structures, two tuples are declared to be $\equiv_0$-equivalent if all quantifier-free formulas agree on them; in the metric situation, we would like to define r_0 as the supremum of the difference of the values that quantifier-free formulas take when evaluated on the tuples. This approach meets an immediate obstacle: the difference can be arbitrarily amplified by scaling formulas by a multiplicative constant. Thus we are led to consider formulas with a fixed modulus of continuity and in order to organize this, we introduce the notion of a *weak modulus of continuity* (denoted by Ω) that controls what formulas we are allowed to use in the definition of r_0. It turns out that the weak modulus is an additional parameter in the construction that has no analogue in the classical setting; by varying it one obtains different equivalence relations E_∞ at the end.

The basic results of the general theory are independent of the choice of a weak modulus. Our first theorem is that the pseudo-distances r_α capture exactly the expressive power of formulas of quantifier rank at most α.

Theorem 1.1. *Let A, B be metric structures in the same signature and let $\bar{a} \in A^n, \bar{b} \in B^n$. Then, for any $\alpha < \omega_1$,*

$$r_\alpha(A\bar{a}, B\bar{b}) = \sup_\phi |\phi^A(\bar{a}) - \phi^B(\bar{b})|,$$

where the sup is taken over all n-ary Ω-formulas of quantifier rank α.

We also obtain analogues of Scott's results from the classical setting.

Theorem 1.2. *For every separable structure A, there exists a $\mathcal{L}_{\omega_1\omega}$-sentence σ_A (the* Scott sentence *of A) such that for every separable structure B,*

$$A \, E_\infty \, B \iff \sigma_A(B) = 0.$$

Moreover, the quantifier rank of σ_A is equal to the Scott rank of A plus ω. In particular, the E_∞-class of A is Borel and its complexity is bounded by the Scott rank of A.

With a certain choice of the weak modulus Ω (which we call *universal*), one obtains isomorphism as E_∞-equivalence, exactly as in the classical case. All separable metric structures for a fixed language can be seen as points in a Polish space, as explained in detail in Section 4. Applying Theorem 1.2 tells us that isomorphism classes of separable metric structures are Borel (which is one of the classical applications of Scott sentences). This result, however, is not new: it can be deduced from [EFP$^+$], where the authors prove that isomorphism is reducible to the orbit equivalence relation of a group action. Our approach, however, gives more detailed information and bounds for the Borel complexity of the equivalence class in terms of the Scott rank.

Our next theorem characterizes exactly when the isomorphism equivalence relation is Borel (again inspired by a similar result in the classical setting).

Theorem 1.3. *Let $\cong$ denote the isomorphism equivalence relation of separable structures (in a fixed signature) and let X be an $\cong$-invariant Borel set of structures. Then the following are equivalent:*

(i) $\cong|_X$ *is Borel;*
(ii) *the supremum of the Scott ranks of the elements of X is bounded below ω_1.*

In particular, this theorem provides a new method to show that certain isomorphism equivalence relations are not Borel, so long as one is able to calculate the Scott ranks.

An important connection between infinitary logic and descriptive set theory is provided by the López-Escobar theorem which asserts that the σ-algebra of Borel, $\cong$-invariant sets in the space of models coincides with the algebra of sets definable by $\mathcal{L}_{\omega_1\omega}$-sentences. We obtain an analogue of this theorem in the continuous setting.

Theorem 1.4. *Let U be a bounded, Borel function on the space of separable models that is invariant under isomorphism. Then there exists a $\mathcal{L}_{\omega_1\omega}$-sentence ϕ such that $U(A) = \phi^A$ for every structure A.*

Our proof of this theorem is based on *Vaught transforms* [V]. However, as we do not have a group action readily available, we develop the transforms in a different setting, better adapted to our situation.

A theorem similar to Theorem 1.4 was independently and simultaneously proved by Coskey and Lupini [CL]. The main difference between their approach

and ours is that they only consider structures with universe the Urysohn sphere $\mathbf{U}_1$. In that case, the equivalence relation of isomorphism is given by the action of $\mathrm{Iso}(\mathbf{U}_1)$ and Vaught transforms can be used directly.

With a different natural choice of a weak modulus (the 1-*Lipschitz* one), the pseudo-distance $r_\infty = \sup_\alpha r_\alpha$ defines a coarser equivalence relation than isomorphism and specializes to Gromov–Hausdorff distance for metric spaces and Kadets distance for Banach spaces. This approach to the Gromov–Hausdorff distance has the advantage that it does not require embeddings into a third structure and is defined for arbitrary metric structures, even ones that do not have amalgamation (and where the original definition is not applicable). Combining this with our general results, we obtain the following.

Corollary 1.5. *For a Polish metric space A, the set of Polish metric spaces such that the Gromov–Hausdorff distance to A is 0 is Borel. A similar fact holds for Banach spaces and the Kadets distance.*

Finally, generalizing a theorem of Gao [G1] from the classical setting, we characterize the separable structures with a Scott sentence that has only separable models (that is, it is *absolutely categorical*).

Theorem 1.6. *Let A be a separable metric structure. Then the following are equivalent:*

(i) *The Scott sentence of A only has separable models;*
(ii) *The left uniformity of the group $\mathrm{Aut}(A)$ is complete.*

Acknowledgements. We would like to thank Christian Rosendal for some useful discussions and Maciej Malicki for providing a reference. Part of the work on this project was carried out at Research Centre Coromandel in New Zealand. Continuing work was carried out at the Erwin Schrödinger Institute in Vienna during the thematic program on *Measured group theory*; we are grateful to the Institute and the organizers for the excellent working conditions they provided. This research was partially supported by the Marsden fund of New Zealand, the ANR project GrupoLoco (ANR-11-JS01-008), the ANR project GAMME (ANR-14-CE25-0004), the ERC grant ANALYTIC (no. 259527), and the Franche–Comté region.

We are also grateful to the anonymous referee for a careful reading of the paper and some useful suggestions.

2. Continuous infinitary logic

2.1. **Moduli of continuity.** An important feature of classical $\mathcal{L}_{\omega_1\omega}$-logic is that when one forms an infinite conjunction (or disjunction) of the formulas $\{\phi_i(\bar{x}) : i \in \mathbf{N}\}$, all ϕ_i share the same finite set of free variables $\bar{x}$; as a result, every $\mathcal{L}_{\omega_1\omega}$-formula has only finitely many free variables. The analogous uniformity condition in infinitary continuous logic is ensured by mandating that all ϕ_i obey the same continuity modulus. This ensures that the interpretations of all formulas are uniformly continuous functions (with a modulus that can be determined syntactically). To formalize this, we start with several basic definitions and facts about moduli of continuity.

Definition 2.1. Let n be a natural number or $\mathbf{N}$. A *modulus of arity n* is a function $\Delta\colon [0,\infty)^n \to [0,\infty)$ that is:

(i) non-decreasing, subadditive, vanishing at zero:
$$\Delta(\delta) \leq \Delta(\delta+\delta') \leq \Delta(\delta)+\Delta(\delta'), \qquad \Delta(0)=0;$$
(ii) continuous.

A *weak modulus* is a function $\Omega\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$ that satisfies (i) and is

(ii′) lower semi-continuous in the product topology and separately continuous in each argument.

The main use of a modulus is to measure the uniform continuity of a function defined on a product of finitely many metric spaces.

Definition 2.2. Let Δ be an n-ary modulus, and let $X = \prod_{i<n} X_i$ be a product of metric spaces. On X^2, define

$$d^\Delta(x,y) = \Delta\big(d^{X_i}(x_i,y_i) : i \in n\big).$$

If Z is another metric space and $f\colon X \to Z$ is a map, we say that f *respects* (or *obeys*) Δ if for all $x, y \in X$, we have

$$d_Z(f(x),f(y)) \leq d^\Delta(x,y).$$

The conditions in the definition of a modulus are chosen in such a way that d^Δ is a continuous pseudo-distance on any product of metric spaces. If Δ is moreover *faithful*, i.e., $\Delta(\delta) = 0$ implies that $\delta = 0$, then d^Δ is a distance compatible with the product uniform structure.

If $K \subseteq \mathbf{R}^n$ is a product of compact intervals and $f\colon K \to \mathbf{R}$ is a continuous function, we define its modulus of continuity Δ_f by

$$\Delta_f(\delta) = \sup\{|f(x) - f(y)| : x, y \in K, |x_i - y_i| \leq \delta_i\}. \tag{2.1}$$

Δ_f is the least modulus of continuity that f obeys.

The main purpose of weak moduli is to control the uniform continuity of formulas and organize together an infinite collection of moduli of different arities. A weak modulus will never be used directly but rather via its traces on finite products. If $\Omega\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$ is a weak modulus and $n \in \mathbf{N}$, define the *truncation* $\Omega|_n\colon [0,\infty)^n \to [0,\infty]$ by

$$\Omega|_n(\delta_0,\dots,\delta_{n-1}) = \Omega(\delta_0,\dots,\delta_{n-1},0,0,\dots).$$

A weak modulus Ω is *shift-increasing* if for every sequence $i_0 < i_1 < \cdots$ of natural numbers and every $\delta \in [0,\infty)^{\mathbf{N}}$, we have $\Omega(\delta) \leq \Omega(\delta')$, where $\delta'_{i_p} = \delta_p$ and $\delta'_k = 0$ if $k \notin \{i_0, i_1, \dots\}$. All natural weak moduli that we have in mind satisfy this condition; however, it is only used in one place in the general theory (Proposition 3.4) and we have preferred to keep it as a separate hypothesis where necessary rather than make it part of the definition of a weak modulus.

The following lemma clarifies the connection between weak moduli and moduli.

Lemma 2.3. *Let $\Omega\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$ be a weak modulus. Then all truncations of Ω are moduli and Ω is determined by its truncations:*

$$\Omega(\delta_0,\delta_1,\dots) = \sup_n \Omega|_n(\delta_0,\dots,\delta_{n-1}). \tag{2.2}$$

Proof. Since $\Omega|_n$ is lower semi-continuous, it suffices to show that it is upper semi-continuous. Fix $\delta \in [0,\infty)^n$ and $s \in \mathbf{R}$ such that $\Omega|_n(\delta) < s$. Using that $\Omega|_n$ is separately continuous in each variable, find consecutively $\gamma_0,\dots,\gamma_{n-1}$ such that

$$\Omega|_n(\delta) \leq \Omega|_n(\delta_0+\gamma_0,\delta_1,\dots,\delta_{n-1}) \leq \cdots \leq \Omega|_n(\delta_0+\gamma_0,\dots,\delta_{n-1}+\gamma_{n-1}) < s.$$

As $\Omega|_n$ is monotone, this shows that $\{\delta \in [0,\infty)^n : \Omega|_n(\delta) < s\}$ is open, completing the proof.

For (2.2), the inequality $\geq$ follows from the monotonicity and $\leq$ follows from lower semi-continuity. □

Thus, given a metric space (X, d), each weak modulus Ω defines a family of distances $d^{\Omega|_n}$ on powers of X; we will often abuse notation and write d^Ω instead when n is clear from the context. Similarly, we will say that a function $f\colon X^n \to \mathbf{R}$ respects Ω rather than that it respects $\Omega|_n$.

Two examples of weak moduli that will be important for us are the following. The 1-*Lipschitz weak modulus* $\Omega_{\mathrm{L}}\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$ is defined by

$$\Omega_{\mathrm{L}}(\delta) = \sup_i \delta_i, \quad \text{where } \delta = (\delta_0, \delta_1, \ldots). \tag{2.3}$$

The *universal weak modulus for Lipschitz languages* $\Omega_{\mathrm{U}}(\mathrm{Lip})\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$ is defined by

$$\Omega_{\mathrm{U}}(\mathrm{Lip})(\delta) = \sum_{i=0}^{\infty} i \cdot \delta_i. \tag{2.4}$$

Both of those weak moduli are shift-increasing.

2.2. **Infinitary logic.** Continuous infinitary logic was first introduced by Ben Yaacov and Iovino in [BI]. The definitions we give here are compatible with theirs.

A *metric language* (or signature) is a collection L of symbols. For each symbol $s \in L$, the language also determines its *kind* (function or predicate), its arity (a natural number n_s), a n_s-ary modulus of continuity Δ_s, and, for predicates, a compact interval $I_s \subseteq \mathbf{R}$ of allowed values for s that we will refer to as a *bound*. The language always contains, implicitly, a binary predicate symbol d with $\Delta_d(\delta_1, \delta_2) = \delta_1 + \delta_2$. The bound for d is determined by the language.

An *L-structure* A is a complete metric space equipped with interpretations of the symbols:

- Each function symbol F is interpreted by a map $F^A\colon A^{n_F} \to A$ respecting the modulus Δ_F;
- Each predicate symbol P is interpreted by a function $P^A\colon A^{n_P} \to \mathbf{R}$ respecting the modulus Δ_P and bound I_P (i.e., $P^A(\bar{a}) \in I_P$ for all $\bar{a} \in A^{n_P}$);
- The symbol d is always interpreted by the distance. It must respect the bound I_d.

Let L be a metric signature. The logic $\mathcal{L}_{\omega_1\omega}(L)$ is defined as follows. First, we fix a family $\{x_i : i \in \mathbf{N}\}$ of distinct *variable symbols*. The syntactic objects of the logic are terms and formulas; terms come equipped with a modulus of continuity that they respect and formulas have a modulus of continuity and a bound.

Terms, atomic formulas, and *basic formulas* are constructed inductively as follows.

- Each x_i is a term that respects the $\mathbf{N}$-ary modulus $\Delta_{x_i}(\delta) = \delta_i$;
- If τ_i, $i < n$ are terms and F is a function symbol of arity n, then $\sigma = F(\bar{\tau})$ is a term that respects $\Delta_\sigma = \Delta_F \circ (\Delta_{\tau_i} : i < n)$;
- If P a predicate symbol of arity n and $\bar{\tau}$ are terms, then $\phi = P(\bar{\tau})$ is an *atomic formula* that respects the modulus $\Delta_\phi = \Delta_P \circ (\Delta_{\tau_i} : i < n)$ and the bound I_P;
- If $\{\phi_i : i < n\}$ are atomic formulas with moduli of continuity Δ_{ϕ_i} and bounds I_{ϕ_i}, and $f\colon \prod_i I_{\phi_i} \to \mathbf{R}$ is continuous, then $\psi = f(\bar{\phi})$ is a *basic formula* that respects the modulus $\Delta_\psi = \Delta_f \circ (\Delta_{\phi_i} : i < n)$, where Δ_f is as per (2.1), and the bound $I_\phi = f(\prod_i I_{\phi_i})$.

Next, we define general $\mathcal{L}_{\omega_1\omega}$-formulas starting from atomic formulas and combining them using finitary connectives, quantifiers, and countable infima and suprema (also called *infinitary connectives*). As before, every formula ϕ respects

some modulus of continuity Δ_ϕ and a bound I_ϕ. If ϕ respects Δ and I, and $\Delta' \geq \Delta$, $I' \supseteq I$, then we will also say that ϕ respects Δ' and I'.

- Every atomic formula is a formula;
- If $\phi_0, \dots, \phi_{n-1}$ are formulas that respect $\Delta_{\phi_0}, \dots, \Delta_{\phi_{n-1}}$ and $I_{\phi_0}, \dots, I_{\phi_{n-1}}$ and $f\colon \prod_i I_{\phi_i} \to \mathbf{R}$ is a continuous function, then $f(\phi_0, \dots, \phi_{n-1})$ is a formula that respects $\Delta_f \circ (\Delta_{\phi_i} : i < n)$ and $f(\prod_i I_{\phi_i})$.
- If ϕ is a formula that respects Δ and I and $i \in \mathbf{N}$, then $\sup_{x_i} \phi$ and $\inf_{x_i} \phi$ are formulas that respect $\hat{\Delta}$ and I, where

$$\hat{\Delta}(\delta_0, \dots, \delta_n) = \Delta(\delta_0, \dots, \delta_{i-1}, 0, \delta_{i+1}, \dots, \delta_n).$$

- If $\{\phi_i : i \in \mathbf{N}\}$ are formulas, Δ is a modulus, and I is a bound such that each ϕ_i respects Δ and I then $\bigvee_i \phi_i$, $\bigwedge_i \phi_i$ are also formulas that respect Δ and I. $\bigvee_i \phi_i$ is interpreted as $\sup_i \phi_i$ and $\bigwedge \phi_i$ is interpreted as $\inf_i \phi_i$.

