



SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The Compactness Theorem: Proofs and Applications

av

Samuel Jansson

2026 - No K18

The Compactness Theorem: Proofs and Applications

Samuel Jansson

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Peter LeFanu Lumsdaine

2026

Abstract

The Compactness Theorem is a landmark result in mathematical logic that bridges the gap between the finite and the infinite. It states that a vast collection of logical statements can be true simultaneously if and only if every finite selection of those statements is consistent on its own. This thesis provides a comprehensive study of the theorem, presenting three distinct proofs: an inductive approach for the countable case, a general proof via Zorn's Lemma, and a model-theoretic construction using ultraproducts. By exploring its applications in logic, topology, and combinatorics, this work demonstrates the theorem's versatility and its ability to solve problems across seemingly unrelated branches of mathematics.

Sammanfattning

Kompakthetsatsen är ett banbrytande resultat inom matematisk logik som överbryggar klyftan mellan det ändliga och det oändliga. Den säger att en omfattande samling logiska påståenden kan vara sanna samtidigt om och endast om varje ändligt urval av dessa påståenden är konsistent i sig självt. Denna uppsats ger en omfattande studie av satsen och presenterar tre distinkta bevis: ett induktivt tillvägagångssätt för det räknebara fallet, ett generellt bevis via Zorns lemma, och en modellteoretisk konstruktion med hjälp av ultraprodukt. Genom att utforska dess tillämpningar inom logik, topologi och kombinatorik demonstrerar detta arbete satsens mångsidighet och dess förmåga att lösa problem inom tillsynes orelaterade grenar av matematiken.

Contents

1	Introduction	9
1.1	Structure of the Thesis	9
1.2	Acknowledgements	10
2	Definitions and Notation	11
2.1	Foundations of First-Order Model Theory	11
2.1.1	Language	11
2.1.2	Structures and models	13
2.1.3	Deduction and Peano's axioms	14
2.1.4	Ultraproducts	16
2.2	Mathematical Background	18
2.2.1	Ordering	18
2.2.2	Topology	18
3	Versions of Compactness	20
3.1	The Model-Theoretic Perspective	20
3.1.1	Zorn's lemma	21
3.1.2	A general proof	22
3.2	The Ultraproduct Perspective	24
3.2.1	Łoś's Theorem	25
4	Applications of Compactness	28
4.1	Non-standard models of PA	28
4.2	The Topological Interpretation	29
4.3	Tiling of the plane	30

1 Introduction

Working with infinity is common in mathematics, but it often presents subtle challenges. In general, mathematical reasoning is simpler in the finite than in the infinite; we can often verify properties of finite systems through direct exhaustion or simple induction. The Compactness Theorem provides a fundamental bridge between these realms. It states: “A theory has an interpretation if and only if every finite subset of that theory has an interpretation.”

To put it in more concrete terms, this means that an infinite list of mathematical statements is “satisfiable”, i.e. “is true” as long as no finite part of that list contradicts itself. It tells us that in first-order logic, “infinity” has no hidden surprises: if you can successfully build any finite piece of a structure, the theorem guarantees you can build the whole infinite thing.

This captures a remarkable property of first-order logic: that a “global” property (the satisfiability of an infinite set of formulas) is entirely determined by its “local” properties (the satisfiability of its finite subsets).

1.1 Structure of the Thesis

To provide a logical progression from theory to practice, this thesis is organized into the following sections:

- **Section 2: Definitions and Notation.** We establish the necessary notation and formal definitions for first-order languages, structures, and models. This section serves as the technical foundation required to understand the proofs that follow.
- **Section 3: Versions of Compactness.** Here, we state the theorem and present three rigorous proofs: an inductive approach for the countable case, a general proof utilizing Zorn’s Lemma, and a model-theoretic construction via ultraproducts and Łoś’s Theorem.
- **Section 4: Applications of Compactness.** Finally, we demonstrate the theorem’s versatility through a variety of applications. We begin within the realm of logic by exploring the construction of non-standard models of arithmetic, before extending our scope to examine the theorem’s implications in topology and tiling theory. This shows how a single result in logic can provide insight into both logic as well as other areas of mathematics.

1.2 Acknowledgements

I am very grateful to my supervisor, Peter LeFanu Lumsdaine, for all his guidance and support throughout this project. From helping shape the main ideas to offering countless tips on writing in \LaTeX , structuring the thesis, and tackling tricky mathematical problems, Peter's advice has been invaluable. I have learned a great deal from his insights, and this thesis would not have been possible without his help. I am also indebted to the contributors of **ProofWiki**, whose collaborative work served as a helpful reference and a frequent source of inspiration.

2 Definitions and Notation

We begin by establishing the definitions and notation that will be used throughout this thesis. This ensures that the reader has the necessary background to understand the concepts presented in the following chapters. While no prior formal training in logic is strictly assumed, the reader should be comfortable with general mathematical thinking and standard notation, such as sets, set operations and the use of Π for products.

