

# Structurable equivalence relations and $\mathcal{L}_{\omega_1\omega}$ interpretations

Rishi Banerjee and Ruiyuan Chen

## Abstract

We show that the category of countable Borel equivalence relations (CBERs) is dually equivalent to the category of countable  $\mathcal{L}_{\omega_1\omega}$  theories which admit a one-sorted interpretation of a particular theory we call  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  that witnesses embeddability into  $2^{\mathbb{N}}$  and the Lusin–Novikov uniformization theorem. This allows problems about Borel combinatorial structures on CBERs to be translated into syntactic definability problems in  $\mathcal{L}_{\omega_1\omega}$ , modulo the extra structure provided by  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , thereby formalizing a folklore intuition in locally countable Borel combinatorics. We illustrate this with a catalogue of the precise interpretability relations between several standard classes of structures commonly used in Borel combinatorics, such as Feldman–Moore  $\omega$ -colorings and the Slaman–Steel marker lemma. We also generalize this correspondence to locally countable Borel groupoids and theories interpreting  $\mathcal{T}_{\text{LN}}$ , which admit a characterization analogous to that of Hjorth–Kechris for essentially countable isomorphism relations.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.A	Structurability	2
1.B	Interpretations	4
1.C	Comparing the strengths of Borel combinatorial structures	6
1.D	Locally countable Borel groupoids	7
1.E	Essential countability and imaginaries	8
<b>2</b>	<b>Preliminaries</b>	<b>9</b>
2.A	Descriptive set theory	9
2.B	Countable first-order logic $\mathcal{L}_{\omega_1\omega}$	11
2.C	Spaces of countable models	11
2.D	Spaces of $\mathcal{L}_{\omega_1\omega}$ types	12
2.E	Interpretations in $\mathcal{L}_{\omega_1\omega}$	14
2.F	Operations on theories	18
<b>3</b>	<b>Structurability of CBERs</b>	<b>19</b>
3.A	Structurability	19
3.B	Scott theories of CBERs	21
3.C	Lusin–Novikov functions	24
3.D	Structurability via interpretations	27

---

2020 *Mathematics Subject Classification*: 03E15, 03C15.

*Key words and phrases*: countable Borel equivalence relation, Borel groupoid, infinitary logic, interpretation, Lusin–Novikov, Lopez-Escobar.

<b>4</b>	<b>Examples of (non-)interpretations</b>	<b>29</b>
4.A	Variants of the Feldman–Moore theorem . . . . .	30
4.B	Edge-colorings and the intersection graph on finite subsets . . . . .	33
4.C	Other examples . . . . .	37
<b>5</b>	<b>Structurability of groupoids</b>	<b>40</b>
5.A	Groupoids . . . . .	40
5.B	Simplicial nerves . . . . .	42
5.C	Scott theories of groupoids . . . . .	44
5.D	Connection to essential countability and imaginaries . . . . .	46
5.E	Structurability . . . . .	49
5.F	Some examples . . . . .	51
	<b>References</b>	<b>53</b>

# 1 Introduction

This paper is a contribution to the global theory of locally countable Borel combinatorics and equivalence relations and connections with countable model theory. A **countable Borel equivalence relation (CBER)**  $E$  on a standard Borel space  $X$  is a Borel equivalence relation  $E \subseteq X^2$  such that all  $E$ -classes are countable. Over the past thirty years, CBERs have been widely studied in descriptive set theory and adjacent areas, such as group theory, ergodic theory, and combinatorics. See [Kec24] for a comprehensive survey.

## 1.A Structurability

An important aspect of the theory of CBERs consists of analyzing the Borel ways of assigning a combinatorial structure to each  $E$ -class. For instance, a CBER is called **smooth** iff there is a Borel way to pick a single point from each class. A CBER is called **hyperfinite** iff it is the orbit equivalence relation of a Borel action  $\mathbb{Z} \curvearrowright X$  (see [DJK94]); this amounts to a *transitive*  $\mathbb{Z}$ -action on each class. Smooth and hyperfinite CBERs are the “simplest” CBERs. On the other hand, every class of every CBER may be given the structure of a locally finite connected graph in a Borel way [JKL02]; and every such Borel locally finite graph admits a Borel  $\omega$ -coloring [KST99].

In general, given a CBER  $(X, E)$ , a Borel family of first-order structures (group actions, graphs, etc.) on each  $E$ -class  $C \in X/E$  is called a **structuring** of  $E$ ; if one exists,  $E$  is called **structurable** (by group actions, graphs, etc.). See Definition 3.1 for the precise definition, including the meaning of “Borel family”. Structurability provides a global framework for comparing and organizing “all (locally countable) Borel combinatorics problems”; see [Kec91], [JKL02], [CK18].

It is a basic folklore intuition that doing locally countable Borel combinatorics often amounts to “countable combinatorics, done canonically/in a uniform Borel way”; see e.g., [CPTT23, §1.E]. For instance, not every CBER is hyperfinite, even though obviously every nonempty countable set admits a transitive  $\mathbb{Z}$ -action, because there is no “canonical” way to choose said action. On the other hand, after picking a single distinguished point, it becomes possible to define a “sufficiently canonical” such action (or indeed any other structure, subject to cardinality constraints).

In this paper, extending the work of [CK18], we develop a correspondence between the theory of structurability of CBERs and countable model theory that allows such intuitions to be made precise. A weak form of this correspondence concerns existence of structurings:

**Theorem 1.1** (see Corollary 3.27). *Let  $\mathcal{K}, \mathcal{K}'$  be two classes of countable first-order structures, axiomatized by respective countable  $\mathcal{L}_{\omega_1\omega}$  theories  $\mathcal{T}, \mathcal{T}'$ . The following are equivalent:*

- (i) *Every CBER structurable by structures in  $\mathcal{K}$  is also structurable by structures in  $\mathcal{K}'$ .*
- (ii) *We may uniformly define a countable  $\mathcal{K}'$ -structure  $\mathcal{M}'$  from every countable  $\mathcal{K}$ -structure  $\mathcal{M}$ , equipped with two additional pieces of structure, namely (1) a countable family  $(U_i)_{i \in \mathbb{N}}$  of subsets separating distinct points and (2) a countable family  $(f_i)_{i \in \mathbb{N}}$  of unary functions whose graphs cover  $M^2$ . These definitions are expressible by  $\mathcal{L}_{\omega_1\omega}$  formulas independent of  $\mathcal{M}$ .*

Here  $\mathcal{L}_{\omega_1\omega}$  is the countably infinitary first-order logic, with countable Boolean connectives  $\bigwedge, \bigvee, \neg$  as well as the usual (finitary) quantifiers  $\forall, \exists$ ; see [Mar16] and Section 2.B below. For a class of structures to be  $\mathcal{L}_{\omega_1\omega}$ -definable means equivalently, in semantic terms, that said class is Borel in the standard Borel space of structures on a fixed countable set, by the Lopez-Escobar Theorem 2.8. Given a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , by a  **$\mathcal{T}$ -structuring** we mean a structuring by models of  $\mathcal{T}$ .

**Example 1.2.** The class  $\mathcal{K}'$  of sets equipped with a single distinguished point may be axiomatized by the empty (always true) theory  $\mathcal{T}_{\text{pt}} = \emptyset$  in the first-order language  $\{c\}$  with a single constant. The class  $\mathcal{K}$  of sets equipped with a finite nonempty subset  $D$  may be axiomatized by the theory

$$\mathcal{T}_{\text{finsub}} := \bigvee_{n \geq 1} \exists z_0, \dots, z_{n-1} \forall w (D(w) \leftrightarrow \bigvee_{i < n} (w = z_i)),$$

where now  $D$  is a unary relation. Clearly, given a set  $M$  equipped with a finite nonempty set  $D \subseteq M$ , there is in general no way to canonically pick a single distinguished element  $c \in M$ ; thus we cannot uniformly define a model  $(M, c) \in \mathcal{K}'$  of  $\mathcal{T}_{\text{pt}}$  from a model  $(M, D) \in \mathcal{K}$  of  $\mathcal{T}_{\text{finsub}}$ . (Formally, this is because  $(M, D)$  may have automorphisms with no fixed points.)

However, it is well-known that given a CBER  $E$ , if there is a Borel way to pick a finite nonempty set in each class, i.e.,  $E$  is  $\mathcal{T}_{\text{finsub}}$ -structurable, then in fact  $E$  is smooth, i.e.,  $\mathcal{T}_{\text{pt}}$ -structurable. This is because, on the one hand, given a finite nonempty set  $D$  as well as a linear order  $<$ , we may canonically pick the least element  $c$  of  $D$ :

$$(1.3) \quad c = x \iff D(x) \wedge \forall y ((y < x) \rightarrow \neg D(y)).$$

In other words, we may define a model of  $\mathcal{T}_{\text{pt}}$  given a model of  $\mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LO}}$ , where  $\mathcal{T}_{\text{LO}}$  is the theory of linear orders. On the other hand, every CBER  $E \subseteq X^2$  is  $\mathcal{T}_{\text{LO}}$ -structurable, because e.g., we may assume the standard Borel space  $X$  to be embedded in  $2^{\mathbb{N}}$ , and then take the lexicographical order on each  $E$ -class. The embedding  $X \rightarrow 2^{\mathbb{N}}$  is given by the indicator functions of countably many  $U_i \subseteq X$  separating points (the subbasic clopens); the lexicographical order is then given by

$$(1.4) \quad x < y \iff \bigvee_{i \in \mathbb{N}} (\neg U_i(x) \wedge U_i(y) \wedge \bigwedge_{j < i} (U_j(x) \leftrightarrow U_j(y))).$$

By substituting this latter formula (1.4) for the symbol  $<$  in (1.3), we obtain a formula  $\phi(x)$  uniformly defining a single point from a finite nonempty set  $D$  as well as a countable separating family  $(U_i)_{i \in \mathbb{N}}$  as in Theorem 1.1(ii), thereby witnessing that  $\mathcal{T}_{\text{finsub}}$ -structurability implies  $\mathcal{T}_{\text{pt}}$ -structurability.

Based on anecdotal evidence, Theorem 1.1 seems to be more-or-less folklore among the Borel combinatorics community; at least, it seems widely believed that some statement along these lines ought to be true. However, to our knowledge, such a formal statement or proof has not appeared before in the literature.

## 1.B Interpretations

Our main results in this paper show that there is a “complete correspondence” between semantic concepts in the realm of CBERs and structurability, and “syntactic definitions” of such concepts in the realm of the infinitary logic  $\mathcal{L}_{\omega_1\omega}$  and countable model theory. From this correspondence, Theorem 1.1 and similar results will follow as immediate consequences.

Theorem 1.1(ii) says that “in every model of  $\mathcal{T}$ , we may uniformly define a model of  $\mathcal{T}'$ ”. In general, given two  $\mathcal{L}_{\omega_1\omega}$  theories  $\mathcal{T}, \mathcal{T}'$  (in two different languages, which we assume to be relational for simplicity), by an **interpretation**  $\alpha : \mathcal{T}' \rightarrow \mathcal{T}$ , we mean a family of  $\mathcal{L}_{\omega_1\omega}$  formulas  $\alpha(R)(x_0, \dots, x_{n-1})$  in the language of  $\mathcal{T}$ , for each  $n$ -ary relation symbol  $R$  in the language of  $\mathcal{T}'$ , such that these formulas define a model of  $\mathcal{T}'$  in every model of  $\mathcal{T}$ . See Section 2.E for details and alternate formulations.<sup>1</sup> For example, the above formula (1.3) yields an interpretation

$$(1.5) \quad \mathcal{T}_{\text{pt}} \longrightarrow \mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LO}}.$$

Now Theorem 1.1(ii) does not assert the existence of an interpretation  $\mathcal{T}' \rightarrow \mathcal{T}$ , but rather the existence of an interpretation  $\mathcal{T}' \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  into the theory  $\mathcal{T}$  expanded with two additional pieces of structure. Namely,  $\mathcal{T}_{\text{sep}}$  is the **theory of countable separating families**, in the language with countably many unary relations  $\{U_i\}_{i \in \mathbb{N}}$  asserting that they separate distinct points (see Definition 3.18); and  $\mathcal{T}_{\text{LN}}$  is the **theory of Lusin–Novikov functions**, in the language with countably many unary functions  $\{f_i\}_{i \in \mathbb{N}}$  asserting that they cover all pairs (see Definition 3.23). Intuitively, the theory  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  describes combinatorial structure which is available “for free” on every CBER, by the Lusin–Novikov theorem and standard Borelness of the underlying space. For example, the above formula (1.4) yields an interpretation

$$(1.6) \quad \mathcal{T}_{\text{LO}} \longrightarrow \mathcal{T}_{\text{sep}},$$

hence also  $\mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{sep}}$ , which may be composed with the above interpretation (1.5) to yield an interpretation  $\mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{sep}}$  describing how a Borel transversal of a CBER may be constructed from a Borel choice of a finite subset in each class (and a countable separating family, which is available for free on a CBER).

We thus have a category of countable  $\mathcal{L}_{\omega_1\omega}$  theories, with interpretations between them as morphisms. Within this category, there is a distinguished theory  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , such that the theories admitting an interpretation from it capture structurability of CBERs. We show that in fact, such theories are “equivalent” to CBERs:

**Theorem 1.7** (see Theorem 3.26). *We have a dual equivalence of categories*

$$\begin{aligned} \{ \text{CBERs, class-bijective homomorphisms} \} &\simeq \{ \text{countable } \mathcal{L}_{\omega_1\omega} \text{ theories interpreting } \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \} \\ E &\mapsto \mathcal{T}_E. \end{aligned}$$

Moreover,  $\mathcal{T}$ -structurings of a CBER  $E$  are in natural bijection with interpretations  $\mathcal{T} \rightarrow \mathcal{T}_E$ .

This is illustrated in Figure 1.9. Here, a **(Borel) class-bijective homomorphism**  $f : (X, E) \rightarrow (Y, F)$  between two CBERs is a Borel map  $f : X \rightarrow Y$  whose restriction to each  $E$ -class is a bijection

<sup>1</sup>In this paper, we focus on *one-sorted* interpretations (sometimes called *definitions* in model theory), which define a uniform construction of a  $\mathcal{T}'$ -model on the same underlying set of a  $\mathcal{T}$ -model. The more general model-theoretic interpretations into imaginaries will play a minor role only; see Section 1.E.

with some  $F$ -class. The theory  $\mathcal{T}_E$  corresponding to a CBER  $E$  is the **Scott theory** of  $E$ , which is defined uniquely up to bi-interpretability by declaring its models on any countable set  $Y$  to be bijections between  $Y$  and some  $E$ -class; see Definition 3.10.

The Scott theory was introduced in [CK18] (there called *Scott sentences*), which also established much of its significance with respect to structurability. In particular, the construction of  $\mathcal{T}_E$ , the last assertion of Theorem 1.7, as well as the full faithfulness of the functor  $E \mapsto \mathcal{T}_E$  (i.e., that it is a bijection on morphisms) appeared in that paper in some form. However, that paper was focused on the CBER-theoretic aspects of structurability, and as such, discussed the model-theoretic significance of the Scott theory in somewhat roundabout terms; in particular, only a terse overview of  $\mathcal{L}_{\omega_1\omega}$  interpretations was given, in the unpublished Appendix B of the arXiv preprint of [CK18]. (Henceforth, all references to [CK18, Appendix B] are to the arXiv preprint.)

In this paper, we give a detailed, self-contained introduction to  $\mathcal{L}_{\omega_1\omega}$  interpretations in Section 2.E, with connections to and equivalent formulations in terms of Borel  $S_{\leq\infty}$ -spaces of countable models, Borel spaces of  $\mathcal{L}_{\omega_1\omega}$  types, and Boolean  $\sigma$ -algebras of formulas. After reviewing several equivalent definitions of structurability in Section 3.A, we then give in Section 3.B an alternate treatment of Scott theories  $\mathcal{T}_E$ , based on the conceptual definition up to bi-interpretability given above, rather than an explicit coding as in [CK18]; and we show how all of the fundamental properties of Scott theories have simple derivations from this perspective.

In Section 3.C, we prove Theorem 1.7 (Theorem 3.26). The bulk of the proof is devoted to characterizing the essential image of the functor  $E \mapsto \mathcal{T}_E$ , i.e., showing that every theory  $\mathcal{T}$  interpreting  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is bi-interpretable with the Scott theory of a CBER  $E$ . This is the main new technical contribution of Theorem 3.26, and answers a question from [CK18, sentence after B.4]. To briefly outline the proof: the CBER  $E$  is constructed on the space  $\mathcal{S}_1(\mathcal{T})$  of  $\mathcal{L}_{\omega_1\omega}$  1-types of  $\mathcal{T}$  (equivalently, the space of isomorphism types of pointed models), two such types being equivalent iff they are realized in isomorphic models. The interpretation of  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is used to verify that  $E$  is indeed a CBER with Scott theory  $\mathcal{T}_E \cong \mathcal{T}$ , using a quantifier-elimination argument.

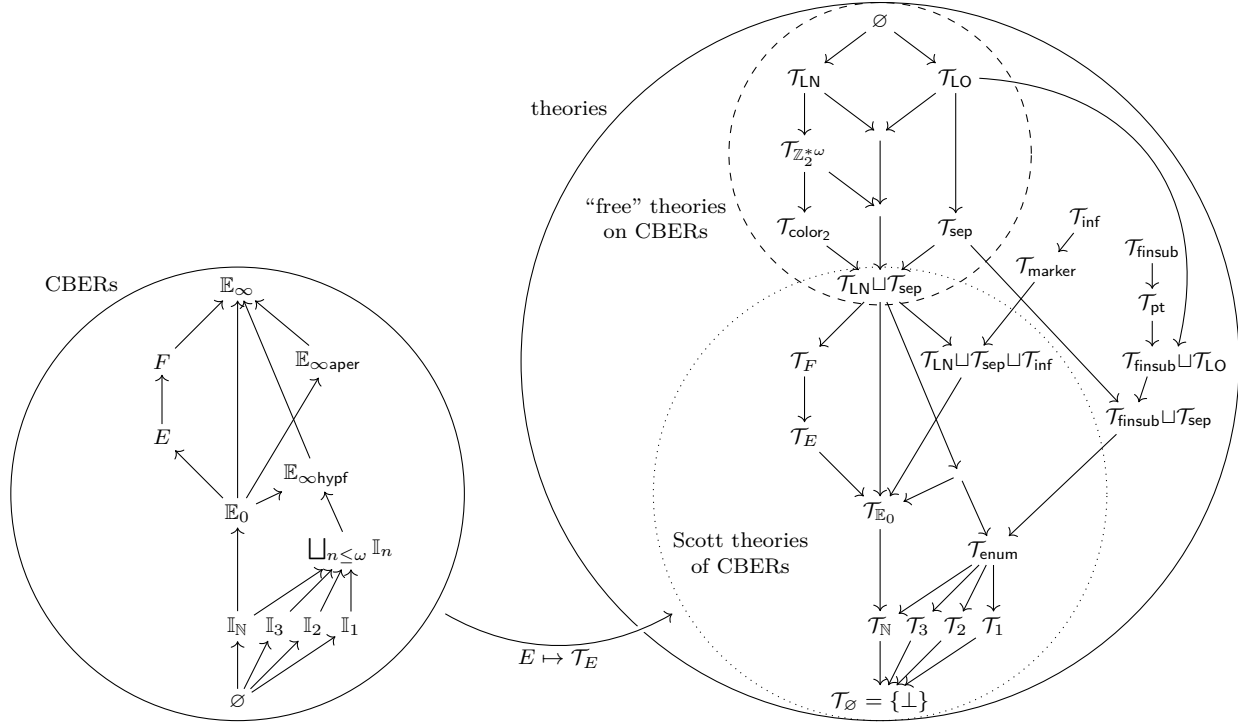
In Section 3.D, we show how Theorem 1.1, as well as syntactic formulations of all other results and constructions in structurability theory, easily follow from Theorem 1.7. We stress that this latter result, about individual *structurings* and *interpretations*, contains much more information than the former result about *structurability* and *interpretability*.

For instance, whereas a typical problem concerning CBERs asks whether a CBER may be structured by e.g., graphs of a certain kind, a typical problem in Borel combinatorics asks whether e.g., a given structuring  $\mathcal{M}$  of a CBER by graphs admits a Borel coloring. This may be abstractly phrased in terms of an *expansion* of  $\mathcal{M}$  to a structuring by graphs equipped with a coloring. Using Theorem 1.7, we may likewise show that all such “Borel expandability problems” amount to “countable expandability, done uniformly”, thereby generalizing Theorem 1.1:

**Theorem 1.8** (see Corollary 3.29). *Let  $\mathcal{T} \subseteq \mathcal{T}'$  be countable  $\mathcal{L}_{\omega_1\omega}$  theories in languages  $\mathcal{L} \subseteq \mathcal{L}'$ . The following are equivalent:*

- (i) *Every  $\mathcal{T}$ -structuring of a CBER admits an expansion to a  $\mathcal{T}'$ -structuring.*
- (ii) *There exists an interpretation  $\mathcal{T}' \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  whose restriction to the language  $\mathcal{L}$  is ( $\mathcal{T}$ -provably equivalent to) the identity.*

We illustrate this in the aforementioned example of graph colorings in Proposition 4.29.



**Figure 1.9:** Hierarchies of CBERs under class-bijective homomorphisms, and  $\mathcal{L}_{\omega_1\omega}$  theories under interpretations, with embedding from former to latter via Scott theories.

### 1.C Comparing the strengths of Borel combinatorial structures

By Theorem 1.1, every positive result relating two classes of CBERs defined by structurability may in principle be formulated entirely in terms of countable structures, without mentioning Borelness or CBERs. In fact, in practice many well-known constructions in Borel combinatorics already amount to  $\mathcal{L}_{\omega_1\omega}$  interpretations between the relevant theories. We illustrate this in Section 4 with several examples along the lines of Example 1.2, reformulating the standard proofs of e.g., the Feldman–Moore theorem [FM77] (Proposition 4.4), its generalization by Kechris–Miller [KM04] to an  $\omega$ -coloring of the intersection graph on finite subsets (Proposition 4.7), and the Slaman–Steel marker lemma [SS16] (Proposition 4.18) into explicit  $\mathcal{L}_{\omega_1\omega}$  interpretations.

Along the way, we also make a catalogue of *non-interpretability* between the relevant theories. When the theories in question, such as the three aforementioned, describe structure which is available “for free” on every CBER, this means that the construction of said structure on a CBER must necessarily make use of the “free” structure on CBERs provided by  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ . By comparing different theories in this way, we may make precise the idea that certain results in Borel combinatorics are “stronger”, or “more specific to CBERs”, than others.

To illustrate this, consider the Feldman–Moore theorem [FM77], which states that every CBER is the orbit equivalence relation of a countable Borel group action. However, the original proof showed the *a priori* stronger statement that every CBER is a countable union of Borel involutions, in which form the result is often used in Borel combinatorics (see e.g., [Kec24, 3.4]). We give the following analysis of these and other forms of the Feldman–Moore theorem, which we state here in somewhat informal terms; see the quoted results for more precise statements:

**Theorem 1.10** (see Propositions 4.4, 4.5, 4.14 and 4.15 and Remark 4.13). *Let  $E \subseteq X^2$  be a CBER. Consider the following 3 forms of the Feldman–Moore theorem:*

- (i)  *$E$  is induced by a Borel action of the free group  $\mathbb{F}_\omega \curvearrowright X$ .*
- (ii)  *$E$  is induced by a Borel action of a group generated by involutions  $\mathbb{Z}_2^{*\omega} \curvearrowright X$ .*
- (iii)  *$E = \bigcup_i f_i$  can be covered by the graphs of countably many Borel involutions  $f_i : X \rightarrow X$ .*

*Then:*

- (a) *It is possible to prove (i) using only the Lusin–Novikov theorem applied to  $E$ .*
- (b) *It is not possible to prove (ii) or (iii) using only Lusin–Novikov (equivalently, using (i)), without also using a countable separating family of Borel subsets  $U_i \subseteq X$ .*
- (c) *It is also not possible to prove (iii) using only (ii). Thus, the 3 versions of Feldman–Moore are “strictly increasing” in strength.*

Here, for example, (c) means precisely that the theory  $\mathcal{T}_{\mathbb{Z}_2^{*\omega}}$  of transitive  $\mathbb{Z}_2^{*\omega}$ -actions (structurability by which formalizes the statement of (ii)) does not interpret the theory  $\mathcal{T}_{\text{color}_2}$  of edge  $\omega$ -colorings of the complete graph (which formalizes (iii)). Note that, as (i)–(iii) are all true (by the Feldman–Moore theorem), by Theorem 1.1, the three respective theories do interpret each other when combined with the theory  $\mathcal{T}_{\text{sep}}$ . Nonetheless, the above result allows us to make precise the idea that e.g., (iii) is a “strictly stronger” statement of the Feldman–Moore theorem than (i).

Figure 1.9 shows various other theories describing common Borel combinatorial structures, including several available “for free” on a CBER, whose strengths we may distinguish in a similar manner; a more detailed diagram is shown in Figure 4.17.

## 1.D Locally countable Borel groupoids

A **groupoid**  $(X, G)$  is, in short, a category with invertible morphisms, consisting of a space of objects  $X$  and a space of morphisms  $G$ , together with operations (maps) specifying the domain and codomain of a morphism, as well as composition, identity, and inverse morphisms. A groupoid can be viewed as a common generalization of a group (when  $X = 1$ ), an equivalence relation (when  $G \subseteq X^2$ ), and a group action (when equipped with a “fibration” to the acting group  $\Gamma$ ).

A **locally countable Borel groupoid** is a groupoid in which  $X, G$  are standard Borel spaces, all of the groupoid operations are Borel maps, and each object is the domain of only countably many morphisms; see Definition 5.1. Such groupoids generalize CBERs and countable Borel group actions, and admit a largely analogous structural theory, with natural analogues of concepts such as amenability (see e.g., [TDW21]), treeability, graphings and cost (see [Ued06], [Alv08], [Car11]). Nonetheless, to our knowledge, a general theory of structurability for groupoids, akin to [CK18] for CBERs, has not been considered before in the literature.

In Section 5, we develop the beginnings of such a theory. Given a locally countable Borel groupoid  $(X, G)$  and an  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , by a  **$\mathcal{T}$ -structuring** of  $G$ , we mean a Borel right-translation-invariant family  $(\mathcal{M}_x)_{x \in X}$  of models of  $\mathcal{T}$  on the set of morphisms with domain  $x$ , for each  $x \in X$ ; see Definition 5.26. This is the natural generalization of structurability for CBERs, with the expected examples: e.g, a groupoid will be “treeable” (structurable by trees) iff it is the free groupoid generated by a Borel multigraph (so in particular, a “treeable” group in this sense is a free group; see Example 5.35). We prove the following analogue of Theorem 1.7:

**Theorem 1.11** (see Theorem 5.15 and Corollary 5.32). *We have a dual equivalence of categories*

$$\begin{aligned} \{\text{locally countable Borel groupoids}\} &\simeq \{\text{countable } \mathcal{L}_{\omega_1\omega} \text{ theories interpreting } \mathcal{T}_{\text{LN}}\} \\ G &\mapsto \mathcal{T}_G. \end{aligned}$$

*Moreover,  $\mathcal{T}$ -structurings of a locally countable Borel groupoid  $G$  are in natural bijection with interpretations  $\mathcal{T} \rightarrow \mathcal{T}_G$ .*

Here the morphisms in the left category are Borel **fibrations**  $G \rightarrow H$  between groupoids, i.e., functors exhibiting  $G$  as the action groupoid of a Borel action of  $H$ ; see Example 5.6. The **Scott theory**  $\mathcal{T}_G$  associated to a locally countable Borel groupoid  $G$  is defined by declaring its models on a set  $Y$  to be fibrations  $\mathbb{I}_Y \rightarrow G$  from the indiscrete groupoid  $\mathbb{I}_Y = Y^2$ ; see Definition 5.11.

Just as Theorem 1.1 shows that  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  describes all of the combinatorial structure available “for free” on a CBER, it follows analogously from Theorem 1.11 that  $\mathcal{T}_{\text{LN}}$  describes precisely the structure available “for free” on a locally countable Borel groupoid. In other words, every Borel combinatorial construction on groupoids may in principle be done “countably, in a uniform way”, where “uniform” here has a more restrictive meaning than in the case of CBERs (namely, we still have access “for free” to Lusin–Novikov functions  $\{f_i\}_{i \in \mathbb{N}}$ , but no longer have a countable separating family  $\{U_i\}_{i \in \mathbb{N}}$ ). See Corollary 5.33 for the precise statement.

For instance, recall that by Theorem 1.10, the Feldman–Moore theorem in the weak form of “generated by a countable group action” can be proved using only Lusin–Novikov, but the stronger form “generated by involutions” cannot. It follows that we have a weak “Feldman–Moore theorem for groupoids”, which says that every locally countable Borel groupoid  $(X, G)$  admits an action of a countable group  $\Gamma \curvearrowright G$  on the space of morphisms, which commutes with right multiplication and whose orbits are precisely all morphisms with a fixed domain. Equivalently, such data correspond to a countable subgroup of the **Borel full group**  $[G]$  which covers  $G$ ; in this form, the “Feldman–Moore theorem for groupoids” was first stated (without proof) in [TDW21]. See Example 5.37.

The route we take to Theorem 1.11 (Theorem 5.15) is somewhat more involved than in the case of CBERs, but we believe it provides some insight that may be of independent interest. Given a theory  $\mathcal{T}$  interpreting  $\mathcal{T}_{\text{LN}}$ , we consider the sequence of  $\mathcal{L}_{\omega_1\omega}$  type spaces  $\mathcal{S}_n(\mathcal{T})$  for each  $n \geq 1$ . We show that these form a **locally countable standard Borel simplicial set**, satisfying the so-called *Grothendieck–Segal condition*, that ensures it is the simplicial nerve of a groupoid. We will review the relevant concepts from simplicial homotopy theory in Section 5.B.

## 1.E Essential countability and imaginaries

In Section 5.D, we briefly discuss the significance of Scott theories of CBERs and locally countable Borel groupoids within the broader context of countable model theory.

In [HK96], Hjorth–Kechris gave a model-theoretic characterization of those countable  $\mathcal{L}_{\omega_1\omega}$  theories  $\mathcal{T}$  whose isomorphism relation  $\cong$  is Borel bireducible to a CBER (such  $\cong$  are called **essentially countable**). Namely, every countable model  $\mathcal{M}$  of  $\mathcal{T}$  must be  $\aleph_0$ -categorical with respect to a countable fragment  $\mathcal{F}$  of  $\mathcal{L}_{\omega_1\omega}$  after fixing finitely many constants  $\vec{a} \in M^n$ .

In [Che19a, Che19b], following earlier work by Harrison–Trainor–Miller–Montalbán [HTMM18], Chen showed that *arbitrary* countable  $\mathcal{L}_{\omega_1\omega}$  theories are “equivalent” to their Borel groupoids of countable models. More precisely, we have an equivalence of 2-categories

$$(1.12) \quad \begin{aligned} \{\text{theories}\} &\simeq \{\text{groupoids}\} \\ \mathcal{T} &\mapsto \{\text{countable models of } \mathcal{T}\}, \end{aligned}$$



where the left-hand side consists of countable  $\mathcal{L}_{\omega_1\omega}$  theories and interpretations between them as morphisms (and definable isomorphisms as 2-cells), and the right-hand side consists of standard Borel groupoids obeying a certain topological condition, with Borel functors between them (and Borel natural isomorphisms between those). Here the *interpretations* between theories  $\mathcal{T}' \rightarrow \mathcal{T}$  are allowed to be more general than those considered above (see Section 1.B): they describe a uniform construction of a model of  $\mathcal{T}'$  from each model  $\mathcal{M}$  of  $\mathcal{T}$ , but living not necessarily on the same underlying set, but rather on a set “uniformly defined” from  $\mathcal{M}$ . This “uniformly defined set” is the model-theoretic notion of an **imaginary** (see [Hod93, Ch. 7]), adapted to the  $\mathcal{L}_{\omega_1\omega}$  setting (namely, a definable quotient of a countable disjoint union of definable sets).

In light of (1.12), the aforementioned Hjorth–Kechris result may be regarded as characterizing those theories  $\mathcal{T}$  whose groupoid of models is Borel equivalent to a groupoid each of whose connected components has only countably many objects (namely, isomorphic models). Note that such a groupoid still need not be a locally countable Borel groupoid; it will be one iff each model moreover has only countably many automorphisms. From this and Theorem 5.15, we derive:

**Theorem 1.13** (see Corollary 5.24). *The above correspondence (1.12) restricts to an equivalence between theories obeying the following equivalent conditions, and locally countable Borel groupoids:*

- (i) *Every countable model has only countably many automorphisms, and the isomorphism relation  $\cong$  is essentially countable.*
- (ii) *There is a countable fragment  $\mathcal{F}$  of  $\mathcal{L}_{\omega_1\omega}$ , such that every countable model becomes rigid and  $\mathcal{F}$ -categorical after fixing finitely many constants  $\vec{a}$ .*

*The inverse is given (up to bi-interpretability) by the Scott theory  $G \mapsto \mathcal{T}_G$  of a groupoid.*

*Under this correspondence, the CBERs correspond to theories all of whose models are rigid.*

We expect that  $\mathcal{L}_{\omega_1\omega}$  imaginaries and the more general notion of interpretations into them ought to play a broader role in Borel structurability theory as well. Indeed, many well-known “uniform Borel combinatorial constructions” produce structures on a new underlying set derived from an existing structure. For example, this is the case for arguments showing closure of a given class of structurable CBERs under Borel reducibility (e.g., the treeable CBERs [JKL02, 3.3]). It would be interesting to have analogues of Theorems 1.1 and 1.7 for Borel reducibility, inclusions of subequivalence relations, class-injective homomorphisms (see [CK18, §5]), etc.; such analogues would presumably involve interpretations into imaginaries. We leave such questions for future work.