The *finitary fragment* $\mathcal{L}_{\omega\omega}$ is defined as the set of all $\mathcal{L}_{\omega_1\omega}$-formulas, where the infinitary connectives (the last item above) are not used. Two frequent binary connectives are $\wedge$ (min) and $\vee$ (max).

Finally, we need the notion of (Ω, I)-*formulas* for some given weak modulus Ω and bound $I \subseteq \mathbf{R}$. This definition is more restrictive than that of general formulas in several ways: first, we require that all (Ω, I)-formulas respect Ω and I; second, we only allow 1-Lipschitz connectives in the inductive definition; and third, we keep track of the variables used and quantifiers are allowed only in a certain order. The last restriction is needed when we compute quantifier ranks. Here, the base of the inductive construction are the basic rather than the atomic formulas; that is, we allow applying an arbitrary connective in the beginning. The formal definition of *an n-ary (Ω, I)-formula* is by induction as follows.

- All basic formulas $\phi(x_0, \dots, x_{n-1})$ that only depend on the first n variables and respect Ω and I are n-ary (Ω, I)-formulas.
- If $\{\phi_i : i \in \mathbf{N}\}$ are n-ary (Ω, I)-formulas, then $\bigvee_i \phi_i$ and $\bigwedge_i \phi_i$ are n-ary (Ω, I)-formulas.
- If ϕ is an $(n+1)$-ary (Ω, I)-formula, then $\inf_{x_n} \phi$ and $\sup_{x_n} \phi$ are n-ary (Ω, I)-formulas.
- If $\phi_0, \dots, \phi_{k-1}$ are n-ary (Ω, I)-formulas and $f\colon \mathbf{R}^k \to \mathbf{R}$ is a 1-Lipschitz function (for the max distance on $\mathbf{R}^k$), then $f(\phi_0, \dots, \phi_{k-1})$ is a n-ary $(\Omega, f(I^k))$-formula.
- An n-ary Ω-formula is an n-ary (Ω, I)-formula for some I. An Ω-*formula* is an n-ary Ω-formula for some n. An Ω-*sentence* is a 0-ary Ω-formula.

Note that an n-ary (Ω, I)-formula automatically respects the modulus $\Omega|_n$ and the bound I. Thus the collection of all n-ary (Ω, I)-formulas is equicontinuous and uniformly bounded; in particular, we do not need any further equicontinuity and boundedness requirements in the second item of the definition. Note, however, that, as Ω is not required to be symmetric, our variables are not necessarily interchangeable. If, however, Ω is symmetric (as is the case with Ω_{L}), then we can quantify over any variable and not only over the one with the largest index. The condition that Ω is shift-increasing translates into the fact that we are allowed to substitute variables with bigger indices for free variables in formulas: if $\phi(x_0, \dots, x_n)$ is an Ω-formula and $i_0 < i_1 < \cdots < i_{n-1}$, then $\phi(x_{i_0}, \dots, x_{i_n})$ is also an Ω-formula. This property turns out to be very convenient when one tries to write actual formulas.

The notion of an Ω-formula becomes more permissive as Ω becomes larger; we will see later (Corollary 6.5) that for a certain choice of Ω, every $\mathcal{L}_{\omega_1\omega}$-sentence

is equivalent to an Ω-sentence. On the other hand, the expressive power of Ω_{L}-formulas is strictly weaker than that of full $\mathcal{L}_{\omega_1\omega}$-logic (see Section 8). This distinction is purely an infinitary phenomenon: if one restricts to $\mathcal{L}_{\omega\omega}$, then it follows from [BY2, Corollary 1.7] that any formula can be uniformly approximated by a Lipschitz formula and thus the values of 1-Lipschitz formulas completely determine the values of all $\mathcal{L}_{\omega\omega}$-formulas.

Terms and formulas in $\mathcal{L}_{\omega_1\omega}(L)$ can naturally be interpreted in any L-structure A: every term σ is interpreted as a function $\sigma^A\colon A^{\mathbf{N}} \to A$ and every formula ϕ is interpreted as a function $\phi^A\colon A^{\mathbf{N}} \to \mathbf{R}$ that obeys its modulus and bound. If ϕ is an n-ary formula, then it only depends on the first n variables, so its interpretation can be considered as a function $A^n \to \mathbf{R}$. Sometimes we will write $A \models (\phi = r)$ instead of $\phi^A = r$.

Formulas of fixed arity n are naturally equipped with a seminorm defined as follows:

$$\|\phi\| = \sup\{|\phi(\bar{a})| : A \text{ is a structure and } \bar{a} \in A^n\}. \tag{2.5}$$

The norm $\|\phi\|$ is always finite because an interpretation of a formula is required to obey its bound I_ϕ. The following basic fact will be needed later.

Lemma 2.4. *Let κ be an infinite cardinal. If the language L has size at most κ, then the space of $\mathcal{L}_{\omega\omega}$-formulas, equipped with the norm given by (2.5) has density character at most κ.*

Proof sketch. The only possible problem is that we allow arbitrary continuous functions as connectives and there are uncountably many of them. However, as the space of continuous functions defined on a *compact subset of* $\mathbf{R}^n$ is separable in the uniform norm, we can use a countable collection of connectives and thus obtain a dense set of size κ. (In fact, it is possible to use only finitely many connectives.) □

Note, however, that the space of (even quantifier-free) $\mathcal{L}_{\omega_1\omega}$-formulas has density character $2^{|L|}$.

Remark 2.5. Our framework also allows us to treat unbounded predicates. If P is an unbounded predicate, we replace it with an infinite family $\{P_n : n \in \mathbf{N}\}$ of predicate symbols interpreted as $P_n = \min(P, n)$. In the special case where the distance d is unbounded, we take d_1 to be the "official" distance required by the language. This does not change much as d and d_1 are uniformly equivalent. Note also that isomorphism is preserved by this procedure.

3. The back-and-forth hierarchy and Scott ranks

3.1. **The back-and-forth pseudo-distances.** Throughout, we fix a signature L and a weak modulus $\Omega\colon [0,\infty)^{\mathbf{N}} \to [0,\infty]$.

The r_α pseudo-distances that we define in this subsection are the continuous analogue of the back-and-forth equivalence relations for classical structures. Note that r_α take values in $[0,\infty]$.

Definition 3.1. Let α be an ordinal or the symbol ∞ greater than all ordinals. Let $n \in \mathbf{N}$, let A and B be structures and let $\bar{a} \in A^n$, $\bar{b} \in B^n$. We define the *back-and-forth pseudo-distance (of rank α and arity n, with respect to Ω)*, denoted by $r_{\alpha,n}^{A,B,\Omega}(\bar{a},\bar{b})$ (or simply $r_\alpha^{A,B}(\bar{a},\bar{b})$) by induction on α as follows. For $\alpha = 0$, we set

$$r_0^{A,B}(\bar{a},\bar{b}) = \sup_\phi \left|\phi^A(\bar{a}) - \phi^B(\bar{b})\right|,$$

where ϕ varies over all basic n-ary Ω-formulas. For α limit (or ∞),

$$r_\alpha^{A,B}(\bar{a},\bar{b}) = \sup_{\beta<\alpha} r_\beta^{A,B}(\bar{a},\bar{b}).$$

Finally, for the successor step,

$$r_{\alpha+1}^{A,B}(\bar{a},\bar{b}) = \sup_{c\in A,\, d\in B} \inf_{c'\in A,\, d'\in B} r_\alpha^{A,B}(\bar{a}c,\bar{b}d') \vee r_\alpha^{A,B}(\bar{a}c',\bar{b}d).$$

We may also write $r_{\alpha,n}(A\bar{a},B\bar{b})$ instead of $r_{\alpha,n}^{A,B}(\bar{a},\bar{b})$, allowing A and B to vary together with $\bar{a}$ and $\bar{b}$. In case $n=0$, we write just $r_\alpha(A,B)$.

For the rest of this section, fix a signature L and a weak modulus Ω.

Lemma 3.2.

(i) *For fixed α and n, $r_{\alpha,n}$ is a pseudo-distance on the class of all pairs $A\bar{a}$.*
(ii) *For every α, A and $\bar{a},\bar{b}\in A^n$, we have $r_\alpha^{A,A}(\bar{a},\bar{b}) \le d^\Omega(\bar{a},\bar{b})$.*
(iii) *For fixed α, n, A, and B, the function $r_\alpha^{A,B}$ is uniformly continuous on $A^n\times B^n$, respecting the modulus $\Omega|_n$ on each side. In particular, if $r_\alpha(A\bar{a},B\bar{b})<\infty$ for some $\bar{a}\in A^n$, $\bar{b}\in B^n$, then $r_\alpha(A\bar{c},B\bar{d})<\infty$ for all $\bar{c}\in A^n$, $\bar{d}\in B^n$.*

Proof. All three items are proved by induction on α.

(i) The only non-obvious property is the triangle inequality. For $\alpha=0$ and α limit, this is easy. For the successor step, assume that $r_{\alpha+1}(A\bar{a},B\bar{b})<s$ and $r_{\alpha+1}(B\bar{b},C\bar{c})<t$ in order to show that $r_{\alpha+1}(A\bar{a},C\bar{c})\le s+t$. Fix $d\in A$. Since $s>r_{\alpha+1}(A\bar{a},B\bar{b})$, there exists $e\in B$ such that $r_\alpha(A\bar{a}d,B\bar{b}e)<s$. Similarly, there exists f such that $r_\alpha(B\bar{b}e,\bar{c}f)<t$. By the induction hypothesis, $r_\alpha(A\bar{a}d,C\bar{c}f)<s+t$. Similarly, for all f there exists d such that the same holds, so $r_{\alpha+1}(A\bar{a},C\bar{c})\le s+t$ and we are done.

(ii) For $\alpha=0$, this holds since the interpretation of an n-ary Ω-formula respects $\Omega|_n$. For limit steps, this is clear, and at the successor step, take $d'=c$ and $c'=d$ and note that by the definition of d^Ω, $d^\Omega(\bar{a}c,\bar{b}c)=d^\Omega(\bar{a},\bar{b})$.

(iii) follows from (i) and (ii). □

The next lemma shows that the r_α stabilize at a certain point.

Lemma 3.3. *The following statements hold:*

(i) *If $\beta<\alpha$ then $r_\beta\le r_\alpha$ (i.e., $r_{\beta,n}\le r_{\alpha,n}$ for all n);*
(ii) *If κ is an infinite cardinal and A and B are structures of density character at most κ, then there exists $\alpha<\kappa^+$ such that $r_{\alpha+1}^{A,B}=r_\alpha^{A,B}$. Moreover, in this case, the sequence of $r^{A,B}$ stabilizes beyond α, i.e., $r_\infty^{A,B}=r_\alpha^{A,B}$.*

Proof. (i) We argue by induction on α. For $\alpha=0$ and α limit, there is nothing to show. We now prove the statement for $\alpha+1$ assuming that it holds for α. By the induction hypothesis, it will suffice to show that $r_\alpha\le r_{\alpha+1}$, which we do by distinguishing different cases.

If $\alpha=0$, then $r_0\le r_1$ because a formula $\phi(x_0,\dots x_{n-1},x_n)$ that respects $\Omega|_{n+1}$ and does not depend on x_n also respects $\Omega|_n$. On the other hand, if $\beta<\alpha$, then $r_\beta\le r_\alpha$ by the induction hypothesis, so $r_{\beta+1}\le r_{\alpha+1}$. From this, for both α limit and α successor, we deduce that $r_\alpha\le r_{\alpha+1}$.

(ii) For $\beta<\kappa^+$, $q\in\mathbf{Q}$, and $n\in\mathbf{N}$, let

$$U_{\beta,q,n} = \{(\bar{a},\bar{b})\in A^n\times B^n : r_\beta(\bar{a},\bar{b})>q\}.$$

If we keep q and n fixed, $\{U_{\beta,q,n}:\beta<\kappa^+\}$ is an increasing sequence of open sets in the space $A^n\times B^n$ which has weight κ; therefore the sequence must stabilize at some $\beta(q,n)<\kappa^+$. Finally, set $\alpha=\sup\{\beta(q,n):q\in\mathbf{Q},n\in\mathbf{N}\}$. □

The pseudo-distances r_α define naturally equivalence relations E_α:

$$A\bar{a}\, E_\alpha\, B\bar{b} \iff r_\alpha(A\bar{a}, B\bar{b}) = 0.$$

In view of Lemma 3.3, we naturally have that $E_\alpha \supseteq E_{\alpha+1}$ and $E_\infty = \bigcap_\alpha E_\alpha$.

If A is a separable structure, say that the sequence $(a_i)_{i\in\mathbf{N}}$ of elements of A is *tail-dense* if for every k, $\{a_i : i > k\}$ is dense in A. It is easy to see that a sequence is tail-dense iff it is dense and it hits every isolated point of A infinitely many times. The following is the key back-and-forth fact that will be used throughout the paper.

Proposition 3.4. *Suppose that Ω is shift-increasing. Let A and B be separable structures and let $\bar{v} \in A^k$, $\bar{w} \in B^k$. Then we have that $r_\infty(A\bar{v}, B\bar{w}) < t$ if and only if there exist tail-dense sequences $a \in A^{\mathbf{N}}$ and $b \in B^{\mathbf{N}}$ such that $a|_k = \bar{v}$, $b|_k = \bar{w}$, and*

$$\sup_n r_{0,n}^{A,B}(a|_n, b|_n) < t.$$

Proof. We start with the "only if" part. Let $\mathcal{B}_A$ be a countable basis for A and let $(U_n^A : n \in \mathbf{N})$ be a sequence of open sets such that every element of $\mathcal{B}_A$ appears infinitely often; similarly, define $(U_n^B : n \in \mathbf{N})$ for B.

We construct the desired sequences by a back-and-forth argument; we only describe the "forth" step. Let $\bar{a} = (a_0, \dots, a_{n-1})$ and $\bar{b} = (b_0, \dots b_{n-1})$ be given (for some even $n \geq k$) and suppose that $r_\infty(A\bar{a}, B\bar{b}) < t' < t$. Let a_n be an arbitrary element of $U_{n/2}^A$. We are looking for $b_n \in B$ such that $r_\infty(A\bar{a}a_n, B\bar{b}b_n) < t'$. By Lemma 3.3, there exists $\alpha = \alpha_{A,B}$, so that $r_\infty^{A,B} = r_\alpha^{A,B} = r_{\alpha+1}^{A,B}$. We have

$$\begin{aligned} t' &> r_\alpha(A\bar{a}, B\bar{b}) \\ &= r_{\alpha+1}(A\bar{a}, B\bar{b}) \\ &\geq \sup_{c\in A} \inf_{d'\in B} r_\alpha(A\bar{a}c, B\bar{b}d') \\ &\geq \inf_{d'\in B} r_\alpha(A\bar{a}a_n, B\bar{b}d'). \end{aligned}$$

We obtain that there exists $b_n \in B$ such that $r_\alpha(A\bar{a}a_n, B\bar{b}b_n) < t'$, which allows us to continue. The fact that $a_{2n} \in U_n^A$ and $b_{2n+1} \in U_n^B$ for every n ensures that both sequences are tail-dense. In the end, we have

$$\sup_n r_0(a|_n, b|_n) \leq \sup_n r_\infty(a|_n, b|_n) \leq t' < t,$$

and we are done.