2.1 Foundations of First-Order Model Theory

The following development of first-order model theory is based primarily on the approach taken by Carlström [Car13] regarding the introductory logical framework. For the more advanced treatment of the construction of ultraproducts, we follow the methodology of Cori and Lascar [CL01].

2.1.1 Language

The language of propositional logic typically includes seven standard connectives and constants: *truth* (\top), *falsum* (\perp), *negation* (\neg), *conjunction* (\wedge), *disjunction* (\vee), *implication* (\rightarrow), and the *biconditional* (\leftrightarrow). While these provide a rich vocabulary, they are not all logically independent.

A set of connectives is called **complete** if it generates, under composition, the set of all propositional connectives. Furthermore, a complete set is considered **minimal** when no proper subset of it is itself complete (see [CL00, Definition 1.32]). For the purposes of this work, we adopt the minimal complete set $\{\perp, \rightarrow\}$. The remaining connectives are constructed from this minimal set as follows:

$$\begin{aligned}\neg p &:= p \rightarrow \perp \\ p \wedge q &:= \neg(p \rightarrow \neg q) \\ p \vee q &:= \neg p \rightarrow q \\ \top &:= \neg \perp \\ p \leftrightarrow q &:= (p \rightarrow q) \wedge (q \rightarrow p)\end{aligned}$$

Here, p and q are propositions, and parentheses are used to clarify grouping.

Remark 2.1. Adopting a minimal set of connectives significantly streamlines the

inductive proofs in later chapters. We will rely on this specific construction in the proof of Łoś's Theorem (see Theorem 3.5). In all other contexts, we treat the remaining connectives as standard abbreviations and use them freely throughout the thesis.

First-order logic can be seen as an expansion of propositional logic. It incorporates all propositional connectives and, in addition, the universal and existential quantifiers, \forall and \exists , respectively. Analogous to the propositional connectives, these quantifiers are duals; one may be defined in terms of the other via negation. Specifically, the universal quantifier can be expressed as $\forall x\varphi := \neg\exists x\neg\varphi$ (see [Car13, Chapter 11.1]).

Definition 2.2. A *first-order language* \mathcal{L} consists of:

- (i) a set of constant symbols \mathcal{C} ,
- (ii) a set of function symbols \mathcal{F} ,
- (iii) a set of relation (predicate) symbols \mathcal{R} ,
- (iv) and the logical symbols: the connectives \rightarrow and \perp (from which the standard connectives can be defined), the quantifiers \forall and \exists , and parentheses for grouping.

We then write $\mathcal{L} = \langle \mathcal{C}, \mathcal{F}, \mathcal{R} \rangle$, with the understanding that the logical symbols are always included.

Example 2.3. A standard example of a first-order language is that of arithmetic:

$$\mathcal{L}_A = \langle 0, S, +, \times, = \rangle,$$

where 0 is a constant symbol, S is a unary function symbol (successor), $+$ and \times are binary function symbols, and $=$ is a binary relation symbol.

While the symbols of \mathcal{L} provide the basic vocabulary, we require a systematic way to combine them into expressions that can denote individuals within a structure. These syntactic building blocks are known as *terms*.

Definition 2.4. The set of \mathcal{L} -terms is the smallest set containing all variables and all constant symbols in \mathcal{C} that is closed under the following property: if $f \in \mathcal{F}$ is an n -ary function symbol and t_1, \dots, t_n are in the set, then $f(t_1, \dots, t_n)$ is also in the set.

Example 2.5. The definition of terms allows us to construct complex expressions from the basic symbols of a language. In \mathcal{L}_{PA} , for instance, the variables x, y and the constant 0 are terms; by applying function symbols, we can form more complex terms such as $S(0)$, $x + S(0)$, or $(S(S(0)) \times y)$.

A particularly important class of terms in the language of arithmetic are the *canonical numerals* \bar{n} for each $n \in \mathbb{N}$, defined by recursion:

$$\begin{aligned}\bar{0} &:= 0, \\ \overline{n+1} &:= S(\bar{n}).\end{aligned}$$

These numerals serve as the formal syntactic names for the natural numbers within \mathcal{L}_A .

2.1.2 Structures and models

To determine whether a first-order formula is true or false, we must evaluate it within a specific mathematical setting. We now formalize the notion of such a setting, known as a *structure*, and define what it means for a structure to serve as a *model* for a set of formulas.

Definition 2.6. A *structure* \mathcal{M} of a first-order language $\mathcal{L} = \langle \mathcal{C}, \mathcal{F}, \mathcal{R} \rangle$ is a tuple

$$\mathcal{M} = \langle M, c_1^{\mathcal{M}}, c_2^{\mathcal{M}}, \dots, f_1^{\mathcal{M}}, f_2^{\mathcal{M}}, \dots, R_1^{\mathcal{M}}, R_2^{\mathcal{M}}, \dots \rangle$$

where:

- (i) M is a nonempty set called the *universe*,
- (ii) each $c_i \in \mathcal{C}$ is interpreted as $c_i^{\mathcal{M}} \in M$,
- (iii) each $f_i \in \mathcal{F}$ is interpreted as a function $f_i^{\mathcal{M}} : M^{n_i} \rightarrow M$ and
- (iv) each $R_i \in \mathcal{R}$ is interpreted as a relation $R_i^{\mathcal{M}} \subseteq M^{m_i}$.