**Acknowledgments** We would like to thank Alexander Kechris and Anush Tserunyan for encouraging us to pursue this line of work, Robin Tucker-Drob for pointing out the connection to [TDW21], and Matthew Harrison-Trainor and Alexander Kechris for helpful comments and suggestions. R.C. was supported by NSF grant DMS-2224709.

## 2 Preliminaries

### 2.A Descriptive set theory

For background on the descriptive set-theoretic notions we will use, see [Kec95], [Gao09].

By a **Borel space**, we will mean what is often called a **measurable space**: a set  $X$  equipped with a  $\sigma$ -algebra  $\mathcal{B}(X) \subseteq \mathcal{P}(X)$ . A **standard Borel space** is one which is Borel isomorphic to a

Borel subspace of the Cantor space  $2^{\mathbb{N}}$  with the product topology (or equivalently, any Polish space). A **Borel map** between Borel spaces is one such that the preimage of every Borel set is Borel.

For a group action  $G \curvearrowright X$  on a standard Borel space  $X$ , the **orbit equivalence relation**  $\mathbb{E}_G^X$  is

$$x \mathbb{E}_G^X y \iff \exists g \in G (g \cdot x = y).$$

All of the Borel spaces we consider will be quotients of standard Borel spaces  $X$  by orbit equivalence relations  $E = \mathbb{E}_G^X \subseteq X^2$  of Borel actions of Polish groups  $G \curvearrowright X$ . Recall that the quotient Borel structure  $\mathcal{B}(X/E)$  consists of the subsets whose lift in  $X$  is  $E$ -invariant Borel.

An important fact about maps between such quotient spaces is

**Proposition 2.1** (see [Gao09, 5.4.6, 5.2.3, 5.4.12], [Che19b, 2.8]). *Let  $X/E, Y/F$  be two quotient spaces of Polish group actions, and  $f : X \rightarrow Y$  be a Borel map which descends to an injection  $X/E \hookrightarrow Y/F$  (a **Borel reduction**).*

- (a) *Then  $f$  descends to a Borel embedding  $X/E \hookrightarrow Y/F$ . In other words, each  $E$ -invariant Borel set in  $X$  is the preimage of an  $F$ -invariant Borel set in  $Y$ .*
- (b) *If moreover  $F \subseteq Y^2$  is Borel, then the image of  $X/E$  in  $Y/F$  is Borel, i.e., the image saturation  $[f(X)]_F \subseteq Y$  is Borel. Moreover, the inverse map between the quotient spaces  $[f(X)]_F/F \cong X/E$  has a Borel lift  $[f(X)]_F \rightarrow X$ .*

An equivalence relation  $E \subseteq X^2$  on a standard Borel space  $X$  is **smooth** if  $X/E$  has countably many Borel sets which separate points, or equivalently,  $X/E$  is Borel isomorphic to an analytic set in a standard Borel space. A **Borel transversal** of  $E$  is a Borel subset  $D \subseteq X$  containing exactly one point from each  $E$ -class. If  $E$  is an analytic equivalence relation, and has a Borel transversal  $D$ , then the quotient space  $X/E$  is standard Borel, being Borel isomorphic to  $D$ . If  $X/E$  is standard Borel, then clearly  $E$  is smooth. Neither of these two implications reverses in general; however, they do reverse if  $E$  is induced by a Polish group action, as a consequence of Proposition 2.1.

Another special fact about orbit equivalence relations of Polish group actions  $E \subseteq X^2$  is that each orbit is Borel, i.e., each point in  $X/E$  is Borel; see [Kec95, 15.14], [Gao09, 3.3.2].

A **countable Borel equivalence relation (CBER)** is a Borel equivalence relation  $E \subseteq X^2$  all of whose classes  $[x]_E \subseteq X$  are countable. Equivalently, they are precisely orbit equivalence relations of Borel actions of countable groups, by the **Feldman–Moore theorem** [FM77, Th. 1] (see also [Gao09, 7.1.4]). See [Kec24] for a comprehensive survey of the theory of CBERS.

The fundamental tool underlying the theory of CBERS is

**Theorem 2.2** (Lusin–Novikov; see [Kec95, 18.10]). *Let  $f : X \rightarrow Y$  be a countable-to-1 Borel map between standard Borel spaces. Then  $X = \bigcup_i X_i$  for countably many Borel sets  $X_i \subseteq X$  such that  $f$  restricts to a bijection  $f : X_i \cong f(X)$ .*

**Corollary 2.3.** *Let  $E \subseteq X^2$  be a CBERS on a standard Borel space  $X$ .*

- (a) *There are countably many Borel functions  $f_i : X \rightarrow X$  such that  $E = \bigcup_i f_i$ .*
- (b) *For each  $N \leq \omega$ ,  $X_N := \{x \in X \mid |[x]_E| = N\}$  is Borel; and there is a Borel family of enumerations  $(g_x : |[x]_E| \cong [x]_E)$ , meaning that  $(x, i) \mapsto g_x(i) : \bigsqcup_{N \leq \omega} (X_N \times N) \rightarrow X$  is Borel.*

Here (a) follows from Lusin–Novikov applied to the first projection  $E \rightarrow X$ , while (b) is an easy modification; see [Kec95, 18.15]. The Lusin–Novikov theorem, especially (a) above, will also play a central role in our analysis of Borel structurability in this paper; see Section 3.C.

## 2.B Countable first-order logic $\mathcal{L}_{\omega_1\omega}$

Let  $\mathcal{L}$  be a first-order language. The infinitary logic  $\mathcal{L}_{\omega_1\omega}$  is the extension of the usual finitary first-order logic by allowing countable conjunctions  $\bigwedge$  and disjunctions  $\bigvee$  in formulas (in addition to the usual connectives  $\rightarrow, \neg$ , quantifiers  $\forall, \exists$ , and equality  $=$ ). See [Gao09, Ch. 11], [Mar16].

Throughout, when we say “formula”, “theory”, etc., we mean in  $\mathcal{L}_{\omega_1\omega}$ . By convention, we require each  $\mathcal{L}_{\omega_1\omega}$  formula  $\phi(x_0, \dots, x_{n-1})$  to contain only finitely many free variables  $x_0, \dots, x_{n-1}$ , and we identify formulas up to renaming of bound variables.

There is a completeness theorem for  $\mathcal{L}_{\omega_1\omega}$  (for various deductive systems [Kar64], [LE65]), which say that an  $\mathcal{L}_{\omega_1\omega}$  sentence has a proof from a *countable* theory iff it is true in every *countable* model. Thus, we may interchangeably speak of provable or semantic truth.

Unless otherwise specified, by default we assume languages  $\mathcal{L}$ , theories  $\mathcal{T}$ , and models  $\mathcal{M}$  to be countable. Note that we might as well then replace theories  $\mathcal{T}$  with single sentences  $\bigwedge \mathcal{T}$ ; however, we find it more convenient for expositional purposes to allow theories with multiple axioms.

**Convention 2.4.** We will assume that formally, all languages are relational. In examples, we will often want to consider languages with function symbols; formally, we understand this to mean that each  $n$ -ary function symbol  $f$  is replaced with an  $(n+1)$ -ary relation symbol  $F$ , with axioms added to the theory under consideration that say that  $F$  is the graph of a function. Thus for example, a constant (nullary function)  $c$  is replaced with a unary relation  $C$  subject to the axiom  $\exists!x C(x)$ .

**Notation 2.5.** For a language  $\mathcal{L}$ , we write  $\mathcal{L}_{\omega_1\omega}$  for the set of all  $\mathcal{L}_{\omega_1\omega}$  formulas. Occasionally, we will also write  $\mathcal{L}_{\omega_1\omega}^n$  for the set of formulas  $\phi(x_0, \dots, x_{n-1})$  with  $n$  free variables, and similarly  $\mathcal{L}^n$  for the set of  $n$ -ary relation symbols in  $\mathcal{L}$ .

## 2.C Spaces of countable models

**Definition 2.6.** Fix a countable language  $\mathcal{L}$ , assumed to be relational as per Convention 2.4. For a countable set  $Y$ , the **space of  $\mathcal{L}$ -structures on  $Y$**  is the product standard Borel space

$$\text{Mod}_Y(\mathcal{L}) := \prod_{\substack{n \in \mathbb{N} \\ R \in \mathcal{L}^n}} 2^{Y^n} = \{\mathcal{M} = (R^{\mathcal{M}})_{R \in \mathcal{L}} \mid \forall R \in \mathcal{L}^n (R^{\mathcal{M}} \subseteq Y^n)\}.$$

More generally, for an  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ ,

$$\text{Mod}_Y(\mathcal{T}) = \text{Mod}_Y(\mathcal{L}, \mathcal{T}) := \{\mathcal{M} \in \text{Mod}_Y(\mathcal{L}) \mid \mathcal{M} \models \mathcal{T}\}.$$

By an easy induction, for each  $\mathcal{L}_{\omega_1\omega}$  formula  $\phi(x_0, \dots, x_{n-1})$ ,  $\text{Mod}_Y^n(\phi) \subseteq \text{Mod}_Y^n(\mathcal{L})$  is Borel. Thus, for a countable theory  $\mathcal{T}$ ,  $\text{Mod}_Y(\mathcal{T}) \subseteq \text{Mod}_Y(\mathcal{L})$  is Borel. See [Kec95, §16.C], [Gao09, Ch. 11].

Note that the space of models of  $\mathcal{T}$  expanded with  $n$  constants is then

$$\text{Mod}_Y^n(\mathcal{T}) := \text{Mod}_Y(\mathcal{L} \sqcup \{x_0, \dots, x_{n-1}\}, \mathcal{T}) \cong \text{Mod}_Y(\mathcal{T}) \times Y^n = \{(\mathcal{M}, \vec{a}) \mid \mathcal{M} \in \text{Mod}_Y(\mathcal{T}), \vec{a} \in Y^n\}.$$

We call  $(\mathcal{M}, \vec{a}) \in \text{Mod}_Y^n(\mathcal{T})$  an  **$n$ -pointed model**, or a **pointed model** for  $n = 1$ .

**Definition 2.7.** For two sets  $Y, Z$ , let  $\text{Sym}(Y, Z) \subseteq Z^Y$  denote the set of all bijections  $g : Y \cong Z$ ; when  $Y = Z$ , this is the **symmetric group**  $\text{Sym}(Y) := \text{Sym}(Y, Y)$  (often denoted  $S_Y$ ). For  $Y, Z$  countable,  $\text{Sym}(Y, Z)$  is a  $G_\delta$  subset of  $Z^Y$  (with  $Z$  discrete); hence  $\text{Sym}(Y)$  is a Polish group.

The **logic action** on the space of structures is given by, for any  $Y, Z$ :<sup>2</sup>

$$\begin{aligned} \text{Sym}(Y, Z) \times \text{Mod}_Y(\mathcal{L}) &\longrightarrow \text{Mod}_Z(\mathcal{L}) \\ (g : Y \cong Z, \mathcal{M} = (R^{\mathcal{M}})_{R \in \mathcal{L}}) &\longmapsto g \cdot \mathcal{M} := (g(R^{\mathcal{M}}))_{R \in \mathcal{L}}. \end{aligned}$$

That is, for a structure  $\mathcal{M}$  on  $Y$ ,  $g \cdot \mathcal{M}$  is the unique structure on  $Z$  such that  $g : \mathcal{M} \cong g \cdot \mathcal{M}$ .

Thus, a class  $\mathcal{K}$  of countable structures, on various underlying sets  $Y$ , is invariant under the logic action (between those  $\text{Mod}_Y(\mathcal{L})$ 's) iff it is closed under isomorphisms.

Clearly, for any theory  $\mathcal{T}$ ,  $\text{Mod}_Y(\mathcal{T}) \subseteq \text{Mod}_Y(\mathcal{L})$  is closed under isomorphisms. Conversely:

**Theorem 2.8** (Lopez-Escobar [LE65]; see also [Kec95, 16.8], [Gao09, 11.3.6]). *For any countable set  $Y$ , a subset  $\mathcal{K} \subseteq \text{Mod}_Y(\mathcal{L})$  is axiomatizable by an  $\mathcal{L}_{\omega_1\omega}$  sentence (equivalently, a countable theory) iff it is Borel and closed under isomorphisms.*

A special case is **Scott's isomorphism theorem**: the isomorphism class of a single countable structure, being Borel (as an orbit of a Polish group action), is axiomatized by a sentence.

We also note the following easy generalized formulations of Lopez-Escobar's theorem:

**Corollary 2.9.** *For any countable set  $Y$  and  $n \in \mathbb{N}$ , a subset  $\mathcal{K} \subseteq \text{Mod}_Y^n(\mathcal{L})$  is axiomatizable by a formula  $\phi(x_0, \dots, x_{n-1})$  iff it is Borel and closed under isomorphisms (of  $n$ -pointed models).*

**Corollary 2.10.** *A class  $\mathcal{K}$  of countable models on all countable underlying sets  $Y$  is axiomatizable by a sentence (or countable theory) iff it is closed under isomorphisms and  $\mathcal{K} \cap \text{Mod}_Y(\mathcal{L}) \subseteq \text{Mod}_Y(\mathcal{L})$  is Borel for each countable set  $Y$ . Similarly for a class  $\mathcal{K}$  of  $n$ -pointed models.*

This follows from applying Theorem 2.8 to  $Y = 0, 1, 2, \dots, \mathbb{N}$  to get sentences  $\phi_Y$ , and then combining them into a sentence  $\bigvee_{Y \leq \omega}$  ("there are exactly  $Y$  many elements"  $\wedge \phi_Y$ ).

## 2.D Spaces of $\mathcal{L}_{\omega_1\omega}$ types

**Definition 2.11.** Let  $(\mathcal{L}, \mathcal{T})$  be a countable theory, and let  $n \in \mathbb{N}$ . For  $\mathcal{M} \models \mathcal{T}$  and  $\vec{a} \in M^n$ , the  $(\mathcal{L}_{\omega_1\omega})$  **type** of  $\vec{a}$  will mean the complete  $\mathcal{L}_{\omega_1\omega}$  theory of the  $n$ -pointed structure<sup>3</sup>

$$\text{tp}(\mathcal{M}, \vec{a}) := \{ \phi \in \mathcal{L}_{\omega_1\omega}^n \mid \phi^{\mathcal{M}}(\vec{a}) \}.$$

The **Borel space of  $(\mathcal{L}_{\omega_1\omega})$   $n$ -types**  $\mathcal{S}_n(\mathcal{T})$  is the set of all  $\text{tp}(\mathcal{M}, \vec{a})$  over all  $n$ -tuples  $\vec{a}$  in all countable models  $\mathcal{M}$  of  $\mathcal{T}$ , equipped with the Borel  $\sigma$ -algebra

$$\mathcal{B}(\mathcal{S}_n(\mathcal{T})) := \{ \llbracket \phi \rrbracket \mid \phi \in \mathcal{L}_{\omega_1\omega}^n \}, \quad \text{where } \llbracket \phi \rrbracket := \{ p \in \mathcal{S}_n(\mathcal{T}) \mid \phi \in p \}.$$

Note that by the completeness theorem, we thus have an isomorphism of  $\sigma$ -algebras

$$\llbracket - \rrbracket : \mathcal{L}_{\omega_1\omega}^n / \mathcal{T} := \{ \text{formulas } \phi(x_0, \dots, x_{n-1}) \text{ mod } \mathcal{T}\text{-provable equivalence} \} \cong \mathcal{B}(\mathcal{S}_n(\mathcal{T}));$$

an alternate definition of  $\mathcal{S}_n(\mathcal{T})$  is the space of principal  $\sigma$ -ultrafilters in  $\mathcal{L}_{\omega_1\omega}^n / \mathcal{T}$ .<sup>4</sup>

<sup>2</sup>Formally, this is an action of the groupoid of countable sets, with hom-sets  $\text{Sym}(Y, Z)$ , on the bundle of  $\text{Mod}_Y(\mathcal{L})$ 's.

<sup>3</sup>We will never consider finitary first-order (i.e.,  $\mathcal{L}_{\omega\omega}$ ) types in this paper.

<sup>4</sup>In this paper, we only consider principal types realized in countable models, since they are all we will need. In general  $\mathcal{L}_{\omega_1\omega}$  theories, there is a more general notion of type, namely an arbitrary  $\sigma$ -ultrafilter of formulas, which is not necessarily realizable in any model due to the failure of the completeness theorem for uncountable theories.

**Corollary 2.12** (of Lopez-Escobar). *The maps  $\text{tp}$  induce an isomorphism of Borel spaces*

$$\text{tp} : (\bigsqcup_Y \text{Mod}_Y^n(\mathcal{T})) / (\cong) \cong \mathcal{S}_n(\mathcal{T}).$$

Here, the disjoint union may be taken over any representative family of sets  $Y$  of each countable cardinality, e.g., the ordinals  $Y \leq \omega$ . The left-hand side is then the disjoint union of the quotient Borel spaces  $\text{Mod}_Y^n(\mathcal{T}) / \text{Sym}(Y)$ , for  $Y \leq \omega$ , each of which is a quotient of a Polish group action.

Recall (Section 2.A) that such a quotient is standard Borel iff the  $\sigma$ -algebra of invariant Borel sets in  $\text{Mod}_Y^n(\mathcal{T})$  is countably generated. By Lopez-Escobar, this means that there are countably many formulas  $\phi(x_0, \dots, x_{n-1})$  separating all non-isomorphic  $n$ -pointed models.

**Definition 2.13.** For a set  $\mathcal{F} \subseteq \mathcal{L}_{\omega_1\omega}$  of formulas, an  $n$ -pointed model  $(\mathcal{M}, \vec{a})$  is  **$\mathcal{F}$ -categorical** if it is the only countable model of  $\text{tp}(\mathcal{M}, \vec{a}) \cap \mathcal{F}$ , up to isomorphism.

**Corollary 2.14** (of Lopez-Escobar; see [Gao09, 11.5.8]). *For any countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$  and  $n \in \mathbb{N}$ , the following are equivalent:*

- (i) *The space of  $n$ -types  $\mathcal{S}_n(\mathcal{T})$  is standard Borel.*
- (ii) *Isomorphism of  $n$ -pointed models of  $\mathcal{T}$  is a smooth equivalence relation.*
- (iii) *There are countably many formulas  $\mathcal{F} \subseteq \mathcal{L}_{\omega_1\omega}^n$  such that every  $n$ -pointed model is  $\mathcal{F}$ -categorical.*

**Theorem 2.15** (Becker–Kechris). *For a countable theory  $\mathcal{T}$ , isomorphism  $\cong \subseteq \text{Mod}_Y^n(\mathcal{T})^2$  of  $n$ -pointed models is a Borel equivalence relation iff isomorphism  $\cong \subseteq \text{Mod}_Y^{n+1}(\mathcal{T})^2$  of  $(n+1)$ -pointed models is.*

*Proof.* The forward direction is [BK96, 7.1.4]. For the converse, when  $n \geq 1$ , the logic action on  $\text{Mod}_Y(\mathcal{T}) \times Y^n$  diagonally embeds into the logic action on  $\text{Mod}_Y(\mathcal{T}) \times Y^{n+1}$  (by duplicating the last coordinate, say). When  $n = 0$ , use instead that

$$\mathcal{M} \cong \mathcal{N} \iff \exists a \in M, b \in N ((\mathcal{M}, a) \cong (\mathcal{N}, b)) \text{ or } (M = N = \emptyset \ \& \ \forall P \in \mathcal{L}^0 (P^{\mathcal{M}} \iff P^{\mathcal{N}})). \square$$

**Corollary 2.16.** *If  $\mathcal{S}_n(\mathcal{T})$  is standard Borel, then so is  $\mathcal{S}_{n+1}(\mathcal{T})$ .*

*Proof.* Suppose isomorphism  $\cong \subseteq \text{Mod}_Y^n(\mathcal{T})^2$  of  $n$ -pointed models is smooth; we show that isomorphism  $\cong \subseteq \text{Mod}_Y^{n+1}(\mathcal{T})^2$  of  $(n+1)$ -pointed models is smooth. We may assume  $Y \leq \omega$ . Let  $D \subseteq \text{Mod}_Y^n(\mathcal{T})$  be a Borel transversal of  $\cong \subseteq \text{Mod}_Y^n(\mathcal{T})^2$ . Then

$$\{(\mathcal{M}, a_0, \dots, a_n) \mid (\mathcal{M}, a_0, \dots, a_{n-1}) \in D \ \& \ \forall b_n < a_n ((\mathcal{M}, a_0, \dots, a_{n-1}, b_n) \not\cong (\mathcal{M}, a_0, \dots, a_n))\}$$

is a Borel transversal of  $\cong \subseteq \text{Mod}_Y^{n+1}(\mathcal{T})^2$ .  $\square$

**Remark 2.17.** The Borel spaces  $(\mathcal{S}_n(\mathcal{T}))_{n < \omega}$  are related by Borel **projection maps**

$$\dots \xrightarrow{\partial_3} \mathcal{S}_2(\mathcal{T}) \xrightarrow{\partial_2} \mathcal{S}_1(\mathcal{T}) \xrightarrow{\partial_1} \mathcal{S}_0(\mathcal{T})$$

where, for an  $n$ -type  $p \in \mathcal{S}_n(\mathcal{T})$ ,  $\partial_n(p)$  consists of the formulas  $\phi(x_0, \dots, x_{n-2})$  in  $n-1$  variables which belong to  $p$  when regarded as having an extra variable  $x_{n-1}$ ; thus

$$(2.18) \quad \partial_n^{-1}(\llbracket \phi(x_0, \dots, x_{n-2}) \rrbracket) = \llbracket \phi(x_0, \dots, x_{n-1}) \rrbracket.$$

In other words, via Corollary 2.12,  $\partial_n$  descends from the coordinate projection

$$\pi_n : \bigsqcup_Y \text{Mod}_Y^n(\mathcal{T}) \longrightarrow \bigsqcup_Y \text{Mod}_Y^{n-1}(\mathcal{T})$$

omitting the last coordinate, which is clearly countable-to-1 and maps (invariant) Borel sets to (invariant) Borel sets; thus  $\partial_n$  is also countable-to-1 and maps Borel sets onto Borel sets. Namely,

$$(2.19) \quad \partial_n(\llbracket \phi(x_0, \dots, x_{n-1}) \rrbracket) = \llbracket \exists x_{n-1} \phi(x_0, \dots, x_{n-1}) \rrbracket.$$

In fact there are other canonical “projection maps” between the  $\mathcal{S}_n(\mathcal{T})$ ’s. Instead of regarding an  $(n-1)$ -ary formula as  $n$ -ary via the inclusion of variables  $n-1 \hookrightarrow n$ , we may consider a variable substitution from  $m$  variables to  $n$  variables given by an arbitrary map  $s : m \rightarrow n$ , which induces

$$\begin{aligned} \partial_s : \mathcal{S}_n(\mathcal{T}) &\longrightarrow \mathcal{S}_m(\mathcal{T}) \\ p &\longmapsto \{ \phi(x_0, \dots, x_{m-1}) \mid \phi(x_{s(0)}, \dots, x_{s(m-1)}) \in p \} \end{aligned}$$

with

$$\partial_s^{-1}(\llbracket \phi(x_0, \dots, x_{m-1}) \rrbracket) = \llbracket \phi(x_{s(0)}, \dots, x_{s(m-1)}) \rrbracket.$$

This descends from the corresponding map  $\pi_s : \bigsqcup_Y \text{Mod}_Y^n(\mathcal{T}) \rightarrow \bigsqcup_Y \text{Mod}_Y^m(\mathcal{T})$ , and is again countable-to-1 and maps Borel sets to Borel sets. When  $s$  is the inclusion  $n-1 \hookrightarrow n$ ,  $\partial_s$  is the coordinate projection  $\partial_n$  from above. When  $s : n \twoheadrightarrow n-1$  is instead the surjection collapsing  $n-1 \mapsto n-2$  (and fixing other elements), then  $\partial_s$  is the diagonal embedding duplicating the last two coordinates, and image under  $\partial_s$  is given by

$$(2.20) \quad \partial_s(\llbracket \phi(x_0, \dots, x_{n-2}) \rrbracket) = \llbracket \phi(x_0, \dots, x_{n-2}) \wedge (x_{n-2} = x_{n-1}) \rrbracket.$$

For a general  $s : m \rightarrow n$  between finite ordinals  $m, n < \omega$ , we may write  $s$  as a composition of such inclusions and surjections as well as bijective permutations; thus  $\partial_s$  may be understood as a combination of the above cases, using the obvious fact

$$\partial_{s \circ t} = \partial_t \circ \partial_s \quad (\text{also } \partial_{\text{id}} = \text{id}).$$

(Category-theoretically, we get that  $n \mapsto \mathcal{S}_n(\mathcal{T})$  forms a contravariant functor from the category of finite ordinals to the category of Borel spaces and countable-to-1 Borel maps with Borel images. Such a functor is sometimes called an *augmented symmetric simplicial set* in relation to homotopy theory; we will make use of this connection in Section 5 below. The family  $(\mathcal{S}_n(\mathcal{T}))_n$  is also dual to the Lindenbaum–Tarski hyperdoctrine from [CK18, §B.4]; see Remarks 2.23 and 2.25.)

## 2.E Interpretations in $\mathcal{L}_{\omega_1\omega}$

**Definition 2.21.** Let  $\mathcal{L}, \mathcal{L}'$  be languages, assumed to be relational as per Convention 2.4.

An **interpretation**  $\alpha : \mathcal{L} \rightarrow \mathcal{L}'$  is a map taking each  $n$ -ary relation  $R \in \mathcal{L}$  to an  $n$ -ary formula  $\alpha(R)(x_0, \dots, x_{n-1}) \in \mathcal{L}'_{\omega_1\omega}$ .<sup>1</sup> Such  $\alpha$  provides a syntactic recipe for constructing  $\mathcal{L}$ -structures from  $\mathcal{L}'$ -structures: given an  $\mathcal{L}'$ -structure  $\mathcal{M}$ , its  **$\alpha$ -reduct** is the  $\mathcal{L}$ -structure  $\alpha^*(\mathcal{M})$  given by

$$R^{\alpha^*(\mathcal{M})} := \alpha(R)^{\mathcal{M}}$$

for each relation symbol  $R \in \mathcal{L}$ . (The name “reduct” comes from the case when  $\alpha$  is the inclusion of a sublanguage  $\mathcal{L} \subseteq \mathcal{L}'$ .) When the underlying set  $Y$  of  $\mathcal{M}$  is countable, this yields a map

$$\alpha^* = \alpha_Y^* : \text{Mod}_Y(\mathcal{L}') \longrightarrow \text{Mod}_Y(\mathcal{L})$$

which is easily seen to be Borel and equivariant under the logic action:

$$\alpha_Z^*(g \cdot \mathcal{M}) = g \cdot \alpha_Y^*(\mathcal{M})$$

for any  $\mathcal{M} \in \text{Mod}_Y(\mathcal{L}')$  and bijection  $g \in \text{Sym}(Y, Z)$ .

Given  $\alpha : \mathcal{L} \rightarrow \mathcal{L}'$ , we extend  $\alpha$  to all  $\mathcal{L}$ -formulas inductively in the obvious way:

$$\begin{aligned} \alpha(R(x_0, \dots, x_{n-1})) &:= \alpha(R)(x_0, \dots, x_{n-1}), \\ \alpha(\bigvee_i \phi_i) &:= \bigvee_i \alpha(\phi_i), \\ \alpha(\exists x \phi) &:= \exists x \alpha(\phi), \end{aligned}$$

etc. Semantically, this means that  $\alpha(\phi)$  is an  $\mathcal{L}'$ -formula such that in every  $\mathcal{L}'$ -structure  $\mathcal{M}$ ,

$$\alpha(\phi)^{\mathcal{M}} = \phi^{\alpha^*(\mathcal{M})}.$$

Now given an  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$  and  $\mathcal{L}'_{\omega_1\omega}$  theory  $\mathcal{T}'$ , we say that  $\alpha$  is an **interpretation of  $\mathcal{T}$  in  $\mathcal{T}'$** , written  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  or  $\alpha : (\mathcal{L}, \mathcal{T}) \rightarrow (\mathcal{L}', \mathcal{T}')$ , if it preserves provable truth:

$$\mathcal{T}' \vdash \alpha(\mathcal{T}).$$

For  $\mathcal{T}'$  countable, by the completeness theorem this means equivalently that  $\alpha_Y^*$  restricts to a map

$$(2.22) \quad \alpha_Y^* : \text{Mod}_Y(\mathcal{T}') \longrightarrow \text{Mod}_Y(\mathcal{T})$$

for every countable set  $Y$ , which is again Borel and equivariant under the logic action.

Two interpretations  $\alpha, \beta : \mathcal{T} \rightarrow \mathcal{T}'$  are  **$\mathcal{T}'$ -provably equivalent** if  $\mathcal{T}' \vdash \alpha(R) \leftrightarrow \beta(R)$  for every symbol  $R \in \mathcal{L}$ , or equivalently (if  $\mathcal{T}'$  is countable)  $\alpha^* = \beta^*$ . We normally identify interpretations up to provable equivalence; that is, an “interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$ ” will really mean a  $\mathcal{T}'$ -provable equivalence class of interpretations.

If there exists an interpretation  $\mathcal{T} \rightarrow \mathcal{T}'$ , we say that  $\mathcal{T}$  is **interpretable** in  $\mathcal{T}'$ , or that  $\mathcal{T}'$  **interprets  $\mathcal{T}$** . There is an obvious composition of interpretations  $\mathcal{T} \rightarrow \mathcal{T}' \rightarrow \mathcal{T}''$ ; thus interpretability is a preorder on the class of all theories, while interpretations themselves form a category. If an interpretation has an inverse, we call it a **bi-interpretation**; if there is a bi-interpretation between  $\mathcal{T}, \mathcal{T}'$ , we call them **bi-interpretable**, written  $\mathcal{T} \cong \mathcal{T}'$ . Note that this is stronger than the existence of interpretations both ways, which we call **mutually interpretable**.

See the introduction (Example 1.2, (1.5) and (1.6)) for examples of interpretations  $\mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LO}}$  and  $\mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{sep}}$ . In the rest of this subsection, we give some equivalent reformulations of the definition of interpretation, which are more abstract, but easier to work with when developing the general theory.

**Remark 2.23.** Conceptually, an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  may be viewed as a family of maps

$$(\alpha_n : \mathcal{L}_{\omega_1\omega}^n / \mathcal{T} \longrightarrow \mathcal{L}'_{\omega_1\omega} / \mathcal{T}')_{n < \omega},$$

where  $\mathcal{L}_{\omega_1\omega}^n/\mathcal{T}$  denotes  $\mathcal{T}$ -provable equivalence classes of  $n$ -ary  $\mathcal{L}_{\omega_1\omega}$  formulas (Definition 2.11), which commutes with variable substitutions and the logical operations:

$$\begin{aligned}\alpha_n([\phi(x_{s(0)}, \dots, x_{s(m-1)})]) &= [\alpha_m(\phi)(x_{s(0)}, \dots, x_{s(m-1)})] \quad \text{for } s : m \rightarrow n, \\ \alpha_2([x_0 = x_1]) &= [x_0 = x_1], \\ \alpha_n([\bigvee_i \phi_i]) &= \bigvee_i \alpha_n([\phi_i]), \\ \alpha_n([\exists x_n \phi(x_0, \dots, x_n)]) &= \exists x_n \alpha_{n+1}([\phi(x_0, \dots, x_n)]),\end{aligned}$$

etc. This ensures that  $\alpha$  is determined by its values on atomic formulas, i.e., the symbols in  $\mathcal{L}$ .

In other words,  $\alpha$  is a homomorphism of multi-sorted structures  $(\mathcal{L}_{\omega_1\omega}^n/\mathcal{T})_{n<\omega} \rightarrow (\mathcal{L}'_{\omega_1\omega}/\mathcal{T}')_{n<\omega}$ , consisting of sequences of Boolean  $\sigma$ -algebras which are furthermore equipped with operations between the various algebras corresponding to variable substitution (for each  $s : m \rightarrow n$ ),  $\exists$  (from the  $(n+1)$ th algebra to the  $n$ th) and a constant “ $x_0 = x_1$ ” (in the 2nd algebra). These structures are the **Lindenbaum–Tarski hyperdoctrines** of the theories, called “ $\omega_1\omega$ -Boolean algebras” in [CK18, §B.4]; see there as well as [Jac99] for more information on hyperdoctrines.