Conversely, suppose we are given sequences $\bar{a}$ and $\bar{b}$ with

$$\sup_n r_{0,n}^{A,B}(\bar{a}|_n, \bar{b}|_n) < t' < t.$$

We show by induction on α that for any $i_0 < \cdots < i_{n-1}$ and all α, we have

$$r_\alpha(Aa_{i_0} \dots a_{i_{n-1}}, Bb_{i_0} \dots b_{i_{n-1}}) \leq t'. \tag{3.1}$$

First consider the case that $\alpha = 0$ and take some $i_0 < \cdots < i_{n-1}$. Then we have

$$r_0(Aa_{i_0}a_{i_1} \dots a_{i_{n-1}}, Bb_{i_0}b_{i_1} \dots b_{i_{n-1}}) \leq r_0(Aa_0a_1 \dots a_{i_{n-1}}, Bb_0b_1 \dots b_{i_{n-1}}) < t,$$

where the fist inequality follows from the fact that Ω is shift-increasing and the second from the assumption.

Suppose now that $\alpha = \beta + 1$ for some $\beta \geq 0$ for which (3.1) has been proved. Fix again some $i_0 < \cdots < i_{n-1}$. Then we have

$$\begin{aligned} r_{\beta+1}(Aa_{i_0} \dots a_{i_{n-1}}, Bb_{i_0} \dots b_{i_{n-1}}) &\leq \limsup_{m\to\infty} r_\beta(Aa_{i_0} \dots a_{i_{n-1}}a_m, Bb_{i_0} \dots b_{i_{n-1}}b_m) \\ &\leq t'. \end{aligned}$$

Indeed, to see why the first inequality holds, suppose that the right-hand side is smaller than s. Let $\epsilon > 0$. Fix $c \in A$ and let $m_k \to \infty$ be such that $a_{m_k} \to c$ and $r_\beta(Aa_{i_0} \dots a_{i_{n-1}} a_{m_k}, Bb_{i_0} \dots b_{i_{n-1}} b_{m_k}) < s + \epsilon$ for all k. Now taking k big enough so that $r_\beta(Aa_{i_0} \dots a_{i_{n-1}} c, Aa_{i_0} \dots a_{i_{n-1}} a_{m_k}) < \epsilon$ (which exists because r_β is contractive in d^Ω by Lemma 3.2), shows that

$$\inf_d r_\beta(Aa_{i_0} \dots a_{i_{n-1}} a_{m_k}, Bb_{i_0} \dots b_{i_{n-1}} d) < s + 2\epsilon.$$

The other term in the inductive definition of $r_{\beta+1}$ is treated in a similar way.

The second inequality follows from the inductive hypothesis. The limit case is trivial. This completes the induction and the proof of the proposition. □

Note that the assumption that Ω is shift-increasing is only used in the "if" direction of the proposition.

3.2. **Quantifier rank.** The *quantifier rank* of a formula ϕ, denoted by $\operatorname{qr}\phi$, is defined by induction as follows:

- $\operatorname{qr}\phi = 0$ if ϕ is an atomic formula;
- $\operatorname{qr} f(\phi_0, \dots, \phi_{n-1}) = \max_i \operatorname{qr}\phi_i$ if f is a connective;
- $\operatorname{qr} \bigvee_i \phi_i = \operatorname{qr} \bigwedge_i \phi_i = \sup_i \operatorname{qr}\phi_i$;
- $\operatorname{qr}(\sup_x \phi) = \operatorname{qr}(\inf_x \phi) = \operatorname{qr}\phi + 1$.

The following theorem tells us that, as in the classical case, the distance r_α captures exactly the expressive power of the Ω-formulas of quantifier rank at most α.

Theorem 3.5. *Let α be an ordinal, $A, B \in \mathcal{M}$, $\bar{a} \in A^n$ and $\bar{b} \in B^n$. Then*

$$r_\alpha^{A,B}(\bar{a}, \bar{b}) = \sup_\phi \left|\phi^A(\bar{a}) - \phi^B(\bar{b})\right|, \tag{3.2}$$

where ϕ varies over all n-ary Ω-formulas of quantifier rank at most α.

Proof. We prove by induction on α that for all bounds I,

$$r_\alpha^{A,B}(\bar{a}, \bar{b}) \wedge |I| = \sup_\phi \left|\phi^A(\bar{a}) - \phi^B(\bar{b})\right|,$$

where $|I|$ denotes the length of I and ϕ varies over all n-ary (Ω, I)-formulas of quantifier rank at most α. For $\alpha = 0$ and limit this is by definition, so assume this for α and let us prove it for $\alpha + 1$. For simplicity, suppose that $\min I = 0$.

Fix $t > 0$, and assume that $r_{\alpha+1}^{A,B}(\bar{a}, \bar{b}) > t$. Without loss of generality, there exists c such that $r_\alpha^{A,B}(\bar{a}c, \bar{b}d) > t$ for all d. By the induction hypothesis, for each d, there exists an $(n+1)$-ary (Ω, I)-formula ϕ_d of quantifier rank $\leq \alpha$ such that $\left|\phi_d^A(\bar{a}, c) - \phi_d^B(\bar{b}, d)\right| > t$, and possibly replacing ϕ_d with another formula of the same kind, we may assume that $\phi_d^A(\bar{a}, c) > t > 0 = \phi_d^B(\bar{b}, d)$. Now, $\psi = \sup_{x_n} \bigwedge_{d \in \mathbf{N}} \phi_d$ is an n-ary (Ω, I)-formula of quantifier rank $\leq \alpha + 1$ and

$$\psi^A(\bar{a}) \geq t > 0 \geq \psi^B(\bar{b}),$$

which is enough.

Conversely, assume that $\sup_\phi \left|\phi^A(\bar{a}) - \phi^B(\bar{b})\right| > t$. Then $\left|\phi^A(\bar{a}) - \phi^B(\bar{b})\right| > t$ for some n-ary Ω-formula ϕ of quantifier rank $\leq \alpha + 1$. If ϕ is of the form $\bigvee_i \phi_i$, $\bigwedge_i \phi_i$, or $f(\phi_0, \dots, \phi_{k-1})$, where f is a 1-Lipschitz connective, then we can replace ϕ with one of the ϕ_i. If ϕ is basic, use the fact that $r_0 \leq r_{\alpha+1}$ (Lemma 3.3). We are left with the case where $\phi = \sup_{x_n} \psi$ (or $\phi = \inf_{x_n} \psi$ but it is similar), where $\operatorname{qr}\psi = \alpha$. We may assume that $\phi^A(\bar{a}) > t > 0 = \phi^B(\bar{b})$. In other words, there exists c such that $\psi^A(\bar{a}, c) > t$ and yet $\psi^B(\bar{b}, d) \leq 0$ for all d. By the induction

hypothesis, $r_\alpha^{A,B}(\bar{a}c, \bar{b}d) > t$ for this one c and all d, so $r_{\alpha+1}^{A,B}(\bar{a}, \bar{b}) \geq t$ and we are done. □

3.3. **Scott rank and Scott sentence.** In this subsection, given a separable structure A, we describe how to associate to it a countable ordinal, its *Scott rank*, and construct a sentence that describes it up to E_∞-equivalence.

Definition 3.6. We call the least ordinal α for which $r_\alpha^{A,B} = r_{\alpha+1}^{A,B}$ the *(Ω-)Scott rank* of the pair A, B, denoted by $\alpha_{A,B}$ (or $\alpha_{\Omega,A,B}$). If $A = B$, we call it the *(Ω-)Scott rank* of A and denote it by α_A.

Note that by Lemma 3.3, if A is infinite, then $\alpha_A < |A|^+$.

Lemma 3.7. *If the structures A and B are E_∞-equivalent, then $\alpha_{A,C} = \alpha_{B,C}$ for any structure C and, in particular, $\alpha_A = \alpha_{A,B} = \alpha_B$.*

Proof. Let $\alpha = \alpha_{A,C}$. By symmetry, it will suffice to prove that $\alpha_{B,C} \leq \alpha$.

Let $\bar{b} \in B^n$ and $\bar{c} \in C^n$. As $r_\infty(A, B) = 0$, for any $\epsilon > 0$, there exists a tuple $\bar{a} \in A^n$ such that

$$r_\alpha(A\bar{a}, B\bar{b}) \leq r_{\alpha+1}(A\bar{a}, B\bar{b}) \leq r_\infty(A\bar{a}, B\bar{b}) < \epsilon.$$

Since $r_\alpha(A\bar{a}, C\bar{c}) = r_{\alpha+1}(A\bar{a}, C\bar{c})$, we have

$$\left|r_\alpha(B\bar{b}, C\bar{c}) - r_{\alpha+1}(B\bar{b}, C\bar{c})\right| < \left|r_\alpha(A\bar{a}, C\bar{c}) - r_{\alpha+1}(A\bar{a}, C\bar{c})\right| + 2\epsilon = 2\epsilon.$$

As ϵ is arbitrary, $r_\alpha(B\bar{b}, C\bar{c}) = r_{\alpha+1}(B\bar{b}, C\bar{c})$, as desired. □

Next we observe that, analogously to the classical case, for every separable structure A, each $\bar{a} \in A^n$, and each ordinal α, there exists a formula $\phi_{\alpha,A\bar{a}}(\bar{x})$ such that for all structures B,

$$\phi^B_{\alpha,n,A\bar{a}}(\bar{b}) = r^{A,B}_{\alpha,n}(\bar{a}, \bar{b}) \wedge 1. \tag{3.3}$$

As formulas are always uniformly bounded, taking the minimum with 1 (or some other constant) above is necessary.

First, we fix a countable, dense subset $D \subseteq A$. Note that the formulas that we define do depend on this choice of D; however for different choices of D, one obtains equivalent formulas. For a countable ordinal α, $n \in \mathbf{N}$ and $\bar{a} \in A^n$, we define inductively the n-ary Ω-formula $\phi_{\alpha,n,A\bar{a}}$ as follows.

For $\alpha = 0$,

$$\phi_{0,n,A\bar{a}}(x_0, \ldots, x_{n-1}) = \bigvee_\phi \left|\phi^A(\bar{a}) - \phi(x_0, \ldots, x_{n-1})\right|,$$

as ϕ varies over a countable family of basic n-ary $(\Omega, [0,1])$-formulas, dense in the norm given by (2.5) (see Lemma 2.4). For α limit,

$$\phi_{\alpha,n,A\bar{a}} = \bigvee_{\beta<\alpha} \phi_{\beta,n,A\bar{a}}.$$

For a successor,

$$\phi_{\alpha+1,n,A\bar{a}}(x_0, \ldots, x_{n-1}) = \Big(\bigvee_{c\in D} \inf_{x_n} \phi_{\alpha,n+1,A\bar{a}c}\Big) \vee \Big(\sup_{x_n} \bigwedge_{c\in D} \phi_{\alpha,n+1,A\bar{a}c}\Big). \tag{3.4}$$

An easy induction shows that $\phi_{\alpha,n,A\bar{b}}$ is an n-ary $(\Omega, [0,1])$-formula of quantifier rank α, and that (3.3) holds.

Now let α_A be the Scott rank of A and note that, as A is separable, by Lemma 3.3, $\alpha_A < \omega_1$. We define σ_A, the *Scott sentence* of A, as

$$\sigma_A = \phi_{\alpha_A,0,A} \vee \bigvee_{n,\,\bar{a}\in D^n} \sup_{x_0,\ldots,x_{n-1}} \tfrac{1}{2}\left|\phi_{\alpha_A,n,A\bar{a}} - \phi_{\alpha_A+1,n,A\bar{a}}\right|. \tag{3.5}$$

This is an $(\Omega, [0,1])$-sentence; the coefficient $\frac{1}{2}$ is needed because the function $(x_1, x_2) \mapsto |x_1 - x_2|$ is 2-Lipschitz and in Ω-formulas, we only allow 1-Lipschitz connectives.

The main property of the Scott sentence is the following.

Theorem 3.8. *Let B be a separable structure. Then $B \models (\sigma_A = 0)$ iff $r_\infty(A, B) = 0$.*

Proof. Assume first that $B \models (\sigma_A = 0)$. Then the second part of σ_A ensures that $r^{A,B}_{\alpha_A} = r^{A,B}_\infty$, and then the first part ensures that $r_{\alpha_A}(A, B) = 0$. Together, $r_\infty(A, B) = 0$.

Conversely, assume that $r_\infty(A, B) = 0$. Then $r_{\alpha_A}(A, B) = 0$ so the first part of σ_A vanishes on B. By Lemma 3.7, we have $\alpha_A = \alpha_{A,B}$, so the second part of σ_A also vanishes on B. □

4. The space of Polish structures

From now on we will assume that the language L is countable, and we will only consider separable structures. Then it is possible to parametrize all L-structures by elements of a Polish space $\mathcal{M}$, in such a way that the pseudo-distances r_α become Borel functions on $\mathcal{M}$.

We will code function symbols by predicates in the following way. If F is an n-ary function symbol with modulus Δ_F, we replace it by the $(n+1)$-ary predicate D_F defined by

$$D_F(\bar{x}, y) = d(F(\bar{x}), y) \tag{4.1}$$

with modulus of continuity Δ_{D_F} given by

$$\Delta_{D_F}(\delta, \delta') = \Delta_F(\delta) + \delta' \tag{4.2}$$

and bound equal to the bound of the metric d. Call L' the resulting language.

Lemma 4.1. *Every $\mathcal{L}_{\omega\omega}(L)$-formula is equivalent to a $\mathcal{L}_{\omega\omega}(L')$-formula. A similar fact holds for $\mathcal{L}_{\omega_1\omega}$.*

Proof. We show by induction that for every term $\tau(\bar{x})$, there exists an L'-formula $D_\tau(\bar{x}, y)$ such that $D_\tau(\bar{x}, y) = d(\tau(\bar{x}), y)$. Suppose that $\tau = F(\tau_0, \dots, \tau_{n-1})$, where F is a function symbol and the τ_i are terms. Define D_τ by

$$D_\tau(\bar{x}, y) = \inf_{\bar{z}} \{D_F(\bar{z}, y) : \bigvee_i D_{\tau_i}(\bar{x}, z_i) = 0\}.$$

This can be written as a legitimate $\mathcal{L}_{\omega\omega}$-formula by [BBHU, Theorem 9.17].

Similarly, if $P(\bar{\tau}_0, \dots, \tau_{n-1})$ is an atomic formula, the following L'-formula is equivalent to it:

$$P'(\bar{x}) = \inf_{\bar{y}} \{P(\bar{y}) : \bigvee_i D_{\tau_i}(\bar{x}, y_i) = 0\}.$$

Now the lemma follows by induction on formulas. □

Enumerate all predicates in L' as $d = P_0, P_1, P_2, \dots$ and let $n_0, n_1, \dots$ be their respective arities. Let $\mathcal{M}(L) = \mathcal{M}(L')$ be the set of all $p \in \prod_i \mathbf{R}^{\mathbf{N}^{n_i}}$ such that there exists an L-structure A and a tail-dense sequence $(a_i)_{i \in \mathbf{N}}$ of elements of A such that

$$p(i)(j_0, \dots, j_{n_i - 1}) = P_i^A(a_{j_0}, \dots, a_{j_{n_i-1}})$$

for all $i \in \mathbf{N}$, $(j_0, \dots, j_{n_i-1}) \in \mathbf{N}^{n_i}$. We will also often write just $\mathcal{M}$ when the language L is clear from the context.

Proposition 4.2. *$\mathcal{M}$ is a G_δ subset of $\prod_i \mathbf{R}^{\mathbf{N}^{n_i}}$, and therefore a Polish space.*

Proof. It is easy to check that $p \in \mathcal{M}$ iff the following hold: $p(0)$ defines a pseudo-distance on $\mathbf{N}$; the set $\mathbf{N}$ is tail-dense in the metric space $(\mathbf{N}, p(0))$; for every $i > 0$, $p(i)\colon \mathbf{N}^{n_i} \to \mathbf{R}$ respects the modulus Δ_{P_i} and the bound I_{P_i} on $\mathbf{N}$; finally, for every predicate of the form D_F, there exists a function F that satisfies (4.1). Indeed, if p satisfies these conditions, one can just take A to be the completion of $\mathbf{N}$ with respect to the distance $p(0)$, extend all predicates by uniform continuity, and define the functions via (4.1). The first three of these conditions are clearly G_δ; we check the fourth.