Definition 2.7. Let Γ be a set of formulas in a first-order language and let \mathcal{M} be a structure for that language. We say that \mathcal{M} is a *model* of Γ , and write

$$\mathcal{M} \models \Gamma,$$

if every formula in Γ is satisfied in \mathcal{M} ; that is,

$$\mathcal{M} \models \Gamma \iff \text{for all } \varphi \in \Gamma, [[\varphi]]^{\mathcal{M}} = \top.$$

Example 2.8. Let

$$\mathcal{N} = \langle \mathbb{N}, 0^{\mathcal{N}}, S^{\mathcal{N}}, +^{\mathcal{N}}, \times^{\mathcal{N}}, =^{\mathcal{N}} \rangle$$

be a structure for the language of arithmetic in which each symbol is interpreted in the standard way: $0^{\mathcal{N}}$ is 0, $S^{\mathcal{N}}$ is the successor function, and $+^{\mathcal{N}}$ and $\times^{\mathcal{N}}$ are ordinary addition and multiplication, respectively. Let Q be the set of the following axioms, known as Robinson's Axioms:

$$(Q1) \quad \forall x \neg(S(x) = 0)$$

$$(Q2) \quad \forall x \forall y (S(x) = S(y) \rightarrow x = y)$$

$$(Q3) \quad \forall x (x + 0 = x)$$

$$(Q4) \quad \forall x \forall y (x + S(y) = S(x + y))$$

$$(Q5) \quad \forall x (x \times 0 = 0)$$

$$(Q6) \quad \forall x \forall y (x \times S(y) = (x \times y) + x)$$

$$(Q7) \quad \forall x (\neg(x = 0) \rightarrow \exists y S(y) = x)$$

Then $\mathcal{N} \models Q$. Now change the universe of \mathcal{N} from \mathbb{N} to the set of integers \mathbb{Z} , keeping the interpretations of the symbols the same. In this new structure, axiom (Q1) fails, since $S(-1) = 0$. Thus,

$$\langle \mathbb{Z}, 0, S, +, \times, = \rangle \not\models Q.$$

2.1.3 Deduction and Peano's axioms

Having established the relationship between symbols and structures (Semantics), we now turn to the formal rules for manipulating these symbols (Syntax) through the notion of *deduction*.

Definition 2.9. Let Γ be a set of formulas and let φ be a formula. We say that Γ *deduces* φ , written

$$\Gamma \vdash \varphi,$$

if there exists a *derivation* of φ from the set Γ using a given set of rules (such as Modus Ponens or natural deduction rules). This means that φ can be obtained as a logical consequence of the formulas in Γ through a finite sequence of applications of these rules.

Example 2.10. Let \mathcal{N} be as before, i.e.

$$\mathcal{N} = \langle \mathbb{N}, 0^{\mathcal{N}}, S^{\mathcal{N}}, +^{\mathcal{N}}, \times^{\mathcal{N}}, =^{\mathcal{N}} \rangle$$

and let Q be the set of Robinson's axioms (Q1) - (Q7). We can define the natural number 1 as the successor of 0, that is $1 := S(0)$. By (Q1) we can then deduce that $S(0) \neq 0$, so in other words $Q \vdash 1 \neq 0$.

We now present the axioms of Peano Arithmetic (PA), which provide a formal foundation for the arithmetic of natural numbers. For a more comprehensive treatment of these axioms and their properties, we refer the reader to Cori and Lascar [CL01, Chapter 6.1].

Definition 2.11. Let Q be the set of Robinson's Axioms 2.8. The Peano Axioms, denoted by PA , consist of Q along with the *axiom scheme of induction*. For every formula $F(x, y_1, \dots, y_n)$ in the language of arithmetic \mathcal{L}_A , the following sentence is an axiom:

$$\forall y_1 \dots \forall y_n \left(\begin{array}{l} (F(0, y_1, \dots, y_n) \wedge \forall x (F(x, y_1, \dots, y_n) \rightarrow F(S(x), y_1, \dots, y_n))) \\ \rightarrow \forall x F(x, y_1, \dots, y_n) \end{array} \right).$$

Note that since there are infinitely many formulas F , PA is an infinite set of sentences.

The inclusion of the induction scheme allows us to prove general algebraic properties that are unprovable in the weaker Robinson's theory (see [Smi13, Chapter 10.5]).