**Example 2.24.** For a theory  $(\mathcal{L}, \mathcal{T})$ , a countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$  may be regarded as a special case of an interpretation. Indeed, note that by the usual definition of the *interpretation* of first-order logic in a structure,  $\mathcal{M}$  amounts to a map from formulas to relations

$$\begin{aligned}\alpha : \mathcal{L}_{\omega_1\omega}^n &\longrightarrow 2^{Y^n} \\ \phi &\longmapsto \phi^{\mathcal{M}},\end{aligned}$$

preserving countable Boolean connectives, variable substitution, quantifiers and  $=$ , and  $\mathcal{T}$ -truth.

Now the Boolean  $\sigma$ -algebra  $2^{Y^n}$  is canonically isomorphic to the algebra of  $n$ -ary formulas modulo equivalence for the theory  $\mathcal{T}_Y$  of  **$Y$ -enumerated sets**, with language  $\mathcal{L}_Y$  consisting of constant symbols  $c_y$  for each  $y \in Y$  and axioms saying that these form a bijection with  $Y$ :

$$\mathcal{T}_Y := \{ \bigwedge_{y \neq z \in Y} (c_y \neq c_z), \forall x \bigvee_{y \in Y} (x = c_y) \}.$$

Indeed, every  $A \subseteq Y^n$  gives a formula  $\phi_A(x_0, \dots, x_{n-1}) := \bigvee_{\vec{y} \in A} (\vec{x} = c_{\vec{y}})$ ; and it is easy to see that  $\mathcal{T}_Y$  has quantifier elimination, using the above axioms to rewrite  $\exists x \phi(x)$  to  $\bigvee_{y \in Y} \phi(c_y)$ , whence every  $n$ -ary formula is equivalent modulo  $\mathcal{T}_Y$  to a quantifier-free formula, hence to some  $\phi_A$  since the above axioms also easily yield that the formulas  $\vec{x} = c_{\vec{y}}$  are the atoms of the  $\sigma$ -algebra. It is also easy to see that these isomorphisms  $2^{Y^n} \cong (\mathcal{L}_Y)_{\omega_1\omega}^n/\mathcal{T}_Y$  preserve substitution, quantifiers, and  $=$ .

Thus  $\alpha$ , i.e., the model  $\mathcal{M}$ , may be regarded as an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}_Y$ . Semantically,  $\alpha^*$  turns a model  $\mathcal{N} \in \text{Mod}_Z(\mathcal{T}_Y)$ , i.e., a bijection  $g_{\mathcal{N}} : y \mapsto c_y^{\mathcal{N}} : Y \cong Z$ , into the model  $\alpha^*(\mathcal{N}) = g_{\mathcal{N}} \cdot \mathcal{M} \in \text{Mod}_Z(\mathcal{T})$ .

**Remark 2.25.** The Lindenbaum–Tarski hyperdoctrine of a theory  $(\mathcal{L}, \mathcal{T})$  is dual to the sequence of type spaces  $(\mathcal{S}_n(\mathcal{T}))_{n<\omega}$ , with the canonical projection maps  $\partial_s$  (induced by arbitrary substitutions  $s$ ) between them, from Remark 2.17. Indeed, the algebra  $\mathcal{L}_{\omega_1\omega}^n/\mathcal{T}$  of  $\mathcal{T}$ -equivalence classes of formulas is isomorphic to the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathcal{S}_n(\mathcal{T}))$  by Definition 2.11; and by (2.18), (2.19) and (2.20), variable substitution,  $\exists$  and  $=$  correspond to preimage and image under the  $\partial_s$ .

Combined with Remark 2.23, this yields several equivalent ways to describe the concept of an interpretation  $\alpha : (\mathcal{L}, \mathcal{T}) \rightarrow (\mathcal{L}', \mathcal{T}')$  between theories:

- (i) maps  $\alpha : \mathcal{L} \rightarrow \mathcal{L}'_{\omega_1\omega}$  such that  $\mathcal{T}' \vdash \alpha(\mathcal{T})$ , modulo  $\mathcal{T}'$ -provable equivalence;



- (ii) homomorphisms of hyperdoctrines  $(\alpha_n : \mathcal{L}_{\omega_1\omega}^n/\mathcal{T} \rightarrow \mathcal{L}_{\omega_1\omega}^m/\mathcal{T}')_{n < \omega}$ ;
- (iii) families of Borel maps  $(\alpha_n^* : \mathcal{S}_n(\mathcal{T}') \rightarrow \mathcal{S}_n(\mathcal{T}))_{n < \omega}$ , commuting with projections  $\partial_s : \mathcal{S}_n \rightarrow \mathcal{S}_m$ :

$$\alpha_m^* \circ \partial_s = \partial_s \circ \alpha_n^* : \mathcal{S}_n(\mathcal{T}') \rightarrow \mathcal{S}_m(\mathcal{T}),$$

and also obeying

$$(\alpha_m^*)^{-1} \circ \partial_s = \partial_s \circ (\alpha_n^*)^{-1} : \mathcal{B}(\mathcal{S}_n(\mathcal{T})) \rightarrow \mathcal{B}(\mathcal{S}_m(\mathcal{T}'))$$

for each  $s : m \rightarrow n$  (the maps here are image/preimage in the following commutative square);

$$\begin{array}{ccc} \mathcal{S}_n(\mathcal{T}') & \xrightarrow{\partial_s} & \mathcal{S}_m(\mathcal{T}') \\ \alpha_n^* \downarrow & & \downarrow \alpha_m^* \\ \mathcal{S}_n(\mathcal{T}) & \xrightarrow{\partial_s} & \mathcal{S}_m(\mathcal{T}) \end{array}$$

- (iv) families of Borel maps  $(\alpha_Y^* : \text{Mod}_Y(\mathcal{T}') \rightarrow \text{Mod}_Y(\mathcal{T}))_Y$  for all countable sets  $Y$ , equivariant under the logic action (of  $\text{Sym}(Y, Z)$  for all countable  $Y, Z$ ).

The correspondence between (i) and (ii) was described in Remark 2.23. The maps in (ii) are, via  $\mathcal{L}_{\omega_1\omega}^n/\mathcal{T} \cong \mathcal{B}(\mathcal{S}_n(\mathcal{T}))$ , preimage under the maps in (iii); the conditions in (iii) ensure that preimage under those maps preserves substitution,  $\exists$ , and  $=$ . The maps in (iii) are obtained from (iv) by passing to  $\alpha_Y^* \times \text{id}_{Y^n} : \text{Mod}_Y^n(\mathcal{T}') = \text{Mod}_Y(\mathcal{T}') \times Y^n \rightarrow \text{Mod}_Y(\mathcal{T}) \times Y^n = \text{Mod}_Y^n(\mathcal{T})$  and then quotienting by the logic action, i.e., passing to types.

*Proof that these correspondences are inverses of each other.* Starting from  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  as in (i) and (ii), we indeed recover  $\alpha$  up to  $\mathcal{T}'$ -provable equivalence as  $(\alpha_n^*)^{-1}$  by the completeness theorem; thus the composite (i)  $\rightarrow$  (iv)  $\rightarrow$  (iii)  $\rightarrow$  (ii)  $\cong$  (i) is the identity. The passage from (iii) to (ii)  $\cong$  (i) is also injective, i.e., a Borel map  $\mathcal{S}_n(\mathcal{T}') \rightarrow \mathcal{S}_n(\mathcal{T})$  is determined by its preimage map between the Borel  $\sigma$ -algebras, because singletons in  $\mathcal{S}_n(\mathcal{T})$  are Borel by the Scott isomorphism theorem. Finally, (iv)  $\rightarrow$  (iii) is injective: given an arbitrary equivariant family  $(f_Y : \text{Mod}_Y(\mathcal{T}') \rightarrow \text{Mod}_Y(\mathcal{T}))_Y$ , for each  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T}')$ , the  $\mathcal{L}$ -structure  $f_Y(\mathcal{M})$  is determined by the types  $\text{tp}(f_Y(\mathcal{M}), \vec{a}) = \text{tp}((f_Y \times \text{id}_{Y^n})(\mathcal{M}, \vec{a}))$  of all finite tuples  $\vec{a} \in Y^n$ .  $\square$

Remark 2.25 allows us to work with  $\mathcal{L}_{\omega_1\omega}$  theories in a syntax-independent way: a theory  $\mathcal{T}$  is completely determined, up to bi-interpretability, (ii) by the hyperdoctrine  $(\mathcal{L}_{\omega_1\omega}^n/\mathcal{T})_n$ , which is an algebraic structure, or (iii) by the family of type spaces  $(\mathcal{S}_n(\mathcal{T}))_n$  and projections  $\partial_s$  between them, or (iv) by the family of spaces of models  $(\text{Mod}_Y(\mathcal{T}))_Y$  equipped with the logic action. Among these, (iv) has the benefit of living entirely in the familiar category of standard Borel spaces. This makes it possible to not only recover but even abstractly define a theory via its spaces of models:

**Theorem 2.26** (Lopez-Escobar, Becker–Kechris). *We have a dual equivalence of categories from:*

- countable  $\mathcal{L}_{\omega_1\omega}$  theories  $(\mathcal{L}, \mathcal{T})$ , and interpretations between them as morphisms; to
- families of standard Borel spaces  $M_Y$  over all countable sets  $Y$ , equipped with Borel actions  $\text{Sym}(Y, Z) \times M_Y \rightarrow M_Z$ , and equivariant families of Borel maps as morphisms;<sup>5</sup>

taking a theory  $\mathcal{T}$  to the family of spaces  $\text{Mod}_Y(\mathcal{T})$  and an interpretation  $\alpha$  to  $\alpha^*$ .

<sup>5</sup>Formally, these are presheaves on the groupoid of sets, enriched in the category of standard Borel spaces.

*Proof.* Said functor from the former category to the latter is full and faithful by the preceding remark (which boils down to Lopez-Escobar in the form of Corollary 2.12). It remains to show it is essentially surjective, i.e., given such a family of spaces  $(M_Y)_Y$  equipped with Borel actions of  $\text{Sym}(Y, Z)$ , to show that they are equivariantly isomorphic to  $(\text{Mod}_Y(\mathcal{T}))_Y$  for some theory  $\mathcal{T}$ .

By [BK96, 2.7.3], there is a language  $\mathcal{L}_{\mathbb{N}}$  such that  $M_{\mathbb{N}}$  Borel  $\text{Sym}(\mathbb{N})$ -equivariantly embeds into  $\text{Mod}_{\mathbb{N}}(\mathcal{L}_{\mathbb{N}})$ , hence by Lopez-Escobar there is a  $\mathcal{L}_{\mathbb{N}}$ -sentence  $\phi_{\mathbb{N}}$  such that  $\text{Mod}_{\mathbb{N}}(\phi_{\mathbb{N}}) \cong M_{\mathbb{N}}$  as  $\text{Sym}(\mathbb{N})$ -spaces. It is also easy to find languages  $\mathcal{L}_N$  such that  $M_N$   $\text{Sym}(N)$ -equivariantly embeds into  $\text{Mod}_N(\mathcal{L}_N)$  for each finite  $N < \omega$ , hence again by Lopez-Escobar there are sentences  $\phi_N$  such that  $\text{Mod}_N(\phi_N) \cong M_N$ . For example, for each subgroup  $H \leq \text{Sym}(N)$ , we may create a structure  $\mathcal{M}_H$  in a language  $\mathcal{L}_H$  on the set  $N$  such that the  $H$ -orbits in each  $N^n$  become definable, so that  $\text{Aut}(\mathcal{M}_H) = H$ ; then letting  $\mathcal{L}_N := \bigsqcup_{H \leq \text{Sym}(N)} \mathcal{L}_H$  together with countably many nullary relations,  $\text{Mod}_N(\mathcal{L}_N)$  has  $2^{\aleph_0}$  many points with each possible stabilizer  $H \leq \text{Sym}(N)$ ; now by partitioning the  $\text{Sym}(N)$ -orbits of  $M_N$  according to stabilizers, we may easily embed it into  $\text{Mod}_N(\mathcal{L}_N)$ .

So we have found Borel  $\text{Sym}(N)$ -equivariant isomorphisms  $\text{Mod}_N(\phi_N) \cong M_N$  for each  $N \leq \omega$ . By unioning the languages (and modifying  $\phi_N$  to assert the symbols from the other languages are false), we may assume the  $\phi_N$  are over the same language. By adding to  $\phi_N$  a clause “there exist exactly  $N$  elements”, we may assume each  $\phi_N$  has only models of size  $N$ . Then  $\phi := \bigvee_{N \leq \omega} \phi_N$  is such that  $\text{Mod}_N(\phi) \cong M_N$  for each  $N \leq \omega$ ; call these isomorphisms  $f_N$ . Then for any other countable set  $Y$ , define  $f_Y : \text{Mod}_Y(\phi) \cong M_Y$  by  $f_Y(\mathcal{M}) := g \cdot f_{|Y|}(g^{-1} \cdot \mathcal{M})$ , for any bijection  $g : |Y| \cong Y$  (where  $|Y| \leq \omega$  is the cardinality). Using that  $f_{|Y|}$  is  $\text{Sym}(|Y|)$ -equivariant, this is easily seen to be independent of  $g$  and to be equivariant with respect to the actions of all  $\text{Sym}(Y, Z)$ .  $\square$

## 2.F Operations on theories

We also recall here the following simple methods of combining theories; see [CK18, §B.2].

**Definition 2.27.** Given countably many theories  $(\mathcal{L}_i, \mathcal{T}_i)$ , their **coproduct theory** is the disjoint union  $(\bigsqcup_i \mathcal{L}_i, \bigsqcup_i \mathcal{T}_i)$ . Clearly, for any set  $Y$ , we have a canonical  $\text{Sym}(Y, Z)$ -equivariant isomorphism

$$\text{Mod}_Y(\bigsqcup_i \mathcal{T}_i) \cong \prod_i \text{Mod}_Y(\mathcal{T}_i);$$

thus,  $\text{Mod}_Y(\bigsqcup_i \mathcal{T}_i)$  has the universal property of the categorical product (in the category of families of standard Borel spaces with Borel  $\text{Sym}(Y, Z)$ -actions from Theorem 2.26). Dually, this means  $\bigsqcup_i \mathcal{T}_i$  has the universal property of the coproduct: an interpretation  $\bigsqcup_i \mathcal{T}_i \rightarrow \mathcal{T}'$  into another theory  $\mathcal{T}'$  is equivalently a family of interpretations  $\mathcal{T}_i \rightarrow \mathcal{T}'$  for each  $i$ .

In particular, this means that  $\bigsqcup_i \mathcal{T}_i$  is the join (least upper bound) of the  $\mathcal{T}_i$  in the preorder of theories under the interpretability relation  $\rightarrow$ .

**Definition 2.28.** Given countably many theories  $(\mathcal{L}_i, \mathcal{T}_i)$ , their **product theory**  $(\bigoplus_i \mathcal{L}_i, \bigoplus_i \mathcal{T}_i)$  has the dual universal property that an interpretation  $\mathcal{T}' \rightarrow \bigoplus_i \mathcal{T}_i$  is equivalently a family of interpretations  $\mathcal{T}' \rightarrow \mathcal{T}_i$  for each  $i$ . Thus, the space of models should be

$$\text{Mod}_Y(\bigoplus_i \mathcal{T}_i) \cong \bigsqcup_i \text{Mod}_Y(\mathcal{T}_i).$$

In other words, a model of  $\bigoplus_i \mathcal{T}_i$  should be a specification of a unique index  $i$  together with a model of  $\mathcal{T}_i$  (and no other data). By Theorem 2.26, this suffices to define  $\bigoplus_i \mathcal{T}_i$  up to bi-interpretability. For an explicit axiomatization of  $\bigoplus_i \mathcal{T}_i$ , see [CK18, 6.2].

In particular,  $\bigoplus_i \mathcal{T}_i$  is the meet (greatest lower bound) of the  $\mathcal{T}_i$  with respect to interpretability.

### 3 Structurability of CBERs

#### 3.A Structurability

**Definition 3.1** (see [JKL02], [CK18]). Let  $E \subseteq X^2$  be a CBER,  $\mathcal{T}$  be a countable  $\mathcal{L}_{\omega_1\omega}$  theory in a countable language  $\mathcal{L}$ . A **(Borel)  $\mathcal{T}$ -structuring** of  $E$  is a family of models  $\mathcal{M} = (\mathcal{M}_C)_{C \in X/E}$ , where for each  $E$ -class  $C \in X/E$ , we have  $\mathcal{M}_C \in \text{Mod}_C(\mathcal{T})$ ; and “ $x \mapsto \mathcal{M}_{[x]_E}$  is Borel”, meaning the following conditions which are equivalent by Lemma 3.2 below:

- (i) For some (equivalently any) Borel family of enumerations  $(g_x : |[x]_E| \cong [x]_E)_{x \in X}$  (as in Corollary 2.3(b)), the following map is Borel:

$$\begin{aligned} X &\longrightarrow \bigsqcup_{N \leq \omega} \text{Mod}_N(\mathcal{T}) \\ x &\longmapsto g_x^{-1} \cdot \mathcal{M}_{[x]_E}. \end{aligned}$$

- (ii) For any countable set  $Y$ , standard Borel space  $Z$ , Borel map  $f : Z \rightarrow X$ , and Borel family of bijections  $(g_z : Y \cong [f(z)]_E)_{z \in Z}$ , the following map is Borel:

$$\begin{aligned} Z &\longrightarrow \text{Mod}_Y(\mathcal{T}) \\ z &\longmapsto g_z^{-1} \cdot \mathcal{M}_{[f(z)]_E}. \end{aligned}$$

- (iii) For each  $n$ -ary relation symbol  $R \in \mathcal{L}$ , the following set is a Borel subset of  $X^{n+1}$ :

$$\widetilde{R}^{\mathcal{M}} := \{(x, x_0, \dots, x_{n-1}) \in X^{n+1} \mid x E x_0 E \cdots E x_{n-1} \ \& \ R^{\mathcal{M}_{[x]_E}}(x_0, \dots, x_{n-1})\}.$$

- (iv) (assuming  $\mathcal{L}$  has no nullary relation symbols) For each  $n$ -ary  $R \in \mathcal{L}$ , the following is a Borel subset of  $X^n$ :

$$R^{\mathcal{M}} := \{(x_0, \dots, x_{n-1}) \in X^n \mid x_0 E \cdots E x_{n-1} \ \& \ R^{\mathcal{M}_{[x_0]_E}}(x_0, \dots, x_{n-1})\}.$$

(This is often taken as the main definition, e.g., in [JKL02], [CK18].)

- (iii') and (iv') Same as above, but replacing  $R$  with arbitrary  $\phi(x_0, \dots, x_{n-1}) \in \mathcal{L}_{\omega_1\omega}$  (giving rise to Borel  $\widetilde{\phi}^{\mathcal{M}} \subseteq X^{n+1}$ , respectively  $\phi^{\mathcal{M}} \subseteq X^n$  for  $n \geq 1$ ).

Let  $\text{Mod}_E(\mathcal{T})$  denote the set of  $\mathcal{T}$ -structurings of  $E$ . If  $\text{Mod}_E(\mathcal{T}) \neq \emptyset$ , we call  $E$   **$\mathcal{T}$ -structurable**.

**Lemma 3.2.** *The conditions above are equivalent.*

*Proof.* (iii')  $\implies$  (iii)  $\iff$  (iv) is clear, as is (iii')  $\implies$  (iv').

(iv')  $\implies$  (iii'): Regard  $\phi(x_0, \dots, x_{n-1})$  as having an extra free variable.

(iii)  $\implies$  (ii): A generator for the Borel  $\sigma$ -algebra of  $\text{Mod}_Y(\mathcal{T}) \subseteq \text{Mod}_Y(\mathcal{L}) = \prod_{R \in \mathcal{L}^n} 2^{Y^n}$  is  $\{\mathcal{M} \mid R^{\mathcal{M}}(\vec{a})\}$  for some  $n$ -ary  $R \in \mathcal{L}$  and  $\vec{a} \in Y^n$ ; the preimage of this set under the map in (ii) is

$$\{z \in Z \mid R^{g_z^{-1} \cdot \mathcal{M}_{[f(z)]_E}}(\vec{a})\} = \{z \in Z \mid R^{\mathcal{M}_{[f(z)]_E}}(g_z(\vec{a}))\} = \{z \in Z \mid \widetilde{R}^{\mathcal{M}}(f(z), g_z(\vec{a}))\}.$$

(ii)  $\implies$  (i), for any  $(g_x)_x$ : It suffices to verify Borelness of the map in (i) restricted to the union of all  $E$ -classes of each size  $N \leq \omega$ , which is immediate from (ii) (for  $Z =$  said union).

(i) for some  $(g_x)_x \implies$  (iii'):

$$\widetilde{\phi}^{\mathcal{M}} = \{(x, x_0, \dots, x_{n-1}) \mid x E x_0 E \cdots E x_{n-1} \ \& \ R^{g_x^{-1} \cdot \mathcal{M}_{[x]_E}}(g_x^{-1} \cdot x_0, \dots, g_x^{-1} \cdot x_{n-1})\}. \quad \square$$

**Definition 3.3** (see [CK18, 3.1]). Let  $(X, E), (Y, F)$  be two CBERs. A **(Borel) class-bijective homomorphism**  $f : (X, E) \rightarrow (Y, F)$  is a Borel map  $f : X \rightarrow Y$  which restricts to a bijection  $f : [x]_E \cong [f(x)]_F$  for each  $x \in X$ .

Given such  $f$ , and a  $\mathcal{T}$ -structuring  $\mathcal{M} = (\mathcal{M}_D)_{D \in Y/F}$  of  $F$ , the **pullback  $\mathcal{T}$ -structuring**  $f^{-1} \cdot \mathcal{M}$  of  $E$  is defined by

$$(f^{-1} \cdot \mathcal{M})_C := (f|_C : C \cong f(C))^{-1} \cdot \mathcal{M}_{f(C)} \quad \text{for each } C \in X/E.$$

The Borelness conditions 3.1(iii) or (ii) are easily seen.

**Definition 3.4** (see [CK18, §B.3]). Let  $(\mathcal{L}, \mathcal{T}), (\mathcal{L}', \mathcal{T}')$  be two theories,  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  be an interpretation, and  $(X, E)$  be a CBER with a  $\mathcal{T}'$ -structuring  $\mathcal{M} = (\mathcal{M}_C)_{C \in X/E}$ . Recall from Definition 2.21 the notion of  $\alpha$ -reduct of a model of  $\mathcal{T}'$ . The  **$\alpha$ -reduct** of  $\mathcal{M}$  is the  $\mathcal{T}$ -structuring  $\alpha^* \mathcal{M}$  where

$$(\alpha^* \mathcal{M})_C := \alpha^*(\mathcal{M}_C) \quad \text{for each } C \in X/E.$$

The Borelness condition 3.1(ii) is again easily seen (just compose the map  $Z \rightarrow \text{Mod}_Y(\mathcal{T}')$  with  $\alpha_Y^* : \text{Mod}_Y(\mathcal{T}') \rightarrow \text{Mod}_Y(\mathcal{T})$  from (2.22)).

Thus, the operation

$$(3.5) \quad \begin{aligned} \{\text{CBERs}\} \times \{\text{theories}\} &\longrightarrow \{\text{sets}\} \\ (E, \mathcal{T}) &\longmapsto \text{Mod}_E(\mathcal{T}) \end{aligned}$$

is jointly (contravariantly) functorial in both variables, with respect to class-bijective homomorphisms, respectively interpretations. To verify functoriality, note that we have the obvious identities

$$f^{-1} \cdot g^{-1} \cdot \mathcal{K} = (g \circ f)^{-1} \cdot \mathcal{K}, \quad \alpha^* \beta^* \mathcal{K} = (\beta \circ \alpha)^* \mathcal{K}, \quad f^{-1} \cdot \alpha^* \mathcal{K} = \alpha^*(f^{-1} \cdot \mathcal{K})$$

(as well as  $\text{id}^{-1} \cdot \mathcal{K} = \mathcal{K}$  and  $\text{id}^* \mathcal{K} = \mathcal{K}$ ).

**Remark 3.6.** Definition 3.1(i) and (ii) are manifestly syntax-independent, depending only on the spaces of models of  $\mathcal{T}$  and not any particular axiomatization. We now give reformulations of the notion of structuring in terms of the other two ways of presenting a theory from Remark 2.25, namely (ii) the algebra of formulas modulo  $\mathcal{T}$ -equivalence and (iii) the spaces of types.

Recall from Example 2.24 that a countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$  is essentially by definition a homomorphism  $\phi \mapsto \phi^{\mathcal{M}} : \mathcal{L}_{\omega_1 \omega}^n / \mathcal{T} \rightarrow 2^{Y^n}$  taking formulas to relations for each  $n$ , preserving  $\bigwedge, \exists$ , etc. Thus, a structuring  $\mathcal{M} = (\mathcal{M}_C)_{C \in X/E} \in \text{Mod}_E(\mathcal{T})$  is a family of such homomorphisms

$$(3.7) \quad (\mathcal{L}_{\omega_1 \omega}^n / \mathcal{T} \longrightarrow 2^{C^n})_{n < \omega, C \in X/E}$$

or

$$\begin{aligned} (\mathcal{L}_{\omega_1 \omega}^n / \mathcal{T} \longrightarrow 2^{[x]_E^n})_{n < \omega, x \in X} \\ \phi \longmapsto \phi^{\mathcal{M}_{[x]_E}} \end{aligned}$$

which is  $E$ -invariant in  $x$ . Now a family of relations  $\phi^{\mathcal{M}_{[x]_E}} \in 2^{[x]_E^n} \subseteq 2^{X^n}$ , for each  $x \in X$ , is equivalently a single relation in  $2^{X^{n+1}}$  (since  $(2^{X^n})^X \cong 2^{X^{n+1}}$ ), which is contained in

$$\begin{aligned} E_X^n &:= \{(x, x_0, \dots, x_{n-1}) \in X^{n+1} \mid (x_0, \dots, x_{n-1}) \in [x]_E^n\} \\ &= \{(x, x_0, \dots, x_{n-1}) \in X^{n+1} \mid x \ E \ x_0 \ E \ \dots \ E \ x_{n-1}\}. \end{aligned}$$

(Here, the notation  $E_X^n$  refers to the fiber product of  $n$  copies of  $E$  over  $X$  via the first projection.) The family  $(\phi^{\mathcal{M}_{[x]_E}} \in 2^{[x]_E^n})_{x \in X}$  corresponds to the single Borel relation  $\tilde{\phi}^{\mathcal{M}} \subseteq E_X^n$  from 3.1(iii'). Thus, a structuring  $\mathcal{M} \in \text{Mod}_E(\mathcal{T})$  is equivalently a homomorphism

$$(3.8) \quad \begin{aligned} (\mathcal{L}_{\omega_1\omega}^n / \mathcal{T} &\longrightarrow \mathcal{B}^E(E_X^n))_{n < \omega} \\ \phi &\longmapsto \tilde{\phi}^{\mathcal{M}}, \end{aligned}$$

where  $\mathcal{B}^E(E_X^n) \subseteq \mathcal{B}(E_X^n)$  is the  $\sigma$ -algebra of Borel sets  $R \subseteq E_X^n$  which are  $E$ -invariant in the first coordinate; these may be thought of as ‘‘Borel families of  $n$ -ary relations in  $E$ -classes’’.

Note that here, ‘‘homomorphism’’ means that countable Boolean operations are preserved, while variable substitutions, quantifiers, and equality operate on the last  $n$  coordinates only (corresponding to the usual operations in  $2^{C^n}$  from (3.7)). Thus for  $s : m \rightarrow n$ , variable substitution along  $s$  maps

$$\begin{aligned} \mathcal{B}^E(E_X^m) &\longrightarrow \mathcal{B}^E(E_X^n) \\ R &\longmapsto \{(x, x_0, \dots, x_{n-1}) \mid R(x, x_{s(0)}, \dots, x_{s(m-1)})\}; \end{aligned}$$

while  $\exists$  is given by image along the projection  $E_X^{n+1} \rightarrow E_X^n$ , and likewise  $=$  is given by the image of the diagonal embedding  $E = E_X^1 \rightarrow E_X^2$ . In the following subsection, we will show that the family of  $\sigma$ -algebras  $(\mathcal{B}^E(E_X^n))_{n < \omega}$  equipped with these operations is isomorphic to the Lindenbaum–Tarski hyperdoctrine of a theory, analogous to the theory of  $Y$ -enumerated sets from Example 2.24, so that a  $\mathcal{T}$ -structuring, like a countable  $\mathcal{T}$ -model, may be understood as an interpretation.

Note also that  $\mathcal{B}^E(E_X^n)$  is isomorphic to the Borel sets in the quotient space  $E_X^n/E$  (meaning the quotient by the equivalence relation of  $E$ -equivalence in the first coordinate). For  $n \geq 1$ , we have  $E_X^n/E \cong E_X^{n-1}$ , along which the  $E$ -invariant  $\tilde{\phi}^{\mathcal{M}} \subseteq E_X^n$  descends to  $\phi^{\mathcal{M}} \subseteq E_X^{n-1}$  from 3.1(iv'). (For  $n = 0$ , we instead have  $E_X^n/E = X/E$ .) Since  $\mathcal{L}_{\omega_1\omega}^n / \mathcal{T}$  is isomorphic to the Borel sets  $\llbracket \phi \rrbracket$  in the space of types  $\mathcal{S}_n(\mathcal{T})$  (Definition 2.11), (3.8) becomes preimage under the Borel maps

$$(3.9) \quad \begin{aligned} \text{tp}_{\mathcal{M}}^n : E_X^n/E &\longrightarrow \mathcal{S}_n(\mathcal{T}) \\ [(x, x_0, \dots, x_{n-1})] &\longmapsto \text{tp}(\mathcal{M}_{[x]_E}, x_0, \dots, x_{n-1}), \end{aligned}$$

with  $(\text{tp}_{\mathcal{M}}^n)^{-1}(\llbracket \phi \rrbracket) = \tilde{\phi}^{\mathcal{M}}$ . Thus a structuring  $\mathcal{M} \in \text{Mod}_E(\mathcal{T})$  is also equivalently given by a family of such Borel maps  $(\text{tp}_{\mathcal{M}}^n)_{n < \omega}$ , subject to commutativity with image/preimage under the variable projections  $E_X^n/E \rightarrow E_X^m/E$  for each  $s : m \rightarrow n$  (analogous to the conditions in Remark 2.25(iii)).

### 3.B Scott theories of CBERs

We now define an operation  $E \mapsto \mathcal{T}_E$  that associates to each CBER  $E$  a canonical theory  $\mathcal{T}_E$  that ‘‘completely encodes’’  $E$ , thereby allowing the concepts of ‘‘structurability’’, ‘‘class-bijective homomorphism’’, and (as shown in the next subsection) even ‘‘CBER’’ to be subsumed by equivalent model-theoretic concepts. This construction is essentially from [CK18, §4.2, B.2], from which we also adopt the name ‘‘Scott sentence of a CBER’’, tweaked to fit our convention of working with theories instead of sentences. The main novelty of our approach here is to do everything in terms of the following intrinsic definition in terms of  $E$ , rather than a specific syntactic encoding as in [CK18, §4.2] (we will also review a version of that encoding below; see Construction 3.20).

**Definition 3.10.** Let  $(X, E)$  be a CBER. The **Scott theory** of  $E$  is the countable  $\mathcal{L}_{\omega_1\omega}$  theory  $(\mathcal{L}_E, \mathcal{T}_E)$  defined uniquely up to bi-interpretability via Theorem 2.26 by declaring its models on any countable set  $Y$  to be bijections between  $Y$  and some  $E$ -class:

$$(3.11) \quad \begin{aligned} \text{Mod}_Y(\mathcal{T}_E) &:= \{u \in X^Y \mid u : Y \cong C \text{ for some } C \in X/E\} \\ &= \left\{ u \in X^Y \left| \begin{array}{l} \forall y \neq z \in Y (u(y) \neq u(z)), \\ \forall y, z \in Y (u(y) E u(z)), \\ \forall x E u(y_0) \exists y \in Y (u(y) = x) \end{array} \right. \right\} \text{ for arbitrary } y_0 \in Y \end{aligned}$$

(and  $\text{Mod}_Y(\mathcal{T}_E) := \emptyset$  if  $Y = \emptyset$ ). This clearly defines a standard Borel subspace of  $X^Y$ . For two countable sets  $Y, Z$ , the logic action  $\text{Sym}(Y, Z) \times \text{Mod}_Y(\mathcal{T}_E) \rightarrow \text{Mod}_Z(\mathcal{T}_E)$  is given by relabeling: for  $g : Y \cong Z$  and  $u \in \text{Mod}_Y(\mathcal{T}_E)$ ,

$$g \cdot u := u \circ g^{-1}.$$

We have a **tautological  $\mathcal{T}_E$ -structuring**  $\mathcal{H}_E$  of  $E$ , given by the bijections  $(\mathcal{H}_E)_C := \text{id}_C : C \cong C$ .