Suppose that F is a function symbol in L with modulus Δ_F. We claim that a predicate D_F satisfying the modulus (4.2) comes from a function iff it satisfies the conditions

$$d(y_1, y_2) \leq D_F(\bar{x}, y_1) + D_F(\bar{x}, y_2) \tag{4.3}$$

and

$$\forall \bar{x}\ \forall \epsilon > 0\ \exists y \quad D_F(\bar{x}, y) < \epsilon. \tag{4.4}$$

The first one is a closed condition and says that F is a function; the second ensures that F is total. As D_F is uniformly continuous, the quantifiers $\forall \bar{x}$ and $\exists y$ in (4.4) can be taken to range over the distinguished dense subset, so the condition (4.4) is G_δ. Then any predicate that respects Δ_{D_F} and satisfies (4.3) and (4.4) is of the form D_F for some function F that respects Δ_F. To see this, fix $\bar{x}$ and take a sequence (y_n) as given by (4.4) for $\epsilon = 2^{-n}$; then by (4.3), the sequence (y_n) is Cauchy and its limit y satisfies $D_F(\bar{x}, y) = 0$. Define $F(\bar{x})$ to be the (unique by (4.3)) y such that $D_F(\bar{x}, y) = 0$. One then easily checks that F respects Δ_F and that $D_F(\bar{x}, y) = d(F(\bar{x}), y)$. □

We will consider an element $A \in \mathcal{M}$ as a complete structure with a distinguished tail-dense set $\mathbf{N} \subseteq A$; thus we will write $P_i^A(a_0, \dots, a_{n_i-1})$ instead of $A(i)(a_0, \dots, a_{n_i-1})$ for $a_0, \dots, a_{n_i-1} \in \mathbf{N}$. In this way, we also interpret $r_\alpha(A\bar{a}, B\bar{b})$, where $A, B \in \mathcal{M}$ and $\bar{a}, \bar{b} \in \mathbf{N}^n$. Thus $r_{\alpha,n}$ is a pseudo-distance on $\mathcal{M} \times \mathbf{N}^n$. Since $\mathbf{N}$ is dense and $r_\alpha^{A,B}$ is continuous in each variable, we also have

$$r_{\alpha+1}^{A,B}(\bar{a}, \bar{b}) = \sup_{c,d \in \mathbf{N}} \inf_{c',d' \in \mathbf{N}} r_\alpha^{A,B}(\bar{a}c, \bar{b}d') \vee r_\alpha^{A,B}(\bar{a}c', \bar{b}d), \tag{4.5}$$

that is, it is enough to take suprema and infima over the distinguished dense sets.

Proposition 4.3. *The following statements hold:*

(i) *For every formula $\phi(\bar{x})$, the function $\mathcal{M} \times \mathbf{N}^n \to \mathbf{R}$, $A\bar{a} \mapsto \phi^A(\bar{a})$ is Borel.*
(ii) *For every $n \in \mathbf{N}$ and $\alpha < \omega_1$, the function $r_{\alpha,n}\colon (\mathcal{M} \times \mathbf{N}^n)^2 \to \mathbf{R}$ is Borel.*
(iii) *If Ω is shift-increasing, then for every $s \in \mathbf{R}$, the set*

$$\{(A\bar{a}, B\bar{b}) \in (\mathcal{M} \times \mathbf{N}^n)^2 : r_\infty(A\bar{a}, B\bar{b}) < s\}$$

is analytic.

Proof. (i) By Lemma 4.1, it is enough to prove the claim for L'-formulas, which is done by induction. Evaluation of L'-atomic formulas is a continuous function. For the quantifier step, note that it suffices to quantify over the distinguished dense set.

(ii) For $\alpha = 0$, recall that $r_0(A\bar{a}, B\bar{b}) \leq s$ iff for all basic ϕ that obey Ω, $|\phi^A(\bar{a}) - \phi^B(\bar{b})| \leq s$. By (i), all of those are Borel conditions and by Lemma 2.4, it suffices to consider only countably many of them. The rest of the proof is a straightforward induction on α using (4.5).

(iii) This follows from Proposition 3.4, which gives an analytic description of the condition $r_\infty(A\bar{a}, B\bar{b}) < s$. □

We obtain the following corollary from Theorem 3.8.

Corollary 4.4. *Let L be a countable signature and Ω be a weak modulus. Then for every separable L-structure A, the set*

$$\{B \in \mathcal{M}(L) : r_\infty^\Omega(A, B) = 0\}$$

is Borel.

5. The universal weak modulus

In this section, we show that for every countable language L, there exists a *universal weak modulus* Ω such that every $\mathcal{L}_{\omega_1\omega}(L)$-sentence is equivalent to an Ω-sentence (Corollary 6.5) and for which the equivalence relation E_∞ is the finest one possible, that of isomorphism (Theorem 5.5).

What we will need from the universal modulus is that Ω-formulas be sufficiently expressive; see Proposition 5.2 below. The definition is chosen in such a way to make this work.

Definition 5.1. Let L be a countable signature. We say that a weak modulus Ω is *universal for L* if it satisfies the following conditions:

(i) For every atomic formula $\phi(x_0, \dots, x_{k-1})$, there exists n such that

$$\Delta_\phi(\delta_0, \dots, \delta_{k-1}) \leq \Omega|_n(0, \dots, 0, \delta_0, \dots, \delta_{k-1});$$

(ii) For every $k \in \mathbf{N}$ and every $M > 0$, there exists n such that

$$M \cdot \Omega|_k(\delta_0, \dots, \delta_{k-1}) \leq \Omega|_n(0, \dots, 0, \delta_0, \dots, \delta_{k-1});$$

(iii) For every $k, n \in \mathbf{N}$,

$$\Omega|_k(\delta_0, \dots, \delta_{k-1}) + \Omega|_n(\gamma_0, \dots, \gamma_{n-1}) \leq \Omega|_{k+n}(\delta_0, \dots, \delta_{k-1}, \gamma_0, \dots, \gamma_{n-1});$$

(iv) Ω is shift-increasing.

Proposition 5.2. *Let Ω be a universal weak modulus for L. Then the following hold:*

(i) *For every k-ary atomic formula $\phi(\bar{x})$, there exists n such that $\phi(x_n, \dots, x_{n+k-1})$ is an Ω-formula;*
(ii) *For every k-ary Ω-formula $\phi(\bar{x})$ and every $M > 0$, there exists n such that $M \cdot \phi(x_n, \dots, x_{n+k-1})$ is an Ω-formula;*
(iii) *For all tuples*

$$i_0 < \cdots < i_{n-1} < j_0 < \cdots < j_{n-1},$$

$d^\Omega\big((x_{i_0}, \dots, x_{i_{n-1}}), (x_{j_0}, \dots, x_{j_{n-1}})\big)$ is an Ω-formula.

Proof. The items (i) and (ii) follow from the corresponding ones in Definition 5.1. We check (iii). Let A be an L-structure and $\bar{a}, \bar{b}, \bar{c}, \bar{d} \in A^n$. Let $\delta_i = d(a_i, c_i)$ and $\gamma_i = d(b_i, d_i)$. We have:

$$\begin{aligned} |d^\Omega(\bar{a}, \bar{b}) - d^\Omega(\bar{c}, \bar{d})| &\leq d^\Omega(\bar{a}, \bar{c}) + d^\Omega(\bar{b}, \bar{d}) \\ &= \Omega(\delta_0, \dots, \delta_{n-1}) + \Omega(\gamma_0, \dots, \gamma_{n-1}) \\ &\leq \Omega(\delta_0, \dots, \delta_{n-1}, \gamma_0, \dots, \gamma_{n-1}) \\ &\leq \Omega(\dots, \delta_0, \dots, \delta_{n-1}, \dots, \gamma_0, \dots, \gamma_{n-1}, \dots), \end{aligned}$$

where in the last line, $\delta_0, \dots, \delta_{n-1}, \gamma_0, \dots, \gamma_{n-1}$ are in positions $i_0, \dots, i_{n-1}, j_0, \dots, j_{n-1}$, respectively, and the other positions are filled with zeros. The second inequality is condition (iii) in Definition 5.1 and the last one follows from the fact that Ω is shift-increasing. This completes the proof of (iii). □

Proposition 5.3. *For every signature L, a universal modulus $\Omega_{\mathrm{U}}(L)$ for L exists. If L is a* Lipschitz language *(that is, all moduli of continuity for symbols in L are linear functions), then we can take $\Omega = \Omega_{\mathrm{U}}(\mathrm{Lip})$ as defined by (2.4).*

Proof. Let $\{\phi_i\}_{i\in\mathbf{N}}$ be an enumeration of all atomic formulas in L. Let

$$\Omega_{\mathrm{U}}(L)(\delta_0,\delta_1,\ldots) = \sum_{i=0}^{\infty} i \cdot \sup_{k\leq i} \Delta_{\phi_k}(\delta_i,\ldots,\delta_i).$$

One easily checks that all conditions in Definition 5.1 are satisfied. Similarly for $\Omega_{\mathrm{U}}(\mathrm{Lip})$ and Lipschitz languages. □

Remark 5.4. Our definition of a universal weak modulus is somewhat arbitrary: we have put together all conditions that we need (here, as well as in Sections 6 and 7) and what is important for us is the fact that such a universal modulus exists. For other purposes, one might need additional properties.

Theorem 5.5. *Let L be a countable signature and let Ω be a universal weak modulus for L. Let A and B be separable L-structures and $\bar{v} \in A^k$, $\bar{w} \in B^k$. Then*

$$r_\infty^\Omega(A\bar{v}, B\bar{w}) = \inf\{d^\Omega(f(\bar{v}),\bar{w}) : f \text{ is an isomorphism } A \to B\}.$$

In particular,

$$r_\infty^\Omega(A,B) = \begin{cases} 0 & \text{if } A \cong B, \\ \infty & \text{otherwise.} \end{cases}$$

Proof. The inequality $\leq$ is clear: if $f\colon A \to B$ is an isomorphism, then it follows from Lemma 3.2 that for all α,

$$r_\alpha(A\bar{v}, B\bar{w}) \leq d^\Omega(f(\bar{v}),\bar{w}).$$

For the inequality $\geq$, suppose that $r_\infty(A\bar{v}, B\bar{w}) < t$. By Proposition 3.4, there exist tail-dense sequences $a \in A^{\mathbf{N}}$ and $b \in B^{\mathbf{N}}$ such that $a|_k = \bar{v}$, $b|_k = \bar{w}$ and for all n, $r_0^{A,B}(a|_n, b|_n) \leq t$.

Consider an n-ary atomic formula $\phi(\bar{y})$. By Proposition 5.2 (i), (ii) and the fact that Ω is shift-increasing, for every $M > 0$, there exists $l(M,\phi)$ such that for $\sigma \in \mathbf{N}^n$ with $\min\sigma > l$,

$$\phi_\sigma = M \cdot \phi(x_{\sigma_0},\ldots,x_{\sigma_{n-1}})$$

is a basic Ω-formula. It follows that

$$\left|\phi_\sigma(a_{\sigma_0},\ldots,a_{\sigma_{n-1}}) - \phi_\sigma(b_{\sigma_0},\ldots,b_{\sigma_{n-1}})\right| \leq r_0(a|_{\max\sigma}, b|_{\max\sigma}) < t, \tag{5.1}$$

i.e., for all σ with $\min\sigma > l(M,\phi)$,

$$\left|\phi(a_{\sigma_0},\ldots,a_{\sigma_{n-1}}) - \phi(b_{\sigma_0},\ldots,b_{\sigma_{n-1}})\right| < t/M,$$

which implies that

$$\left|\phi(a_{\sigma_0},\ldots,a_{\sigma_{n-1}}) - \phi(b_{\sigma_0},\ldots,b_{\sigma_{n-1}})\right| \to 0 \text{ as } \min\sigma \to \infty. \tag{5.2}$$

Applied to the formula $d(y_0,y_1)$, this means that for a strictly increasing sequence $(m_n) \subseteq \mathbf{N}$, if one of the subsequences (a_{m_n}) and (b_{m_n}) is Cauchy, then so is the other. If $\lim a_{m_n} = \lim a_{m'_n}$, then $(a_{m_n}, a_{m'_n})$ is Cauchy, therefore so is $(b_{m_n}, b_{m'_n})$, and thus $\lim b_{m_n} = \lim b_{m'_n}$ (and vice versa). We therefore obtain a bijection $\theta\colon A \to B$ defined by $\theta(\lim a_{m_n}) = \lim b_{m_n}$. By (5.2) again, θ is an isomorphism. Next we check that $d^\Omega(\theta(a|_k), b|_k) \leq t$. For $i < k$, let $(m_{i,n})_n$ be strictly increasing sequences such that $a_{m_{i,n}} \to a_i$ as $n \to \infty$ and

$$k < m_{0,n} < \cdots < m_{k-1,n} \quad \text{for each } n.$$

Then Proposition 5.2 (iii) and (5.1) give us that

$$\left|d^{\Omega}(a|_k,(a_{m_{0,n}},\dots,a_{m_{k-1,n}}))-d^{\Omega}(b|_k,(b_{m_{0,n}},\dots,b_{m_{k-1,n}}))\right|<t$$

for all n and taking limits yields $d^{\Omega}(b|_k,\theta(a|_k))\leq t$ as desired. □

Applying Corollary 4.4 allows us to recover one of the original applications of Scott sentences.

Corollary 5.6. *Let L be a countable signature and A be a separable L-structure. Then the set $\{B\in\mathcal{M}(L):B\cong A\}$ is Borel.*

Remark 5.7. We should note that Corollary 5.6 also follows from the main result of [EFP+], where the authors prove that isomorphism of Polish metric structures is Borel reducible to a Polish group action (which implies that classes are Borel). However, our result is more precise as it gives a bound of the Borel complexity of the isomorphism class in terms of the Scott rank.

Remark 5.8. Note that in Theorem 5.5, if $r_\infty(A\bar{a},B\bar{b})=0$, this does *not* imply that there is an isomorphism between A and B that sends $\bar{a}$ to $\bar{b}$. An easy example of this is when $A=B$ and the orbits of the action of $\mathrm{Aut}(A)$ on A are not closed. Then, for any a and b in the same orbit closure of $\mathrm{Aut}(A)$, $r_\infty(Aa,Ab)=0$. However, if a and b are not in the same orbit, there is no automorphism that maps a to b. Such a structure can be obtained as follows: if G is a Polish group, its completion under a left-invariant metric G_L can be made into a metric structure in such a way that $\mathrm{Aut}(G_L)=G$ with the action of G on G_L by left translation. If $G\neq G_L$ (this holds, e.g., for $G=S_\infty$), this provides an example of an automorphism group with orbits that are not closed. See Melleray [M] for more details.

Remark 5.9. It is possible to take a slightly different approach that may be more suitable for some purposes when defining the back-and-forth pseudo-distances. In the definition of r_0, one could replace basic formulas by full $\mathcal{L}_{\omega\omega}$-formulas that obey Ω. This has the advantage of being much more robust with respect to syntactical considerations: one can assume from the start, without loss of generality, that the signature contains no function symbols and that all predicates are Lipschitz (see [BY2, Corollary 1.7]), thus allowing a unique universal weak modulus that works for all languages (namely, $\Omega_{\mathrm{U}}(\mathrm{Lip})$). The main disadvantage of this approach is that r_0 becomes very hard to compute, while with our definition, computations are sometimes feasible (see Section 8).

6. A López-Escobar theorem

Next we prove a continuous analogue of the classical López-Escobar theorem: that every Borel set of structures invariant under isomorphism is the set of models of some $\mathcal{L}_{\omega_1\omega}$-sentence. This is a converse to Proposition 4.3 (i).