Example 2.12. While Robinson's axioms are sufficient to compute specific numerical sums (for instance, $Q \vdash \bar{2} + \bar{2} = \bar{4}$), they are insufficient to prove universal laws such as the commutativity of addition:

$$Q \not\vdash \forall x \forall y (x + y = y + x).$$

To prove this in Peano Arithmetic (PA), one applies the induction scheme to the formula $\varphi(x) \equiv \forall y (x + y = y + x)$. The proof typically proceeds in two stages:

- **Base Case:** One proves $\varphi(0)$, which states $\forall y(0 + y = y + 0)$. This itself often requires a “nested” induction on y .
- **Inductive Step:** Assuming the inductive hypothesis $\varphi(k)$ holds, one demonstrates that $\varphi(S(k))$ follows. This involves showing that if k commutes with all y , then its successor $S(k)$ must also satisfy the commutative property.

This highlights the primary difference between Q and PA : while Q describes the basic behavior of the successor, addition, and multiplication operations, PA uses induction to ensure that the universe “behaves” like the standard natural numbers \mathbb{N} in its global properties.

2.1.4 Ultraproducts

Finally, we will introduce the concept of ultraproducts. However, before doing so, we must first establish a few additional definitions.

Definition 2.13. Let \mathcal{L} be a language and let I be a set. The *product* of the family of \mathcal{L} -structures $(\mathcal{M}_i)_{i \in I}$, written

$$\prod_{i \in I} \mathcal{M}_i,$$

is defined as follows:

1. The universe of $\prod_{i \in I} \mathcal{M}_i$ is the Cartesian product

$$\prod_{i \in I} M_i = \{(x_i)_{i \in I} : x_i \in M_i \text{ for each } i \in I\},$$

where M_i is the universe of \mathcal{M}_i for each $i \in I$.

2. The interpretations of function and relation symbols are defined component-wise. If f is an n -ary function symbol and R is an n -ary relation symbol in \mathcal{L} , then

$$f^{\prod_{i \in I} \mathcal{M}_i}((x_i^1)_{i \in I}, \dots, (x_i^n)_{i \in I}) = (f^{\mathcal{M}_i}(x_i^1, \dots, x_i^n))_{i \in I},$$

and

$$((x_i^1)_{i \in I}, \dots, (x_i^n)_{i \in I}) \in R^{\prod_{i \in I} \mathcal{M}_i} \iff \forall i \in I, (x_i^1, \dots, x_i^n) \in R^{\mathcal{M}_i},$$

where $f^{\mathcal{M}_i}$ and $R^{\mathcal{M}_i}$ denote the interpretations of f and R in \mathcal{M}_i .

Definition 2.14. Let I be a set and let \mathcal{F} be a set of subsets of I . We say that \mathcal{F} is a *filter* on I if:

- (i) \mathcal{F} is upward closed, that is if $X \in \mathcal{F}$ and $X \subseteq Y \subseteq I$, then $Y \in \mathcal{F}$,
- (ii) the set I is in \mathcal{F} and
- (iii) \mathcal{F} is closed under intersection, i.e. for all $X, Y \in \mathcal{F}$, $X \cap Y \in \mathcal{F}$.

Furthermore, a set \mathcal{U} is called an *ultrafilter* if:

- (iv) \mathcal{U} is a filter on some set I ,
- (v) the empty set is not in \mathcal{U} and
- (vi) for every subset $X \subseteq I$ exactly one of the following holds:
 - (a) $X \in \mathcal{U}$, or
 - (b) the complement of X is in \mathcal{U} , i.e. $I \setminus X \in \mathcal{U}$.

Combining the direct product of structures with the notion of a filter, we obtain the following construction.

Definition 2.15. Let \mathcal{L} be a first-order language, let I be a set, and let $(\mathcal{M}_i)_{i \in I}$ be a family of \mathcal{L} -structures with universes M_i . Let \mathcal{F} be a filter on I .

Define an equivalence relation $\sim_{\mathcal{F}}$ on the Cartesian product $\prod_{i \in I} M_i$ by

$$(x_i)_{i \in I} \sim_{\mathcal{F}} (y_i)_{i \in I} \quad \text{if and only if} \quad \{i \in I : x_i = y_i\} \in \mathcal{F}$$

The *reduced product of $(\mathcal{M}_i)_{i \in I}$ modulo \mathcal{F}* is the structure

$$\prod_{i \in I} \mathcal{M}_i / \mathcal{F}$$

whose universe is $\prod_{i \in I} M_i / \sim_{\mathcal{F}}$, and where the interpretations of function and relation symbols are defined pointwise modulo \mathcal{F} .

If \mathcal{U} is an ultrafilter on I , the reduced product

$$\prod_{i \in I} \mathcal{M}_i / \mathcal{U}$$

is called the *ultraproduct* of the family $(\mathcal{M}_i)_{i \in I}$ modulo \mathcal{U} .

2.2 Mathematical Background

In this section we aim to provide the necessary definitions for concepts outside of logic. We believe these are niche enough to not fall under the blanket of “general mathematical thinking”.

2.2.1 Ordering

The concept of order is central to the application of Zorn’s lemma (see 3.2). For intuition, one may consider the set of natural numbers \mathbb{N} , where we naturally view 1 as being “less than” 2. This creates an order that looks like this:

$$1 < 2 < 3 < \dots$$

where we can take any two elements of \mathbb{N} and get this comparison. While the order on \mathbb{N} is total, the following definition generalizes this concept to sets where some elements may be incomparable.