Informally speaking, the models of  $\mathcal{T}_E$  are just the  $E$ -classes, with every element labeled. This is made precise by the following alternate characterizations of  $\mathcal{T}_E$ , in terms of Remark 2.25(ii) and (ii):

**Proposition 3.12.** *For each  $n \in \mathbb{N}$ , the  $n$ -types of  $\mathcal{T}_E$  are determined by the Borel isomorphism*

$$(3.13) \quad \begin{aligned} \text{tp}_{\mathcal{H}_E}^n : E_X^n/E &\cong \mathcal{S}_n(\mathcal{T}_E) \\ [(x, x_0, \dots, x_{n-1})] &\mapsto \text{tp}(\text{id}_{[x]_E}, x_0, \dots, x_{n-1}), \end{aligned}$$

where  $E_X^n/E$  is the space of  $n$ -tuples in  $E$ -classes ( $\cong E_X^{n-1}$  for  $n \geq 1$ ) from (3.9).

Thus, the algebra of  $n$ -ary formulas modulo  $\mathcal{T}_E$  is determined by the isomorphism

$$(3.14) \quad \begin{aligned} (\mathcal{L}_E)_{\omega_1\omega}^n / \mathcal{T}_E &\cong \mathcal{B}(\mathcal{S}_n(\mathcal{T}_E)) \cong \mathcal{B}(E_X^n/E) \cong \mathcal{B}^E(E_X^n), \\ \phi &\mapsto \llbracket \phi \rrbracket \quad \mapsto \phi^{\mathcal{H}_E} \quad \mapsto \tilde{\phi}^{\mathcal{H}_E}, \end{aligned}$$

where  $\mathcal{B}^E(E_X^n)$  is the algebra of “Borel families of  $n$ -ary relations in  $E$ -classes” from (3.8).

*Proof.* For any model  $u \in \text{Mod}_Y(\mathcal{T}_E)$ , i.e.,  $u : Y \cong u(Y) \in X/E$ , and  $\vec{a} \in Y^n$ , the unique preimage of  $\text{tp}(u, \vec{a}) \in \mathcal{S}_n(\mathcal{T}_E)$  under (3.13) is easily seen to be  $[(x, u(\vec{a}))]$  for any  $x \in u(Y)$ . So  $\text{tp}_{\mathcal{H}_E}^n$  is a bijection; since  $E_X^n/E, \mathcal{S}_n(\mathcal{T}_E)$  are quotients by Polish group actions, this implies Borel isomorphism.  $\square$

**Corollary 3.15** (see [CK18, B.2]). *For any CBER  $(X, E)$  and theory  $(\mathcal{L}, \mathcal{T})$ , we have a bijection*

$$\begin{aligned} \{\text{interpretations } \mathcal{T} \rightarrow \mathcal{T}_E\} &\cong \text{Mod}_E(\mathcal{T}) = \{\mathcal{T}\text{-structurings of } E\} \\ \alpha &\mapsto \alpha^*(\mathcal{H}_E). \end{aligned}$$

*Proof.* Note that  $\alpha^*(\mathcal{H}_E)$  is the structuring such that the type map (3.9) is, via (3.13),

$$\begin{aligned} \text{tp}_{\alpha^*(\mathcal{H}_E)}^n &= \alpha_n^* \circ \text{tp}_{\mathcal{H}_E}^n : E_X^n/E \cong \mathcal{S}_n(\mathcal{T}_E) \quad \longrightarrow \mathcal{S}_n(\mathcal{T}) \\ [(x, x_0, \dots, x_{n-1})] &\mapsto \text{tp}(\text{id}_{[x]_E}, \vec{x}) \longmapsto \text{tp}(\alpha^*((\mathcal{H}_E)_{[x]_E}), \vec{x}). \end{aligned}$$

Or via (3.14),  $\alpha^*(\mathcal{H}_E)$  interprets formulas as

$$\begin{aligned} \mathcal{L}_{\omega_1\omega}^n/\mathcal{T} &\longrightarrow (\mathcal{L}_E)_{\omega_1\omega}^n/\mathcal{T}_E \cong \mathcal{B}^E(E_X^n) \\ [\phi(x_0, \dots, x_{n-1})] &\longmapsto \alpha(\phi) \qquad \mapsto \widetilde{\alpha(\phi)}^{\mathcal{H}_E}. \end{aligned}$$

Thus the inverse of the bijection takes a structuring  $\mathcal{M}$  to  $\alpha : \mathcal{T} \rightarrow \mathcal{T}_E$  such that  $\widetilde{\alpha(\phi)}^{\mathcal{H}_E} = \tilde{\phi}^{\mathcal{M}}$ , or  $\alpha_n^* \circ \text{tp}_{\mathcal{H}_E}^n = \text{tp}_{\mathcal{M}}^n$ ; this determines a unique interpretation  $\alpha$  by Proposition 3.12 and Remark 2.25.  $\square$

**Proposition 3.16** (see [CK18, 4.7]). *For any CBERs  $(X, E), (Y, F)$ , we have a bijection*

$$\begin{aligned} \{\text{class-bijective homomorphisms } E \rightarrow F\} &\cong \text{Mod}_E(\mathcal{T}_F) = \{\mathcal{T}_F\text{-structurings of } E\} \\ f &\mapsto f^{-1} \cdot \mathcal{H}_F. \end{aligned}$$

*Proof.* A class-bijective homomorphism  $f : E \rightarrow F$  is clearly the same thing as a family of bijections  $(f|C : C \cong f(C) \in Y/F)_{C \in X/E}$ , i.e., a family of models of  $\mathcal{T}_F$

$$(f|C = \text{id}_{f(C)} \circ f|C = (f|C)^{-1} \cdot (\mathcal{H}_F)_{f(C)} = (f^{-1} \cdot \mathcal{H}_F)_{f(C)} \in \text{Mod}_C(\mathcal{T}_F))_{C \in X/E};$$

the Borelness of  $f$  is easily seen to correspond to the Borelness of the structuring.  $\square$

Note that here,  $f^{-1} \cdot \mathcal{H}_F$  is the structuring whose types are, via (3.13),

$$\begin{aligned} \text{tp}_{f^{-1} \cdot \mathcal{H}_F}^n : E_X^n/E &\longrightarrow F_Y^n/F && \cong \mathcal{S}_n(\mathcal{T}_F) \\ [(x, x_0, \dots, x_{n-1})] &\longmapsto [(f(x), f(\vec{x}))] && \mapsto \text{tp}(\text{id}_{[f(x)]_F}, f(\vec{x})), \end{aligned}$$

Combining the two preceding results, we have

**Corollary 3.17** (see [CK18, B.3]). *For any CBERs  $(X, E), (Y, F)$ , we have a bijection*

$$\{\text{class-bijective homomorphisms } E \rightarrow F\} \cong \text{Mod}_E(\mathcal{T}_F) \cong \{\text{interpretations } \mathcal{T}_F \rightarrow \mathcal{T}_E\}.$$

*Namely, each  $f : E \rightarrow F$  corresponds to the unique  $\alpha : \mathcal{T}_F \rightarrow \mathcal{T}_E$  such that  $\alpha^*(\mathcal{H}_E) = f^{-1} \cdot \mathcal{H}_F$ , or such that  $\alpha_n^* : \mathcal{S}_n(\mathcal{T}_E) \cong E_X^n/E \rightarrow F_Y^n/F \cong \mathcal{S}_n(\mathcal{T}_F)$  is the map induced by  $f$ .  $\square$*

In other words, the construction of Scott theories  $E \mapsto \mathcal{T}_E$  extends to a contravariant functor

$$\{\text{CBERs, class-bijective homomorphisms}\} \longrightarrow \{\mathcal{L}_{\omega_1\omega} \text{ theories, interpretations}\}$$

which is a full and faithful embedding (i.e., restricts to a bijection on each hom-set). We will characterize its essential image (i.e., image-up-to-isomorphism) in the following subsection.

We conclude this subsection by briefly recalling the explicit syntactic axiomatization of Scott theories  $\mathcal{T}_E$  from [CK18, §4.2]. We do this in order to keep this paper self-contained, and because some additional light may be shed on this axiomatization from our current perspective.

**Definition 3.18.** The **theory of countable separating families**  $(\mathcal{L}_{\text{sep}}, \mathcal{T}_{\text{sep}})$  is given by

$$\begin{aligned} \mathcal{L}_{\text{sep}} &:= \{U_i\}_{i \in \mathbb{N}} \quad (\text{each } U_i \text{ unary}), \\ \mathcal{T}_{\text{sep}} &:= \{\forall y \forall z [y \neq z \rightarrow \bigvee_{i \in \mathbb{N}} (U_i(y) \leftrightarrow \neg U_i(z))]\}. \end{aligned}$$

Note that for any countable set  $Y$ , we have a canonical bijection

$$(3.19) \quad \begin{aligned} \text{Mod}_Y(\mathcal{L}_{\text{sep}}) &= \prod_{i \in \mathbb{N}} 2^Y \cong (2^{\mathbb{N}})^Y \\ \mathcal{M} &= (U_i^{\mathcal{M}})_{i \in \mathbb{N}} \mapsto (u_{\mathcal{M}} : y \mapsto (U_i^{\mathcal{M}}(y))_i), \end{aligned}$$

which restricts to

$$\text{Mod}_Y(\mathcal{T}_{\text{sep}}) \cong \{u : Y \rightarrow 2^{\mathbb{N}} \mid u \text{ injective}\}.$$

Moreover, the logic action on the left is easily seen to correspond on the right to the relabeling action  $g \cdot u = u \circ g^{-1}$  as in Definition 3.10 of the Scott theory of a CBER.

**Construction 3.20.** Now given a CBER  $(X, E)$ , we may assume without loss that  $X \subseteq 2^{\mathbb{N}}$  is a Borel subspace. Then  $\text{Mod}_Y(\mathcal{T}_E)$ , as defined abstractly in (3.11), embeds via (3.19) into  $\text{Mod}_Y(\mathcal{T}_{\text{sep}})$ , and so by Lopez-Escobar,  $\mathcal{T}_E$  must be axiomatizable over the language  $\mathcal{L}_{\text{sep}}$  by  $\mathcal{T}_{\text{sep}}$  plus some additional axioms. Namely, we need axioms whose interpretations in a model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{L}_{\text{sep}})$  will correspond via the above bijection (3.19) to the second and third conditions in (3.11).

To axiomatize the second condition  $\forall y, z \in Y (u(y) E u(z))$ , note that since  $E \subseteq 2^{\mathbb{N}} \times 2^{\mathbb{N}}$  is Borel, we may write “ $u(y) E u(z)$ ” as a countable Boolean combination of assertions of the form “ $u(y)_i = 1$ ” or “ $u(z)_i = 1$ ” for various  $i \in \mathbb{N}$ , which correspond via (3.19) to  $U_i^{\mathcal{M}}(y)$  or  $U_i^{\mathcal{M}}(z)$  respectively. Thus letting  $\Phi_E(y, z)$  be the corresponding quantifier-free Boolean combination of atomic formulas  $U_i(y), U_i(z)$ , the sentence

$$(3.21) \quad \forall y, z \Phi_E(y, z)$$

axiomatizes those  $\mathcal{M} \in \text{Mod}_Y(\mathcal{L}_{\text{sep}})$  corresponding via (3.19) to  $u$  obeying  $\forall y, z \in Y (u(y) E u(z))$ .

To axiomatize the third condition in (3.11), first write  $E = \bigcup_{i \in \mathbb{N}} f_i$  for Borel functions  $f_i : X \rightarrow X$  by Lusin–Novikov. Then rewrite the third condition in (3.11) as

$$\begin{aligned} &\exists y_0 \in Y \forall x \in X (u(y_0) E x) \iff \exists y \in Y (u(y) = x) \\ \iff &\exists y_0 \in Y \forall i \in \mathbb{N} \quad \exists y \in Y (u(y) = f_i(u(y_0))). \end{aligned}$$

Now similarly to before, using that the graphs of  $f_i \subseteq 2^{\mathbb{N}} \times 2^{\mathbb{N}}$  are Borel, find quantifier-free  $\mathcal{L}_{\text{sep}}$ -formulas  $\phi_i(y_0, y)$  axiomatizing “ $u(y) = f_i(u(y_0))$ ”; then

$$(3.22) \quad \exists y_0 \bigwedge_{i \in \mathbb{N}} \exists y \phi_i(y_0, y)$$

works. So we may axiomatize  $\mathcal{T}_E$  over  $\mathcal{L}_E := \mathcal{L}_{\text{sep}}$  via  $\mathcal{T}_{\text{sep}} \cup \{(3.21), (3.22)\}$ .

### 3.C Lusin–Novikov functions

We now aim to characterize those  $\mathcal{L}_{\omega_1\omega}$  theories bi-interpretable to a Scott theory of some CBER, i.e., the essential image of the full functorial embedding from Corollary 3.17. The key ingredient is to isolate the role played by the formulas  $\phi_i$  obtained from Lusin–Novikov in Construction 3.20.

**Definition 3.23.** The **theory of Lusin–Novikov functions**  $(\mathcal{L}_{\text{LN}}, \mathcal{T}_{\text{LN}})$  is given by

$$\begin{aligned} \mathcal{L}_{\text{LN}} &:= \{f_i\}_{i \in \mathbb{N}} \quad (\text{each } f_i \text{ a unary function}), \\ \mathcal{T}_{\text{LN}} &:= \{\forall x \forall y \bigvee_i (f_i(x) = y)\}. \end{aligned}$$

(Following Convention 2.4, we formally encode each  $f_i$  as its graph relation  $F_i(x, y) \iff f_i(x) = y$ , and add to  $\mathcal{T}_{\text{LN}}$  axioms saying that each  $F_i$  is a graph of a function.)



**Example 3.24.** In Construction 3.20 of the Scott theory  $\mathcal{T}_E$  of a CBER, we have an interpretation  $\alpha : \mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}_E$  mapping  $f_i \mapsto \phi_i$  (formally, mapping the graph  $f_i(x) = y$  to  $\phi_i(x, y)$ ). Indeed, note that the formula  $\Phi_E$  defined there may be given in terms of the  $\phi_i$  as

$$\Phi_E(x, y) := \bigvee_i \phi_i(x, y).$$

Then the axiom (3.21) in  $\mathcal{T}_E$  becomes precisely  $\alpha(\mathcal{T}_{\text{LN}})$ .

**Proposition 3.25.** *A theory  $(\mathcal{L}, \mathcal{T})$  interprets  $\mathcal{T}_{\text{LN}}$  iff the following equivalent conditions hold:*

- (i) *The space of 1-types  $\mathcal{S}_1(\mathcal{T})$  is standard Borel, and for any pointed model  $(\mathcal{M}, a) \in \text{Mod}_Y^1(\mathcal{T})$ , the map  $b \mapsto \text{tp}(\mathcal{M}, a, b) : Y \rightarrow \mathcal{S}_2(\mathcal{T})$  is injective.*
- (ii) *For any countable set  $Y$ , the logic action  $\text{Sym}(Y) \curvearrowright \text{Mod}_Y^1(\mathcal{T})$  is free and smooth.*
- (iii) *There is a countable set  $\mathcal{F} \subseteq \mathcal{L}_{\omega_1\omega}^1$  of formulas in one variable, such that every pointed model of  $\mathcal{T}$  is rigid and  $\mathcal{F}$ -categorical.*

*Proof.* These conditions are equivalent by Corollaries 2.12 and 2.14.

$\implies$ : Let  $(\mathcal{M}, a) \in \text{Mod}_Y(\mathcal{T})$  be a pointed model, and let  $g : \mathcal{M} \cong \mathcal{M}$  be an automorphism fixing  $a$ . Then for any  $b \in Y$ , there is a Lusin–Novikov function  $f_i : Y \rightarrow Y$  that is part of  $\mathcal{M}$  such that  $f_i(a) = b$ , whence  $g(b) = g(f_i(a)) = f_i(g(a)) = f_i(a) = b$ . This shows that  $(\mathcal{M}, a)$  is rigid.

To show categoricity over a countable fragment  $\mathcal{F}$ , we may as well assume that the Lusin–Novikov functions  $f_i$  are part of the language  $\mathcal{L}$  (and  $\mathcal{T} \vdash \mathcal{T}_{\text{LN}}$ ); if not, then consider the expansion of  $\mathcal{T}$  with such unary functions  $f_i$  defined to be equal to the function graphs  $\phi_i$  in the image of the interpretation  $\mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}$ . Then  $\mathcal{T}$  admits quantifier elimination for formulas with at least one free variable, since every existential  $\exists y \psi(x_0, \dots, x_{n-1}, y)$ ,  $n \geq 1$ , is  $\mathcal{T}$ -provably equivalent to  $\bigvee_i \psi(x_0, \dots, x_{n-1}, f_i(x_0))$ . So we may take  $\mathcal{F}$  to consist of just the atomic formulas. (To undo the expansion of  $\mathcal{T}$ , replace these with existential formulas over the function graph relations  $\phi_i$ .)

$\impliedby$ : Suppose the logic action  $\text{Sym}(N) \curvearrowright \text{Mod}_N^1(\mathcal{T})$  is free and smooth for each  $N \leq \omega$ ; let  $D_N \subseteq \text{Mod}_N^1(\mathcal{T})$  be a Borel transversal. For each  $i < N$ , let

$$\begin{aligned} F_{N,i} &:= \text{Sym}(N) \cdot (D_N \times \{i\}) \subseteq \text{Mod}_N^2(\mathcal{T}) \\ &= \{(\mathcal{M}, a, b) \mid g(b) = i \text{ for the unique } g \in \text{Sym}(N) \text{ such that } g(\mathcal{M}, a) \in D_N\}. \end{aligned}$$

These are Borel  $\text{Sym}(N)$ -invariant and form graphs of functions of the first two variables  $\mathcal{M}, a$  covering  $\text{Mod}_N^2(\mathcal{T})$  as  $i$  varies, hence by Lopez-Escobar, are defined by formulas  $\phi_{N,i}(x, y)$  such that  $\mathcal{T} \vdash$  “there are  $N$  elements”  $\rightarrow$  “ $\phi_{N,i}$  is a function”  $\wedge \forall x \forall y \bigvee_i \phi_{N,i}(x, y)$ . For  $N \leq i < \omega$ , let  $\phi_{N,i}(x, y) := (x = y)$  be the identity function, and put  $\phi_i := \bigvee_{N \leq \omega}$  (“there are  $N$  elements”  $\wedge \phi_{N,i}$ ). Then the  $\phi_i$  form an interpretation  $\mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}$ .  $\square$

The above conditions may be verified abstractly for a Scott theory  $\mathcal{T} = \mathcal{T}_E$  of a CBER  $(X, E)$ , without resorting to the explicit axiomatization in Construction 3.20. Indeed, every (unpointed) model of  $\mathcal{T}_E$ , i.e., bijection with an  $E$ -class, is clearly rigid. And by Proposition 3.12,  $\mathcal{S}_1(\mathcal{T}_E) \cong E_X^0 = X$  is standard Borel. Note also that  $\mathcal{S}_0(\mathcal{T}_E) \cong X/E$ , with the projection  $\partial_1 : \mathcal{S}_1(\mathcal{T}_E) \rightarrow \mathcal{S}_0(\mathcal{T}_E)$  (from Remark 2.17) corresponding to the quotient map  $X \twoheadrightarrow X/E$ :

$$\begin{array}{ccc} E \rightrightarrows X & \longrightarrow & X/E \\ \parallel & & \parallel \\ \ker(\partial_1) \rightrightarrows \mathcal{S}_1(\mathcal{T}) & \xrightarrow{\partial_1} & \mathcal{S}_0(\mathcal{T}) \end{array}$$

We may thus recover the CBER  $(X, E)$  canonically up to isomorphism from  $\mathcal{T} = \mathcal{T}_E$ , as the kernel of  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \rightarrow \mathcal{S}_0(\mathcal{T})$ .

**Theorem 3.26.** *For a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the following are equivalent:*

- (i)  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \rightarrow \mathcal{S}_0(\mathcal{T})$  is surjective,  $\mathcal{S}_1(\mathcal{T})$  is standard Borel, and for any model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$ , the map  $\text{tp}_{\mathcal{M}} : Y \rightarrow \mathcal{S}_1(\mathcal{T})$  is injective.
- (ii)  $\mathcal{T}$  has no empty models, and for any countable set  $Y$ , the logic action  $\text{Sym}(Y) \curvearrowright \text{Mod}_Y(\mathcal{T})$  is free and the logic action  $\text{Sym}(Y) \curvearrowright \text{Mod}_Y^1(\mathcal{T})$  on pointed models is smooth.
- (iii) Every model of  $\mathcal{T}$  is nonempty rigid, and there is a countable set  $\mathcal{F} \subseteq \mathcal{L}_{\omega_1\omega}^1$  of formulas in one variable such that every pointed model is  $\mathcal{F}$ -categorical.
- (iv)  $\mathcal{T}$  has no empty models and interprets  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ .
- (v)  $\mathcal{T}$  is bi-interpretable with the Scott theory  $\mathcal{T}_E$  of a CBER  $E$ , namely  $E = \ker(\partial_1)$ .

Thus we have a dual equivalence of categories

$$\{\text{CBERs, class-bijective homomorphisms}\} \simeq \{\mathcal{L}_{\omega_1\omega} \text{ theories obeying (i)–(v), interpretations}\}$$

taking a CBER  $E$  to its Scott theory  $\mathcal{T}_E$ , with the correspondence on morphisms as in Corollary 3.17.

*Proof.* As in Proposition 3.25, the equivalence between (i), (ii) and (iii) follows from Corollaries 2.12 and 2.14, noting that to say  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \rightarrow \mathcal{S}_0(\mathcal{T})$  is surjective means that every model is nonempty.

These conditions are equivalent to (iv) by Proposition 3.25, the fact that models of  $\mathcal{T}_{\text{sep}}$  are clearly rigid, and that conversely, if models are rigid, then the family  $\mathcal{F}$  from (iii) must form a countable separating family of unary relations in every model by Scott's isomorphism theorem.

As noted above, by Construction 3.20 or an easy check, a Scott theory of a CBER obeys (i)–(iv).

Finally, suppose the theory  $\mathcal{T}$  obeys (i); we verify (v) for  $E = \ker(\partial_1)$ . Note first that  $E \subseteq \mathcal{S}_1(\mathcal{T})^2$  is Borel, since its lift in  $\text{Mod}_Y^1(\mathcal{T})^2$  is the preimage under the projection of  $\cong \subseteq \text{Mod}_Y(\mathcal{T})^2$  which is Borel by Theorem 2.15. Clearly  $E$  is also countable, since we are assuming models to be countable. So  $E = \ker(\partial_1) \subseteq \mathcal{S}_1(\mathcal{T})^2$  is a CBER.

For any countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$ , by (i), we have an injection  $\text{tp}_{\mathcal{M}} : Y \hookrightarrow \mathcal{S}_1(\mathcal{T})$ , whose image is clearly the  $E$ -class  $\partial_1^{-1}(\text{tp}(\mathcal{M}))$ . This defines a map

$$\begin{aligned} f_Y : \text{Mod}_Y(\mathcal{T}) &\longrightarrow \text{Mod}_Y(\mathcal{T}_E) \subseteq \mathcal{S}_1(\mathcal{T})^Y \\ \mathcal{M} &\longmapsto \text{tp}_{\mathcal{M}}, \end{aligned}$$

which is easily seen to be Borel (since  $f_Y(\mathcal{M})(a) = \text{tp}(\mathcal{M}, a) \in \llbracket \phi \rrbracket \iff \phi^{\mathcal{M}}(a)$ ) and equivariant under bijections  $g : Y \cong Z$  (since  $f_Y(g \cdot \mathcal{M})(a) = \text{tp}(g \cdot \mathcal{M}, a) = \text{tp}(\mathcal{M}, g^{-1} \cdot a) = (f_Y(\mathcal{M}) \circ g^{-1})(a)$ ). Thus by Remark 2.25(iv),  $f_Y = \alpha_Y^*$  for an interpretation  $\alpha : \mathcal{T}_E \rightarrow \mathcal{T}$ . Since both  $\mathcal{T}, \mathcal{T}_E$  have rigid models, the logic actions on both  $\text{Mod}_Y(\mathcal{T}), \text{Mod}_Y(\mathcal{T}_E)$  are free; thus to check that  $f_Y$  is bijective, it suffices to check that it is bijective on orbits, i.e., it induces a bijection  $\mathcal{S}_0(\mathcal{T}) \cong \mathcal{S}_0(\mathcal{T}_E)$ .

By Proposition 3.12, we have a bijection  $\mathcal{S}_0(\mathcal{T}) \cong \mathcal{S}_1(\mathcal{T})/E \cong \mathcal{S}_0(\mathcal{T}_E)$  taking  $C \mapsto \text{tp}(\text{id}_C)$  for each  $E$ -class  $C \in \mathcal{S}_1(\mathcal{T})/E$ . Let the 1-types in  $C$  be realized in  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$ ; then  $\text{tp}_{\mathcal{M}} : Y \cong C \subseteq \mathcal{S}_1(\mathcal{T})$ . So  $\text{tp}_{\mathcal{M}} : \text{tp}_{\mathcal{M}} \cong \text{id}_C$  as models of  $\mathcal{T}_E$ , and so  $\text{tp}(\text{id}_C) = \text{tp}(\text{tp}_{\mathcal{M}}) = \text{tp}(f_Y(\mathcal{M}))$ .  $\square$

### 3.D Structurability via interpretations

Using Theorem 3.26, as well as the universal properties of Scott theories from Section 3.B, we may now translate questions of structurability into purely model-theoretic terms.

To avoid having to restate this assumption repeatedly, below we will assume that all theories under consideration have no empty models.

**Corollary 3.27.** *For two  $\mathcal{L}_{\omega_1\omega}$  theories  $\mathcal{T}, \mathcal{T}'$  (in respective languages), the following are equivalent:*

- (i) *Every  $\mathcal{T}$ -structurable CBER is  $\mathcal{T}'$ -structurable.*
- (ii) *There exists an interpretation  $\mathcal{T}' \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ .*

*Proof.* We have

$$\begin{aligned} \text{(i)} &\iff \forall \text{ CBERs } E, (\mathcal{T} \rightarrow \mathcal{T}_E \implies \mathcal{T}' \rightarrow \mathcal{T}_E) && \text{by Corollary 3.15} \\ &\iff \forall \text{ theories } \mathcal{T}_E \leftarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}, (\mathcal{T} \rightarrow \mathcal{T}_E \implies \mathcal{T}' \rightarrow \mathcal{T}_E) && \text{by Theorem 3.26;} \end{aligned}$$

since the least theory  $\mathcal{T}_E$  (under the interpretability preorder) interpreting both  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  and  $\mathcal{T}$  is their least upper bound  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  (recall Definition 2.27), this is equivalent to (ii).  $\square$

Recall that in Example 1.2, we showed how the well-known fact that selecting a finite nonempty set in each class of a CBER ( $\mathcal{T}_{\text{finsub}}$ -structurability) implies smoothness ( $\mathcal{T}_{\text{pt}}$ -structurability) amounts to an interpretation  $\mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ . We will exhibit many more such concrete instances of Corollary 3.27 in Section 4.

We note that Corollary 3.27, as a statement only about *existence* of structurings, contains strictly less information than Theorem 3.26 and Corollary 3.15, which precisely characterize the set of all possible structurings. This becomes relevant for problems in Borel combinatorics, rather than just CBERs, where the input data is something richer than an equivalence relation:

**Definition 3.28.** Given two theories  $\mathcal{T}, \mathcal{T}'$ , an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$ , and a  $\mathcal{T}$ -structuring  $\mathcal{M}$  of a CBER  $E$ , an  $\alpha$ -**expansion** of  $\mathcal{M}$  is a  $\mathcal{T}'$ -structuring  $\mathcal{M}'$  such that  $\alpha^*(\mathcal{M}') = \mathcal{M}$ .

For example, consider the problem of Borel coloring a (locally countable) Borel graph  $G \subseteq X^2$ . Letting  $E$  be the connectedness relation generated by  $G$ , this problem may be described as starting with a structuring  $\mathcal{M}$  of  $E$  by the theory  $\mathcal{T}_{\text{graph}}$  of graphs, and seeking an expansion of  $\mathcal{M}$  to a structuring by the theory  $\mathcal{T}_{\text{graph+color}}$  of graphs equipped with a coloring, i.e., an  $\alpha$ -expansion for  $\alpha : \mathcal{T}_{\text{graph}} \hookrightarrow \mathcal{T}_{\text{graph+color}}$  the inclusion. See Proposition 4.29 for a concrete instance of this.

**Corollary 3.29.** *For any interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$ , the following are equivalent:*

- (i) *Every  $\mathcal{T}$ -structuring of a CBER admits an  $\alpha$ -expansion to a  $\mathcal{T}'$ -structuring.*
- (ii) *There exists an interpretation  $\beta : \mathcal{T}' \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  such that  $\beta \circ \alpha : \mathcal{T} \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the inclusion (up to provable equivalence as usual).*

*Proof.* We again have

$$\begin{aligned} \text{(i)} &\iff \forall \text{ CBERs } E, \forall \text{ interpretations } \gamma : \mathcal{T} \rightarrow \mathcal{T}_E, \exists \gamma' : \mathcal{T}' \rightarrow \mathcal{T}_E (\gamma' \circ \alpha = \gamma) && \text{by 3.15} \\ &\iff \forall \text{ theories } \mathcal{T}_E \leftarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}, \forall \gamma : \mathcal{T} \rightarrow \mathcal{T}_E, \exists \gamma' : \mathcal{T}' \rightarrow \mathcal{T}_E (\gamma' \circ \alpha = \gamma) && \text{by 3.26.} \end{aligned}$$

If this holds, then taking  $\mathcal{T}_E := \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  with  $\gamma$  the inclusion yields (ii). Conversely, given  $\beta$  from (ii), then for any other  $\mathcal{T}_E, \gamma$  as above, we may combine the interpretation  $\gamma$  and the interpretation  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  into an interpretation from the coproduct  $\delta : \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_E$  (recall again Definition 2.27); then  $\gamma' := \delta \circ \beta$  works.  $\square$

Theories of the form  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  featured above play a special role in structurability. Note that up to *mutual* interpretability, they are just all the Scott theories of CBERs by Theorem 3.26, since  $\sqcup$  is join in the preorder  $\rightarrow$ . But up to *bi*-interpretability, we may characterize them as follows.

**Definition 3.30** (see [CK18, 4.13]). A class-bijective homomorphism  $f : (X, E) \rightarrow (Y, F)$  between CBERs which is also injective is called an **invariant (Borel) embedding**, denoted  $f : E \sqsubseteq^i F$ .

Given a class  $\mathcal{K}$  of CBERs, we say  $E \in \mathcal{K}$  is **invariantly universal in  $\mathcal{K}$**  if it is  $\sqsubseteq^i$ -largest among (i.e., contains invariant isomorphic copies of all other) CBERs in  $\mathcal{K}$ . By the Borel Schröder–Bernstein theorem [Kec95, 15.7], such an  $E$  is unique up to Borel isomorphism if it exists.

In particular, there is an **invariantly universal CBER** among all CBERs, denoted  $\mathbb{E}_\infty$ .

More generally, for any theory  $\mathcal{T}$ , there is an invariantly universal  $\mathcal{T}$ -structurable CBER, denoted  $\mathbb{E}_{\infty\mathcal{T}}$ .<sup>6</sup> A CBER  $E$  which is isomorphic to some  $\mathbb{E}_{\infty\mathcal{T}}$  is called **universally structurable**. In that case, as follows from Corollary 3.37 below, we may always take  $\mathcal{T} = \mathcal{T}_E$ .

**Remark 3.31.** We may characterize those class-bijective homomorphisms  $f : (X, E) \rightarrow (Y, F)$  which are invariant embeddings model-theoretically, in terms of the corresponding  $\mathcal{T}_F$ -structurings  $\mathcal{M}$  of  $E$  from Proposition 3.16, or the corresponding interpretations  $\alpha : \mathcal{T}_F \rightarrow \mathcal{T}_E$  from Corollary 3.17, as follows. Note that since  $f$  is class-bijective, it is an invariant embedding iff it is also a reduction, i.e., descends to an injection  $X/E \hookrightarrow Y/F$ . Via Proposition 3.16, this means that

- (i) Distinct  $E$ -classes  $C \neq D \in X/E$  receive non-isomorphic  $\mathcal{T}_F$ -structures  $\mathcal{M}_C \not\cong \mathcal{M}_D$ .

Since  $E$ -classes are isomorphism types of  $\mathcal{T}_E$ -models, this means

- (ii) The interpretation  $\alpha$  induces an injection  $\alpha_0^* : \mathcal{S}_0(\mathcal{T}_E) \rightarrow \mathcal{S}_0(\mathcal{T}_F)$ .

Or dually, between algebras of sentences  $\mathcal{B}(\mathcal{S}_0(\mathcal{T}_E)) \cong (\mathcal{L}_E)_{\omega_1\omega}^0 / \mathcal{T}_E$ ,

- (iii)  $\alpha$  is a surjection on provable equivalence classes of sentences.

We may also replace sentences with arbitrary formulas here, or 0-types with  $n$ -types, since  $f$  being injective clearly implies the same for the induced maps  $E_X^n \rightarrow F_Y^n$  which descend to the induced maps  $\alpha_n^* : \mathcal{S}_n(\mathcal{T}_E) \rightarrow \mathcal{S}_n(\mathcal{T}_F)$  by Corollary 3.17. We call such  $\alpha : \mathcal{T}_F \rightarrow \mathcal{T}_E$  a **surjective interpretation**.