Let L be a fixed countable language, and, as before, denote by $\mathcal{M}$ the space of Polish metric L-structures. Let Ω be a universal weak modulus for L. Let $\cong$ be the (analytic) equivalence relation of isomorphism on $\mathcal{M}$ (by Theorem 5.5, this is the same as E_∞).

We proceed to the main theorem of this section. We use the definition of Baire class for real valued, Borel functions given in [K, 24.1]. If θ is the least ordinal such that the function U is of Baire class θ, we write $\mathrm{BC}(U)=\theta$.

Theorem 6.1. *Let $I\subseteq\mathbf{R}$ be a compact interval and $U\colon\mathcal{M}\to I$ be a Borel function that is $\cong$-invariant. Then there exists an (Ω,I)-sentence ϕ such that*

$$U(A)=\phi^A\quad\text{for all } A\in\mathcal{M}.$$

Moreover, we have that $\mathrm{qr}(\phi) \leq \omega \cdot 2 \cdot (2 + \mathrm{BC}(U) + 1)$.

There are two known proofs of the López-Escobar theorem in the classical case: the original one in [LE], based on proof theory, and another one by Vaught [V], based on Baire category and the fact that isomorphism is given by an action of S_∞. It is the latter that we adapt to our situation. In our setting, we do not have a group action around but we do have Baire category and it will turn out that this is sufficient.

If $A \in \mathcal{M}$ is a model, we denote by $[A]$ the $\cong$-equivalence class of A in $\mathcal{M}$ (by Corollary 5.6, this is a Borel set). We also define

$$D(A) = \big\{y \in A^{\mathbf{N}} : \{y(n) : n \in \mathbf{N}\} \text{ is tail-dense in } A\big\}.$$

Here and later, by a slight abuse of notation, we denote by A both the element of $\mathcal{M}$ (which prescribes the values of all predicates on a dense set) and the actual model (the completion of $\mathbf{N}$ with respect to the metric). Note that $D(A)$ is a G_δ set in $A^{\mathbf{N}}$, and therefore a Polish space.

For each $A \in \mathcal{M}$, there is a natural continuous surjection $\pi_A \colon D(A) \to [A]$ defined by

$$P^{\pi_A(y)}(i_0, \dots, i_{n-1}) = P^A(y(i_0), \dots, y(i_{n-1}))$$

for all predicates $P \in L$ of arity n and all $i_0, \dots, i_{n-1} \in \mathbf{N}$. Because of the way we code models, we may assume, without loss of generality, that the language L does not contain function symbols. The map π_A will allow us to push forward the ideal of meager sets on $D(A)$ to $[A]$, which is an essential element of the proof.

If X is a Baire topological space, we will use the category quantifiers "$\exists^*$" to mean "for a non-meagre set of" and "$\forall^*$" to mean "for a comeagre set of". If $f \colon X \to \mathbf{R}$ is a Baire measurable function, define the operators $\sup^*$ (essential supremum) and $\inf^*$ (essential infimum) as follows:

$$\sup_{x \in X}{}^* f(x) > t \iff \exists^* x \in X \; f(x) > t$$
$$\inf_{x \in X}{}^* f(x) < t \iff \exists^* x \in X \; f(x) < t.$$

Note that if f is continuous, then

$$\sup_{x \in X}{}^* f(x) = \sup_{x \in X} f(x) \quad \text{and} \quad \inf_{x \in X}{}^* f(x) = \inf_{x \in X} f(x).$$

If y and z are finite or infinite sequences of elements of a metric space (Y, d) at least one of which is finite, we will abuse notation and write $d^\Omega(y, z)$ instead of $d^\Omega(y|_m, z|_m)$, where $m = \min(|y|, |z|)$. If y or z is the empty sequence, we set $d^\Omega(y, z) = 0$.

We will call a function $F \colon \mathcal{M} \to \mathbf{R}$ a *basic continuous function* if there exists $k \in \mathbf{N}$, predicates P_i of arity n_i for $i < k$, elements $a_{i,j} \in \mathbf{N}$ $(i < k, j < n_i)$, and a Lipschitz (for the max distance on $\mathbf{R}^k$) function $f \colon \mathbf{R}^k \to \mathbf{R}$ such that

$$F(A) = f\big((P_i^A(a_{i,0}, \dots, a_{i,n_i-1}))_{i<k}\big).$$

Lemma 6.2. *Let $I \subseteq \mathbf{R}$ be a compact interval. The class of Borel functions $\mathcal{M} \to I$ is the smallest class that contains the basic continuous functions taking values in I and is closed under countable suprema and infima.*

Proof. First recall that, by [K, 11.6], the class of Borel functions on a Polish space is the smallest class containing all continuous functions and closed under pointwise limits. If $\lim_n f_n$ exists then $\lim_n f_n = \inf_n \sup_{m \geq n} f_m$, so it is also the smallest class containing all continuous functions and closed under countable suprema and infima.

Thus it suffices to prove that all continuous functions can be obtained from the basic ones using pointwise limits. As $\mathcal{M} \subseteq \prod_i I_{P_i}^{\mathbf{N}^{n_{P_i}}}$ is a G_δ set (Proposition 4.2), any continuous function on $\mathcal{M}$ extends to a Baire class 2 function on the compact space $Z = \prod_i I_{P_i}^{\mathbf{N}^{n_{P_i}}}$. Any continuous function on Z is uniformly continuous and can be approximated by a function that only depends on finitely many coordinates, that is, a function of the type $A \mapsto f\big((P_i^A(a_{i,0}, \dots, a_{i,n_i-1}))_{i<k}\big)$ with $f\colon K \to \mathbf{R}$ continuous, where $K = \prod_i I_{P_i}$. By Stone–Weierstrass, f can be approximated by Lipschitz functions, and such functions are permitted in our definition of basic continuous functions. By taking limits (two times), we can obtain any Baire class 2 function on Z from continuous functions. □

If $M > 0$, $U\colon \mathcal{M} \to [0, M]$ is a Borel function, and $k \in \mathbf{N}$, define the function $U^{*k}\colon \mathcal{M} \times \mathbf{N}^k \to [0, M]$ as follows:

$$U^{*k}(A, \bar{u}) = \inf_{y \in D(A)}{}^{*}\, U(\pi_A(y)) \vee k d^{\Omega}(y, \bar{u}).$$

The following result easily implies Theorem 6.1 and is better suited for an inductive argument.

Theorem 6.3. *Let $M > 0$. For every Borel function $U\colon \mathcal{M} \to [0, M]$ and for every $k \in \mathbf{N}$, there exists $l \in \mathbf{N}$ and an $(\Omega, [0, M])$-formula $\phi_{U,k}(x_l, \dots, x_{l+k-1})$ such that*

$$U^{*k}(A, \bar{u}) = \phi_{U,k}^A(u_0, \dots, u_{k-1}) \tag{6.1}$$

for all $(A, \bar{u}) \in \mathcal{M} \times \mathbf{N}^k$. Moreover,

$$\mathrm{qr}(\phi_{U,k}) \leq \omega \cdot 2 \cdot (2 + \mathrm{BC}(U) + 1). \tag{6.2}$$

Proof. For the main statement of the theorem, by Lemma 6.2, it is enough to check that the class of functions U that satisfy the theorem contains the basic continuous functions and is closed under countable infima and the operation $U \mapsto M - U$. (This operation exchanges infima and suprema and preserves the interval $[0, M]$.) Thereafter we will check "moreover" statement by bounding the quantifier rank of the constructed formulas.

First suppose that U is a basic continuous function. Then, by the properties of the universal modulus (Proposition 5.2 (i), (ii) and the fact that Ω is shift-increasing), there exist $n \in \mathbf{N}$, variables $z_0, \dots, z_{n-1}$, and a basic Ω-formula $\theta(z_0, \dots, z_{n-1})$ such that $U(A) = \theta^A(0, \dots, n-1)$. As U is continuous, sup and sup* coincide, and by the properties of the universal modulus again, there exists $l \in \mathbf{N}$ such that

$$\phi_{U,k}(x_l, \dots, x_{l+k-1}) = \inf_{\bar{z}}\, \theta(\bar{z}) \vee k d^{\Omega}(\bar{x}, \bar{z}) \tag{6.3}$$

is an Ω-formula and satisfies (6.1). (Note that $\phi_{U,k}$ is technically not an $(\Omega, [0, M])$-formula as the subformula $k d^{\Omega}(\bar{x}, \bar{z})$ does not respect the bound $[0, M]$. This can be easily fixed by replacing it by $(k d^{\Omega}(\bar{x}, \bar{z})) \wedge M$ but we will not do this in order to avoid cluttering the exposition. This change is completely harmless as it transforms $\phi_{U,k}$ into an equivalent formula. A similar remark also applies to the constructions below.)

Suppose now that $U = \inf_n U_n$ and that the formulas $\phi_{U_n,k}$ have already been constructed. We verify that

$$\phi_{U,k}(\bar{x}) = \bigwedge_n \phi_{U_n,k}(\bar{x})$$

works. Using the inductive hypothesis and the fact that $\inf^*$ commutes with taking countable infima, we obtain

$$\begin{aligned} U^{*k}(A,\bar{u}) &= \inf_{y\in D(A)}{}^{*} \inf_n U_n(\pi_A(y)) \vee kd^{\Omega}(y,\bar{u}) \\ &= \inf_n \inf_{y\in D(A)}{}^{*} U_n(\pi_A(y)) \vee kd^{\Omega}(y,\bar{u}) \\ &= \inf_n \phi^A_{U_n,k}(\bar{u}). \end{aligned}$$

Finally, suppose that $U = M - V$, and that formulas $\phi_{V,m}$ satisfying

$$\phi^A_{V,m}(\bar{z}) = \inf_{v\in D(A)}{}^{*} V(v) \vee md^{\Omega}(v,\bar{z}) \tag{6.4}$$

have already been constructed for every m.

We show that

$$\phi_{U,k}(\bar{x}) = \bigwedge_m \inf_{z_0,\dots,z_{m-1}} (M - \phi_{V,m}(\bar{z})) \vee kd^{\Omega}(\bar{x},\bar{z})$$

satisfies (6.1). Note also that the variables can be chosen in a such a way that $\phi_{U,k}$ is an Ω-formula.

Fix A, $\bar{u}$, and $r \in \mathbf{R}$ in order to show that

$$U^{*k}(A,\bar{u}) < r \iff \phi^A_{U,k}(\bar{u}) < r. \tag{6.5}$$

If $\bar{z} \in A^m$ and $r > 0$, let $\mathcal{B}(\bar{z},r)$ be the open set

$$\mathcal{B}(\bar{z},r) = \{y \in D(A) : d^{\Omega}(\bar{z},y) < r\}.$$

Suppressing π_A from the notation, we have that

$$\begin{aligned} U^{*k}(A,\bar{u}) < r &\iff \exists^* y \in D(A)\ \Big(kd^{\Omega}(\bar{u},y) < r \text{ and } (M - V(y)) < r\Big) \\ &\iff \exists \text{ open } W \subseteq D(A)\ \Big(W \subseteq \mathcal{B}(\bar{u},r/k) \text{ and } \forall^* y \in W\ V(y) > M - r\Big). \end{aligned}$$

On the other hand,

$$\begin{aligned} \phi^A_{U,k}(\bar{u}) < r &\iff \exists m,\bar{z}\ \Big(\big((M - \inf_{v\in D(A)}{}^{*} V(v)) \vee md^{\Omega}(\bar{z},v)\big) < r \\ &\qquad \text{and } kd^{\Omega}(\bar{u},\bar{z}) < r\Big) \\ &\iff \exists m,\bar{z}\ \Big(\big(\forall^* v \in D(A)\ V(v) > M - r \text{ or } md^{\Omega}(v,\bar{z}) > M - r\big)\Big) \\ &\qquad \text{and } kd^{\Omega}(\bar{u},\bar{z}) < r\Big) \\ &\iff \exists m,\bar{z}\ \Big(\big(\forall^* v \in \mathcal{B}(\bar{z},(M-r)/m)\ V(v) > M - r\big) \\ &\qquad \text{and } kd^{\Omega}(\bar{u},\bar{z}) < r\Big). \end{aligned}$$

For the direction ($\Leftarrow$) of (6.5), suppose that z and m are given. By enlarging m if necessary (and prolonging $\bar{z}$ arbitrarily), we may assume that $(M-r)/m < r/k - d^{\Omega}(\bar{u},\bar{z})$. Then it suffices to take $W = \mathcal{B}(\bar{z},(M-r)/m)$ to witness that $U^{*k}(A,\bar{u}) < r$.

For the other direction, suppose that W is given. Let $y \in W$ be arbitrary and take m so big that $\mathcal{B}(y|_m,(M-r)/m) \subseteq W$. Finally, set $z = y|_m$.

This completes the induction.

Now we give bounds on the quantifier ranks. Let F be the function defined by $F(\theta) = \omega \cdot 2 \cdot (2 + \theta + 1)$ for each ordinal θ. Note that if $U = \lim_n U_n$ for

Borel functions U_n, then we may assume that the range of each U_n is contained in $[0, M]$ without increasing the Baire class, and we have

$$U = \limsup_n U_n = \inf_m (M - \inf_{n \geq m} (M - U_n)). \tag{6.6}$$

If U is basic continuous, for $\phi_{U,k}$ defined in (6.3), we have $\mathrm{qr}(\phi_{U,k}) < \omega$.

For the general case, we induct on $\mathrm{BC}(U)$. First suppose that $\mathrm{BC}(U) = 0$, that is, U is continuous. By the proof of Lemma 6.2, we need to take limits two times to obtain U from basic continuous functions. Moreover, each basic continuous function is represented by a formula of quantifier rank $< \omega$; thus (6.6) shows that $\mathrm{qr}\, \phi_{U,k} \leq \omega + \omega \cdot 2 + \omega \cdot 2 = \omega \cdot 5$, whence (6.2) holds for U.

Now suppose that $\mathrm{BC}(U) = \theta > 0$. Then $U = \lim_n U_n$ where $\mathrm{BC}(U_n) < \theta$ for each n. By the inductive hypothesis, $\mathrm{qr}(\phi_{U_n,k}) \leq F(\mathrm{BC}(U_n))$, so there is $\rho < F(\theta)$ such that for each n, k we have $\mathrm{qr}(\phi_{M-U_n,k}) \leq \rho$. Therefore $\mathrm{qr}(\psi_{m,k}) \leq \rho$ where $\psi_{m,k} = \phi_{\inf_{n \geq m} M - U_n, k}$. A similar argument now shows that $\mathrm{qr}(\phi_{U,k}) \leq \rho + \omega \leq F(\theta)$. □

Proof of Theorem 6.1. Suppose $U \colon \mathcal{M} \to I$ is a Borel function invariant under isomorphism. Let $M = |I|$ and let $U' = U - \min I$, so that U' takes values in $[0, M]$. In Theorem 6.3, take $\bar{u} = \emptyset$ and observe that $U'^{*0}(A, \emptyset) = U'(A)$ for every A. Thus, $\phi = \phi_{U',0} + M$ is (equivalent to) an (Ω, I)-sentence such that $U(A) = \phi^A$ for every A. Moreover, such a sentence has quantifier rank at most $\omega \cdot 2 \cdot (2 + \mathrm{BC}(U) + 1)$. □

A standard corollary of the López-Escobar theorem is the Craig interpolation theorem for $\mathcal{L}_{\omega_1\omega}$-logic. (In fact, López-Escobar first proved the interpolation result and then deduced his theorem from it.) Here we note the continuous version.