Definition 2.16. A set P along with the order relation \leq is called *partially ordered* (sometimes called a *poset*) if for all $a, b, c \in P$, we have

- (i) (Reflexivity) $a \leq a$,
- (ii) (Antisymmetry) if $a \leq b$ and $b \leq a$ then $a = b$, and
- (iii) (Transitivity) if $a \leq b$ and $b \leq c$ then $a \leq c$.

Definition 2.17. Let (P, \leq) be a partially ordered set.

- (i) A subset $C \subseteq P$ is called a *chain* if for every $x, y \in C$, either $x \leq y$ or $y \leq x$.
- (ii) An element $u \in P$ is an *upper bound* for a subset $S \subseteq P$ if $s \leq u$ for all $s \in S$.
- (iii) An element $m \in P$ is called a *maximal element* of P if for all $x \in P$, $m \leq x$ implies $m = x$.

2.2.2 Topology

In a later section, we will examine the application of the Compactness Theorem of logic to its namesake: the Compactness Theorem from topology. To understand this deep connection, we must first establish what a *topology* actually is.

Definition 2.18 (Topological Space). A *topology* on a set X is a collection \mathcal{T} of subsets of X (called *open sets*) such that $\emptyset, X \in \mathcal{T}$, \mathcal{T} is closed under arbitrary unions, and \mathcal{T} is closed under finite intersections. The pair (X, \mathcal{T}) is called a *topological space*.

Definition 2.19 (Compactness). A topological space (X, \mathcal{T}) is *compact* if every open cover of X has a finite subcover. That is, if $X = \bigcup_{i \in I} U_i$ where each $U_i \in \mathcal{T}$, there exists a finite subset $J \subseteq I$ such that $X = \bigcup_{j \in J} U_j$.

Another concept that plays a fundamental role in topology is the finite intersection property (FIP). Intuitively, a collection of sets has the finite intersection property if no matter how many sets you pick—so long as you only pick finitely many—their intersection is never empty. To make this more precise, we provide a definition.

Definition 2.20. Let X be a set and let \mathcal{A} be a collection of subsets of X . We say that \mathcal{A} has the *finite intersection property* (FIP) if the intersection of any finite subcollection of \mathcal{A} is non-empty. That is, for any $A_1, A_2, \dots, A_n \in \mathcal{A}$, we have:

$$\bigcap_{i=1}^n A_i \neq \emptyset.$$

3 Versions of Compactness

3.1 The Model-Theoretic Perspective

A very important theorem in mathematical logic is known as the *compactness theorem*. The theorem was first implicitly present in the work of Thoralf Skolem, notably in his 1923 paper *Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre*, but was not explicitly formulated until 1930, following Kurt Gödel's completeness result in *Die Vollständigkeit der Axiome des logischen Funktionenkalküls* [Daw93]. This result established the fundamental relationship between syntactic consistency and semantic satisfiability, culminating in the following modern formulation:

Theorem 3.1 (Compactness Theorem). *Let Γ be a theory, i.e. a set of first-order sentences in a language \mathcal{L} . Then Γ has a model if and only if every finite subset of Γ has a model.*

Proof. The (\Rightarrow) direction is trivial: If \mathcal{M} is a model of Γ , then \mathcal{M} is clearly a model of any subset $\Gamma_0 \subseteq \Gamma$.

We now prove the (\Leftarrow) direction for the case where the set of propositional variables $P = \{p_1, p_2, \dots\}$ is countably infinite. Assume that every finite subset of Γ is satisfiable (i.e., Γ is finitely satisfiable).

Outline of the Proof: We will construct a valuation $v : P \rightarrow \{0, 1\}$ by induction. Specifically, we will define a sequence of truth values $(\Delta_n)_{n \in \mathbb{N}}$ such that for every n , the set $\Gamma \cup \{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}\}$ remains finitely satisfiable. Here, the set $\{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}\}$ uses the notation

$$p^{\Delta_i} = \begin{cases} p & \text{if } \Delta_i = 1, \\ \neg p & \text{if } \Delta_i = 0. \end{cases}$$

Finally, we will show that the valuation defined by these choices satisfies every formula in Γ .

Step 1: The Inductive Construction. We define the values $\Delta_n \in \{0, 1\}$ for $n \in \mathbb{N}$ by induction on n .

- **Base Case ($n = 0$):** By hypothesis, the set $\Gamma_0 := \Gamma$ is finitely satisfiable.

- **Inductive Step:** Suppose we have chosen values $\Delta_1, \dots, \Delta_n$ such that the set $\Gamma_n := \Gamma \cup \{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}\}$ is finitely satisfiable. We define $\Delta_{n+1} = 1$ if $\Gamma \cup \{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}, p_{n+1}\}$ is finitely satisfiable. Otherwise, we set $\Delta_{n+1} = 0$.