**Remark 3.32.** It follows that we have a Schröder–Bernstein theorem for interpretations: if  $\mathcal{T}, \mathcal{T}'$  are both Scott theories of CBERs, and there exist surjective interpretations  $\mathcal{T} \rightarrow \mathcal{T}'$  and  $\mathcal{T}' \rightarrow \mathcal{T}$ , then there exists a bi-interpretation  $\mathcal{T} \cong \mathcal{T}'$ .

**Proposition 3.33.**  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the Scott theory of the invariantly universal CBER  $\mathbb{E}_\infty$ .

*Proof.* For any other CBER  $E$ , the interpretation  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_E$  given by the explicit Construction 3.20 of  $\mathcal{T}_E$  is obviously surjective (on atomic formulas, hence on all formulas), since it is the identity on the  $\mathcal{T}_{\text{sep}}$  part.

Or in other words, we may obviously find a  $(\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}})$ -structuring  $\mathcal{M}$  of  $E$  such that distinct  $E$ -classes  $C, D$  receive non-isomorphic  $\mathcal{M}_C \not\cong \mathcal{M}_D$ , by taking the  $\mathcal{T}_{\text{sep}}$  part of the structuring to come from an embedding into  $2^{\mathbb{N}}$  (and the  $\mathcal{T}_{\text{LN}}$  part from Lusin–Novikov applied to  $E$ ).  $\square$

<sup>6</sup>Note that in this paper, unlike many other works including [JKL02], [CK18],  $\mathcal{T}$  denotes an arbitrary countable  $\mathcal{L}_{\omega_1\omega}$  theory and *not* the class of trees.

**Corollary 3.34.** *For any theory  $\mathcal{T}$ ,  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the Scott theory of the invariantly universal  $\mathcal{T}$ -structurable CBER  $\mathbb{E}_{\infty\mathcal{T}}$ .*

*Proof.* The inclusion  $\mathcal{T} \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  shows that  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the Scott theory of a  $\mathcal{T}$ -structurable CBER  $\mathbb{E}_{\infty\mathcal{T}}$ . For any other  $\mathcal{T}$ -structurable CBER  $E$ , we get an interpretation  $\mathcal{T} \rightarrow \mathcal{T}_E$  by Corollary 3.15; combined with the surjective interpretation  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \twoheadrightarrow \mathcal{T}$  given by the preceding result, we get a surjective interpretation  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \twoheadrightarrow \mathcal{T}_E$ , corresponding to an invariant embedding  $E \sqsubseteq^i \mathbb{E}_{\infty\mathcal{T}}$ .  $\square$

**Example 3.35.** Let  $\mathcal{T}_{\text{inf}}$  be the **theory of infinite sets** (in the empty language):

$$\mathcal{T}_{\text{inf}} := \bigwedge_n \exists x_0 \cdots \exists x_{n-1} \bigwedge_{i \neq j} (x_i \neq x_j).$$

Then  $\mathcal{T}_{\text{inf}} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the Scott theory of the invariantly universal **aperiodic** (i.e., every class is infinite) CBER, denoted  $\mathbb{E}_{\infty\text{aper}}$  in Figures 1.9 and 4.17.

**Remark 3.36.** More generally, for any theory  $\mathcal{T}$  and any Scott theory  $\mathcal{T}_E$ , the coproduct  $\mathcal{T} \sqcup \mathcal{T}_E$  is also a Scott theory, namely of the **universal  $\mathcal{T}$ -structurable CBER over  $E$**  denoted  $E \times \mathcal{T}$  from [CK18, 4.1], with the defining property that a class-bijective homomorphism  $F \rightarrow E \times \mathcal{T}$  is the same thing as a class-bijective homomorphism  $F \rightarrow E$  together with a  $\mathcal{T}$ -structuring of  $F$ .

In particular, if  $\mathcal{T} = \mathcal{T}_F$  is also a Scott theory, then  $\mathcal{T}_E \sqcup \mathcal{T}_F$  is the Scott theory of the **class-bijective product**  $E \otimes F$  from [CK18, 4.17], the product in the category of CBERs and class-bijective homomorphisms; this is clear from the full embedding given by Corollary 3.17.

The following is the model-theoretic counterpart of [CK18, 4.13]:

**Corollary 3.37.** *For a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the following are equivalent:*

- (i)  $\mathcal{T} \cong \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , i.e.,  $\mathcal{T}$  is the Scott theory of  $\mathbb{E}_{\infty\mathcal{T}}$ .
- (ii)  $\mathcal{T} \cong \mathcal{T}' \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  for some theory  $\mathcal{T}'$ , i.e.,  $\mathcal{T}$  is the Scott theory of some universally structurable CBER.
- (iii)  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}$ , and for any  $\mathcal{T}'$  interpreting  $\mathcal{T}$ , there is a surjective interpretation  $\mathcal{T} \rightarrow \mathcal{T}'$ .

*Proof.* (i)  $\implies$  (ii) is obvious, and (ii)  $\implies$  (iii) by Corollary 3.34, since if  $\mathcal{T} \rightarrow \mathcal{T}'$  then  $\mathcal{T}'$  is the Scott theory of a  $\mathcal{T}$ -structurable CBER (by Theorem 3.26 and Corollary 3.15). Now suppose (iii); then from  $\mathcal{T} \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  we get a surjective such interpretation, and we clearly also have a surjective  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \twoheadrightarrow \mathcal{T}$  by combining the identity with the given  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}$ , whence  $\mathcal{T} \cong \mathcal{T} \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  by the Schröder–Bernstein property 3.32.  $\square$

## 4 Examples of (non-)interpretations

Many fundamental theorems about CBERs can be naturally understood in terms of structurability by particular  $\mathcal{L}_{\omega_1\omega}$  theories, and the proofs of these theorems often amount to interpretations between these theories. In this section, we give several examples of such interpretations to illustrate this perspective. These examples include the Feldman–Moore theorem, its generalization to  $\omega$ -coloring the intersection graph on finite subsets of  $E$ -classes, and the Slaman–Steel marker lemma.

Moreover, we will show several *non-interpretability* results between these theories, thereby making precise the idea that certain kinds of Borel combinatorial structures are “more powerful”

than others. For example, we will show that an  $\omega$ -coloring of finite subsets is “more powerful” than the Feldman–Moore theorem, despite both being available on every CBER; see Remark 4.13.

The theories we will consider are listed in Table 4.1. (Recall our Convention 2.4, that function symbols occurring in examples are implicitly encoded as their graph relations.)

#### 4.A Variants of the Feldman–Moore theorem

A key result in the theory of CBERs is the Feldman–Moore theorem [FM77], which states that every CBER is generated by a Borel action of a countable group. In fact, the usual proof of Feldman–Moore yields an *a priori* stronger conclusion, namely that the Borel functions covering  $E = \bigcup_i f_i$  given by the Lusin–Novikov theorem 2.3(a) can be upgraded to a family of Borel involutions that still cover  $E$ . Indeed, this stronger statement is nowadays often understood, especially in the context of Borel combinatorics, as the real content of the Feldman–Moore theorem (see e.g., [Kec24, 3.4]).

We may distinguish between these variants of the Feldman–Moore and Lusin–Novikov theorems, by restating their conclusions as structurability by different  $\mathcal{L}_{\omega_1\omega}$  theories. Recall from Example 3.24 that Lusin–Novikov applied to a CBER  $E$  yields structurability by the theory  $\mathcal{T}_{\text{LN}}$  in the language with countably many unary functions  $\mathcal{L}_{\text{LN}} = \{f_i\}_{i \in \mathbb{N}}$  asserting that they cover all pairs of elements. We may strengthen this by requiring the  $f_i$  to be bijections, or even involutions. On the other hand, the original statement of Feldman–Moore is slightly different in flavor, in that it gives a transitive action on each  $E$ -class of a countable group (which we may assume to be the free group  $\mathbb{F}_\omega$ , without loss of generality). We may strengthen this to an action of a group generated by involutions (which we may assume to be the free product  $\mathbb{Z}_2^{*\omega}$ ), or weaken it to only an action of a monoid (e.g., the free monoid  $\mathbb{N}^{<\omega}$ ). All told, we have 6 theories capturing variants of Lusin–Novikov and Feldman–Moore:

$$(4.2) \quad \begin{array}{ll} \mathcal{T}_{\text{LN}} = \mathcal{T}_{\text{LN}_{\text{fun}}} := \text{covering family of functions,} & \mathcal{T}_{\text{FM}_{\text{fun}}} = \mathcal{T}_{\mathbb{N}^{<\omega}} := \text{transitive } \mathbb{N}^{<\omega} \text{ action,} \\ \mathcal{T}_{\text{LN}_{\text{bij}}} := \text{covering family of bijections,} & \mathcal{T}_{\text{FM}_{\text{bij}}} = \mathcal{T}_{\mathbb{F}_\omega} := \text{transitive } \mathbb{F}_\omega \text{ action,} \\ \mathcal{T}_{\text{LN}_{\text{invol}}} := \text{covering family of involutions,} & \mathcal{T}_{\text{FM}_{\text{invol}}} = \mathcal{T}_{\mathbb{Z}_2^{*\omega}} := \text{transitive } \mathbb{Z}_2^{*\omega} \text{ action.} \end{array}$$

Here by a *covering* family of functions  $f_i$  on a set  $Y$ , we mean that their graphs cover  $Y^2$ ; and by  $\mathcal{T}_\Gamma$  in general for a countable monoid  $\Gamma$ , we mean the theory of sets equipped with a transitive  $\Gamma$ -action. See Table 4.1 for precise axiomatizations of each of these theories.

Our somewhat pedantic naming scheme for these theories is that “ $\text{LN}_{\langle \text{property} \rangle}$ ” refers to a covering family of functions each with said property, while “ $\text{FM}_{\langle \text{property} \rangle}$ ” means that only the closure under composition/inverse of said functions is required to cover, i.e., we have a transitive action of the group/monoid generated by said functions. Thus, the Lusin–Novikov theorem says that every CBER is structurable by  $\mathcal{T}_{\text{LN}_{\text{fun}}}$  ( $= \mathcal{T}_{\text{LN}}$  from Definition 3.23); while the Feldman–Moore theorem asserts, depending on the variant, structurability by  $\mathcal{T}_{\text{FM}_{\text{bij}}}$ ,  $\mathcal{T}_{\text{FM}_{\text{invol}}}$ , or  $\mathcal{T}_{\text{LN}_{\text{invol}}}$ .

Note that these theories are not *a priori* equivalent in strength: for instance, it is not clear if a covering family of bijections may be canonically turned into involutions. The following diagram

$\langle \text{name} \rangle$ (ref)	$\mathcal{L}_{\langle \text{name} \rangle}$	$\mathcal{T}_{\langle \text{name} \rangle}$
pt (1.2)	constant $c$ (unary rel $C$ )	$\emptyset$ ( $\exists!x C(x)$ )
finsub (1.2)	unary relation $D$	$\bigvee_{n \geq 1} \exists z_0, \dots, z_{n-1} \forall w (D(w) \leftrightarrow \bigvee_{i < n} (w = z_i))$
LO (1.2)	binary relation $<$	$\forall x (x \not< x), \forall x \forall y \forall z ((x < y < z) \rightarrow (x < z))$ $\forall x \forall y ((x = y) \vee (x < y) \vee (y < x))$
sep (3.18)	unary relations $\{U_i\}_{i \in \mathbb{N}}$	$\forall x \forall y [x \neq y \rightarrow \bigvee_i (U_i(x) \leftrightarrow \neg U_i(y))]$
LN (3.23)	unary functions $\{f_i\}_{i \in \mathbb{N}}$	$\forall x \forall y \bigvee_i (f_i(x) = y)$
LN <sub>bij</sub> (4.2)	unary functions $\{f_i\}_{i \in \mathbb{N}}$	$\forall x \forall y \bigvee_i (f_i(x) = y), \forall y \bigwedge_i \exists!x (f_i(x) = y)$
LN <sub>invol</sub> (4.2)	unary functions $\{f_i\}_{i \in \mathbb{N}}$	$\forall x \forall y \bigvee_i (f_i(x) = y), \forall x \bigwedge_i (f_i(f_i(x)) = x)$
$\Gamma$ (a monoid) (4.2)	unary functions $\{a_\gamma\}_{\gamma \in \Gamma}$	$\forall x \forall y \bigvee_\gamma (a_\gamma(x) = y)$ $\forall x (a_1(x) = x), \forall x \bigwedge_{\gamma, \gamma'} (a_{\gamma\gamma'}(x) = a_\gamma(a_{\gamma'}(x)))$
FM <sub>fun</sub> (4.2)	unary functions $\{a_\gamma\}_{\gamma \in \mathbb{N}^{<\omega}}$	$\mathcal{T}_{\mathbb{N}^{<\omega}}$
FM <sub>bij</sub> (4.2)	unary functions $\{a_\gamma\}_{\gamma \in \mathbb{F}_\omega}$	$\mathcal{T}_{\mathbb{F}_\omega}$
FM <sub>invol</sub> (4.2)	unary functions $\{a_\gamma\}_{\gamma \in \mathbb{Z}_2^{*\omega}}$	$\mathcal{T}_{\mathbb{Z}_2^{*\omega}}$
color <sub>2</sub> (4.6)	binary relations $\{C_k\}_{k \in \mathbb{N}}$	$\bigwedge_k \forall x \forall y (C_k(x, y) \leftrightarrow C_k(y, x))$ $\forall x \forall y [x \neq y \rightarrow \bigvee_k (C_k(x, y) \wedge \bigwedge_{\ell \neq k} \neg C_\ell(x, y))]$ $\forall x_0, x_1, y_0, y_1$ $[\bigvee_{i, j \in 2} (x_i = y_j) \wedge \bigvee_{i \in 2} \bigwedge_{j \in 2} (x_i \neq y_j)$ $\rightarrow \bigwedge_\ell \neg (C_\ell(x_0, x_1) \wedge C_\ell(y_0, y_1))]$
color <sub>&lt;\omega</sub> (4.7)	$n$ -ary relations $\{C_{nk}\}_{\substack{2 \leq n \in \mathbb{N} \\ k \in \mathbb{N}}}$	$\bigwedge_{n, k} \bigwedge_{\sigma \in \text{Sym}(n)} \forall x_0, \dots, x_{n-1}$ $[C_{nk}(\vec{x}) \leftrightarrow C_{nk}(x_{\sigma(0)}, \dots, x_{\sigma(n-1)})]$ $\bigwedge_n \forall x_0, \dots, x_{n-1} [\bigwedge_{i \neq j < n} (x_i \neq x_j)$ $\rightarrow \bigvee_k (C_{nk}(\vec{x}) \wedge \bigwedge_{\ell \neq k} \neg C_{n\ell}(\vec{x}))]$ $\bigwedge_n \forall x_0, \dots, x_{n-1}, y_0, \dots, y_{n-1}$ $[\bigvee_{i, j < n} (x_i = y_j) \wedge \bigvee_{i < n} \bigwedge_{j < n} (x_i \neq y_j)$ $\rightarrow \bigwedge_k \neg (C_{nk}(\vec{x}) \wedge C_{nk}(\vec{y}))]$
$Y$ (a set) (2.24)	constants $\{c_y\}_{y \in Y}$	$\forall x \bigvee_{y \in Y} (x = c_y), \bigwedge_{y \neq z \in Y} (c_y \neq c_z)$
enum (4.21)	$\bigoplus_{Y \leq \omega} \mathcal{L}_Y$	$\bigoplus_{Y \leq \omega} \mathcal{T}_Y$
inf (3.35)	$\emptyset$	$\bigwedge_n \exists x_0 \dots \exists x_{n-1} \bigwedge_{i \neq j} (x_i \neq x_j)$
marker (4.18)	unary relations $\{A_i\}_{i \in \mathbb{N}}$	$\bigwedge_i \forall x (A_{i+1}(x) \rightarrow A_i(x))$ $\bigwedge_i \exists x A_i(x), \forall x \bigvee_i \neg A_i(x)$
lfgraph (4.29)	binary relation $G$	$\forall x \neg (xE x), \forall x \forall y (xE y \rightarrow yE x)$ $\forall x \bigvee_n \exists y_0 \dots y_{n-1} \forall z (xG z \rightarrow \bigvee_{i < n} (z = y_i))$
lfgraph+ $\omega$ color (4.29)	binary relation $G$ unary relations $\{C_i\}_{i \in \omega}$	$\mathcal{T}_{\text{lfgraph}}$ $\forall x \bigvee_i (C_i(x) \wedge \bigwedge_{j \neq i} \neg C_j(x))$ $\forall x \forall y [xG y \rightarrow \bigwedge_i \neg (C_i(x) \wedge C_i(y))]$

**Table 4.1:** Theories of some common combinatorial structures (with reference to first mention).

shows the *a priori* obvious interpretations (solid arrows) between the above 6 theories:

$$(4.3) \quad \begin{array}{ccc} \mathcal{T}_{\mathbb{Z}_2^{*\omega}} = \mathcal{T}_{\text{FM}_{\text{invol}}} & \longrightarrow & \mathcal{T}_{\text{LN}_{\text{invol}}} \\ \uparrow & & \uparrow \\ \mathcal{T}_{\mathbb{F}_\omega} = \mathcal{T}_{\text{FM}_{\text{bij}}} & \longleftarrow & \mathcal{T}_{\text{LN}_{\text{bij}}} \\ \uparrow & & \uparrow \text{ (dashed)} \\ \mathcal{T}_{\mathbb{N}^{<\omega}} = \mathcal{T}_{\text{FM}_{\text{fun}}} & \longleftarrow & \mathcal{T}_{\text{LN}_{\text{fun}}} = \mathcal{T}_{\text{LN}} \end{array}$$

The vertical arrows on the right are identity interpretations; e.g. a covering family of involutions is already a covering family of bijections. The two right-to-left interpretations are also essentially identity interpretations; e.g., a transitive  $\mathbb{F}_\omega$ -action is already a covering family of bijections. (Note that we do not have a similar interpretation  $\mathcal{T}_{\text{LN}_{\text{invol}}} \rightarrow \mathcal{T}_{\text{FM}_{\text{invol}}}$ , as compositions of involutions may no longer be involutions.) The vertical arrows on the left are induced by surjective homomorphisms between the acting groups/monoids; e.g., a transitive  $\mathbb{Z}_2^{*\omega}$ -action becomes a transitive  $\mathbb{F}_\omega$ -action, by regarding  $\mathbb{Z}_2^{*\omega}$  as a quotient of  $\mathbb{F}_\omega$ . And the left-to-right interpretations define a transitive action by closing a covering family of functions under composition (and inverses for  $\mathcal{T}_{\text{LN}_{\text{bij}}}$ ); e.g., given a covering family of functions  $f_i$ , we get a transitive  $\mathbb{N}^{<\omega}$ -action by defining the action of an element  $s = s_0 s_1 \dots s_{n-1} \in \mathbb{N}^{<\omega}$  to be the function  $f_{s_0} \circ f_{s_1} \circ \dots \circ f_{s_{n-1}}$ .

The proof of the Feldman–Moore theorem [FM77] (see also [Gao09, 7.1.4]) now amounts to:

**Proposition 4.4** (Feldman–Moore). *There is an interpretation  $\mathcal{T}_{\text{LN}_{\text{invol}}} \rightarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ .*

*Proof.* We must uniformly define a model of  $\mathcal{T}_{\text{LN}_{\text{invol}}}$ , i.e., a covering family of involutions, given a family of Lusin–Novikov functions  $(f_i)_{i \in \mathbb{N}} \models \mathcal{T}_{\text{LN}}$  and a family of separating unary predicates  $(U_i)_{i \in \mathbb{N}} \models \mathcal{T}_{\text{sep}}$ . First we define from the  $f_i$ 's a covering family of partial injections  $(f_{ij})_{i,j \in \mathbb{N}}$  by

$$f_{ij} := f_i \cap f_j^{-1}, \quad \text{i.e.,} \quad f_{ij}(x) = y \iff (f_i(x) = y) \wedge (f_j(y) = x).$$

Now we use the separating family of subsets  $U_k$  to define a covering family of partial injections  $(f_{ijk})_{i,j,k \in \mathbb{N}}$ , each with disjoint domain and range:

$$f_{ijk}(x) = y \iff (f_{ij}(x) = y) \wedge U_k(x) \wedge \neg U_k(y).$$

Finally, we extend the domain of each  $f_{ijk}$  to get a covering family of involutions

$$g_{ijk}(x) = y \iff (f_{ijk}(x) = y) \vee (f_{jik}(y) = x) \vee (x \notin \text{dom}(f_{ijk}) \wedge y \notin \text{dom}(f_{jik}) \wedge x = y),$$

where of course  $x \in \text{dom}(f) \iff \exists z(f(x) = z)$ . □

Thus, when combined with  $\mathcal{T}_{\text{sep}}$ , all 6 theories in (4.3) become mutually interpretable. This follows abstractly from Corollary 3.27 and the fact that all 6 theories structure every CBER (by the Feldman–Moore theorem); the point of the above Proposition is that the *proof* of Feldman–Moore is already essentially an interpretation.

Less obviously, the following shows that we may in fact interpret  $\mathcal{T}_{\text{LN}_{\text{bij}}}$  in  $\mathcal{T}_{\text{LN}}$  alone, without  $\mathcal{T}_{\text{sep}}$  (dashed arrow in (4.3)). In other words, the original statement of the Feldman–Moore theorem [FM77], in terms of Borel group actions, may be proved using only the Lusin–Novikov theorem, without also using a countable separating family of Borel sets. (We will show in the following



subsection that this is false for  $\mathcal{T}_{\text{LN}_{\text{invol}}}$ , i.e., the stronger version of Feldman–Moore yielding Borel involutions *cannot* be proved using just Lusin–Novikov.) It follows that the bottom 4 theories in (4.3) are mutually interpretable.

**Proposition 4.5.** *There is an interpretation  $\mathcal{T}_{\text{LN}_{\text{bij}}} \rightarrow \mathcal{T}_{\text{LN}}$ .*

*Proof.* We must define a covering family of bijections given a covering family of functions  $(f_i)_{i \in \mathbb{N}} \models \mathcal{T}_{\text{LN}}$ . As in Proposition 4.4, start by taking the covering family of partial bijections  $f_{ij} := f_i \cap f_j^{-1}$ . It now suffices to show how, from a single partial bijection  $f$ , we may define two total bijections  $g, h$  such that  $f \subseteq g \cup h$  (as graphs of functions); we may then apply this to each of the  $f_{ij}$ 's.

Informally, we consider the orbits  $\{\dots, f^{-1}(x), x, f(x), f^2(x), \dots\}$  of  $f$ . Each orbit is either

- (i) finite and periodic, in which case  $f$  is already a bijection on that orbit; or
- (ii) finite and aperiodic, i.e.,  $\{x, f(x), \dots, f^n(x)\}$  where  $x \notin \text{im}(f)$  and  $f^n(x) \notin \text{dom}(f)$ ; or
- (iii) bi-infinite, with  $f$  already a bijection on that orbit; or
- (iv) forward-infinite, i.e.,  $\{x, f(x), f^2(x), \dots\}$  where  $x \notin \text{im}(f)$ ; or
- (v) backward-infinite, i.e.,  $\{\dots, f^{-2}(x), f^{-1}(x), x\}$  where  $x \notin \text{dom}(f)$ .

On orbits of types (i) and (iii), we take  $g$  to be  $f$ ; on orbits of type (ii), we take  $g$  to be  $f$  extended with  $f^n(x) \mapsto x$ . Then  $g$  already covers  $f$  on these orbits, so we may take  $h$  to be the identity. On orbits of type (iv), we take  $g$  to be the involution swapping  $x$  with  $f(x)$ ,  $f^2(x)$  with  $f^3(x)$ , etc., and  $h$  to be the involution fixing  $x$  and swapping  $f(x)$  with  $f^2(x)$ ,  $f^3(x)$  with  $f^4(x)$ , etc; then  $f \subseteq g \cup h$  on these orbits. On orbits of type (v), we perform the reverse construction.

Formally, the above-described  $g, h$  are defined in terms of  $f$  by the formulas

$$\begin{aligned}
g(x) = y &: \iff \left( \left[ \bigvee_{n \in \mathbb{N}} (x \notin \text{dom}(f^n)) \leftrightarrow \bigvee_{m \in \mathbb{N}} (x \notin \text{im}(f^m)) \right] \right. \\
&\quad \wedge \left[ (f(x) = y) \vee \bigvee_{n \geq 0} ((f^n(y) = x) \wedge (x \notin \text{dom}(f) \wedge y \notin \text{im}(f))) \right] \\
&\quad \vee \left( \bigwedge_{n \in \mathbb{N}} (x \in \text{dom}(f^n)) \wedge \bigvee_{m \in \mathbb{N}} [(x \in \text{im}(f^m) \setminus \text{im}(f^{m+1})) \wedge (f^{(-1)^m}(x) = y)] \right) \\
&\quad \vee \left( \bigwedge_{m \in \mathbb{N}} (y \in \text{im}(f^m)) \wedge \bigvee_{n \in \mathbb{N}} [(y \in \text{dom}(f^n) \setminus \text{dom}(f^{n+1})) \wedge (f^{(-1)^n}(y) = x)] \right), \\
h(x) = y &: \iff \left( \left[ \left( \bigvee_{n \in \mathbb{N}} (x \notin \text{dom}(f^n)) \leftrightarrow \bigvee_{m \in \mathbb{N}} (x \notin \text{im}(f^m)) \right) \vee (x \notin \text{dom}(f) \cap \text{im}(f)) \right] \wedge (x = y) \right) \\
&\quad \vee \left( \bigwedge_{n \in \mathbb{N}} (x \in \text{dom}(f^n)) \wedge \bigvee_{m \in \mathbb{N}} [(x \in \text{im}(f^m) \setminus \text{im}(f^{m+1})) \wedge (f^{(-1)^m}(y) = x)] \right) \\
&\quad \vee \left( \bigwedge_{m \in \mathbb{N}} (y \in \text{im}(f^m)) \wedge \bigvee_{n \in \mathbb{N}} [(y \in \text{dom}(f^n) \setminus \text{dom}(f^{n+1})) \wedge (f^{(-1)^n}(x) = y)] \right).
\end{aligned}$$

The resulting interpretation  $\mathcal{T}_{\text{LN}_{\text{bij}}} \rightarrow \mathcal{T}_{\text{LN}}$  is thus given by  $\omega \times \omega$  many copies  $g_{ij}, h_{ij}$  of the above formulas for each  $i, j < \omega$ , with  $f$  above replaced by the partial bijections  $f_{ij}$ .  $\square$

## 4.B Edge-colorings and the intersection graph on finite subsets

The Feldman–Moore theorem is sometimes stated with a more combinatorial flavor, namely, that any CBER  $E \subseteq X^2$  has a Borel edge-coloring  $c : E \setminus (=X) \rightarrow \omega$  of the complete graph on each  $E$ -class; see e.g., [Kec24, 3.5]. Indeed, such a coloring  $c$  can be obtained from a countable family of

Borel involutions  $g_i$  covering  $E$  by taking  $c(x, y) :=$  the least  $i \in \omega$  such that  $g_i(x) = y$ . Conversely, the color classes of any such coloring  $c$  (extended by the identity to make them total) clearly yield a covering family of involutions.

We formally encode such an edge-coloring  $c$  of a complete graph into a theory  $\mathcal{T}_{\text{color}_2}$ , in a language with binary relations  $C_k$  consisting of all edges colored  $k$ , i.e.,  $c(x, y) = k \iff C_k(x, y)$ ; see Table 4.1 for a precise axiomatization. The equivalence to Feldman–Moore now amounts to

**Proposition 4.6.**  $\mathcal{T}_{\text{color}_2}$  and  $\mathcal{T}_{\text{LN}_{\text{invol}}}$  are mutually interpretable.  $\square$

A well-known lemma due to Kechris–Miller [KM04, 7.3] shows how to extend a  $\mathcal{T}_{\text{color}_2}$ -structuring of a CBER  $E$  to a Borel  $\omega$ -coloring  $c$  of the intersection graph on *all* finite subsets of  $E$ -classes. We encode such a coloring into a theory  $\mathcal{T}_{\text{color}_{<\omega}}$ , formally consisting again of  $n$ -ary relations  $C_{nk}$  satisfied by subsets of cardinality  $n$  that are colored  $k$ , i.e.,  $c(\{x_0, \dots, x_{n-1}\}) = k \iff C_{nk}(x_0, \dots, x_{n-1})$ ; see Table 4.1 for a precise axiomatization. The Kechris–Miller argument amounts to

**Proposition 4.7** (Kechris–Miller). *There is an interpretation  $\mathcal{T}_{\text{color}_{<\omega}} \rightarrow \mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$ .*

*Proof.* Let  $c : Y^2 \setminus (=Y) \rightarrow \omega$  be an edge coloring of the complete graph on a set  $Y$  equipped with a linear order. We define the color of a finite set  $\{x_0, \dots, x_{n-1}\}$  by first arranging the elements in increasing order (i.e., choose the unique  $\sigma \in \text{Sym}(n)$  such that  $x_{\sigma(i)} < x_{\sigma(j)}$  for all  $i < j$ ), and then recording the colors of all pairs, keeping track of this order:

$$c(\{x_0, \dots, x_{n-1}\}) := (c(x_{\sigma(i)}, x_{\sigma(j)}))_{(i,j) \in n^2} \in \mathbb{N}^{n^2}.$$

To verify that this is indeed a coloring, let  $(a_0 < \dots < a_{n-1}), (b_0 < \dots < b_{n-1})$  be tuples that intersect (so  $a_i = b_j$  for some  $i, j \in n$ ) but have the same color (so  $c(a_i, a_j) = c(b_i, b_j)$  for all  $i, j \in n$ ). Then  $a_i = b_j \implies c(a_i, a_j) = c(b_i, b_j) = c(b_i, a_i) \implies a_j = b_i$ , so  $i = j$ . But then for any  $k \in n$ ,  $c(a_i, a_k) = c(b_i, b_k) = c(a_i, b_k) \implies a_k = b_k$ , so in fact  $\vec{a} = \vec{b}$ .

Fixing bijections  $b_n : \mathbb{N}^{n^2} \cong \mathbb{N}$  for each  $n$ , we can therefore interpret  $\mathcal{T}_{\text{color}_{<\omega}}$  in  $\mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$  by

$$C_{n, b_n(\gamma)}(x_0, \dots, x_{n-1}) :\iff \bigvee_{\sigma \in \text{Sym}(n)} \left( \bigwedge_{i < j} (x_{\sigma(i)} < x_{\sigma(j)}) \wedge \bigwedge_{i, j \in n} (c(x_{\sigma(i)}, x_{\sigma(j)}) = \gamma_{ij}) \right)$$

for every  $n \in \mathbb{N}, \gamma \in \mathbb{N}^{n^2}$ .  $\square$

Together with Propositions 4.6 and 4.4 and (1.6), this yields  $\mathcal{T}_{\text{color}_{<\omega}} \rightarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , i.e., every CBER  $E$  is  $\mathcal{T}_{\text{color}_{<\omega}}$ -structurable, i.e., has a Borel  $\omega$ -coloring of the intersection graph of finite subsets of  $E$ -classes, as the result is usually stated.

Note that the argument above uses  $\mathcal{T}_{\text{LO}}$  rather than  $\mathcal{T}_{\text{sep}}$ ; the former theory is strictly weaker than the latter, by Proposition 4.14 below. Nonetheless, in the presence of  $\mathcal{T}_{\text{color}_2}$ , they turn out to be mutually interpretable:

**Proposition 4.8.** *There is an interpretation  $\mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$ .*

*Proof.* Define for each  $k \in \mathbb{N}$  a formula

$$U_k(x) :\iff \exists z((x < z) \wedge (c(x, z) = k)).$$

It suffices to show that the  $U_k$ 's separate points in any model  $(c, \leq)$  of  $\mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$  (where  $c$  is the coloring of pairs). So, let  $x \neq y$  and say without loss of generality that  $x < y$ . Then for  $k = c(x, y)$  we have  $U_k(x)$  (choose  $z = y$ ), but  $\neg U_k(y)$ , as  $c(y, z) = k = c(x, y) \implies z = x < y$ .  $\square$

In fact, it is “almost” true that  $\mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_{\text{color}_{<\omega}}$  (which implies the above by Proposition 4.7). The only counterexample is a model of  $\mathcal{T}_{\text{color}_{<\omega}}$  with only 2 indistinguishable points; see Proposition 4.11. But in models of size 3 or larger, we may define a countable separating family from a coloring of finite subsets, or indeed a coloring of subsets of size  $\leq 3$ , whose theory we call  $\mathcal{T}_{\text{color}_{\leq 3}}$  (which can be axiomatized similarly to  $\mathcal{T}_{\text{color}_{<\omega}}$  in Table 4.1).