Corollary 6.4 (Interpolation)**.** *Suppose that L_1 and L_2 are two countable signatures and ϕ_1 and ϕ_2 are $\mathcal{L}_{\omega_1\omega}(L_1)$ and $\mathcal{L}_{\omega_1\omega}(L_2)$ sentences respectively. Suppose that $\phi_1^A \leq \phi_2^A$ for every separable model A of $\mathcal{L}_{\omega_1\omega}(L_1 \cup L_2)$. Then there is an interpolating sentence θ in $\mathcal{L}_{\omega_1\omega}(L_1 \cap L_2)$ such that $\models \phi_1 \leq \theta \leq \phi_2$.*

Proof. Let $L_0 = L_1 \cap L_2$ and let $\pi_1 \colon \mathcal{M}(L_1) \to \mathcal{M}(L_0)$ and $\pi_2 \colon \mathcal{M}(L_2) \to \mathcal{M}(L_0)$ be the natural restriction maps. For every $r \in \mathbf{Q}$, consider the two analytic sets

$$\{A \in \mathcal{M}(L_0) : \exists B \in \mathcal{M}(L_1)\ \pi_1(B) = A \text{ and } \phi_1^B > r\} \text{ and}$$
$$\{A \in \mathcal{M}(L_0) : \exists B \in \mathcal{M}(L_2)\ \pi_2(B) = A \text{ and } \phi_2^B < r\}.$$

They are $\cong$-invariant and by hypothesis, they are disjoint. By [K, Exercise 14.14], there exists an invariant Borel set C_r that separates them. Define $U \colon \mathcal{M}(L_0) \to \mathbf{R}$ by

$$U(A) = \sup\{r : A \in C_r\}.$$

The function U is $\cong$-invariant and Borel. So by Theorem 6.1, there exists a sentence θ such that $U(A) = \theta^A$ and then for all separable A, $\phi_1^A \leq \theta^A \leq \phi_2^A$. Using the downwards Löwenheim–Skolem theorem, this implies that $\phi_1 \leq \theta \leq \phi_2$ is universally valid. □

Another corollary is that our universal modulus Ω is indeed universal.

Corollary 6.5. *Let L be a countable language and Ω be a universal modulus for L. Then every $\mathcal{L}_{\omega_1\omega}(\mathcal{L})$-sentence ϕ is equivalent to an Ω-sentence ϕ'.*

Proof. Apply Theorem 6.1 to the Borel function $U(A) = \phi^A$. □

7. Bounded rank and Borelness of isomorphism

In this section, we characterize when the isomorphism relation restricted to an invariant Borel subset of $\mathcal{M}$ is Borel and prove Theorem 1.3. This is again analogous to the classical setting (see, e.g., [G2, Theorem 12.2.4]). Our proof follows the same general outline, but certain new features appear.

For the remainder of the section, we fix a countable language L and a weak modulus Ω universal for L. Recall that a *fragment* F of $\mathcal{L}_{\omega_1\omega}(L)$ is a *separable* set of formulas containing all atomic formulas and closed under subformulas, substitution of terms for variables, quantifiers, and finitary connectives. An *F-theory* is a collection of statements of the form $\phi = 0$, where ϕ is a sentence in F. If A is a structure, the *theory of A*, denoted by $\mathrm{Th}(A)$, is the collection of all statements $\phi = 0$ that are true in A.

We will use a homogeneity result for atomic models. In order to state it correctly, we will need to define a topometric structure (see [BY1]) on the (non-compact) type spaces, which we proceed to do. Let T be an F-theory. Define a seminorm $\|\cdot\|_T$ on the (unital) algebra of n-ary F-formulas by:

$$\|\phi\|_T = \sup\{|\phi^A(\bar{a})| : A \models T, \bar{a} \in A^n\}.$$

Denote by $\widehat{\mathrm{S}_n(T)}$ the Gelfand space of the Hausdorff completion of the algebra of all n-ary F-formulas equipped with this seminorm. (See, for example, [F, Chapter 1] for details on Gelfand theory.) This is the compact space of *finitely consistent n-types*. In what follows, we will identify the algebra of formulas with $C(\widehat{\mathrm{S}_n(T)})$ via the Gelfand transform and simply use the notation $\phi(p)$ for the value that the type p gives to the formula ϕ. Alternatively, if $\phi(p) = r$, we will also write $p(\bar{x}) \models \phi(\bar{x}) = r$.

Say that a tuple $\bar{a} \in A^n$ *realizes* a type p (notation $\bar{a} \models p$ or $\mathrm{tp}(\bar{a}) = p$) if $\phi(p) = \phi^A(\bar{a})$ for all n-ary formulas $\phi \in F$. If there exists $\bar{a} \in A^n$ such that $\bar{a} \models p$, we will say that A *realizes* p. Otherwise, say that A *omits* p. Denote by $\mathrm{S}_n(T)$ the set of *realizable* types (or just *types* for short):

$$\mathrm{S}_n(T) = \{p \in \widehat{\mathrm{S}_n(T)} : \exists A \models T \text{ such that } A \text{ realizes } p\}.$$

The *logic topology* on $\mathrm{S}_n(T)$ is the one inherited from $\widehat{\mathrm{S}_n(T)}$. The compactness theorem tells us that if $F = \mathcal{L}_{\omega\omega}(L)$, then $\mathrm{S}_n(T) = \widehat{\mathrm{S}_n(T)}$. While this fails for general fragments, $\mathrm{S}_n(T)$ is always a G_δ subset of $\widehat{\mathrm{S}_n(T)}$ and therefore a Polish space. (We will not prove this fact as we will not use it.)

Defining the distance on types is more delicate because of the lack of compactness. The definition we give is inspired by Caicedo and Iovino [CI]. As in finitary continuous logic, the topology defined by the distance on types is finer than the logic topology and the distance is lower semicontinuous in the logic topology.

We will need to fix distances on powers of A and for most purposes, any distance will do. However, in order to obtain exact equalities in the two propositions below, it will be most convenient to take $d = d^\Omega$; in the remainder of the section, when we write $d(\bar{a}, \bar{b})$ for $\bar{a}, \bar{b} \in A^n$, we mean $d^\Omega(\bar{a}, \bar{b})$ (and similarly, in formulas).

Recall that the operation $\dot{-}$ is defined by $x \dot{-} y = 0 \vee (x - y)$. Let the distance ∂_F on $\mathrm{S}_n(T)$ be given by:

$$\partial_F(p, q) \leq s \iff \forall \phi \in F \; q(\bar{x}) \models \inf_{\bar{y}} (d(\bar{x}, \bar{y}) \dot{-} s) \vee |\phi(\bar{y}) - \phi(p)| = 0. \tag{7.1}$$

In words, $\partial_F(p, q) \leq s$ iff for every $\phi \in F$ and every $\epsilon > 0$, for every realization $\bar{a} \in A^n$ of q, there exists $\bar{a}' \in A^n$ such that $d(\bar{a}, \bar{a}') < s + \epsilon$ and $|\phi(\bar{a}') - \phi(p)| < \epsilon$. For $F = \mathcal{L}_{\omega\omega}$, using compactness, this definition is equivalent to the usual one

(see [BBHU, Section 8] for the definition). When the fragment F is clear from the context, we will simply write ∂ instead of ∂_F.

We check that ∂ is symmetric. Suppose that $\partial(p,q) \leq s$ and fix $\phi \in F$, and a realization $\bar{a} \in A^n$ of q. Let $\psi(\bar{x}) = \inf_{\bar{y}} (d(\bar{y},\bar{x}) \dot{-} s) \vee |\phi(\bar{y}) - \phi(q)|$ and suppose, for contradiction, that $\psi(p) = r > 0$. Using (7.1) for ψ, we obtain that for every $\epsilon > 0$, there exists $\bar{a}' \in A^n$ with $d(\bar{a},\bar{a}') < s + \epsilon$ and $\psi(\bar{a}') > r - \epsilon$. However, $\psi(\bar{a}') \leq d(\bar{a},\bar{a}') \dot{-} s$ (as $\phi(\bar{a}) = \phi(q)$), which yields a contradiction for $\epsilon < r/2$.

Next we verify the triangle inequality. Suppose that $\partial(p_1,p_2) \leq s_1$ and $\partial(p_2,p_3) \leq s_2$ in order to show that $\partial(p_1,p_3) \leq s_1 + s_2$. For simplicity of notation, we assume that $n = 1$. Let $\phi \in F$ be a formula such that $\phi(p_1) = 0$. We know that

$$p_2(y) \models \inf_w (d(y,w) \dot{-} s_1) \vee \phi(w) = 0,$$

whence

$$p_3(x) \models \inf_y (d(x,y) \dot{-} s_2) \vee \big(\inf_w (d(y,w) \dot{-} s_1) \vee \phi(w)\big) = 0.$$

Simplifying yields

$$p_3(x) \models \inf_w (d(x,w) \dot{-} (s_1 + s_2)) \vee \phi(w) = 0,$$

as desired.

Another property of ∂ that we will need, easily checked from the definition, is that for any model $A \models T$ and all $\bar{a}, \bar{b} \in A^n$,

$$\partial(\operatorname{tp}\bar{a}, \operatorname{tp}\bar{b}) \leq d(\bar{a},\bar{b}). \tag{7.2}$$

Finally, we check that if ϕ is an n-ary Ω-formula that is in F, $A, B \models T$, and $\bar{a} \in A^n, \bar{b} \in B^n$, then

$$|\phi(\bar{a}) - \phi(\bar{b})| \leq \partial(\operatorname{tp}\bar{a}, \operatorname{tp}\bar{b}). \tag{7.3}$$

Indeed, suppose that $\partial(\operatorname{tp}\bar{a}, \operatorname{tp}\bar{b}) \leq s$ and fix a formula ϕ. Then for every $\epsilon > 0$, there exists $\bar{b}' \in B^n$ such that $d(\bar{b}',\bar{b}) < s + \epsilon$ and $|\phi(\bar{b}') - \phi(\bar{a})| < \epsilon$. By the choice of the metric d on products, ϕ is contractive in d, so we obtain that $|\phi(\bar{a}) - \phi(\bar{b})| \leq s + 2\epsilon$, which is enough.

If $X \subseteq \mathrm{S}_n(T)$ and $\delta > 0$, denote by $(X)_\delta$ the *δ-fattening* of X:

$$(X)_\delta = \{p \in \mathrm{S}_n(T) : \exists q \in X \; \partial(p,q) < \delta\}.$$

If $X = \{p\}$ is a singleton, write $(p)_\delta$ instead of $(\{p\})_\delta$. Say that a type $p \in \mathrm{S}_n(T)$ is *principal* if for every $\delta > 0$, p belongs to the interior of $(p)_\delta$ in the logic topology on $\mathrm{S}_n(T)$. A model A is called *F-atomic* if every type it realizes is principal (for $\mathrm{Th}(A)$). One of the important properties of atomic models is that they are homogeneous.

Proposition 7.1. *Let F be a fragment, let A be an F-atomic model, and $\bar{u}, \bar{v} \in A^k$. Then*

$$r_\infty(A\bar{u}, A\bar{v}) = \partial_F(\operatorname{tp}\bar{u}, \operatorname{tp}\bar{v}).$$

In particular, A is homogeneous, *i.e., for all $\bar{u}, \bar{v} \in A^k$ with $\operatorname{tp}\bar{u} = \operatorname{tp}\bar{v}$ and $\epsilon > 0$, there is an automorphism $f \in \mathrm{Aut}(A)$ such that $d(\bar{u}, f(\bar{v})) < \epsilon$.*

Proof. The inequality $\partial \leq r_\infty$ follows from Theorem 5.5 and (7.2). For the other direction, suppose that $\partial(\operatorname{tp}\bar{u}, \operatorname{tp}\bar{v}) < t$. We will build inductively tail-dense sequences $a, b \in A^{\mathbf{N}}$ that satisfy $a|_k = \bar{u}$, $b|_k = \bar{v}$, and

$$\partial\big(\operatorname{tp}(a|_n), \operatorname{tp}(b|_n)\big) < t \quad \text{for all } n. \tag{7.4}$$

We start by setting $a|_k = \bar{u}$ and $b|_k = \bar{v}$. The rest of the construction is carried by a back-and-forth argument of which we only describe the forth step. Suppose that $a|_n$, $b|_n$ are given such that $\partial(\operatorname{tp} a|_n, \operatorname{tp} b|_n) < s < t$ and let $a_n \in A$ be arbitrary. We

will find $b_n \in A$ such that $\partial(\operatorname{tp} a|_{n+1}, \operatorname{tp} b|_{n+1}) < t$. As $p = \operatorname{tp}(a|_{n+1})$ is principal, there exists a formula ϕ taking non-negative values such that $\phi(a|_{n+1}) = 0$ and $\phi(\bar{w}) < 1$ implies that $\partial(\operatorname{tp} \bar{w}, p) < (t-s)/2$. As $\partial\big(\operatorname{tp}(a|_n), \operatorname{tp}(b|_n)\big) < s$, by (7.1), we have that

$$A \models \inf_{\bar{y}} \big((d(\bar{y}, b|_n) \dot{-} s) \vee \inf_z \phi(\bar{y}, z)\big) = 0.$$

Let $\epsilon > 0$ be arbitrary and let $\bar{c} \in A^n$ and $d \in A$ be such that $d(b|_n, \bar{c}) < s + \epsilon$ and $\phi(\bar{c}d) < 1$. Then $\partial(\operatorname{tp}(\bar{c}d), p) < (t-s)/2$ and we have

$$\begin{aligned} \partial(p, \operatorname{tp}(b|_n d)) &\leq \partial(p, \operatorname{tp}(\bar{c}d)) + \partial(\operatorname{tp}(\bar{c}d), \operatorname{tp}(\bar{b}d)) \\ &\leq (t-s)/2 + s + \epsilon, \end{aligned}$$

which is less than t as long as $\epsilon < (t-s)/2$. This means that we can take $b_n = d$.

Now it only remains to observe that (7.3) implies that

$$r_0(a|_n, b|_n) \leq \partial\big(\operatorname{tp}(a|_n), \operatorname{tp}(b|_n)\big)$$

and apply Proposition 3.4. □

The following proposition bounds the Scott rank of F-atomic models.

Proposition 7.2. *Let F be a fragment such that the quantifier rank of formulas in F is bounded by α. Let A be an F-atomic model. Then for every $\bar{u}, \bar{v} \in A^k$,*

$$\partial_F(\operatorname{tp} \bar{u}, \operatorname{tp} \bar{v}) = r_\alpha(A\bar{u}, A\bar{v}) = r_\infty(A\bar{u}, A\bar{v}).$$

In particular, the Scott rank of A is at most α.

Proof. In view of Proposition 7.1, we only need to prove that $\partial(\operatorname{tp} \bar{u}, \operatorname{tp} \bar{v}) \leq r_\alpha(\bar{u}, \bar{v})$. Suppose that $r_\alpha(\bar{u}, \bar{v}) \leq s$ and let $\epsilon > 0$ be arbitrary. As $\operatorname{tp} \bar{v}$ is principal, there exists an n-ary Ω-formula $\psi \in F$ and $\delta > 0$ such that $\psi(\bar{v}) = 0$ and $\psi(q) < \delta \implies \partial(q, \operatorname{tp} \bar{v}) < \epsilon$. Let $M > s/\delta$ and let

$$\phi(\bar{x}) = \inf_{\bar{z}} d(\bar{x}, \bar{z}) \vee M \cdot \psi(\bar{z}),$$

where the variables $\bar{z}$ are taken in such a way that ϕ is an n-ary Ω-formula (this can be done by Proposition 5.2). Then $\phi(\bar{v}) = 0$ and as $\phi \in F$, by Theorem 3.5, we have that $\phi(\bar{u}) \leq s$. Then there exists $\bar{w} \in A^k$ such that $d(\bar{u}, \bar{w}) \leq s + \epsilon$ and $\psi(\bar{w}) < \delta$. Thus

$$\begin{aligned} \partial(\operatorname{tp} \bar{u}, \operatorname{tp} \bar{v}) &\leq \partial(\operatorname{tp} \bar{u}, \operatorname{tp} \bar{w}) + \partial(\operatorname{tp} \bar{w}, \operatorname{tp} \bar{v}) \\ &\leq d(\bar{u}, \bar{w}) + \partial(\operatorname{tp} \bar{w}, \operatorname{tp} \bar{v}) \\ &\leq s + 2\epsilon. \end{aligned}$$

As ϵ was arbitrary, this completes the proof. □

The following omitting types theorem will allow us to deduce that F-categorical models are F-atomic. The version for infinitary continuous logic was proved in Eagle [E, Theorem 4.14].