Step 2: Proof of Consistency. We must show that our choice of Δ_{n+1} always preserves finite satisfiability. Suppose for the sake of contradiction that it does not. Then neither choice for Δ_{n+1} preserves the property. This implies two things:

- (i) There is a finite subset $\Gamma_A \subseteq \Gamma \cup \{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}\}$ such that $\Gamma_A \cup \{p_{n+1}\}$ is unsatisfiable.
- (ii) There is a finite subset $\Gamma_B \subseteq \Gamma \cup \{p_1^{\Delta_1}, \dots, p_n^{\Delta_n}\}$ such that $\Gamma_B \cup \{\neg p_{n+1}\}$ is unsatisfiable.

Consider the finite union $\Gamma^* = \Gamma_A \cup \Gamma_B$. By the inductive hypothesis, Γ^* is satisfiable. Let w be a valuation satisfying Γ^* . If $w(p_{n+1}) = 1$, then w satisfies $\Gamma_A \cup \{p_{n+1}\}$, a contradiction. If $w(p_{n+1}) = 0$, then w satisfies $\Gamma_B \cup \{\neg p_{n+1}\}$, also a contradiction. Thus, Δ_{n+1} is always well-defined.

Step 3: Model Verification. Define $v(p_n) = \Delta_n$ for all $n \in \mathbb{N}$. We claim $v \models \Gamma$. Let $\varphi \in \Gamma$. Since φ is a formula, it contains only a finite set of variables, say $\{p_1, \dots, p_k\}$.

By our construction in Step 1, the set $\{\varphi, p_1^{\Delta_1}, \dots, p_k^{\Delta_k}\}$ is a finite subset of $\Gamma \cup \{p_1^{\Delta_1}, \dots, p_k^{\Delta_k}\}$, which is finitely satisfiable. Thus, there exists a valuation v^* that satisfies both φ and the literals $\{p_1^{\Delta_1}, \dots, p_k^{\Delta_k}\}$. This means $v^*(p_i) = \Delta_i$ for all $i \leq k$. Since v and v^* agree on all variables present in φ , and $v^* \models \varphi$, it follows that $v \models \varphi$. \square

3.1.1 Zorn's lemma

To prove the (\Leftarrow) direction for the general case where the set of propositional variables P may be uncountably infinite, we provide a direct semantic proof employing Zorn's Lemma.¹

Lemma 3.2. *Let P be a partially ordered set such that every chain has an upper bound. Then P has at least one maximal element.*

¹One may also derive this result as a corollary of the Completeness Theorem: if every finite subset of Γ is satisfiable, then Γ is consistent, and by completeness, Γ has a model. However, we opt for a direct construction following the method of Cori and Lascar [CL00, Chapter 1.5].

Proof. Suppose P is chain-complete, i.e. a partially ordered set where every chain has an upper bound. By the Axiom of Choice, we can define a successor function $S : P \rightarrow P$ such that for each $x \in P$, $S(x) > x$ if x is not maximal, and $S(x) = x$ if x is maximal. Furthermore, let U be a function that assigns to every chain $C \subseteq P$ an upper bound $U(C) \in P$.

Now, let κ be an ordinal with cardinality strictly greater than $|P|$. We define a function $f : \kappa \rightarrow P$ by transfinite recursion:

- **Successor Step:** $f(\alpha + 1) = S(f(\alpha))$.
- **Limit Step:** For a limit ordinal λ , $f(\lambda) = U(\{f(\gamma) : \gamma < \lambda\})$.

Claim: The function f is non-decreasing; specifically, if $\gamma < \beta < \kappa$, then $f(\gamma) \leq f(\beta)$.

Proof of Claim: This follows by transfinite induction on β . If $\beta = \alpha + 1$ is a successor, $f(\beta) = S(f(\alpha)) \geq f(\alpha)$. If β is a limit, $f(\beta)$ is an upper bound for all previous values by definition.

If P had no maximal element, then by our definition of S , we would have $S(x) > x$ for all x . This would make f a *strictly increasing* function ($f(\gamma) < f(\beta)$ for all $\gamma < \beta$).

However, a strictly increasing function $f : \kappa \rightarrow P$ would be an injection from κ into P . This is impossible because we chose κ such that $|\kappa| > |P|$. Therefore, there must exist some ordinal $\alpha < \kappa$ such that $f(\alpha + 1) = f(\alpha)$. By the definition of S , this only happens if $f(\alpha)$ is a maximal element of P . \square

3.1.2 A general proof

After stating and proving Zorn's Lemma, we now define a set that will be central to the proof of the Compactness Theorem.

Definition 3.3. Let E_{map} be the set of valuations $u : S \rightarrow \{0, 1\}$ (where $S \subseteq P$) such that for every finite subset $\Gamma_0 \subseteq \Gamma$, there exists a total assignment $v : P \rightarrow \{0, 1\}$ that extends u and satisfies Γ_0 . By extending u we simply mean that $v(p) = u(p)$ for all $p \in S$.