**Proposition 4.9.** *There is an interpretation  $\mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_{\text{color}_{\leq 3}} \sqcup \{\exists x_0, x_1, x_2 \bigwedge_{i \neq j \in \{0,1,2\}} (x_i \neq x_j)\}$ .*

*Proof.* Define for each  $i, j \in \mathbb{N}$  a formula

$$U_{ij}(x) := \exists z, z' ((c(x, z) = i) \wedge (c(x, z, z') = j)).$$

Let  $c$  be an  $\omega$ -coloring of the subsets of size  $\leq 3$  of some underlying set with  $\geq 3$  elements. Let  $x \neq y$  and fix any  $w \neq x, y$ . Put  $i = c(x, w)$  and  $j = c(x, w, y)$ , so that  $U_{ij}(x)$ . If  $U_{ij}(y)$  as well, then there exist  $z, z'$  such that  $c(y, z) = i$  and  $c(y, z, z') = j$ . But then  $c(y, z, z') = j = c(x, w, y) \implies \{z, z'\} = \{x, w\}$ , so  $\{x, w\}$  intersects  $\{y, z\}$ , and  $c(x, w) = i = c(y, z) \implies \{x, w\} = \{y, z\}$ , contradicting that  $x, y, w$  are all distinct.  $\square$

**Corollary 4.10.**  *$\mathcal{T}_{\text{color}_{\leq 3}}$  and  $\mathcal{T}_{\text{color}_{<\omega}}$  are mutually interpretable.*

*Proof.* Clearly  $\mathcal{T}_{\text{color}_{\leq 3}} \rightarrow \mathcal{T}_{\text{color}_{<\omega}}$ ; conversely, we have  $\mathcal{T}_{\text{color}_{<\omega}} \rightarrow \mathcal{T}_{\text{color}_{\leq 3}}$  using the composite of  $\mathcal{T}_{\text{color}_{<\omega}} \rightarrow \mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$  from Proposition 4.7,  $\mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{sep}}$  from (1.4), and  $\mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_{\text{color}_{\leq 3}}$  from Proposition 4.9 in models of size  $\geq 3$ , and the trivial  $\mathcal{T}_{\text{color}_{<\omega}}$ -coloring in models of size  $\leq 2$ .  $\square$

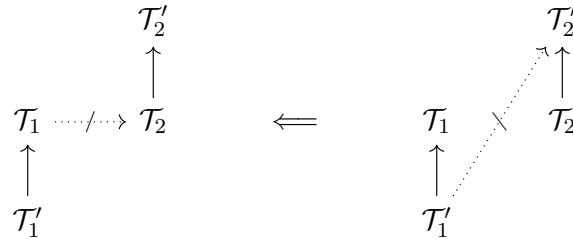
Figure 4.17 shows the interpretability relations between the theories we have considered thus far. Note that these theories all describe structures available “for free” on every CBER. The strongest such theory is  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , by Corollary 3.27. It follows from Proposition 4.8 that this theory is mutually interpretable with  $\mathcal{T}_{\text{color}_2} \sqcup \mathcal{T}_{\text{LO}}$ , which hence also interprets every theory available for free on every CBER. Moreover, by Proposition 4.9, the theory  $\mathcal{T}_{\text{color}_{<\omega}}$  of countable colorings of finite subsets is “almost” equivalent in strength as well; we indicate this with the ‘ $\approx$ ’ in Figure 4.17.

We now verify that no other interpretability relations (not implied by Figure 4.17 and transitivity) hold between these theories. Our primary tool for showing the non-existence of an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  is to examine the possible automorphisms of models: recall (2.22) that  $\alpha$  induces a “reduct” map  $\alpha^* : \text{Mod}(\mathcal{T}') \rightarrow \text{Mod}(\mathcal{T})$  equivariant under the logic action, hence in particular

$$\text{Aut}(\mathcal{M}) \subseteq \text{Aut}(\alpha^* \mathcal{M}) \quad \forall \mathcal{M} \models \mathcal{T}'.$$

So to show  $\mathcal{T} \not\rightarrow \mathcal{T}'$ , it suffices to produce a model  $\mathcal{M}$  of  $\mathcal{T}'$  along with an automorphism  $g \in \text{Aut}(\mathcal{M})$  that cannot be an automorphism of any model of  $\mathcal{T}$ . More generally, it suffices to have a model  $\mathcal{M}$  of  $\mathcal{T}'$  such that for any model  $\mathcal{N}$  of  $\mathcal{T}$  on the same set, there is some  $g \in \text{Aut}(\mathcal{M}) \setminus \text{Aut}(\mathcal{N})$ .

We minimize the number of non-interpretations needed to verify Figure 4.17 by noting that it is enough to show non-interpretations from weak theories to strong theories. That is, to prove  $\mathcal{T}_1 \not\rightarrow \mathcal{T}_2$ , it suffices to prove  $\mathcal{T}'_1 \not\rightarrow \mathcal{T}'_2$ , for any  $\mathcal{T}'_1 \rightarrow \mathcal{T}_1$  and any  $\mathcal{T}'_2 \leftarrow \mathcal{T}_2$ .



Here and in Figure 4.17, dotted arrows between theories indicate non-existence of interpretations. We also do not need to check non-interpretations from coproduct theories, as (by the universal property of the coproduct, see Definition 2.27) an interpretation  $\mathcal{T}_1 \sqcup \mathcal{T}_2 \rightarrow \mathcal{T}$  is equivalent to a pair of interpretations  $\mathcal{T}_1 \rightarrow \mathcal{T} \leftarrow \mathcal{T}_2$ . It suffices then to verify the six non-interpretations shown in Figure 4.17 (below the dashed line, and excluding  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}}$  for which see Section 4.C):

**Proposition 4.11.** *There is no interpretation  $\mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{color}_{<\omega}}$ .*

*Proof.* We may trivially color the subsets of  $2 = \{0, 1\}$  so that we have an automorphism flipping 0 and 1, which clearly is not an automorphism of any linear order on 2.  $\square$

The preceding model of size 2 is the only possible witness to  $\mathcal{T}_{\text{LO}} \not\rightarrow \mathcal{T}_{\text{color}_{<\omega}}$ , by Proposition 4.9. Since finite equivalence classes of CBERs are usually considered “trivial”, in order for a non-interpretation to be viewed as comparing the relative strengths of Borel combinatorial structures, it would be preferable to have infinite models as counterexamples, which we provide in the following.

**Proposition 4.12.** *There is no interpretation  $\mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{color}_2}$ .*

*Proof.* The complete graph on  $\mathbb{Q} - \{0\}$  with the coloring  $c(x, y) := xy$  has an automorphism  $x \mapsto -x$ , which is not an automorphism of any linear order on  $\mathbb{Q} - \{0\}$ .  $\square$

**Remark 4.13.** Since the above counterexample is an infinite model, it follows from Proposition 4.9 that also  $\mathcal{T}_{\text{color}_{\leq 3}} \not\rightarrow \mathcal{T}_{\text{color}_2}$ .

**Proposition 4.14.**  *$\mathcal{T}_{\text{FM}_{\text{invol}}} \sqcup \mathcal{T}_{\text{LO}}$  does not interpret  $\mathcal{T}_{\text{sep}}$  or  $\mathcal{T}_{\text{color}_2}$ .*

*Proof.* The connected graph  $xGy : \iff |x - y| = 1$  on  $(\mathbb{Z}, \leq)$  with the edge coloring  $c(n, m) := \min(n, m) \bmod 2$  yields a model of  $\mathcal{T}_{\text{FM}_{\text{invol}}} \sqcup \mathcal{T}_{\text{LO}}$ , where the  $i$ th generator in  $\mathbb{Z}_2^{*\omega}$  flips edges of color  $i$ , and it has a nontrivial automorphism  $h : x \mapsto x + 2$ . But models of  $\mathcal{T}_{\text{sep}}$  are rigid, so there is no interpretation  $\mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}_{\text{FM}_{\text{invol}}} \sqcup \mathcal{T}_{\text{LO}}$ .

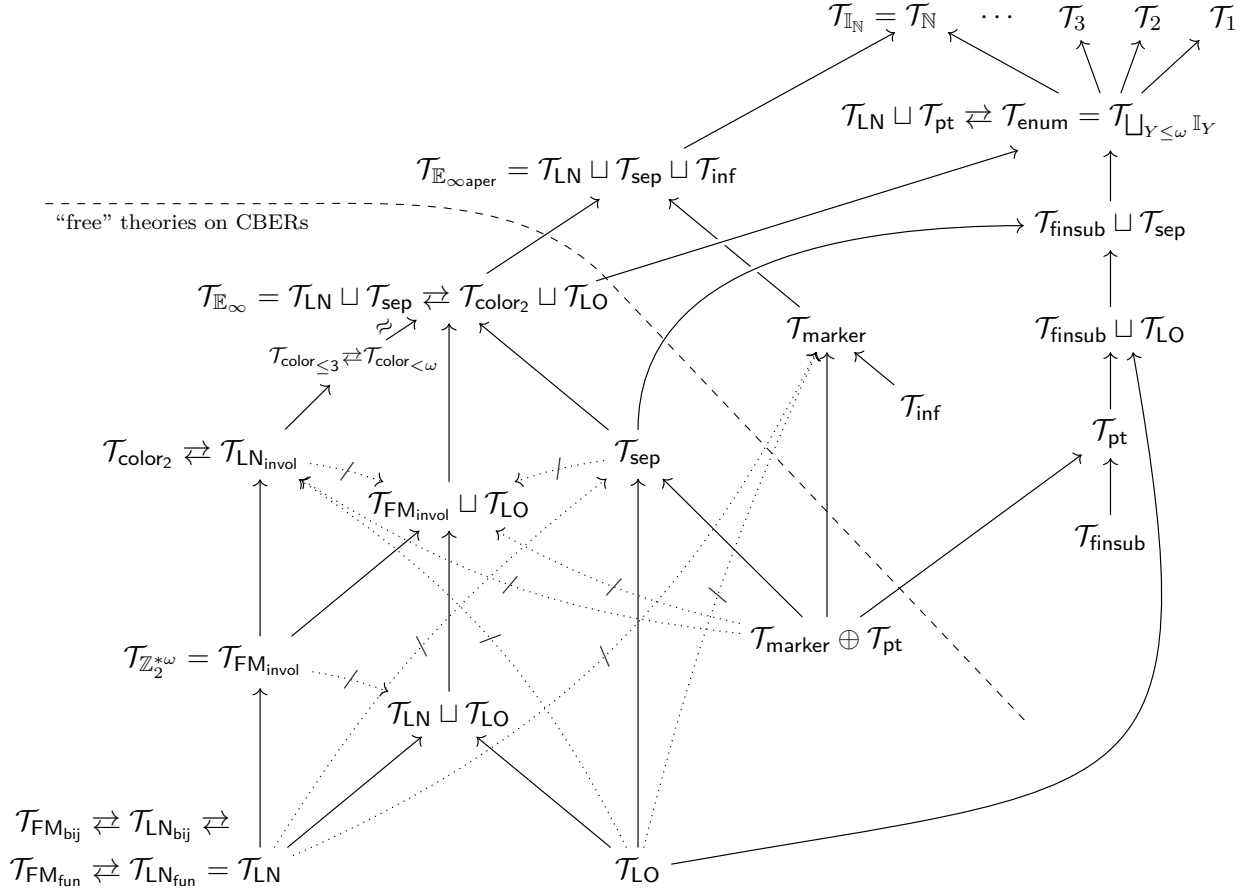
Moreover,  $h$  cannot be an automorphism of an edge coloring  $c'$  of the complete graph on  $\mathbb{Z}$ , since we would have  $c'(h(0), h(2)) = c'(2, 4) \neq c'(0, 2)$ . So there is also no interpretation  $\mathcal{T}_{\text{color}_2} \rightarrow \mathcal{T}_{\text{FM}_{\text{invol}}} \sqcup \mathcal{T}_{\text{LO}}$ .  $\square$

**Proposition 4.15.** *There is no interpretation  $\mathcal{T}_{\text{FM}_{\text{invol}}} \rightarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{LO}}$ .*

*Proof.* Consider  $\mathbb{Z}$  equipped with the usual linear order  $\leq$  and the translation action  $a : (\mathbb{Z}, +) \curvearrowright \mathbb{Z}$ . Then  $\mathcal{M} = (\mathbb{Z}, a, \leq)$  is a model of  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{LO}}$ . Note that automorphisms of  $\mathcal{M}$  are translations by elements of  $\mathbb{Z}$  and therefore all automorphism orbits are infinite. On the other hand, for any transitive action  $b : \mathbb{Z}_2^{*\omega} \curvearrowright \mathbb{Z}$ , there is some order-2 generator  $\gamma \in \mathbb{Z}_2^{*\omega}$  with non-trivial action  $b_\gamma \neq \text{id}$ . So in particular, there exists  $x \in \mathbb{Z}$  with  $b_\gamma(x) \neq x$ , but then  $b_\gamma$  is not preserved by the automorphism of  $\mathcal{M}$  that translates by  $b_\gamma(x) - x$ .  $\square$

**Proposition 4.16.** *There is no interpretation  $\mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}_{\text{sep}}$ .*

*Proof.*  $\mathcal{T}_{\text{sep}}$  has  $2^{\mathbb{N}}$  as a model, whereas all models of  $\mathcal{T}_{\text{LN}}$  are countable.  $\square$



**Figure 4.17:** Interpretability relations between some  $\mathcal{L}_{\omega_1\omega}$  theories commonly used to structure CBERs (flipped version of right half of Figure 1.9), along with non-interpretabilities witnessed by models (see text).

#### 4.C Other examples

Another important result in the theory of CBERs is the Slaman–Steel marker lemma [SS16, Lemma 1] (see also [Gao09, 7.1.5]), which says that every aperiodic CBER  $E$  (i.e., every  $E$ -class is infinite) has a **vanishing sequence of markers**: a decreasing sequence of Borel subsets  $X \supseteq A_0 \supseteq A_1 \supseteq A_2 \supseteq \dots$  which are nonempty in each  $E$ -class but have  $\bigcap_{n \in \omega} A_n = \emptyset$ . We axiomatize such a sequence by a theory  $\mathcal{T}_{\text{marker}}$ ; see Table 4.1. The proof of the marker lemma amounts to:

**Proposition 4.18** (Slaman–Steel). *There is an interpretation  $\mathcal{T}_{\text{marker}} \rightarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \sqcup \mathcal{T}_{\text{inf}}$ .*

Here, as in Example 3.35,  $\mathcal{T}_{\text{inf}}$  asserts that the underlying set is infinite; see again Table 4.1.

*Proof.* Let  $\mathcal{M} = (M, (f_i)_{i \in \mathbb{N}}, (U_i)_{i \in \mathbb{N}}) \models \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \sqcup \mathcal{T}_{\text{inf}}$ . Recall from (1.4) that the  $U_i$ ’s correspond to an injection  $u : M \hookrightarrow 2^{\mathbb{N}}$ , along which we can pull back the lexicographical ordering to  $M$ :

$$x <_{\text{lex}} y \iff \bigvee_i (\neg U_i(x) \wedge U_i(y) \wedge \bigwedge_{j < i} (U_j(x) \leftrightarrow U_j(y))).$$

Similarly, for any  $n \in \omega$  we may define a lexicographical pre-order comparing the first  $n$  bits:

$$x <_{\text{lex}}^n y \iff \bigvee_{i < n} (\neg U_i(x) \wedge U_i(y) \wedge \bigwedge_{j < i} (U_j(x) \leftrightarrow U_j(y))).$$

Now, there are two cases to consider. In Case 1,  $M$  contains a  $\leq_{\text{lex}}$ -least element, i.e., an element  $y \in M$  such that  $\mathcal{M} \models \forall z (y \leq_{\text{lex}} z)$ . Then the Lusin–Novikov functions  $f_i$  give an enumeration of  $M$ , namely  $(f_i(y))_{i \in \omega}$ , which we may disjointify into a bijective enumeration  $(g_i(y))_{i \in \omega}$  using  $\mathcal{T}_{\text{inf}}$ :

$$(4.19) \quad g_i(y) = x \iff \bigvee_{j < \omega} ((f_j(y) = x) \wedge \text{“}j \text{ is the } i\text{th natural s.t. } \bigwedge_{k < j} (f_j(y) \neq f_k(y))\text{”})$$

The final segments of this enumeration then form a vanishing sequence of markers  $(A_n)_n$ :

$$A_n(x) \iff \exists y (\forall z (y \leq_{\text{lex}} z) \wedge \bigvee_{i \geq n} (g_i(y) = x)).$$

In Case 2,  $M$  contains no  $\leq_{\text{lex}}$ -least element. Then there is some  $\leq_{\text{lex}}$ -least element  $\xi \in \overline{u(M)}$ , which is not in  $u(M)$ . We may then define the marker sequence  $(A_n)_n$  by taking the basic clopen neighborhoods of  $\xi \in 2^{\mathbb{N}}$ , pulled back along  $u$ . The  $n$ th basic neighborhood of  $\xi$  consists of strings agreeing with  $\xi$  up to the first  $n$  bits; this holds for  $u(x)$  iff  $u(x)|n \leq_{\text{lex}} u(z)|n$  for every  $z \in M$ . So

$$A_n(x) \iff \forall z (x \leq_{\text{lex}}^n z).$$

Putting the two cases together, we can interpret  $\mathcal{T}_{\text{marker}}$  in  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \sqcup \mathcal{T}_{\text{inf}}$  by

$$\begin{aligned} A_n(x) \iff & \exists y (\forall z (y \leq_{\text{lex}} z) \wedge \bigvee_{i \geq n} (g_i(y) = x)) \\ & \vee [\neg \exists y \forall z (y \leq_{\text{lex}} z) \wedge \forall z (x \leq_{\text{lex}}^n z)]. \end{aligned} \quad \square$$

Note that in this argument, we only used the Lusin–Novikov functions in Case 1 in order to define an enumeration from a point. If we stop Case 1 after defining the  $\leq_{\text{lex}}$ -least point, the argument now shows that from  $\mathcal{T}_{\text{sep}}$  alone, we may define *either* a marker sequence, *or* a single point, i.e., a model of the product theory  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}}$  (recall Definition 2.28), thereby allowing us to encompass the main content of the marker lemma within the “free” region of Figure 4.17:

**Corollary 4.20** (of proof). *There is an interpretation  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{sep}}$ .*  $\square$

The formulas (4.19) in Case 1 above show more generally that we may define from a single distinguished point  $y$  and a family of Lusin–Novikov functions  $(f_i)_i$  a bijective  $\omega$ -enumeration, provided the underlying set is infinite; in other words,  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{pt}} \sqcup \mathcal{T}_{\text{inf}}$  interprets the theory  $\mathcal{T}_\omega$  of  $\omega$ -enumerated sets from Example 2.24. It is also easily seen that if the underlying set has finite size  $Y < \omega$ , then the same formulas yield an interpretation of  $\mathcal{T}_Y$ . Thus, letting  $\mathcal{T}_{\text{enum}} := \bigoplus_{Y \leq \omega} \mathcal{T}_Y$  be the **theory of enumerated sets**, Case 1 of the above argument essentially shows that

**Proposition 4.21.** *There is an interpretation  $\mathcal{T}_{\text{enum}} \rightarrow \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{pt}}$ .*  $\square$

This formalizes a common pattern in many Borel combinatorics arguments, of which the above proof of the marker lemma is an example: one attempts to define some kind of structure (axiomatized by a theory  $\mathcal{T}$ ) on all classes of a CBER, which fails on a set on which the CBER is smooth (so one has only defined an interpretation of  $\mathcal{T} \oplus \mathcal{T}_{\text{pt}}$ ), on which one instead enumerates every class using Lusin–Novikov (yielding an interpretation of  $\mathcal{T} \oplus \mathcal{T}_{\text{enum}}$  by composing with the preceding result), allowing one to easily define any structure in a Borel way. This last step is made precise by

**Proposition 4.22.**  *$\mathcal{T}_{\text{enum}}$  interprets any theory  $\mathcal{T}$  with models of every countable cardinality.*

*Proof.* For each  $Y \leq \omega$ , take a model of  $\mathcal{T}$  on  $Y$ , which gives by Example 2.24 an interpretation  $\mathcal{T} \rightarrow \mathcal{T}_Y$ ; these combine into an interpretation into the product theory  $\mathcal{T} \rightarrow \bigoplus_{Y \leq \omega} \mathcal{T}_Y = \mathcal{T}_{\text{enum}}$ .  $\square$

It follows that  $\mathcal{T}_{\text{enum}}$  is mutually interpretable with  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{pt}}$ , as shown in Figure 4.17.

**Remark 4.23.** The four theories immediately below  $\mathcal{T}_{\text{enum}}$  in Figure 4.17 all yield smoothness on CBERs, as discussed in the introduction (see Example 1.2). As noted there, in the absence of  $\mathcal{T}_{\text{LN}}$ , they are strictly increasing in strength, and strictly weaker than  $\mathcal{T}_{\text{enum}}$ , by arguments like those in the preceding subsection: for example,  $\mathcal{T}_{\text{finsub}} \sqcup \mathcal{T}_{\text{LO}} \not\Leftarrow \mathcal{T}_{\text{pt}} \sqcup \mathcal{T}_{\text{LO}}$  by Example 1.2, but does not interpret  $\mathcal{T}_{\text{enum}}$  since it has non-rigid models.

Figure 4.17 also shows how  $\mathcal{T}_{\text{marker}}$  and  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}}$  relate to the other theories describing “free” structures on CBERs considered in the preceding subsections. Note that no such “free” theory interprets  $\mathcal{T}_{\text{marker}}$ , since not every CBER is aperiodic. Conversely, none of the free theories in Figure 4.17 are interpreted by  $\mathcal{T}_{\text{marker}}$  (hence also not by  $\mathcal{T}_{\text{marker}} \sqcup \mathcal{T}_{\text{pt}}$ ): it suffices to show

**Proposition 4.24.** *There are no interpretations  $\mathcal{T}_{\text{LO}} \rightarrow \mathcal{T}_{\text{marker}}$  or  $\mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}_{\text{marker}}$ .*

*Proof.* Note that in a model  $\mathcal{M} = (M, A_0, A_1, \dots)$  of  $\mathcal{T}_{\text{marker}}$ , there may be points that are not in the first marker set  $A_0$ . Any permutation  $g \in \text{Sym}(M \setminus A_0)$  extends to an automorphism of  $\mathcal{M}$  (e.g., define  $g(x) = x$  for  $x \in A_0$ ). If there are at least 2 points  $x \neq y$  in  $M \setminus A_0$ , then the automorphism swapping  $x, y$  shows that  $\mathcal{T}_{\text{LO}} \not\Leftarrow \mathcal{T}_{\text{marker}}$ . Similarly, if there are 3 distinct points  $x, y, z \in M \setminus A_0$ , then an automorphism  $g$  fixing  $x$  and swapping  $y, z$  cannot be an automorphism of a model of  $\mathcal{T}_{\text{LN}}$ , since a Lusin–Novikov function  $f_i$  cannot map  $x$  to both  $y, z$ .  $\square$

**Remark 4.25.** A model  $\mathcal{M} \models \mathcal{T}_{\text{marker}}$  is essentially just a labeled countable partition (the differences of adjacent marker sets) with infinitely many nonempty pieces. These pieces are the automorphism orbits of  $\mathcal{M}$ ; within each piece, an automorphism can act arbitrarily. So more generally, in order for a theory  $\mathcal{T}$  to be interpretable in  $\mathcal{T}_{\text{marker}}$ , it must admit a model  $\mathcal{M}$  with a countable partition within which all permutations are automorphisms of  $\mathcal{M}$ .

Using this observation, we can also see that, while the marker lemma can essentially (modulo  $\mathcal{T}_{\text{pt}}$ ) be proved from  $\mathcal{T}_{\text{sep}}$ , it cannot be proved from any of the other “free” theories in Figure 4.17:

**Proposition 4.26.** *There is no interpretation  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{FM}_{\text{invol}}} \sqcup \mathcal{T}_{\text{LO}}$ .*

*Proof.* Consider the structure  $\mathcal{M} = (\mathbb{Z}, \leq, c)$  described in Proposition 4.14, which has only two automorphism orbits and therefore cannot have a model of  $\mathcal{T}_{\text{marker}}$  as a reduct. Moreover,  $\mathcal{M}$  has automorphisms with no fixed points, so cannot have a model of  $\mathcal{T}_{\text{pt}}$  as a reduct.  $\square$

**Proposition 4.27.** *There is no interpretation  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{color}_2}$ .*

*Proof.* Consider the complete graph on  $\mathbb{Z}_2^{\oplus \omega}$  with edge coloring  $c(x, y) := x + y$ . This is a model of  $\mathcal{T}_{\text{color}_2}$  with only one automorphism orbit, since translation by each  $a \in \mathbb{Z}_2^{\oplus \omega}$  is an automorphism, so its reducts can model neither  $\mathcal{T}_{\text{pt}}$  nor  $\mathcal{T}_{\text{marker}}$ .  $\square$

**Remark 4.28.** Similarly, there is no interpretation  $\mathcal{T}_{\text{marker}} \oplus \mathcal{T}_{\text{pt}} \rightarrow \mathcal{T}_{\text{finsub}}$  (as shown in Figure 4.17), since any model  $\mathcal{M}$  of  $\mathcal{T}_{\text{finsub}}$  has only two automorphism orbits, and if the finite set  $D$  contains at least two points then  $\mathcal{M}$  will have automorphisms with no fixed points.

It is also easy to see  $\mathcal{T}_{\text{LO}} \not\Leftarrow \mathcal{T}_{\text{pt}}$ . Together with the aforementioned non-interpretations, it follows that (the transitive closure of) Figure 4.17 gives all interpretabilities between the shown theories.

We close this section with a simple example of a different nature: a Borel combinatorial construction encoded not by structurability via a theory  $\mathcal{T}$ , but rather *expandability* along an interpretation  $\mathcal{T} \rightarrow \mathcal{T}'$ , as in Corollary 3.29.

A basic result in Borel combinatorics, due to Kechris–Solecki–Todorcevic [KST99, 4.5], shows that every locally finite Borel graph has a Borel  $\omega$ -coloring (of the vertices). We can formalize this by defining a theory  $\mathcal{T}_{\text{ifgraph}}$  of locally finite graphs (in the language  $\mathcal{L}_{\text{graph}}$  with an edge relation  $G$ ), as well as an expanded theory  $\mathcal{T}_{\text{ifgraph}+\omega\text{color}}$  of  $\omega$ -colored locally finite graphs; see Table 4.1. The result now states that given a Borel locally finite graph  $G \subseteq X^2$ , which we may treat as a  $\mathcal{T}_{\text{ifgraph}}$ -structuring of any CBER  $E \supseteq G$  (e.g., the connectedness relation of  $G$ ), there exists an expansion to a  $\mathcal{T}_{\text{ifgraph}+\omega\text{color}}$ -structuring. Via Corollary 3.29, this is witnessed by

**Proposition 4.29** (Kechris–Solecki–Todorcevic). *There is an interpretation  $\mathcal{T}_{\text{ifgraph}+\omega\text{color}} \rightarrow \mathcal{T}_{\text{ifgraph}} \sqcup \mathcal{T}_{\text{sep}}$ , whose restriction to the language  $\mathcal{L}_{\text{graph}}$  is the identity  $\mathcal{T}_{\text{ifgraph}} \rightarrow \mathcal{T}_{\text{ifgraph}} \sqcup \mathcal{T}_{\text{sep}}$ .*

*Proof.* Let  $\mathcal{M} = (G, U_i)_{i \in \omega} \models \mathcal{T}_{\text{ifgraph}} \sqcup \mathcal{T}_{\text{sep}}$ , a locally finite graph  $G$  together with a separating family  $(U_i)_{i \in \omega}$ . Then for any  $xGy$  there is some least  $i \in \mathbb{N}$  such that  $U_i$  separates  $x$  and  $y$ , call it  $i(x, y)$ . For each  $x \in M$ , let  $A_x = \{i(x, y) \mid yGx, x \in U_{i(x, y)}\}$ , a finite subset of  $\mathbb{N}$  by  $\mathcal{T}_{\text{ifgraph}}$ . If  $xGy$ , then clearly  $i(x, y) \in A_x \triangle A_y$ . So fixing a bijection  $b : \mathcal{P}_{\text{fin}}(\mathbb{N}) \cong \mathbb{N}$ , we can define a coloring of  $G$  by

$$C_{b(A)}(x) : \iff \bigwedge_{i \in \mathbb{N}} ((i \in A) \leftrightarrow \exists y Gx [U_i(x) \wedge \neg U_i(y) \wedge \bigwedge_{j < i} (U_j(x) \leftrightarrow U_j(y))])$$

(where “ $i \in A$ ” denotes the propositional constant  $\top$  if  $i \in A$ , else  $\perp$ ). □

**Remark 4.30.** The above interpretation is only interesting given the last condition on the restriction to  $\mathcal{L}_{\text{graph}}$ , which ensures that it specifies a coloring of the *original* graph. (There is trivially an interpretation  $\mathcal{T}_{\text{ifgraph}+\omega\text{color}} \rightarrow \emptyset$ , since we may take the empty graph.) Thus, it does not make sense to draw this interpretation as part of Figure 4.17.

## 5 Structurability of groupoids

We now generalize the correspondence given by Theorem 3.26 between CBERs and theories interpreting  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , to theories interpreting just  $\mathcal{T}_{\text{LN}}$ . We will show that these correspond to locally countable Borel groupoids, which admit a corresponding theory of “structurability”. Since this is a less well-studied concept than for CBERs, and since the correspondence between  $\mathcal{T}_{\text{LN}}$  and groupoids is a bit more involved, we will approach things in a different order here than in Section 3. First we will introduce groupoids and their “Scott theories”, and show that they are precisely the theories interpreting  $\mathcal{T}_{\text{LN}}$ , and then we will introduce structurability and discuss some examples.

### 5.A Groupoids

**Definition 5.1.** A **groupoid**  $(X, G, \text{dom}, \text{cod}, \text{id}, \circ, {}^{-1})$  is a category with inverses, consisting of:

- a collection  $X$  of **objects**;
- a collection  $G$  of **morphisms** or **arrows**;
- two maps  $\text{dom}, \text{cod} : G \rightrightarrows X$  (**domain** and **codomain**), where for  $g \in G$  with  $\text{dom}(g) = x$  and  $\text{cod}(g) = y$ , we write  $g : x \rightarrow y$ , and we write  $G(x, y) := \text{dom}^{-1}(x) \cap \text{cod}^{-1}(y)$  for the **hom-set** of all such  $g : x \rightarrow y$ ;



- a map  $x \mapsto \text{id}_x = 1_x : X \rightarrow G$  (**identity**), such that  $1_x : x \rightarrow x$ ;
- a map  $\circ : G \times_X G := \{(g, h) \in G \times G \mid \text{dom}(g) = \text{cod}(h)\} \rightarrow G$  (**composition**), taking  $g : y \rightarrow z$  and  $h : x \rightarrow y$  to  $g \circ h : x \rightarrow z$ , obeying the associativity and identity laws;
- a map  $g \mapsto g^{-1} : G \rightarrow G$  (**inverse**), such that for  $g : x \rightarrow y$ ,  $g^{-1} \circ g = 1_x$  and  $g \circ g^{-1} = 1_y$ .

A **standard Borel groupoid** is a groupoid such that  $X, G$  are standard Borel spaces and  $\text{dom}, \text{cod}, \circ, 1, ^{-1}$  are Borel maps. A **locally countable Borel groupoid** is a standard Borel groupoid such that  $\text{dom}$  is countable-to-1; equivalently, each hom-set as well as each connected component is countable.<sup>7</sup>

A **functor**  $f : (X, G) \rightarrow (Y, H)$  between groupoids is a homomorphism of groupoids, i.e., a pair of maps  $f : X \rightarrow Y$  and  $f : G \rightarrow H$  preserving all of the groupoid structure.

For background on category theory and groupoids, see [ML98], [Lei14]; for groupoids in the topological and Borel contexts, see [Ram90], [Alv08], [Car11], [Bow14], [Che19b], [TDW21].

**Example 5.2.** Each CBER  $(X, E)$  is a locally countable Borel groupoid, with  $\text{dom}, \text{cod} : E \rightrightarrows X$  given by the projections. Up to isomorphism, these are precisely the groupoids such that each hom-set has at most one element (sometimes called *thin* groupoids), or equivalently, each isotopy group  $G(x, x)$  is trivial.

**Example 5.3.** A one-object locally countable groupoid is just a countable group.

**Example 5.4.** More generally, given a countable group  $\Gamma$  and a Borel action  $\Gamma \curvearrowright X$ , the **action groupoid**  $\Gamma \ltimes X$  on the space of objects  $X$  has, informally, morphisms  $x \rightarrow y$  consisting of all group elements  $\gamma \in \Gamma$  such that  $\gamma \cdot x = y$ . Formally, such a morphism is encoded by the pair  $(\gamma, x)$  with  $y = \gamma \cdot x$ . We thus take  $\Gamma \ltimes X := \Gamma \times X$ , with  $\text{dom} : \Gamma \ltimes X \rightarrow X$  given by the projection and  $\text{cod}$  given by the action map. The groupoid operations are given by

$$(5.5) \quad \text{id}_x := (1_\Gamma, x), \quad (\delta, \gamma x) \circ (\gamma, x) := (\delta\gamma, x), \quad (\gamma, x)^{-1} := (\gamma^{-1}, \gamma x).$$

When  $X = 1$  with the trivial  $\Gamma$ -action, this recovers the previous example of a one-object groupoid.