Theorem 7.3 (Omitting types)**.** *Let T be a consistent F-theory and $p \in \mathrm{S}_n(T)$ be a type which is not principal. Then there exists a separable model of T which omits p.*

Eagle's definition of a *metrically principal* type [E, Definition 4.12] which is used in his [E, Theorem 4.14] is not quite the same as our definition of a principal type. The following lemma shows that they are equivalent.

Lemma 7.4. *Let $p \in \mathrm{S}_n(T)$. Then the following are equivalent:*

(i) *p is principal;*
(ii) *for every $\delta > 0$, $(p)_\delta$ has non-empty interior in the logic topology.*

Proof. We only need to prove (ii) ⇒ (i). Fix $1 > \delta > 0$ and suppose that q_0 is in the interior of $(p)_{\delta/2}$, i.e., there exists a formula $\phi(\bar{x})$ such that $\phi(q_0) = 0$ and $\phi(q) < 1 \implies \partial(q,p) < \delta/2$. Let

$$\psi(\bar{x}) = \inf_{\bar{y}} \phi(\bar{y}) \vee (d(\bar{x},\bar{y}) \dot{-} \delta/2).$$

First we check that $\psi(p) = 0$. Let $\bar{a} \models p$. Then, as $\partial(p,q_0) < \delta/2$ and $\phi(q_0) = 0$, we have that $\psi(p) = 0$. Next we verify that $\psi(q) < \delta/2 \implies q \in (p)_\delta$, thus showing that p is in the interior of $(p)_\delta$. Suppose that $\psi(q) < \delta/2$ and let $\bar{a} \models q$. Then there exists $\bar{b}$ such that $\phi(\bar{b}) < \delta/2 < 1$ and $d(\bar{a},\bar{b}) < \delta/2$. In particular, $\partial(p, tpb) < \delta/2$. We have

$$\partial(p,q) \leq \partial(p, \operatorname{tp}\bar{b}) + \partial(\operatorname{tp}\bar{b}, \operatorname{tp}\bar{a}) < \delta/2 + \delta/2 = \delta. \qquad \square$$

Combining everything we have so far, we obtain the following theorem.

Theorem 7.5. *Let F be a fragment such that the quantifier rank of formulas in F is bounded by α and let T be an F-theory which has a unique separable model A. Then the Scott rank of A is at most α.*

Proof. By Theorem 7.3, A is F-atomic: if $\operatorname{tp}\bar{a}$ is non-principal for some $\bar{a} \in A^n$, there exists a separable model of T that omits it and is therefore not isomorphic to A. Now Proposition 7.2 implies the conclusion. $\square$

We are finally ready to prove the main theorem of this section.

Theorem 7.6. *Let L be a countable language, let $\cong$ denote the isomorphism relation on $\mathcal{M}(L)$ and let $X \subseteq \mathcal{M}(L)$ be an $\cong$-invariant Borel subset. Let Ω be a universal modulus for L. Then the following are equivalent:*

(i) $\cong|_X$ *is Borel;*
(ii) *The Ω-Scott rank of elements of X is uniformly bounded below ω_1.*

Proof. (i) ⇒ (ii). Suppose that $\cong|_X$ is $\mathbf{\Pi}^0_\alpha$ for some $\alpha < \omega_1$. Then each isomorphism class contained in X is $\mathbf{\Pi}^0_\alpha$; by Theorem 6.1, for every $A \in X$, there exists a sentence ψ_A of quantifier rank at most $\alpha' = \omega \cdot 2 \cdot (2 + \alpha + 1)$ such that $\psi_A^B = 0$ if $A \cong B$ and $\psi_A^B = 1$ otherwise. Now for each $A \in X$, apply Theorem 7.5 to the fragment generated by ψ_A and the theory $\{\psi_A = 0\}$ to obtain that the rank of A is at most α'.

(ii) ⇒ (i). Suppose that the Scott rank of all structures in X is bounded by $\alpha < \omega_1$. Then the Scott sentence of each of those structures has quantifier rank at most $\alpha + \omega$. By Theorems 3.8 and 3.5, for $A, B \in X$, $A\, E_{\alpha+\omega}\, B$ iff $A\, E_\infty\, B$ and by Theorem 5.5, $A\, E_\infty\, B$ iff $A \cong B$. Thus, on X, the Borel relation $E_{\alpha+\omega}$ coincides with isomorphism. $\square$

Example 7.7. Recall that a complete metric space is called *proper* (or Heine–Borel) if all closed bounded sets are compact. The Euclidean spaces and, more generally, all complete Riemannian manifolds are examples of proper metric spaces. Clearly it suffices to require the Heine–Borel condition for closed balls centered at points in a countable dense sequence. Then, since a complete metric space is compact iff it is totally bounded, properness is a Borel property of Polish metric spaces.

By a theorem of Hjorth (see [GK, Thm. 3]), isometry of proper metric spaces is Borel bireducible with the universal countable Borel equivalence relation, and in particular Borel. Hence by Theorem 7.6, the Scott rank of proper metric spaces is bounded.

8. Two examples for the 1-Lipschitz modulus

8.1. **Gromov–Hausdorff distance between metric spaces.** We apply the tools we have developed in a particular example: calculating the Gromov–Hausdorff distance between metric spaces. Throughout this subsection, we fix a signature L containing only the distance symbol d, say, bounded by 1, and we let $\Omega = \Omega_{\mathrm{L}}$ be the 1-Lipschitz weak modulus defined by (2.3). All metric spaces that we consider below have diameter bounded by 1. (Everything goes through if 1 is replaced by an arbitrary positive constant.)

Recall that if (C, d) is a metric space and A, B are closed subsets of C, the *Hausdorff distance* between A and B, denoted by $d_{\mathrm{H}}(A, B)$ is defined by

$$d_{\mathrm{H}}(A, B) = \sup_{x \in A} d(x, B) \vee \sup_{y \in B} d(y, A).$$

If (A, d) and (B, d) are now abstract metric spaces, the *Gromov–Hausdorff* distance [G3, Def. 3.4] between A and B, denoted by $d_{\mathrm{GH}}(A, B)$, is defined by

$$d_{\mathrm{GH}}(A, B) = \inf_{f_1, f_2} d_{\mathrm{H}}(f_1(A), f_2(B)),$$

where f_1 and f_2 vary over all isometric embeddings of A and B in a third space C.

Similarly, we define the *enumerated Gromov–Hausdorff distance* between enumerated metric spaces as follows. Let A and B be metric spaces and $\{a_i : i \in I\}$, $B = \{b_i : i \in I\}$ be sequences of elements of A and B, respectively. We define the *enumerated Gromov–Hausdorff distance* $d_{\mathrm{eGH}}(a, b)$ as

$$d_{\mathrm{eGH}}(a, b) = \inf_{f_1, f_2} \sup_i d\big(f_1(a_i), f_2(b_i)\big),$$

where f_1 and f_2 vary over all isometric embeddings of $\{a_i : i \in I\}$ and $\{b_i : i \in I\}$ in a third metric space C.

Theorem 8.1. *Let $L = \{d\}$ and $\Omega = \Omega_{\mathrm{L}}$ as above. Then for any two metric spaces A, B of diameter at most* 1*,*

$$r_\infty(A, B) = d_{\mathrm{GH}}(A, B).$$

We start by calculating d_{eGH} between finite tuples.

Lemma 8.2. *Let A, B be metric spaces and let $\bar{a} \in A^n$, $\bar{b} \in B^n$ be finite tuples. Then*

$$d_{\mathrm{eGH}}(\bar{a}, \bar{b}) = \tfrac{1}{2} \sup_{i,j} \big|d(a_i, a_j) - d(b_i, b_j)\big| = r_0(A\bar{a}, B\bar{b}).$$

Proof. The first equality is proved in Uspenskij [U, Proposition 7.1]; we proceed to show the second. As basic formulas are 1-Lipschitz and they are preserved by embeddings, it is clear that $r_0(A\bar{a}, B\bar{b}) \leq d_{\mathrm{eGH}}(\bar{a}, \bar{b})$.

On the other hand, we have

$$\Delta_{\frac{1}{2}d(x_i, x_j)}(\delta) = \tfrac{1}{2}(\delta_i + \delta_j) \leq \delta_i \vee \delta_j \leq \Omega_{\mathrm{L}}(\delta),$$

so $\frac{1}{2}d(x_i, x_j)$ is an Ω_{L}-formula. This implies the other inequality. □

Proof of Theorem 8.1. Let A, B be metric spaces and $s \in \mathbf{R}$. We have:

$$\begin{aligned}
d_{\mathrm{GH}}(A, B) < s &\iff \exists a \in A^{\mathbf{N}}, b \in B^{\mathbf{N}} \text{ tail-dense such that } d_{\mathrm{eGH}}(a, b) < s \\
&\iff \exists a \in A^{\mathbf{N}}, b \in B^{\mathbf{N}} \ \sup_n d_{\mathrm{eGH}}(a|_n, b|_n) < s \\
&\iff \exists a \in A^{\mathbf{N}}, b \in B^{\mathbf{N}} \ \sup_n r_0(a|_n, b|_n) < s \\
&\iff r_\infty(A, B) < s.
\end{aligned}$$

The second equivalence follows by compactness (or, alternatively, from the fact that the first equality in Lemma 8.2 also holds for infinite tuples), the third is Lemma 8.2, and the fourth is given by Proposition 3.4. □

Theorem 8.1 and Corollary 4.4 give us the following.

Corollary 8.3. *Let $L = \{d\}$ and let $A \in \mathcal{M}(L)$. Then the set*

$$\{B \in \mathcal{M}(L) : d_{\mathrm{GH}}(A, B) = 0\}$$

is Borel.

Question 8.4. Let $\mathcal{M}$ be the space of Polish metric spaces with distance bounded by 1, as above. Let $A \in \mathcal{M}$ be fixed. Is $d_{\mathrm{GH}}(A, \cdot)$ a Borel function on $\mathcal{M}$?

Remark 8.5. It is not clear what the exact complexity of the equivalence relation E_{GH} ($d_{\mathrm{GH}} = 0$ on bounded metric spaces) is. However, Christian Rosendal pointed out to us that it is above the universal orbit equivalence relation of a Polish group action. This can be seen as follows. For a compact metrizable space X, denote by $C(X)$ the space of continuous functions on X equipped with the sup norm. Dutrieux and Kalton show in [DK] that if X and Y are not homeomorphic, then $d_{\mathrm{GH}}(C(X), C(Y)) \geq 1/16$. Thus the map $X \mapsto C(X)$ is a reduction from homeomorphism of compact spaces to E_{GH}. As the former is universal for orbit equivalence relations of Polish group actions (Zielinski [Z]), this yields the claim.

One way to see that this reduction is Borel is as follows. In [Z], compact metrizable spaces are parametrized by the elements of the hyperspace $K([0,1]^{\mathbf{N}})$ of closed subspaces of the Hilbert cube. An easy way to find a Borel map $\Phi\colon K([0,1]^{\mathbf{N}}) \to \mathcal{M}$ such that $\Phi(X)$ is isometric to $C(X)$ is the following. Let $z_i\colon [0,1]^{\mathbf{N}} \to [0,1]$ denote the projection on the i-th coordinate. Then instead of $\mathbf{N}$, we can take as a distinguished countable dense set in the definition of $\mathcal{M}$ the set of all polynomials with rational coefficients in infinitely many variables (all but finitely many coefficients of which are zero), and calculate the distance by

$$d(P_1, P_2) = \sup_{x \in X} |P_1(z_0(x), z_1(x), \ldots) - P_2(z_0(x), z_1(x), \ldots)|,$$

which is easily a Borel function of X. The Stone–Weierstrass theorem tells us that the resulting element of $\mathcal{M}$ is isometric to $C(X)$.

Question 8.6. Is E_{GH} Borel reducible to an orbit equivalence relation of a Polish group action?

Next we provide a simple example of two discrete metric spaces that have Gromov–Hausdorff distance 0 but are not isometric. Note however, that for compact metric spaces as well as for metric spaces with discrete set of distances, E_{GH} and isometry coincide.

Example 8.7. Let A be a countable subset of the open interval $(0,1)$. We define a countable metric space (X_A, d_A) as follows: $X_A = \{x_a, y_a : a \in A\}$, and for $x \neq y \in X_A$ we set

$$d_A(x,y) = \begin{cases} a & \text{if } x = x_a, y = y_a \text{ or vice versa;} \\ 1 & \text{otherwise.} \end{cases}$$

It is easy to check that (X_A, d_A) is a complete, ultrametric space. Now let $A, B \subseteq (0,1)$ be two distinct, dense, countable subsets of $(0,1)$. Then clearly X_A and X_B are not isometric but $d_{\mathrm{GH}}(X_A, X_B) = 0$. Indeed, fix $\epsilon > 0$ and let $f\colon A \to B$ be a bijection such that for all $a \in A$, $|f(a) - a| < \epsilon$. Let Y be the disjoint union of X_A and X_B and define the distance on Y by $d(x_a, y_a) = a$, $d(x_b, y_b) = b$, for $a \in A$, $b \in B$; $d(x_a, x_{f(a)}) = d(y_a, y_{f(a)}) = \epsilon$, $d(x_a, y_{f(a)}) = d(y_a, x_{f(a)}) =$

$\min((a+b)/2+\epsilon, 1)$ for all $a \in A$ and $d(x,y) = 1$ in all other cases. This Y witnesses that $d_{\mathrm{GH}}(X_A, X_B) \leq \epsilon$.

8.2. **Kadets distance between Banach spaces.** Our second example is Banach spaces. Here the language is $\{0, +, \|\cdot\|_n\} \cup \{M_\lambda : \lambda \in \mathbf{R}\}$, where 0 is a constant symbol, $+$ is a binary operation with $\Delta_+(\delta_1, \delta_2) = \delta_1 + \delta_2$, $\|\cdot\|_n$ is a 1-Lipschitz predicate symbol (interpreted as $\|x\|_n = \|x\| \wedge n$), and M_λ is a function symbol representing multiplication by λ with $\Delta_{M_\lambda}(\delta_0) = |\lambda|\delta_0$. We will denote by $\mathcal{M}$ (and restrict our considerations to) the Borel set of all separable Banach spaces rather than the space of all structures in this signature. Note that this is not relevant for the definition of the pseudo-distances r_α: they are computed in a manner independent of the ambient space. We can also restrict to a countable sublanguage by only keeping M_λ for $\lambda \in \mathbf{Q}$ without loss of generality. The distance function is given by $d(x,y) = \|x - y\|_1$.

Every atomic formula is equivalent (on $\mathcal{M}$) to a formula of the type $\|\sum_{i=0}^{n-1} \lambda_i x_i\|_k$ for some $k, n \in \mathbf{N}$, $\lambda_i \in \mathbf{R}$. An easy calculation yields that the modulus of continuity of this formula is $\Delta(\delta) = \sum_{i=0}^{n-1} |\lambda_i|\delta_i$. In particular, it obeys Ω_{L} iff $\sum_i |\lambda_i| \leq 1$.

Let A, B be two Banach spaces. The Kadets distance between A and B is defined analogously to the Gromov–Hausdorff distance by

$$d_{\mathrm{K}}(A, B) = \inf_{f_1, f_2} d_{\mathrm{H}}(f_1(\mathcal{B}_A), f_2(\mathcal{B}_B)),$$

where $\mathcal{B}_A$ and $\mathcal{B}_B$ denote the unit balls of A and B and f_1 and f_2 vary over all *linear* isometric embeddings of A and B in a third Banach space C.

Let now A and B be two Banach spaces and $\bar{a}$ and $\bar{b}$ be two sequences in $\mathcal{B}_A$ and $\mathcal{B}_B$ with dense span. The *enumerated Kadets distance* is defined as

$$d_{\mathrm{eK}}(A\bar{a}, B\bar{b}) = \inf_{f_1, f_2} \sup_i \|f_1(a_i) - f_2(b_i)\|_C,$$

where f_1 and f_2 vary over all linear isometric embeddings of A and B in a third space C. The following is the analogue of Lemma 8.2 for Banach spaces.