To help understand this set we provide a little example. Let $P = \{p_1, p_2, \dots\}$ (note that this set is countable) and let $\Gamma = \{\neg(p_i \wedge p_j) : i \neq j\}$. Now define two

functions u and v from the subset $\{p_1, p_2\}$ in the following way:

$$\begin{aligned} u(p_1) &= 1, \quad u(p_2) = 0 \\ v(p_1) &= 1, \quad v(p_2) = 1. \end{aligned}$$

In this case, u would be in E_{map} since one can easily find an extension, simply take $u(p_i) = 0$ for all $i > 2$. However, the same thing is not true for v since any extension would not satisfy $\neg(p_1 \wedge p_2)$. It is worth noting that, in the general case, E_{map} is not empty since it contains the empty mapping. This is only natural since one can view any valuation that satisfies Γ_0 as an extension of the empty mapping.

Having defined E_{map} , we outline the proof of the general (\Leftarrow) direction of the Compactness Theorem. First, we define a partial order \leq on E_{map} . We then show that every non-empty chain in E_{map} has an upper bound. By Zorn's lemma, this yields a maximal element γ such that $\text{dom}(\gamma) = P$, completing the proof.

Proof. Let P be the set of propositional variables and Γ be a set of formulas such that every finite subset of Γ is satisfiable. Let $\varphi, \psi \in E_{\text{map}}$. We now define the relation \leq on E_{map} in the following way:

$$\begin{aligned} \varphi \leq \psi \quad &\text{if and only if} \quad \text{dom}(\varphi) \subseteq \text{dom}(\psi) \quad \text{and} \\ &\forall x \in \text{dom}(\varphi) \quad \text{we have} \quad \varphi(x) = \psi(x). \end{aligned}$$

We say that ψ is an extension of φ if $\varphi \leq \psi$. This relation constitutes an order relation on E_{map} , making (E_{map}, \leq) a poset.

Let $C \subseteq E_{\text{map}}$ be a non-empty chain. Let λ be a map whose domain is the union of the domains of all $x \in C$, i.e.

$$\text{dom}(\lambda) = \bigcup_{x \in C} \text{dom}(x),$$

and for all $x \in \text{dom}(\lambda)$ and all $\varphi \in C$, if $x \in \text{dom}(\varphi)$, then $\lambda(x) = \varphi(x)$.

Claim: λ is an element of E_{map} .

Proof of claim: By construction, $\text{dom}(\lambda) = \bigcup_{x \in C} \text{dom}(x)$, which is clearly a subset of P . Let $\lambda \models \Gamma_0$ for some finite $\Gamma_0 \subseteq \Gamma$. For any finite $\Gamma_0 \subseteq \Gamma$, the set S of variables occurring in Γ_0 is finite. We can therefore enumerate these variables as $S = \{p_1, \dots, p_n\}$. Each of these variables belongs to some domain of a map in C . Let $\varphi_1, \dots, \varphi_n \in C$ such that $p_1 \in \text{dom}(\varphi_1), \dots, p_n \in \text{dom}(\varphi_n)$. Since C is

a chain, the finite set $\{\varphi_1, \dots, \varphi_n\}$ has a maximum element $\varphi_{\text{ext}} \in C$ such that $\text{dom}(\varphi_i) \subseteq \text{dom}(\varphi_{\text{ext}})$ for all i . Since $\varphi_{\text{ext}} \in E_{\text{map}}$, there exists an extension ψ_{ext} of φ_{ext} to the whole of P . This gives us the following natural extension of λ :

$$\mu(p) = \begin{cases} \lambda(p) & \text{if } p \in \text{dom}(\lambda), \\ \psi_{\text{ext}}(p) & \text{otherwise.} \end{cases}$$

We have shown that for any non-empty chain $C \in E_{\text{map}}$, there exists an upper bound λ of C . By Zorn's lemma we can then choose an element γ that is maximal in E_{map} .

Claim: The domain of γ is all of P .

Proof of claim: Suppose for contradiction that $\text{dom}(\gamma) \neq P$. Then there exists some $q \in P \setminus \text{dom}(\gamma)$. Consider two possible extensions of γ to the variable q : $\gamma_0 = \gamma \cup \{(q, 0)\}$ and $\gamma_1 = \gamma \cup \{(q, 1)\}$. Since γ is maximal, neither γ_0 nor γ_1 can be in E_{map} .

By the definition of E_{map} , if $\gamma_0 \notin E_{\text{map}}$, there exists a finite subset $\Gamma_0 \subseteq \Gamma$ that cannot be satisfied by any total valuation extending γ_0 . Similarly, if $\gamma_1 \notin E_{\text{map}}$, there exists a finite subset $\Gamma_1 \subseteq \Gamma$ that cannot be satisfied by any total valuation extending γ_1 .