In general, for any countable group action  $\Gamma \curvearrowright X$ , we have two canonical functors

$$\begin{array}{ccc} & (X, \Gamma \ltimes X) & \\ \swarrow \scriptstyle \gamma \mapsto (\gamma, x) & & \searrow \scriptstyle (\gamma, x) \mapsto (x, \gamma x) \\ (1, \Gamma) & & (X, \mathbb{E}_\Gamma^X) \end{array}$$

The latter functor is bijective on objects and surjective on morphisms, hence exhibits the orbit equivalence relation  $\mathbb{E}_\Gamma^X$  as a quotient of the action groupoid; it is an isomorphism iff the action is free. The first functor is sometimes called the *cocycle associated with the action*  $\Gamma \curvearrowright X$ .

**Example 5.6.** A **(left) action** of an arbitrary (Borel) groupoid  $(X, G)$  consists of a (Borel) space  $Y$  equipped with a (Borel) map  $p : Y \rightarrow X$ , thought of as a “bundle” over  $X$ , as well as a map

$$a : G \times_X Y := \{(g, y) \in G \times Y \mid \text{dom}(g) = p(y)\} \longrightarrow Y$$

---

<sup>7</sup>Note that this conflicts with another common usage of “locally” in category theory, to mean that each hom-set obeys said condition (e.g., *locally small category*).

taking  $g : x \rightarrow x' \in G$  and  $y \in p^{-1}(x)$  to  $a(g, y) = g \cdot y \in p^{-1}(x')$  and obeying the usual associativity and identity laws. A **right action** is defined analogously using  $Y \times_X G := \{(y, g) \mid p(y) = \text{cod}(g)\}$ .

Given a left action  $a$  on  $p : Y \rightarrow X$  as above, the **action groupoid** has space of objects  $Y$  and space of morphisms  $G \times Y := G \times_X Y$ , with domain map  $p$  and codomain map  $a$  and groupoid operations as in (5.5). We again have canonical functors

$$\begin{array}{ccc} & (Y, G \times Y) & \\ g \leftarrow (g, y) \swarrow & & \searrow (g, y) \mapsto (y, gy) \\ (X, G) & & (Y, \mathbb{E}_G^Y) \end{array}$$

where  $\mathbb{E}_G^Y$  is the connectedness relation of the groupoid  $G \times Y$ , called the **orbit equivalence relation** of the action; the latter functor to it is bijective on objects and surjective on morphisms.

The first functor  $p : (Y, G \times Y) \rightarrow (X, G)$  is a **(discrete) fibration** of groupoids, meaning that it restricts to a bijection  $\text{dom}^{-1}(y) \cong \text{dom}^{-1}(p(y))$  for each object  $y \in Y$ , i.e., each morphism in the codomain has a unique lift given any lift of its domain. In fact, the data of a fibration over  $G$  is essentially equivalent to an action of  $G$ , in that for any bundle  $p : Y \rightarrow X$ , the action groupoid construction yields a bijection between actions of  $G$  on  $Y$ , and isomorphism classes of groupoids  $H$  on  $Y$  equipped with an extension of  $p$  to a fibration  $p : (Y, H) \rightarrow (X, G)$ .

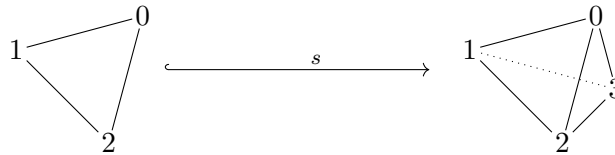
## 5.B Simplicial nerves

**Definition 5.7.** A **symmetric simplicial set**  $S = (S_n)_{0 < n < \omega}$  consists of a sequence of sets  $S_1, S_2, S_3, \dots$ , equipped with, for each function  $s : m \rightarrow n$  between  $0 < m, n < \omega$ , a map

$$\partial_s : S_n \rightarrow S_m$$

which is contravariantly functorial:  $\partial_{s \circ t} = \partial_t \circ \partial_s$  and  $\partial_{\text{id}} = \text{id}$ . In short,  $S$  is a contravariant functor from the category of positive finite ordinals (with arbitrary maps) to the category of sets. A **simplicial map**  $f : S \rightarrow S'$  between symmetric simplicial sets is a family of maps  $(f_n : S_n \rightarrow S'_n)_n$  commuting with the  $\partial_s$ , i.e., a natural transformation between functors.

We may think of each  $n + 1 = \{0, \dots, n\}$  as the vertices of a combinatorial  $n$ -simplex, and a function  $s : m + 1 \rightarrow n + 1$  as a simplicial map. For example, the inclusion  $s : 2 + 1 \hookrightarrow 3 + 1$ :



Now for a symmetric simplicial set  $S$ , we think of  $S$  as a complex obtained by gluing  $n$ -simplices, where each  $S_{n+1}$  is the set of  $n$ -simplices in  $S$ . We thus call  $S_1$  the **vertices** of  $S$ ,  $S_2$  the **edges**, etc. For  $s : m + 1 \rightarrow n + 1$ , the map  $\partial_s : S_{n+1} \rightarrow S_{m+1}$  restricts each  $n$ -simplex along  $s$  to an  $m$ -simplex. For example, when  $s = (0, 1, 2) \in (3 + 1)^{2+1}$  is the above inclusion,  $\partial_{012} : S_{3+1} \rightarrow S_{2+1}$  takes each 3-simplex in  $S$  to its face spanned by vertices 0, 1, 2. In general, when  $s : n \hookrightarrow n + 1$  is an injection, we call  $\partial_s : S_{n+1} \rightarrow S_n$  a **face map**. When  $s : n + 1 \twoheadrightarrow n$  is a surjection, we call  $\partial_s : S_n \rightarrow S_{n+1}$  a **degeneracy map**; these provide a way to regard an  $(n - 1)$ -simplex as a degenerate or “flat”  $n$ -simplex where one of the edges (namely between the two vertices collapsed by  $s$ ) is really a point.

Note that every  $s : m \rightarrow n$  is a composite of one-step injections, surjections, and permutations; thus the face and degeneracy maps, along with vertex permutations, determine all of the  $\partial_s$ .

For background on (symmetric) simplicial sets, which play a fundamental role in abstract homotopy theory, see [GJ99], [Cis19]. For our purposes in this paper, simplicial sets are relevant, because on the one hand, they are precisely the structure on the (positive-arity) type spaces  $\mathcal{S}_n(\mathcal{T})$  of a theory required to determine it up to bi-interpretability; see Remarks 2.17 and 2.25. On the other hand, a simplicial set can also encode a groupoid, via the following well-known construction:

**Definition 5.8.** For each  $0 < n < \omega$ , let  $\mathbb{I}_n = n^2$  denote the indiscrete equivalence relation on  $n$ . Given a groupoid  $(X, G)$ , its **simplicial nerve**  $\mathcal{N}(G)$  is the symmetric simplicial set with

$$\mathcal{N}(G)_n := \{\text{functors } \mathbb{I}_n \rightarrow G\}$$

and  $\partial_s : \mathcal{N}(G)_n \rightarrow \mathcal{N}(G)_m$  for  $s : m \rightarrow n$  given by precomposition with  $s$ . See e.g., [Cis19, §1.4].

Thus, the vertices  $\mathcal{N}(G)_1$  are functors  $\mathbb{I}_1 \rightarrow G$ , which are just objects  $x \in X$ . The edges  $\mathcal{N}(G)_2$  are functors  $f : \mathbb{I}_2 \rightarrow G$ , which are uniquely determined by the single morphism  $f(0, 1) \in G$ . The triangles  $f : \mathbb{I}_3 \rightarrow G$  are commuting triangles

$$(5.9) \quad \begin{array}{ccc} & f(1) & \\ g=f(0,1) \nearrow & & \searrow f(1,2)=hg^{-1} \\ f(0) & \xrightarrow{f(0,2)=h} & f(2) \end{array}$$

which are uniquely determined by a pair of morphisms  $f(0, 1), f(0, 2)$  sharing a domain.

More generally, it is easily seen that for each  $n \geq 2$ , we have a bijection

$$\begin{aligned} (\partial_{01}, \dots, \partial_{0n}) : \mathcal{N}(G)_{n+1} &\cong (\mathcal{N}(G)_2)_{\mathcal{N}(G)_1}^n \cong G_X^n = \{(g_1, \dots, g_n) \in G^n \mid \text{dom}(g_1) = \dots = \text{dom}(g_n)\} \\ f &\mapsto (f(0, 1), \dots, f(0, n)) \end{aligned}$$

between  $n$ -simplices and  $n$ -tuples of morphisms sharing a domain ( $G_X^n$  denotes the  $n$ -fold fiber product of  $\text{dom} : G \rightarrow X$ ). A symmetric simplicial set  $S$  for which  $(\partial_{01}, \dots, \partial_{0n}) : S_{n+1} \rightarrow (S_2)_{S_1}^n$  is a bijection for each  $n$  is said to satisfy the **Grothendieck–Segal condition**.<sup>8</sup>

**Proposition 5.10** (Grothendieck). *The operation  $\mathcal{N}$  is an equivalence of categories*

$$\mathcal{N} : \{\text{groupoids}\} \simeq \{\text{symmetric simplicial sets satisfying the Grothendieck–Segal condition}\}.$$

The inverse-up-to-isomorphism takes a symmetric simplicial set  $S$  to the groupoid with objects  $S_1$ , morphisms  $S_2$ , identity given by the degeneracy map  $\partial_0 : S_2 \rightarrow S_1$  (where 0 here denotes the constant map  $0 : 2 \rightarrow 1$ ), and inverse and composition both given by the face map  $\partial_{12} : S_2 \times_{S_1} S_2 \cong S_3 \rightarrow S_2$  (by taking  $h = \text{id}$  in the above triangle (5.9) to get  $g^{-1}$ , and then replacing  $g$  with  $g^{-1}$  to get  $hg$ ). For details, see [Cis19, 1.4.11].

All of the above also makes sense in the Borel context, yielding an equivalence between locally countable Borel groupoids and **locally countable standard Borel symmetric simplicial sets**  $S$ , meaning that each  $S_n$  is standard Borel and each face map is Borel and countable-to-1, which satisfy the Grothendieck–Segal condition.

<sup>8</sup>There are different versions of the Grothendieck–Segal condition, all equivalent for groupoids, where instead of  $n$  morphisms sharing a domain, we may specify  $f : \mathbb{I}_n \rightarrow G$  by its restriction to any spanning tree of  $\mathbb{I}_n$ .

## 5.C Scott theories of groupoids

**Definition 5.11.** Let  $(X, G)$  be a locally countable Borel groupoid. The **Scott theory** of  $G$  is the countable  $\mathcal{L}_{\omega_1\omega}$  theory  $(\mathcal{L}_G, \mathcal{T}_G)$  defined uniquely up to bi-interpretability via Theorem 2.26 by declaring its models on any nonempty countable set  $Y$  to be the standard Borel space

$$\text{Mod}_Y(\mathcal{T}_G) := \{u : (Y, \mathbb{I}_Y) \rightarrow (X, G) \mid u \text{ is a fibration}\},$$

where  $\mathbb{I}_Y$  is the indiscrete equivalence relation on  $Y$ ; put also  $\text{Mod}_\emptyset(\mathcal{T}_G) := \emptyset$ . The logic action  $\text{Sym}(Y, Z) \times \text{Mod}_Y(\mathcal{T}_G) \rightarrow \text{Mod}_Z(\mathcal{T}_G)$  is given by relabeling: for  $h : Y \cong Z$  and  $u \in \text{Mod}_Y(\mathcal{T}_G)$ ,

$$h \cdot u := u \circ h^{-1}.$$

**Remark 5.12.** Recalling (Example 5.6) that fibrations over  $G$  correspond to actions, the fibrations from an indiscrete equivalence relation correspond to the free transitive  $G$ -actions, sometimes called *principal  $G$ -actions* or  *$G$ -torsors*;  $\mathcal{T}_G$  is the theory of all such.

For each  $x \in X$ , we have a canonical model  $(\mathcal{H}_G)_x$  of  $\mathcal{T}_G$  on  $\text{dom}^{-1}(x)$ , given by the fibration

$$\begin{aligned} (\mathcal{H}_G)_x &:= \text{cod}_x : (\text{dom}^{-1}(x), \mathbb{I}_{\text{dom}^{-1}(x)}) \longrightarrow (X, G) \\ \text{dom}^{-1}(x)^2 \ni (g_0, g_1) &\longmapsto g_1 g_0^{-1} : \text{cod}(g_0) \rightarrow \text{cod}(g_1). \end{aligned}$$

Any other  $u \in \text{Mod}_Y(\mathcal{T}_G)$  is isomorphic to such a canonical model; namely, for any  $y_0 \in Y$ ,

$$\begin{aligned} u_{y_0} : Y &\longrightarrow \text{dom}^{-1}(u(y_0)) \\ y &\longmapsto u(y_0, y) \end{aligned}$$

is an isomorphism  $u_{y_0} : u \cong (\mathcal{H}_G)_{u(y_0)}$ . Moreover, all isomorphisms  $h : u \cong (\mathcal{H}_G)_x$  are of this form for a unique  $y_0 \in Y$ , namely  $y_0 = h^{-1}(1_x)$ ; this fact is the **Yoneda lemma**.

Below in Definition 5.26, we will define a “structuring” of a groupoid  $G$ , and then explain how the models  $(\mathcal{H}_G)_x$  together form a “tautological  $\mathcal{T}_G$ -structuring”  $\mathcal{H}_G$ , with the same universal properties as in the case of CBERs (Definition 3.10).

Following Proposition 3.12, we now compute the  $n$ -types of  $\mathcal{T}_G$ :

**Proposition 5.13.** *The  $n$ -types of  $\mathcal{T}_G$  are determined by the Borel isomorphism*

$$\begin{aligned} \text{tp}_{\mathcal{H}_G}^n : G_X^n / G &\cong \mathcal{S}_n(\mathcal{T}_G) \\ [(x, g_0, \dots, g_{n-1})] &\mapsto \text{tp}(\text{cod}_x, g_0, \dots, g_{n-1}), \end{aligned}$$

where  $G_X^n$  is the set of  $n$ -tuples in some  $\text{dom}^{-1}(x)$ , equipped with the right translation  $G$ -action:

$$\begin{aligned} G_X^n &= \{(x, g_0, \dots, g_{n-1}) \in X \times G^n \mid x = \text{dom}(g_0) = \dots = \text{dom}(g_{n-1})\}, \\ (x, g_0, \dots, g_{n-1}) \cdot (g : x' \rightarrow x) &= (x', g_0 g, \dots, g_{n-1} g). \end{aligned}$$

(The first coordinate  $x$  is only needed for the case  $n = 0$ , yielding  $G_X^0 / G = X / G$ .)

*Proof.* For any model  $u \in \text{Mod}_Y(\mathcal{T}_G)$ , i.e., fibration  $u : \mathbb{I}_Y \rightarrow G$ , and tuple  $\vec{a} \in Y^n$ , we have  $u_{y_0} : (u, \vec{a}) \cong (\text{cod}_{u(y_0)}, u_{y_0}(\vec{a}))$  for any  $y_0 \in Y$  by the preceding remark; and all isomorphisms between such  $n$ -pointed models are realized by a right translation by the Yoneda lemma, hence they become identified in  $G_X^n / G$ . Thus  $\text{tp}_{\mathcal{H}_G}^n$  is a bijection; since  $G_X^n / G, \mathcal{S}_n(\mathcal{T}_G)$  are quotients by Polish group actions (the former being a quotient by a CBER  $\mathbb{E}_G^{G_X^n}$ ), this implies Borel isomorphism.  $\square$

**Remark 5.14.** For  $n \geq 1$ , we have a bijection

$$\begin{aligned} \mathcal{N}(G)_n = \{\text{functors } \mathbb{I}_n \rightarrow G\} &\cong G_X^{n-1} && \cong G_X^n/G \\ f &\mapsto (f(0), f(0, 1), \dots, f(0, n-1)) && \mapsto [(f(0), f(0, 0), \dots, f(0, n-1))] \\ ((i, j) &\mapsto g_j g_i^{-1}) && \longleftarrow \text{-----} \longrightarrow [(x, g_0, \dots, g_{n-1})] \end{aligned}$$

where the first bijection witnesses the Grothendieck–Segal condition of the nerve (Definition 5.8) and the second bijection is given by quotienting out the first factor of  $G$ . It is easily seen that the simplicial boundary maps  $\partial_s : \mathcal{N}(G)_n \rightarrow \mathcal{N}(G)_m$  of  $\mathcal{N}(G)$  correspond to the variable substitutions  $G_X^n \rightarrow G_X^m$  for each  $s : m \rightarrow n$ ; in other words, we have an isomorphism of symmetric simplicial sets. Combined with the preceding result, we get a composite isomorphism

$$\begin{aligned} (\mathcal{N}(G)_n \cong \mathcal{S}_n(\mathcal{T}_G))_{n \geq 1} \\ f \mapsto \text{tp}(\text{cod}_{f(0)}, f(0, 0), \dots, f(0, n-1)). \end{aligned}$$

In particular, we get that the positive-arity type spaces  $(\mathcal{S}_n(\mathcal{T}_G))_{n \geq 1}$  of a Scott theory  $\mathcal{T}_G$  satisfy the Grothendieck–Segal condition.

**Theorem 5.15.** *For a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the following are equivalent:*

- (i)  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \rightarrow \mathcal{S}_0(\mathcal{T})$  is surjective,  $\mathcal{S}_1(\mathcal{T})$  is standard Borel, and for any pointed model  $(\mathcal{M}, a) \in \text{Mod}_Y^1(\mathcal{T})$ , the map  $b \mapsto \text{tp}(\mathcal{M}, a, b) : Y \rightarrow \mathcal{S}_2(\mathcal{T})$  is injective.
- (ii)  $\mathcal{T}$  has no empty models, and for any countable set  $Y$ , the logic action  $\text{Sym}(Y) \curvearrowright \text{Mod}_Y^1(\mathcal{T})$  is free and smooth.
- (iii)  $\mathcal{T}$  has no empty models, and there is a countable set  $\mathcal{F} \subseteq \mathcal{L}_{\omega_1\omega}^1$  of formulas in one variable, such that every pointed model of  $\mathcal{T}$  is rigid and  $\mathcal{F}$ -categorical.
- (iv)  $\mathcal{T}$  has no empty models and interprets  $\mathcal{T}_{\text{LN}}$ .
- (v)  $\mathcal{T}$  is bi-interpretable with the Scott theory  $\mathcal{T}_G$  of a locally countable Borel groupoid  $G$ , namely the groupoid whose nerve is  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$ .

*Proof.* (i)–(iv) are equivalent by Proposition 3.25.

(v)  $\implies$  (i): By Proposition 5.13,  $\partial_1 : \mathcal{S}_1(\mathcal{T}_G) \rightarrow \mathcal{S}_0(\mathcal{T}_G)$  is isomorphic to the projection  $G_X^1/G \twoheadrightarrow G_X^0/G$  (i.e., to  $X \twoheadrightarrow X/G$  by Remark 5.14). By Remark 5.14,  $\mathcal{S}_1(\mathcal{T}_G) \cong G_X^1 = X$  is standard Borel. By Remark 5.12, every pointed model of  $\mathcal{T}_G$  is isomorphic to some  $\text{cod}_x \in \text{Mod}_{\text{dom}^{-1}(x)}(\mathcal{T}_G)$  equipped with some  $a \in \text{dom}^{-1}(x)$ , and we may assume  $a = 1_x$  by otherwise applying the isomorphism  $(-)^{a^{-1}} : \text{cod}_x \cong \text{cod}_{\text{cod}(a)}$ ; then the model  $(\text{cod}_x, a) = (\text{cod}_x, 1_x)$  is rigid by the Yoneda lemma.

(i)  $\implies$  (v): Note first that rigidity of pointed models implies that the symmetric simplicial set  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$  obeys the Grothendieck–Segal condition. Indeed, to check that the map

$$(\partial_{01}, \dots, \partial_{0n}) : \mathcal{S}_{n+1}(\mathcal{T}) \longrightarrow \mathcal{S}_2(\mathcal{T})_{\mathcal{S}_1(\mathcal{T})}^n$$

from Definition 5.8 is a bijection, since this map commutes with the projection  $\partial_0$  to  $\mathcal{S}_1(\mathcal{T})$  in the first coordinate, it suffices to restrict both domain and codomain to those types realized in a fixed pointed model  $(\mathcal{M}, a) \in \text{Mod}_Y^1(\mathcal{T})$ ; but since pointed models are rigid, said restriction becomes simply the canonical bijection

$$\{a\} \times Y^n \longrightarrow (\{a\} \times Y)^n.$$

Also by Corollary 2.16,  $\mathcal{S}_n(\mathcal{T})$  is standard Borel for all  $n \geq 1$ . Thus by Proposition 5.10,  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$  is isomorphic to the nerve  $\mathcal{N}(G)$  of a locally countable Borel groupoid  $(X, G)$ , with objects  $X = \mathcal{S}_1(\mathcal{T})$  and morphisms  $G = \mathcal{S}_2(\mathcal{T})$ . Together with Remark 5.14, we get an isomorphism with  $(\mathcal{S}_n(\mathcal{T}_G))_{n \geq 1}$ :

$$\begin{array}{ccccccc}
\cdots & & \mathcal{S}_3(\mathcal{T}) & \xrightarrow{\quad \cong \quad} & \mathcal{S}_2(\mathcal{T}) & \xrightarrow{\quad \cong \quad} & \mathcal{S}_1(\mathcal{T}) & \xrightarrow{\partial_1} & \mathcal{S}_0(\mathcal{T}) \\
& & \text{5.10} \parallel & & \parallel & & \parallel & & \\
\cdots & & \mathcal{N}(G)_3 & \xrightarrow{\quad \cong \quad} & \mathcal{N}(G)_2 = G & \xrightarrow[\text{cod}]{\text{dom}} & \mathcal{N}(G)_1 = X & \longrightarrow & X/G \\
& & \text{5.14} \parallel & & \parallel & & \parallel & & \parallel \text{5.13} \\
\cdots & & \mathcal{S}_3(\mathcal{T}_G) & \xrightarrow{\quad \cong \quad} & \mathcal{S}_2(\mathcal{T}_G) & \xrightarrow{\quad \cong \quad} & \mathcal{S}_1(\mathcal{T}_G) & \longrightarrow & \mathcal{S}_0(\mathcal{T}_G)
\end{array}$$

By Proposition 5.13, we also have a canonical isomorphism  $X/G \cong \mathcal{S}_0(\mathcal{T}_G)$  (of nonstandard Borel spaces), commuting with the two face maps  $G \rightrightarrows X$  and  $\mathcal{S}_2(\mathcal{T}_G) \rightrightarrows \mathcal{S}_1(\mathcal{T}_G)$ , so that both  $X/G, \mathcal{S}_0(\mathcal{T}_G)$  are the spaces of connected components of the respective simplicial sets. But the surjection  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \twoheadrightarrow \mathcal{S}_0(\mathcal{T})$  also exhibits  $\mathcal{S}_0(\mathcal{T})$  as the connected components of  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$ , since any two 1-types projecting to the same 0-type are realized in isomorphic models, hence may be amalgamated into a 2-type realized in a single model. We thus get a Borel isomorphism of type spaces  $(\mathcal{S}_n(\mathcal{T}) \cong \mathcal{S}_n(\mathcal{T}_G))_n$  for all  $n < \omega$ , hence a bi-interpretation  $\mathcal{T} \cong \mathcal{T}_G$  by Remark 2.25.  $\square$

**Remark 5.16.** If  $(X, G)$  is a locally countable Borel groupoid, then  $G$  is a CBER (up to isomorphism) iff all models of  $\mathcal{T}_G$  are rigid, since a non-identity endomorphism  $g : x \rightarrow x \in G$  gives rise to a non-identity right translation automorphism of the canonical model  $\text{cod}_x$  (Remark 5.12).

Thus for a theory  $\mathcal{T}$  interpreting both  $\mathcal{T}_{\text{LN}}$  and  $\mathcal{T}_{\text{sep}}$ , the groupoid  $G$  constructed in the above proof must be a CBER, hence by Theorem 3.26 must be the same CBER  $E$  constructed in that earlier proof. Note however that we previously constructed  $E$  on  $X = \mathcal{S}_1(\mathcal{T})$  as the kernel of  $\partial_1 : \mathcal{S}_1(\mathcal{T}) \twoheadrightarrow \mathcal{S}_0(\mathcal{T})$ , whereas the above proof instead shows that the two maps  $\mathcal{S}_2(\mathcal{T}) \rightrightarrows \mathcal{S}_1(\mathcal{T})$  yield an injection  $\mathcal{S}_2(\mathcal{T}) \hookrightarrow \mathcal{S}_1(\mathcal{T})^2$  whose image is then  $E$ . (For a general groupoid, said image will not be injective, but will still be  $\ker(\partial_1)$ , which yields the connectedness relation of the groupoid  $G$ .)

**Remark 5.17.** The simplicial set of types  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$  and (weakenings of) the Grothendieck–Segal condition have appeared before in the finitary  $(\mathcal{L}_{\omega\omega})$  model theory literature, in the context of *n-amalgamation* of types; see [GKK15], [Kru19].

In our context, we may give a homotopy-theoretic explanation of  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$  as follows, using the equivalence of 2-categories between theories and standard Borel groupoids of models from [Che19a] (see the following subsection). Given a theory  $\mathcal{T}$  obeying the conditions of Theorem 5.15, let  $\mathcal{T}_1$  be  $\mathcal{T}$  expanded with a constant, so that  $\text{Mod}(\mathcal{T}_1) \cong \text{Mod}^1(\mathcal{T})$  with standard Borel  $\cong$ -quotient  $\mathcal{S}_1(\mathcal{T})$ . The inclusion  $\mathcal{T} \hookrightarrow \mathcal{T}_1$  induces an essentially surjective forgetful functor between groupoids of models  $\text{Mod}(\mathcal{T}_1) \twoheadrightarrow \text{Mod}(\mathcal{T})$ , thereby covering  $\text{Mod}(\mathcal{T})$  with the essentially discrete standard Borel groupoid  $\text{Mod}(\mathcal{T}_1) \simeq \mathcal{S}_1(\mathcal{T})$ . The Čech nerve of this functor (see [Lur09, 6.1.2.11]) is  $(\mathcal{S}_n(\mathcal{T}))_{n \geq 1}$ , from which we recover  $\text{Mod}(\mathcal{T})$  as the realization (2-categorical quotient).

## 5.D Connection to essential countability and imaginaries

In this subsection, which is tangential to the rest of the paper, we discuss connections between Theorem 5.15 above and some other notions and results that have appeared in the countable model theory literature.

**Definition 5.18.** For a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the isomorphism relation  $\cong \subseteq \text{Mod}_Y(\mathcal{T})$  is said to be **essentially countable** if it is Borel and there is a Borel set  $D \subseteq \text{Mod}_Y(\mathcal{T})$  picking at least one and only countably many elements from each isomorphism class; see [Kec24, §4].

**Theorem 5.19** (Hjorth–Kechris [HK96, 4.3]). *For a countable  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the following are equivalent:*

- (i) *For any countable set  $Y$ , the isomorphism relation  $\cong \subseteq \text{Mod}_Y(\mathcal{T})$  is essentially countable.*
- (ii) *There is a countable set  $\mathcal{F}$  of  $\mathcal{L}_{\omega_1\omega}$  formulas (of various arities), such that for any countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$ , there is a finite tuple  $\vec{a} \in Y^n$  such that  $(\mathcal{M}, \vec{a})$  is  $\mathcal{F}$ -categorical.*

**Remark 5.20.** By the proof of [HK96, 4.3(iii)  $\implies$  (ii)], if the above conditions hold, then in fact (ii) may be “uniformly witnessed”: there are formulas  $\phi_n(x_0, \dots, x_{n-1}) \in \mathcal{F}$  for each  $n$ , such that

- (a) any  $n$ -pointed model  $(\mathcal{M}, \vec{a})$  of  $\mathcal{T}$  satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$  must be  $\mathcal{F}$ -categorical; and
- (b) any model  $\mathcal{M}$  contains an  $n$ -tuple  $\vec{a}$  (for some  $n$ ) satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$ .

Moreover, whenever  $(\mathcal{M}, \vec{a})$  satisfies (a), then so does  $(\mathcal{M}, \vec{b})$  for any tuple  $\vec{b}$  containing all the elements of  $\vec{a}$ .

Note the similarity between Theorem 5.19 and our Proposition 3.25 and Theorem 5.15. There are two essential differences between the two characterizations. First, in our results we only consider single-pointed models, while Theorem 5.19 considers models with some finite tuple fixed. Second, we impose rigidity in addition to categoricity of models with respect to some countable fragment  $\mathcal{F}$ . We now give a detailed explanation of these two differences: the first will amount to a change of language (see Remark 5.23). The second, our rigidity requirement, is more substantial:

**Remark 5.21.** For a theory  $\mathcal{T}$  and  $n$ -pointed model  $(\mathcal{M}, \vec{a}) \in \text{Mod}_Y^n(\mathcal{T})$ , we have

$$\text{Aut}(\mathcal{M}) = \bigcup_{\vec{b} \in Y^n} \text{Iso}((\mathcal{M}, \vec{a}), (\mathcal{M}, \vec{b})),$$

where each nonempty set of isomorphisms  $\text{Iso}((\mathcal{M}, \vec{a}), (\mathcal{M}, \vec{a}))$  is a coset of  $\text{Aut}(\mathcal{M}, \vec{a}) \subseteq \text{Aut}(\mathcal{M})$ . Thus, the following are equivalent, for any theory  $\mathcal{T}$ :

- (i) Every countable model of  $\mathcal{T}$  has only countably many automorphisms.
- (ii) Every countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$  becomes rigid after fixing a finite tuple  $\vec{a} \in Y^n$ .

**Lemma 5.22.** *If the above conditions hold for a theory  $\mathcal{T}$ , then for each  $n < \omega$ , the class of  $n$ -pointed rigid models is definable by an  $\mathcal{L}_{\omega_1\omega}$  formula  $\phi_n(x_0, \dots, x_{n-1})$ .*

*Proof.* For any countable set  $Y$  and  $n < \omega$ , the isomorphism relation  $\cong \subseteq \text{Mod}_Y^n(\mathcal{T})^2$  is Borel by Lusin–Novikov, since the logic action  $\text{Sym}(Y) \times \text{Mod}_Y^n(\mathcal{T}) \rightarrow \text{Mod}_Y^n(\mathcal{T})$  is countable-to-1, since each  $n$ -pointed model has only countably many automorphisms. Now the set of rigid  $n$ -pointed models is

$$\{(\mathcal{M}, \vec{a}) \mid \forall b, c \in Y (b \neq c \implies (\mathcal{M}, \vec{a}, b) \not\cong (\mathcal{M}, \vec{a}, c))\} \subseteq \text{Mod}_Y^n(\mathcal{T}),$$

which is Borel and obviously isomorphism-invariant, so the claim follows from Lopez-Escobar.  $\square$

**Remark 5.23.** It follows from this lemma and Remark 5.20 that if a theory  $\mathcal{T}$  has both essentially countable  $\cong$  and also countable automorphism groups, i.e., the analogue of our characterization 3.25 of theories interpreting  $\mathcal{T}_{\text{LN}}$ , with pointed models replaced by models with some fixed finite tuple, then we may find formulas  $\phi_n(x_0, \dots, x_{n-1})$  uniformly defining these tuples:

- (a) any  $n$ -pointed model  $(\mathcal{M}, \vec{a}) \in \text{Mod}_Y^n(\mathcal{T})$  satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$  is rigid and  $\mathcal{F}$ -categorical; and
- (b) any model  $\mathcal{M}$  contains an  $n$ -tuple  $\vec{a}$  satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$ .

Thus, any model  $\mathcal{M} \in \text{Mod}_Y^n(\mathcal{T})$  becomes rigid and  $\mathcal{F}$ -categorical after fixing an element of (not the underlying set  $Y$  but) the “uniformly definable set”

$$Y \times \bigsqcup_n \phi_n^{\mathcal{M}} =: A^{\mathcal{M}}.$$

Note that we may recover the underlying set  $Y$  from this set  $A^{\mathcal{M}}$ , by quotienting out the second coordinate. Thus, if we also lift each of the relations  $R^{\mathcal{M}}$  of  $\mathcal{M}$  (for  $R \in \mathcal{L}$ ) to a relation on  $A^{\mathcal{M}}$ , we obtain a structure  $\alpha^*(\mathcal{M})$  “uniformly defined” from  $\mathcal{M}$ , from which  $\mathcal{M}$  can in turn be “uniformly” recovered up to definable isomorphism, and such that  $\alpha(\mathcal{M})$  obeys the conditions of Proposition 3.25 characterizing theories interpreting  $\mathcal{T}_{\text{LN}}$ . Note, however, that this operation  $\mathcal{M} \mapsto \alpha^*(\mathcal{M})$  is not given by an interpretation  $\alpha$  in the sense we have considered in this paper (Definition 2.21), since  $\alpha^*(\mathcal{M})$  lives not on the original underlying set  $Y$ , but rather on the above set  $A^{\mathcal{M}}$ , which is “uniformly defined” from  $\mathcal{M}$  by the formal expression

$$A := \top(x) \times \bigsqcup_n \phi_n(x_0, \dots, x_n)$$

(here  $\top(x)$  denotes a tautology in one variable, whose interpretation in  $\mathcal{M}$  is the underlying set  $Y$ ).