Lemma 8.8. *Let A and B be Banach spaces and $\bar{a} \in \mathcal{B}_A^n$, $\bar{b} \in \mathcal{B}_B^n$ be finite tuples. Then*

$$\begin{aligned} d_{\mathrm{eK}}(A\bar{a}, B\bar{b}) &= \sup\{\big|\|\sum_i \lambda_i a_i\| - \|\sum_i \lambda_i b_i\|\big| : \sum_i |\lambda_i| \leq 1\} \\ &= r_0(A\bar{a}, B\bar{b}). \end{aligned}$$

Proof. The first equality is proved in [BY3, Fact 3.4]. (There, the sup is taken over all λ_i with $\sum \lambda_i = 1$; however, scaling up in order to make the sum equal to 1 only increases the values inside the sup.) For the second, exactly as in Lemma 8.2, it suffices to recall that the formula $\|\sum_i \lambda_i x_i\|$ obeys Ω_{L} if $\sum_i |\lambda_i| \leq 1$. □

In the same way as before, we obtain the following theorem.

Theorem 8.9. *For any two separable Banach spaces A and B, we have*

$$r_\infty^{\Omega_{\mathrm{L}}}(A, B) = d_{\mathrm{K}}(A, B).$$

Question 8.4, Remark 8.5, and Question 8.6 also apply to the Kadets distance.

9. A characterization of CLI Polish groups

Recall that a Polish group is called a *CLI group* if it admits a complete, compatible, left-invariant metric. In [G1], Gao proves that if M is a classical countable structure, the automorphism group $\mathrm{Aut}(M)$ is a CLI group iff the classical Scott sentence of M does not admit uncountable models. In this section, we note that a similar result holds for general separable structures.

We start with a simple lemma (well-known in the classical setting) that establishes the connection between the left completion of $\mathrm{Aut}(M)$ and $\mathcal{L}_{\omega_1\omega}$-logic.

Lemma 9.1. *Let M be a separable metric structure and $\mathrm{Aut}(M)$ be its automorphism group equipped with the pointwise convergence topology. Then the left completion of $\mathrm{Aut}(M)$ can be identified with the monoid of $\mathcal{L}_{\omega_1\omega}$-elementary embeddings of M into itself. In particular, $\mathrm{Aut}(M)$ is a CLI group iff every such elementary embedding is surjective.*

Proof. The left uniformity of $\mathrm{Aut}(M)$ is defined by the pseudo-distances ρ_a given by $\rho_a(g,h) = d(g\cdot a, h\cdot a)$ for $a \in M$. If $(g_n)_n$ is a left-Cauchy sequence in $\mathrm{Aut}(M)$, then $g_n \cdot a$ converges for every $a \in M$; denote by f the limit map. As the interpretation of any $\mathcal{L}_{\omega_1\omega}$-formula $\phi(\bar{x})$ is continuous, we have that

$$\phi^M(f(\bar{a})) = \lim_n \phi^M(g_n \cdot \bar{a}) = \phi^M(\bar{a})$$

for any tuple $\bar{a} \in M^k$. We conclude that f is elementary.

Conversely, let $f\colon M \to M$ be $\mathcal{L}_{\omega_1\omega}$-elementary; we will show that it can be arbitrarily well approximated by automorphisms. Let $\bar{a} \in M^k$ be any tuple. By elementarity and Theorem 3.8, we have that $r_\infty^{\Omega_{\mathrm{U}}(L)}(M\bar{a}, Mf(\bar{a})) = 0$; now Theorem 5.5 implies that for any $\epsilon > 0$, there exists an automorphism g of M such that $d(f(\bar{a}), g\cdot\bar{a}) < \epsilon$.

For the second claim of the lemma, note that $\mathrm{Aut}(M)$ is CLI iff it coincides with its left completion. □

Theorem 9.2. *Let M be a separable structure in a countable language. Then the following are equivalent:*

(i) *The Scott sentence of M is non-zero on non-separable models;*
(ii) *$\mathrm{Aut}(M)$ is a CLI group.*

Proof. The proof is similar to the one in [G1], so we only give a brief sketch. By Lemma 9.1, we can replace (ii) by the condition that every $\mathcal{L}_{\omega_1\omega}$-elementary embedding of M into itself is surjective.

(i) ⇒ (ii). Suppose that $f\colon M \to M$ is an elementary embedding which is not surjective. By iterating f ω_1 times, we obtain an elementary chain every element of which is isomorphic to M; its union is non-separable and elementarily equivalent to M.

(ii) ⇒ (i). Conversely, suppose that N is a non-separable model elementarily equivalent to M. Let F denote the fragment of $\mathcal{L}_{\omega_1\omega}$ generated by the Scott sentence of M. Then by the downward Löwenheim–Skolem theorem, M can be embedded F-elementarily into N. Let $a \in N \setminus M$ be arbitrary. Again by Löwenheim–Skolem, there exists a separable $M' \prec_F N$ such that $M \cup \{a\} \subseteq M'$. Then $M' \cong M$ (as M' satisfies the Scott sentence of M) and the embedding $M \subseteq M'$ is not surjective. □

As isometry groups of proper metric spaces (see Example 7.7 for the definition) are locally compact, and locally compact groups are CLI, we have the following corollary.

Corollary 9.3. *Let M be a separable structure such that the underlying metric space is proper. Then M is the unique model of its Scott sentence.*

As noted in [G1], Hjorth and Solecki proved that all solvable groups are CLI. This provides a further setting where Theorem 9.2 applies.

10. Connections with classical logic

One might ask about the connection between the continuous $\mathcal{L}_{\omega_1\omega}$-logic considered in this paper and classical ($\{0,1\}$-valued) $\mathcal{L}_{\omega_1\omega}$-logic with regard to Polish metric structures. The latter has already been considered in the literature (see [D1] and the references therein). The main difference is that, while continuous logic treats separable structures as essentially countable objects, in classical logic, which does not allow for approximations, they are basically discrete structures of size continuum and as a result, quite unmanageable from a descriptive set theoretic point of view. Classically $\mathcal{L}_{\omega_1\omega}$-definable sets of Polish structures are usually not Borel sets. For a natural example, take local compactness of Polish metric spaces: by [NS], the class of such spaces is properly co-analytic, and at the same time definable in classical $\mathcal{L}_{\omega_1\omega}$.

In this section, we observe that, with appropriate coding, the expressive power of classical $\mathcal{L}_{\omega_1\omega}$-logic is strictly greater than that of continuous $\mathcal{L}_{\omega_1\omega}$-logic. Each continuous signature L yields a classical signature $\hat{L}$ defined as follows. For every predicate symbol P of arity n_P and for every $q \in \mathbf{Q}$, we put into L' an n_P-ary relation symbol P_q. (For simplicity, we assume that the language L has no function symbols; as we saw in Section 4, this entails no loss of generality.) To every L-structure A, we associate a classical $\hat{L}$-structure $\hat{A}$ with the same domain, where the interpretations of the symbols are given by:

$$\hat{A} \models P_q(\bar{a}) \iff P^A(\bar{a}) < q$$

for each $q \in \mathbf{Q}$ and $\bar{a} \in A^n$.

Proposition 10.1. *For every n-ary $\mathcal{L}_{\omega_1\omega}(L)$-formula $\phi(\bar{x})$ and every $q \in \mathbf{Q}$, there exists a formula $\hat{\phi}_q(\bar{x})$ in $\mathcal{L}_{\omega_1\omega}(\hat{L})$ such that for every L-structure A, $\bar{a} \in A^n$ and each q, we have*

$$\hat{A} \models \hat{\phi}_q(\bar{a}) \iff \phi^A(\bar{a}) < q.$$

Moreover, the quantifier rank of $\hat{\phi}_q$ is equal to that of ϕ.

Proof. First note that if we have formulas $\hat{\phi}_q$ for every q, then we can easily find a formula $\hat{\phi}^q$ that expresses $\phi > q$, namely $\hat{\phi}^q = \bigvee_{p>q} \neg\hat{\phi}_p$.

The proof of the proposition proceeds by induction on ϕ. For atomic formulas this is true by definition. If $\phi(\bar{x}) = \inf_y \psi(\bar{x}, y)$, we can take $\hat{\phi}_q(\bar{x}) = \exists y\ \psi_q(\bar{x}, y)$. If $\phi = \bigwedge_i \psi_i$, we can take $\phi_q = \bigvee_i \psi_{i,q}$.

Finally, suppose that $\phi = f(\psi_0, \dots, \psi_{k-1})$, where $f\colon \mathbf{R}^k \to \mathbf{R}$ is a connective. Write

$$f^{-1}\big((-\infty, q)\big) = \bigcup_{i \in \mathbf{N}} \prod_{j<k} (p_{i,j}, q_{i,j}),$$

where $p_{i,j}, q_{i,j} \in \mathbf{Q}$. This can be done because $f^{-1}\big((-\infty, q)\big)$ is an open set and products of rational intervals form a basis for $\mathbf{R}^k$. Then we can take

$$\hat{\phi}_q = \bigvee_i \bigwedge_{j<k} \hat{\psi}_j^{p_{i,j}} \wedge \hat{\psi}_{j,q_{i,j}}.$$ □

Thus, using Theorem 6.1, one obtains for every invariant Borel set of models a classical sentence that describes it. However, there is no converse to this: the set of models of a classical $\mathcal{L}_{\omega_1\omega}$-sentence is in general not Borel, as mentioned above. From Corollary 5.6, we obtain the following.

Corollary 10.2. *For every separable structure A, there exists a sentence σ_A in classical $\mathcal{L}_{\omega_1\omega}$-logic such that for every separable structure B,*

$$\hat{B} \models \sigma_A \iff B \cong A.$$

Finally, note that the classical Scott rank (as discussed in [D1]) and the continuous Scott rank from this paper are different. For example, consider the structure G_L from Remark 5.8: as it is ultrahomogeneous, its continuous Scott rank is 0. However, as orbits of the automorphism group are not closed, its classical Scott rank is non-zero. It was proved in [D1] that the classical Scott rank of Polish metric spaces is at most ω_1 but it remains an open question whether it must be countable.

References

[BBHU] I. Ben Yaacov, A. Berenstein, C. W. Henson, and A. Usvyatsov, *Model theory for metric structures*, Model theory with applications to algebra and analysis. Vol. 2, 2008, pp. 315–427.

[BI] I. Ben Yaacov and J. Iovino, *Model theoretic forcing in analysis*, Ann. Pure Appl. Logic **158** (2009), no. 3, 163–174.

[BY1] I. Ben Yaacov, *Topometric spaces and perturbations of metric structures*, Log. Anal. **1** (2008), no. 3-4, 235–272.

[BY2] ———, *Lipschitz functions on topometric spaces*, J. Log. Anal. **5** (2013), Paper 8, 21.

[BY3] ———, *The linear isometry group of the Gurarij space is universal*, Proc. Amer. Math. Soc. **142** (2014), no. 7, 2459–2467.

[CI] X. Caicedo and J. N. Iovino, *Omitting uncountable types and the strength of* $[0,1]$*-valued logics*, Ann. Pure Appl. Logic **165** (2014), no. 6, 1169–1200. MR3183322

[CL] S. Coskey and M. Lupini, *A López-Escobar theorem for metric structures, and the topological Vaught conjecture*, Fund. Math. **234** (2016), no. 1, 55–72. MR3509816

[D1] M. Doucha, *Scott rank of Polish metric spaces*, Ann. Pure Appl. Logic **165** (2014), no. 12, 1919–1929. Erratum [D2].

[D2] ———, *Erratum to: "Scott rank of Polish metric spaces" [Ann. Pure Appl. Logic 165 (12) (2014) 1919–1929] [MR3256743]*, Ann. Pure Appl. Logic **168** (2017), no. 7, 1490. MR3638899

[DK] Y. Dutrieux and N. J. Kalton, *Perturbations of isometries between* $C(K)$*-spaces*, Studia Math. **166** (2005), no. 2, 181–197. MR2109588

[E] C. J. Eagle, *Omitting types for infinitary* $[0,1]$*-valued logic*, Ann. Pure Appl. Logic **165** (2014), no. 3, 913–932. MR3142393

[EFP$^+$] G. A. Elliott, I. Farah, V. I. Paulsen, C. Rosendal, A. S. Toms, and A. Törnquist, *The isomorphism relation for separable C*-algebras*, Math. Res. Lett. **20** (2013), no. 6, 1071–1080.

[F] G. B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995.

[FHS] I. Farah, B. Hart, and D. Sherman, *Model theory of operator algebras II: model theory*, Israel J. Math. **201** (2014), no. 1, 477–505.

[G1] S. Gao, *On automorphism groups of countable structures*, J. Symbolic Logic **63** (1998), no. 3, 891–896.

[G2] ———, *Invariant descriptive set theory*, Pure and Applied Mathematics (Boca Raton), vol. 293, CRC Press, Boca Raton, FL, 2009.

[G3] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, English, Modern Birkhäuser Classics, Birkhäuser Boston Inc., Boston, MA, 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. MR2307192 (2007k:53049)

[GK] S. Gao and A. S. Kechris, *On the classification of Polish metric spaces up to isometry*, Mem. Amer. Math. Soc. **161** (2003), no. 766, viii+78. MR1950332 (2004b:03067)

[H] G. Hjorth, *Classification and orbit equivalence relations*, Mathematical Surveys and Monographs, vol. 75, American Mathematical Society, Providence, RI, 2000.

[HK] G. Hjorth and A. S. Kechris, *Borel equivalence relations and classifications of countable models*, Ann. Pure Appl. Logic **82** (1996), no. 3, 221–272.

[HKL] G. Hjorth, A. S. Kechris, and A. Louveau, *Borel equivalence relations induced by actions of the symmetric group*, Ann. Pure Appl. Logic **92** (1998), no. 1, 63–112.

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

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

[M] J. Melleray, *A note on Hjorth's oscillation theorem*, J. Symbolic Logic **75** (2010), no. 4, 1359–1365. MR2767973

[NS] A. Nies and S. Solecki, *Local compactness for computable polish metric spaces is* $\mathbf{\Pi^1_1}$*-complete*, Conference on Computability in Europe, 2015, pp. 286–290.

[S] D. Scott, *Logic with denumerably long formulas and finite strings of quantifiers*, Theory of Models (Proc. 1963 Internat. Sympos. Berkeley), 1965, pp. 329–341. MR0200133
[U] V. Uspenskij, *On subgroups of minimal topological groups*, Topology Appl. **155** (2008), no. 14, 1580–1606.
[V] R. Vaught, *Invariant sets in topology and logic*, Fund. Math. **82** (1974/75), 269–294. Collection of articles dedicated to Andrzej Mostowski on his sixtieth birthday, VII.
[Z] J. Zielinski, *The complexity of the homeomorphism relation between compact metric spaces*, Adv. Math. **291** (2016), 635–645. MR3459026

Itaï Ben Yaacov, Univ Lyon, Université Claude Bernard Lyon 1, Institut Camille Jordan, CNRS UMR 5208, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne cedex, France
URL: http://math.univ-lyon1.fr/~begnac/

Michal Doucha, Laboratoire de Mathématiques de Besançon, Université de Franche-Comté, 16, route de Gray, 25030 Besançon cedex, France
E-mail address: doucha@math.cas.cz
Current address: Institute of Mathematics CAS, Žitná 25, 115 67 Praha 1, Czech Republic

André Nies, Department of Computer Science, University of Auckland, Private Bag 92019, Auckland, New Zealand
E-mail address: andre@cs.auckland.ac.nz

Todor Tsankov, Institut de Mathématiques de Jussieu–PRG, Université Paris 7, case 7012, 75205 Paris cedex 13, France
E-mail address: todor@math.univ-paris-diderot.fr