Recall that a (model-theoretic) **imaginary sort** over a theory  $\mathcal{T}$  is an expression such as this one, that denotes a set uniformly definable from any model  $\mathcal{M} \models \mathcal{T}$ , but that may be external to the underlying set of  $\mathcal{M}$ ; see [Hod93, Ch. 7]. In  $\mathcal{L}_{\omega_1\omega}$ , the natural notion of imaginary is built from formulas by taking formal Cartesian products, countable disjoint unions, and definable quotients; see [Che19a, §4]. A (model-theoretic) **interpretation**  $\alpha : (\mathcal{L}, \mathcal{T}) \rightarrow (\mathcal{L}', \mathcal{T}')$  between two (one-sorted)  $\mathcal{L}_{\omega_1\omega}$  theories maps the single (base) sort of  $\mathcal{T}$  to an imaginary sort  $A$  of  $\mathcal{T}'$ , and each  $\mathcal{L}$ -formula  $\phi(x_0, \dots, x_{n-1})$  to a definable subsort of the  $n$ -fold Cartesian product of imaginaries  $A^n$ ; see [Che19a, §10] (the notion was essentially introduced in [HTMM18]). We will also call these **imaginary interpretations**, to disambiguate from the more restrictive notion from Definition 2.21 we have considered thus far, consisting of imaginary interpretations which fix the base sort, which we will call **one-sorted interpretations** for contrast.

**Corollary 5.24** (of Theorem 5.19, Proposition 3.25, and Theorems 5.15 and 3.26). *For a countable (one-sorted)  $\mathcal{L}_{\omega_1\omega}$  theory  $\mathcal{T}$ , the following are equivalent:*

- (i) *Every countable model of  $\mathcal{T}$  has only countably many automorphisms, and for any countable set  $Y$ , the isomorphism relation  $\cong \subseteq \text{Mod}_Y(\mathcal{T})^2$  is essentially countable.*
- (ii) *There is a countable set  $\mathcal{F}$  of  $\mathcal{L}_{\omega_1\omega}$  formulas, such that for any countable model  $\mathcal{M} \in \text{Mod}_Y(\mathcal{T})$ , there is a finite tuple  $\vec{a} \in Y^n$  such that  $(\mathcal{M}, \vec{a})$  is rigid and  $\mathcal{F}$ -categorical.*
- (iii) *There is a countable set  $\mathcal{F}$  of  $\mathcal{L}_{\omega_1\omega}$  formulas, and  $\phi_n(x_0, \dots, x_{n-1}) \in \mathcal{F}$  for each  $n < \omega$ , such that (a) any  $n$ -pointed model  $(\mathcal{M}, \vec{a}) \in \text{Mod}_Y^n(\mathcal{T})$  satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$  is rigid and  $\mathcal{F}$ -categorical; and (b) any model  $\mathcal{M}$  contains an  $n$ -tuple  $\vec{a}$  satisfying  $\phi_n^{\mathcal{M}}(\vec{a})$ .*
- (iv)  *$\mathcal{T}$  admits an imaginary interpretation from  $\mathcal{T}_{\text{LN}}$ , taking the single base sort of  $\mathcal{T}_{\text{LN}}$  to an imaginary sort of  $\mathcal{T}$  which is nonempty in every model and from which the base sort of  $\mathcal{T}$  may be recovered as a definable quotient.*



(v)  $\mathcal{T}$  admits an imaginary bi-interpretation with the Scott theory  $\mathcal{T}_G$  of a locally countable Borel groupoid  $G$ .

If all models of  $\mathcal{T}$  are rigid, then  $\mathcal{T}_{\text{LN}}$  above may be replaced with  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$ , and the groupoid  $G$  will be a CBER.

*Proof.* (i)  $\iff$  (ii)  $\iff$  (iii) by Theorem 5.19 and Remarks 5.21 and 5.23.

(iii)  $\implies$  (iv): Let the interpretation take the base sort of  $\mathcal{T}_{\text{LN}}$  to  $\top(x) \times \bigsqcup_n \phi_n(x_0, \dots, x_{n-1})$ , as in Remark 5.23.

(iv)  $\implies$  (v): Let  $\alpha : \mathcal{T}_{\text{LN}} \rightarrow \mathcal{T}$  be an imaginary interpretation, taking the base sort of  $\mathcal{T}_{\text{LN}}$  to the imaginary  $A$ , which has a definable equivalence relation  $\sim \subseteq A^2$  whose quotient is definably isomorphic to the base sort of  $\mathcal{T}$ . Consider the expanded language  $\mathcal{L}' := \mathcal{L}_{\text{LN}} \sqcup \mathcal{L} \sqcup \{\sim\}$ . We may extend  $\alpha$  to an interpretation  $\alpha' : \mathcal{L}' \rightarrow (\mathcal{L}, \mathcal{T})$ , by taking  $\sim$  to the definable equivalence relation  $\sim$ , and  $\alpha'(R) \subseteq A^n$  for each  $n$ -ary  $R \in \mathcal{L}$  to be a subsort whose interpretation in each  $\mathcal{M} \models \mathcal{T}$  is the lift of  $R^{\mathcal{M}} \subseteq ((A/\sim)^{\mathcal{M}})^n$ . Then for each  $\mathcal{M} \models \mathcal{T}$ ,  $\alpha'^*(\mathcal{M})$  will be an  $\mathcal{L}'$ -structure with a definable equivalence relation  $\sim$  and  $\sim$ -invariant relations for each  $R \in \mathcal{L}$ , such that  $\mathcal{M}$  is recovered as the quotient structure  $\alpha'^*(\mathcal{M})/\sim$ ; this quotient structure is in turn specified by the interpretation  $\beta : \mathcal{L} \rightarrow \mathcal{L}'$  taking the base sort to the quotient by  $\sim$ . So letting

$$\mathcal{T}' := \mathcal{T}_{\text{LN}} \sqcup \beta(\mathcal{T}) \sqcup \{\text{“}\sim \text{ is an equivalence relation and each } R \in \mathcal{L} \text{ is } \sim\text{-invariant”}\},$$

we have that  $\alpha, \beta$  form an imaginary bi-interpretation  $\mathcal{T}' \leftrightarrow \mathcal{T}$ , where  $\mathcal{T}'$  includes  $\mathcal{T}_{\text{LN}}$  and has every model nonempty, hence is in turn one-sorted bi-interpretable with a Scott theory  $\mathcal{T}_G$  of a locally countable Borel groupoid  $G$  by Theorem 5.15.

(v)  $\implies$  (i):  $\mathcal{T}_G$  clearly obeys (i) (its groupoid of models is Borel equivalent to  $G$ , by Remark 5.12); and (i) is preserved under imaginary bi-interpretations  $\mathcal{T} \leftrightarrow \mathcal{T}_G$  (since such a bi-interpretation induces an Borel equivalence of groupoids of models; see [Che19a]).

If every model of  $\mathcal{T}$  is rigid, then models of the new theory  $\mathcal{T}'$  defined above will still be rigid (since they are bi-interpretable with models of  $\mathcal{T}$ ); thus by Theorem 3.26 we have  $\mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}} \rightarrow \mathcal{T}'$  and so  $\mathcal{T}' \cong \mathcal{T}_E$  for a CBER  $E$  (which must be the same as  $G$  by Remark 5.16).  $\square$

**Remark 5.25.** A less direct “soft” proof of (i)  $\implies$  (v) above may be given by taking  $(X, G)$  to be the locally countable Borel subgroupoid of the groupoid of models  $\bigsqcup_{N \leq \omega} (\text{Sym}(N) \times \text{Mod}_N(\mathcal{T}))$  restricted to a countable complete section, by essential countability (Definition 5.18). We then have Borel equivalences of groupoids between the models of  $\mathcal{T}_G$ ,  $G$ , and the models of  $\mathcal{T}$ , whence  $\mathcal{T}_G, \mathcal{T}$  are imaginary bi-interpretable by [Che19a].

Thus, in some sense the abstract results of this paper may be regarded as a specialization of [Che19a], which concerns arbitrary  $\mathcal{L}_{\omega_1\omega}$  theories and their standard Borel groupoids of models, to the case of theories with countable automorphism groups and essentially countable isomorphism relations, which correspond instead to locally countable Borel groupoids. In that case, a simpler recovery of the theory from the groupoid of models is possible, using the Scott theory, rather than the Becker–Kechris topological realization argument as in [Che19a].

## 5.E Structurability

Let  $(X, G)$  be a locally countable Borel groupoid,  $\mathcal{T}$  be a countable  $\mathcal{L}_{\omega_1\omega}$  theory. Imitating Corollary 3.15, we would like a  $\mathcal{T}$ -structuring of  $G$  to correspond to an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}_G$ , with the **tautological**  $\mathcal{T}_G$ -structuring  $\mathcal{H}_G$  corresponding to  $\alpha = \text{id}_{\mathcal{T}_G}$ .

To reformulate this in explicit combinatorial terms, note that (by Remark 2.25)  $\alpha$  amounts to a Borel way to put models of  $\mathcal{T}$  on models of  $\mathcal{T}_G$ , which by Remark 5.12 are up to isomorphism just the  $G$ -torsors  $\text{dom}^{-1}(x)$  for each  $x \in X$ , which is equivariant with respect to isomorphisms  $\text{dom}^{-1}(x) \cong \text{dom}^{-1}(x')$ , which by the Yoneda lemma (Remark 5.12) are given by right multiplication by  $g : x' \rightarrow x \in G$ . We thus adopt the following (cf. Definition 3.1)

**Definition 5.26.** A  $\mathcal{T}$ -**structuring** of a locally countable Borel groupoid  $(X, G)$  is a family of models  $\mathcal{M} = (\mathcal{M}_x)_{x \in X}$ , where each  $\mathcal{M}_x \in \text{Mod}_{\text{dom}^{-1}(x)}(\mathcal{T})$ , which is right- $G$ -equivariant:

$$\mathcal{M}_x = \mathcal{M}_{x'} \cdot (g : x \rightarrow x')$$

(where  $\mathcal{M}_{x'} \cdot g$  denotes the logic action of  $(- )g : \text{dom}^{-1}(x') \rightarrow \text{dom}^{-1}(x)$ ), and “Borel”, meaning:

- (i) For some (equivalently any) Borel family of enumerations  $(h_x : |\text{dom}^{-1}(x)| \cong \text{dom}^{-1}(x))_{x \in X}$ , the following map is Borel:

$$\begin{aligned} X &\longrightarrow \bigsqcup_{N \leq \omega} \text{Mod}_N(\mathcal{T}) \\ x &\longmapsto h_x^{-1} \cdot \mathcal{M}_x. \end{aligned}$$

- (ii) For any countable set  $Y$ , standard Borel space  $Z$ , Borel map  $f : Z \rightarrow X$ , and Borel family of bijections  $(h_z : Y \cong \text{dom}^{-1}(f(z)))_{z \in Z}$ , the following map is Borel:

$$\begin{aligned} Z &\longrightarrow \text{Mod}_Y(\mathcal{T}) \\ z &\longmapsto h_z^{-1} \cdot \mathcal{M}_{f(z)}. \end{aligned}$$

- (iii) For each  $n$ -ary relation symbol  $R \in \mathcal{L}$  (or more generally formula  $\phi$ ),

$$\widetilde{R}^{\mathcal{M}} := \{(x, g_0, \dots, g_{n-1}) \in G_X^n \mid R^{\mathcal{M}_x}(g_0, \dots, g_{n-1})\}$$

is a Borel subset of  $G_X^n$  (recall Proposition 5.13).

- (iv) For each  $n$ -ary  $R \in \mathcal{L}$  (assuming  $\mathcal{L}$  has no nullary relation symbols), or more generally formula  $\phi$  in at least one variable,

$$R^{\mathcal{M}} := \{\text{functors } f : \mathbb{I}_n \rightarrow G \mid R^{\mathcal{M}_{f(0)}}(f(0, 0), \dots, f(0, n-1))\}$$

is a Borel subset of  $G^{\mathbb{I}_n}$  (recall Remark 5.14).

Let  $\text{Mod}_G(\mathcal{T})$  denote the set of  $\mathcal{T}$ -structurings of  $G$ . If  $\text{Mod}_G(\mathcal{T}) \neq \emptyset$ , we call  $G$   **$\mathcal{T}$ -structurable**.

It is easily seen that the above Borelness conditions are equivalent, following Lemma 3.2.

**Definition 5.27** (cf. Definition 3.10). The **tautological  $\mathcal{T}_G$ -structuring**  $\mathcal{H}_G$  of  $G$  is given by  $(\mathcal{H}_G)_x := \text{cod}_x : \text{dom}^{-1}(x) \rightarrow G$  as in Remark 5.12.

**Definition 5.28** (cf. Definition 3.3). For a Borel fibration  $f : (X, G) \rightarrow (Y, H)$  between locally countable Borel groupoids, and a  $\mathcal{T}$ -structuring  $\mathcal{M}$  of  $H$ , the **pullback  $\mathcal{T}$ -structuring**  $f^{-1} \cdot \mathcal{M}$  of  $G$  is defined by

$$(f^{-1} \cdot \mathcal{M})_x := (f|_{\text{dom}^{-1}(x)} : \text{dom}^{-1}(x) \cong \text{dom}^{-1}(f(x)))^{-1} \cdot \mathcal{M}_{f(x)} \quad \text{for each } x \in X.$$

**Definition 5.29** (cf. Definition 3.4). For an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  between theories, and a  $\mathcal{T}'$ -structuring  $\mathcal{M}$  of  $G$ , the  $\alpha$ -**reduct**  $\alpha^*\mathcal{M}$  is the  $\mathcal{T}$ -structuring given by

$$(\alpha^*\mathcal{M})_x := \alpha^*(\mathcal{M}_x) \quad \text{for each } x \in X.$$

Following the discussion preceding Definition 5.26 and the analogous results in Section 3.B, we now easily have:

**Corollary 5.30** (cf. Corollary 3.15). *For any locally countable Borel groupoid  $(X, G)$  and theory  $(\mathcal{L}, \mathcal{T})$ , we have a bijection*

$$\begin{aligned} \{\text{interpretations } \mathcal{T} \rightarrow \mathcal{T}_G\} &\cong \text{Mod}_G(\mathcal{T}) = \{\mathcal{T}\text{-structurings of } G\} \\ \alpha &\mapsto \alpha^*(\mathcal{H}_G). \end{aligned} \quad \square$$

**Proposition 5.31** (cf. Proposition 3.16). *For any locally countable Borel groupoids  $(X, G), (Y, H)$ , we have a bijection*

$$\begin{aligned} \{\text{fibrations } G \rightarrow H\} &\cong \text{Mod}_G(\mathcal{T}_H) = \{\mathcal{T}_H\text{-structurings of } G\} \\ f &\mapsto f^{-1} \cdot \mathcal{H}_H. \end{aligned} \quad \square$$

**Corollary 5.32** (of Theorem 5.15; cf. Corollary 3.17). *We have a dual equivalence of categories*

$$\{\text{locally countable Borel groupoids, fibrations}\} \simeq \{\mathcal{L}_{\omega_{1\omega}} \text{ theories obeying 5.15, interpretations}\}$$

*taking a groupoid  $G$  to its Scott theory  $\mathcal{T}_G$ , and a fibration  $f : G \rightarrow H$  to the unique  $\alpha : \mathcal{T}_H \rightarrow \mathcal{T}_G$  such that  $\alpha^*(\mathcal{H}_G) = f^{-1} \cdot \mathcal{H}_H$ , or such that  $\alpha_n^* : \mathcal{S}_n(\mathcal{T}_G) \cong G_X^n/G \rightarrow H_Y^n/H \cong \mathcal{S}_n(\mathcal{T}_H)$  (via 5.13) is the map induced by  $f$ .*  $\square$

**Corollary 5.33** (cf. Corollary 3.27). *For two  $\mathcal{L}_{\omega_{1\omega}}$  theories  $\mathcal{T}, \mathcal{T}'$ , the following are equivalent:*

- (i) *Every  $\mathcal{T}$ -structurable locally countable Borel groupoid is  $\mathcal{T}'$ -structurable.*
- (ii) *There exists an interpretation  $\mathcal{T}' \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}}$ .*

*More generally, expandability of every  $\mathcal{T}$ -structuring along an interpretation  $\alpha : \mathcal{T} \rightarrow \mathcal{T}'$  is equivalent to factorizability of  $\alpha$  through the inclusion  $\mathcal{T} \rightarrow \mathcal{T} \sqcup \mathcal{T}_{\text{LN}}$  (cf. Corollary 3.29).*  $\square$

## 5.F Some examples

We conclude with some simple examples of structurings of groupoids. There is undoubtedly much of the general theory that remains to be developed (e.g., much of [CK18] for groupoids); our goal here is just to demonstrate the utility of the concept, by showing how it relates to some familiar notions.

**Example 5.34.** If  $G$  is a locally countable Borel groupoid on one object  $X = 1 = \{*\}$ , i.e., a countable group, then a  $\mathcal{T}$ -structuring  $\mathcal{M}$  of  $G$  is just a single model  $\mathcal{M}_* \in \text{Mod}_G(\mathcal{T})$  which is invariant under right translation.

For example, if  $\mathcal{T}$  is the theory of directed trees, i.e., connected acyclic graphs with a chosen orientation of each edge, then  $G$  is  $\mathcal{T}$ -structurable (“treeable”) iff it is a free group, in which case a  $\mathcal{T}$ -structuring  $\mathcal{M}$  amounts to a choice of free generating set, namely all those  $g \in G$  for which there is an edge  $1 \rightarrow g$  in  $\mathcal{M}$ , with  $\mathcal{M}$  then given by the Cayley graph. (If  $\mathcal{T}$  is the theory of undirected trees, then a free product of copies of  $\mathbb{Z}_2$  would also be  $\mathcal{T}$ -structurable.)

More generally, if  $\mathcal{T}$  is the theory of connected graphs, then a  $\mathcal{T}$ -structuring of a countable group is just a Cayley graph.

**Example 5.35.** For the theory of trees  $\mathcal{T}$  as above, we may also consider “treeings”  $\mathcal{M}$  of a general locally countable Borel groupoid  $(X, G)$ . By definition, this consists of a Borel right-translation-invariant family  $(\mathcal{M}_x)_{x \in X}$  of trees on each fiber  $\text{dom}^{-1}(x)$  of  $\text{dom} : G \rightarrow X$ .

We may represent this in more concrete (but less structurally transparent) terms, by interpreting the edge relation  $E$  as the set of morphisms  $E^{\mathcal{M}} \subseteq G_X^1 = G$  from Definition 5.26(iv) and Remark 5.14; the correspondence with the fiberwise relations  $E^{\mathcal{M}_x} \subseteq \text{dom}^{-1}(x)^2$  is given by

$$E^{\mathcal{M}_x}(g : x \rightarrow y, h : x \rightarrow z) \iff E^{\mathcal{M}_y}(1_y : y \rightarrow y, hg^{-1} : y \rightarrow z) \iff E^{\mathcal{M}}(hg^{-1}).$$

To say that the tree  $\mathcal{M}_x$  on each fiber is connected means that  $1_x \in \text{dom}^{-1}(x)$  is joined by a path

$$1_x = g_0 (E^{\mathcal{M}_x})^{\pm 1} g_1 (E^{\mathcal{M}_x})^{\pm 1} \dots (E^{\mathcal{M}_x})^{\pm 1} g_n = g$$

to any  $g \in \text{dom}^{-1}(x)$ , which means that we have morphisms  $h_i = (g_{i+1}g_i^{-1})^{\pm 1}$  with

$$g = (g_n g_{n-1}^{-1})(g_{n-1} g_{n-2}^{-1}) \dots (g_1 g_0^{-1}) = h_{n-1}^{\pm 1} h_{n-2}^{\pm 1} \dots h_0^{\pm 1} \quad \text{where } h_0, \dots, h_{n-1} \in E^{\mathcal{M}};$$

in other words,  $E^{\mathcal{M}} \subseteq G$  generates the groupoid  $G$ . To say that each  $\mathcal{M}_x$  is moreover acyclic means that each  $g \in G$  may be so written in a unique way, i.e.,  $G$  is the **free groupoid** (or *path groupoid*, or *fundamental groupoid*) generated by the directed multigraph  $E^{\mathcal{M}}$ . See e.g., [Alv08], [Car11].

Likewise, a “graphing” (structuring by connected graphs) of  $G$  just amounts to any subgraph generating  $G$ , not necessarily freely.

**Remark 5.36.** In the literature (see e.g., [KM04, §30]), the term **treeable group** usually refers to not just a group structurable by trees as above, i.e., a free group, but more generally a group  $G$  admitting a probability-measure-preserving free action generating a treeable CBER.

This is equivalent to the existence of a free groupoid  $H$  with an invariant measure and a fibration  $H \rightarrow G$ , by Example 5.6 and the coinduced action (Bernoulli shift) construction (see Example 5.38).

**Example 5.37.** Every locally countable Borel groupoid  $G$  is  $\mathcal{T}_{\text{LN}}$ -structurable, by Theorem 5.15. For example, if  $G$  is a countable group, a sequence of Lusin–Novikov functions is given by left multiplication  $g_n(-)$  by each group element  $g_n \in G$ .

However, not every countable group  $G$  is  $\mathcal{T}_{\text{FM}_{\text{invol}}}$ -structurable, e.g.,  $\mathbb{Z}$ , by the argument in Proposition 4.15. In other words, we do not have a “Feldman–Moore theorem for group(oid)s”, in the form of a translation-invariant transitive action of  $\mathbb{Z}_2^{*\omega}$ .

Nonetheless, by Proposition 4.5, we *do* have a “Feldman–Moore theorem for groupoids”, in the form of a translation-invariant transitive action of  $\mathbb{F}_\omega$ . That is, for a locally countable Borel groupoid  $(X, G)$ , we have a Borel action  $\mathbb{F}_\omega \curvearrowright G$ , whose orbits are the dom-fibers, such that

$$\gamma \cdot gh = (\gamma \cdot g)h \quad \text{for } g : x \rightarrow y \text{ and } h : x' \rightarrow x.$$

As in Example 5.35, such a family of actions on the dom-fibers may by right-invariance be represented more concretely by just the action on the identities  $1_x$ , yielding for each  $\gamma \in \mathbb{F}_\omega$  a Borel family of morphisms  $\gamma \cdot \text{id}_X := \{\gamma \cdot 1_x\}_{x \in X} \subseteq G$  in each dom-fiber. In fact it is easily seen that each such  $\gamma \cdot \text{id}_X \subseteq G$  also picks one morphism in each cod-fiber, i.e., is a **(total) Borel bisection** of  $G$ ; and that  $\delta \gamma \cdot \text{id}_X = (\delta \cdot \text{id}_X)(\gamma \cdot \text{id}_X)$ , so that we have a group homomorphism  $\gamma \mapsto \gamma \cdot \text{id}_X$  from  $\mathbb{F}_\omega$  into the group of Borel bisections of  $G$ , known as the **(Borel) full group** of  $G$ . Transitivity of the fiberwise  $\mathbb{F}_\omega$ -actions means that  $G = \bigcup_{\gamma \in \mathbb{F}_\omega} (\gamma \cdot \text{id}_X)$ . Thus, the “Feldman–Moore theorem for groupoids” says that *every locally countable Borel groupoid can be covered by a countable subgroup of its full group*; in this form, the result was stated (without proof) in [TDW21, 4.1].

**Example 5.38.** For a fixed locally countable Borel groupoid  $G$ , a structuring by its Scott theory  $\mathcal{T}_G$  of another groupoid  $H$  is by Proposition 5.31 a fibration  $H \rightarrow G$ , which is equivalently by Example 5.6 a representation of  $H$  as the action groupoid of a Borel action of  $G$ . If  $H$  is a CBER, then the action is free. Thus by Corollary 3.34, the theory  $\mathcal{T}_G \sqcup \mathcal{T}_{\text{LN}} \sqcup \mathcal{T}_{\text{sep}}$  is the Scott theory of the **invariantly universal CBER  $\mathbb{E}_{\infty G}$  generated by a free Borel action of  $G$ .**

For example, if  $G$  is a countable group, it is well-known that such a universal free Borel  $G$ -action may be constructed as the free part of the shift action on  $(2^{\mathbb{N}})^G$ , giving an explicit realization of  $\mathbb{E}_{\infty G}$ . A “fiberwise” version of this works for arbitrary  $G$ , by taking the fibration  $\mathbb{E}_{\infty G} \rightarrow G$  whose fiber over each  $x \in X$  is  $(2^{\mathbb{N}})^{\text{dom}^{-1}(x)}$ , equipped with a natural Borel structure and action.<sup>9</sup>

**Remark 5.39.** It is not true that more generally, for a locally countable Borel groupoid  $G$ ,  $\mathcal{T}_G \sqcup \mathcal{T}_{\text{LN}}$  is the Scott theory of the invariantly universal locally countable Borel groupoid equipped with a Borel fibration to  $G$  (i.e., the invariantly universal action groupoid of a  $G$ -action).

More generally, it is not true that (analogously to Corollary 3.34) for a theory  $\mathcal{T}$ ,  $\mathcal{T} \sqcup \mathcal{T}_{\text{LN}}$  is the Scott theory of the invariantly universal  $\mathcal{T}$ -structurable locally countable Borel groupoid. This would follow as in Corollary 3.34 if  $\mathcal{T}_{\text{LN}}$  were the Scott theory of the invariantly universal locally countable Borel groupoid  $\mathbb{G}_{\infty}$ . But this is false, for the rather trivial reason that the groupoid whose Scott theory is  $\mathcal{T}_{\text{LN}}$  only has a single trivial connected component (since there is only a single model of  $\mathcal{T}_{\text{LN}}$  on a singleton set), whereas  $\mathbb{G}_{\infty}$  must have  $2^{\aleph_0}$  such components.

(Note that  $\mathbb{G}_{\infty}$  exists, by Corollary 3.34, since we may encode a locally countable Borel groupoid  $(X, G)$  as a structuring by a suitable theory of the connectedness CBER on the underlying space  $G$ . It is possible to explicitly construct a Scott theory for it, by combining  $\mathcal{T}_{\text{LN}}$  with a  $\mathcal{T}_{\text{sep}}$ -like theory applied to pairs. Such a theory would be mutually interpretable with  $\mathcal{T}_{\text{LN}}$ , and is arguably a more canonical representative for a “theory characterizing locally countable Borel groupoids” in results such as Theorem 5.15; however, it seems subjectively less clean than  $\mathcal{T}_{\text{LN}}$ .)

## References

- [Alv08] Aurélien Alvarez, *Une théorie de Bass–Serre pour les relations d’équivalence et les groupoïdes boréliens*, Ph.D. thesis, École Normale Supérieure, Lyon, 2008.
- [BK96] Howard Becker and Alexander S. Kechris, *The descriptive set theory of Polish group actions*, London Mathematical Society Lecture Note Series, vol. 232, Cambridge University Press, Cambridge, 1996.
- [Bow14] Lewis Bowen, *Entropy theory for sofic groupoids I: The foundations*, J. Anal. Math. **124** (2014), 149–233.
- [Car11] Alessandro Carderi, *Cost for discrete measured groupoids*, Master’s thesis, Università degli Studi di Roma “Sapienza”, 2011.
- [Che19a] Ruiyuan Chen, *Borel functors, interpretations, and strong conceptual completeness for  $\mathcal{L}_{\omega_1\omega}$* , Trans. Amer. Math. Soc. **372** (2019), no. 12, 8955–8983.
- [Che19b] ———, *Representing Polish groupoids via metric structures*, preprint (2019), <https://arxiv.org/abs/1908.03268>.

---

<sup>9</sup>This is the *coinduced action* or *right Kan extension* of the universal bundle  $X \times 2^{\mathbb{N}} \rightarrow X$  along the inclusion  $(X, =) \rightarrow (X, G)$ .

- [Cis19] Denis-Charles Cisinski, *Higher categories and homotopical algebra*, Cambridge Studies in Advanced Mathematics, vol. 180, Cambridge University Press, Cambridge, 2019.
- [CK18] Ruiyuan Chen and Alexander S. Kechris, *Structurable equivalence relations*, *Fund. Math.* **242** (2018), no. 2, 109–185, arXiv version: <https://arxiv.org/abs/1606.01995v5>.
- [CPTT23] Ruiyuan Chen, Antoine Poulin, Ran Tao, and Anush Tserunyan, *Tree-like graphings, wallings, and median graphings of equivalence relations*, preprint (2023), <https://arxiv.org/abs/2308.13010>.
- [DJK94] R. Dougherty, S. Jackson, and A. S. Kechris, *The structure of hyperfinite Borel equivalence relations*, *Trans. Amer. Math. Soc.* **341** (1994), no. 1, 193–225.
- [FM77] Jacob Feldman and Calvin C. Moore, *Ergodic equivalence relations, cohomology, and von Neumann algebras. I*, *Trans. Amer. Math. Soc.* **234** (1977), no. 2, 289–324.
- [Gao09] Su Gao, *Invariant descriptive set theory*, Pure and Applied Mathematics (Boca Raton), vol. 293, CRC Press, Boca Raton, FL, 2009.
- [GJ99] Paul G. Goerss and John F. Jardine, *Simplicial homotopy theory*, Progress in Mathematics, vol. 174, Birkhäuser Verlag, Basel, 1999.
- [GKK15] John Goodrick, Byunghan Kim, and Alexei Kolesnikov, *Type-amalgamation properties and polygroupoids in stable theories*, *J. Math. Log.* **15** (2015), no. 1, 1550004, 45.
- [HK96] Greg Hjorth and Alexander S. Kechris, *Borel equivalence relations and classifications of countable models*, *Ann. Pure Appl. Logic* **82** (1996), no. 3, 221–272.
- [Hod93] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993.
- [HTMM18] Matthew Harrison-Trainor, Russell Miller, and Antonio Montalbán, *Borel functors and infinitary interpretations*, *J. Symb. Log.* **83** (2018), no. 4, 1434–1456.
- [Jac99] Bart Jacobs, *Categorical logic and type theory*, Studies in Logic and the Foundations of Mathematics, vol. 141, North-Holland Publishing Co., Amsterdam, 1999.
- [JKL02] S. Jackson, A. S. Kechris, and A. Louveau, *Countable Borel equivalence relations*, *J. Math. Log.* **2** (2002), no. 1, 1–80.
- [Kar64] Carol R. Karp, *Languages with expressions of infinite length*, North-Holland Publishing Co., Amsterdam, 1964.
- [Kec91] Alexander S. Kechris, *Amenable equivalence relations and Turing degrees*, *J. Symbolic Logic* **56** (1991), no. 1, 182–194.
- [Kec95] ———, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995.
- [Kec24] ———, *The theory of countable Borel equivalence relations*, Cambridge University Press, 2024, to appear, <https://pma.caltech.edu/documents/5921/CBER.pdf>.

- [KM04] Alexander S. Kechris and Benjamin D. Miller, *Topics in orbit equivalence*, Lecture Notes in Mathematics, vol. 1852, Springer-Verlag, Berlin, 2004.
- [Kru19] Alex Kruckman, *Disjoint  $n$ -amalgamation and pseudofinite countably categorical theories*, Notre Dame J. Form. Log. **60** (2019), no. 1, 139–160.
- [KST99] A. S. Kechris, S. Solecki, and S. Todorcevic, *Borel chromatic numbers*, Adv. Math. **141** (1999), no. 1, 1–44.
- [LE65] E. G. K. Lopez-Escobar, *An interpolation theorem for denumerably long formulas*, Fund. Math. **57** (1965), 253–272.
- [Lei14] Tom Leinster, *Basic category theory*, Cambridge Studies in Advanced Mathematics, vol. 143, Cambridge University Press, Cambridge, 2014.
- [Lur09] Jacob Lurie, *Higher topos theory*, Annals of Mathematics Studies, vol. 170, Princeton University Press, Princeton, NJ, 2009.
- [Mar16] David Marker, *Lectures on infinitary model theory*, Lecture Notes in Logic, vol. 46, Association for Symbolic Logic, Chicago, IL; Cambridge University Press, Cambridge, 2016.
- [ML98] Saunders Mac Lane, *Categories for the working mathematician*, second ed., Graduate Texts in Mathematics, vol. 5, Springer-Verlag, New York, 1998.
- [Ram90] Arlan Ramsay, *The Mackey-Glimm dichotomy for foliations and other Polish groupoids*, J. Funct. Anal. **94** (1990), no. 2, 358–374.
- [SS16] Theodore A. Slaman and John R. Steel, *Definable functions on degrees*, Ordinal definability and recursion theory: The Cabal Seminar. Vol. III, Lect. Notes Log., vol. 43, Assoc. Symbol. Logic, Ithaca, NY, 2016, Reprint of MR0960895, pp. 458–475.
- [TDW21] Robin Tucker-Drob and Konrad Wróbel, *Cost of inner amenable groupoids*, Proc. Amer. Math. Soc. **149** (2021), no. 10, 4303–4315.
- [Ued06] Yoshimichi Ueda, *Notes on treeability and costs for discrete groupoids in operator algebra framework*, Operator Algebras: The Abel Symposium 2004, Abel Symp., vol. 1, Springer, Berlin, 2006, pp. 259–279.

Department of Mathematics, Statistics, and Computer Science  
 University of Illinois Chicago  
 Chicago, IL, USA  
 rbane8@uic.edu

Department of Mathematics  
 University of Michigan  
 Ann Arbor, MI, USA  
 ruiyuan@umich.edu