Let $\Gamma^* = \Gamma_0 \cup \Gamma_1$. Since Γ^* is a finite subset of Γ and $\gamma \in E_{\text{map}}$, there exists a total valuation v that satisfies Γ^* and agrees with γ on $\text{dom}(\gamma)$. However, $v(q)$ must be either 0 or 1. If $v(q) = 0$, then v is a total valuation extending γ_0 that satisfies Γ_0 , which is a contradiction. If $v(q) = 1$, then v is a total valuation extending γ_1 that satisfies Γ_1 , also a contradiction. Thus, either $\gamma_0 \in E_{\text{map}}$ or $\gamma_1 \in E_{\text{map}}$, contradicting the maximality of γ . Therefore, $\text{dom}(\gamma) = P$.

Thus we see that γ is a distribution of truth values on P and that any extension of γ to P is equal to γ . So by definition of E_{map} , every finite subset Γ_0 of Γ is satisfied by γ . In particular, this is true for every singleton subset, which means that every formula $F \in \Gamma$ is satisfied by γ . So Γ is satisfiable. \square

3.2 The Ultraproduct Perspective

We have now presented two proofs of the Compactness Theorem: one using induction to prove the version for countable sets of propositional variables, and another employing Zorn's Lemma for the general, potentially uncountable case. Our third approach utilizes the theory of ultraproducts (see Section 2.1.4). This method relies

on two foundational results: the Ultrafilter Lemma, which guarantees the existence of sufficiently “large” filters, and Łoś’s Theorem, which relates the properties of an ultraproduct to the properties of the individual models from which it is constructed. We begin with the former.

Lemma 3.4. *Let I be a non-empty set and let \mathcal{F} be a filter on I . Then there exists an ultrafilter \mathcal{U} on I such that $\mathcal{F} \subseteq \mathcal{U}$.*

Proof. Let \mathcal{P} be the set of all filters on I containing \mathcal{F} , ordered by inclusion. Since the union of any chain of filters in \mathcal{P} is again a filter containing \mathcal{F} , every chain in \mathcal{P} has an upper bound. By Zorn’s Lemma 3.2, \mathcal{P} has a maximal element \mathcal{U} , which is by definition an ultrafilter containing \mathcal{F} . \square

3.2.1 Łoś’s Theorem

Having established the existence of ultrafilters, we now turn to the fundamental construction of the ultraproduct. This construction allows us to “average” a collection of models into a single structure. The bridge between the logic of the individual models and the logic of the resulting ultraproduct is provided by Łoś’s Theorem, often referred to as the Fundamental Theorem of Ultraproducts.

Theorem 3.5 (Łoś’s Theorem). *Let \mathcal{L} be a language, I a set, $(\mathcal{M}_i)_{i \in I}$ a family of \mathcal{L} -structures, $\varphi(v_1, \dots, v_n)$ an \mathcal{L} -formula and let \mathcal{U} be an ultrafilter on I . Let $\mathcal{M} = \prod_{i \in I} \mathcal{M}_i / \mathcal{U}$ be the ultraproduct.*

Then, for any elements $m_1, \dots, m_n \in \mathcal{M}$, where each m_j is the equivalence class $[(x_i^j)_{i \in I}]$ of a sequence in $\prod_{i \in I} \mathcal{M}_i$, we have:

$$\mathcal{M} \models \varphi(m_1, \dots, m_n) \iff \{i \in I : \mathcal{M}_i \models \varphi(x_i^1, \dots, x_i^n)\} \in \mathcal{U}$$

Proof. The proof proceeds by induction on the complexity of the formula φ . We utilize the fact that the full set of logical connectives and quantifiers can be defined using only the minimal set $\{\perp, \rightarrow, \exists\}$. Establishing the theorem for this restricted set of primitives is sufficient to prove the result for all formulas of first-order logic (see [CL00, Remark 3.57]).

Throughout the proof, let $m_j = [(x_i^j)_{i \in I}]$ for each $j \in \{1, \dots, n\}$.

Case 1: Atomic formulas. By the definition of the ultraproduct, the interpretations of function symbols and relation symbols are defined component-wise modulo \mathcal{U} . It follows that for any term t , we have $t^{\mathcal{M}}(m_1, \dots, m_n) = [(t^{\mathcal{M}_i}(x_i^1, \dots, x_i^n))_{i \in I}]$.

- If φ is $t_1 \doteq t_2$, then $\mathcal{M} \models \varphi$ if and only if $t_1^{\mathcal{M}} = t_2^{\mathcal{M}}$. By the definition of the equivalence relation $\sim_{\mathcal{U}}$, this holds if and only if $\{i \in I : t_1^{\mathcal{M}_i} = t_2^{\mathcal{M}_i}\} \in \mathcal{U}$, which is exactly $\{i \in I : \mathcal{M}_i \models \varphi\} \in \mathcal{U}$.
- If φ is $R(t_1, \dots, t_k)$, the result follows immediately from the definition of the interpretation of relation symbols in the reduced product: $(m_1, \dots, m_k) \in R^{\mathcal{M}}$ if and only if $\{i \in I : (x_i^1, \dots, x_i^k) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$.

Case 2: The formula φ is \perp . By the definition of satisfaction, $\mathcal{M} \models \perp$ is always false. Similarly, the set $\{i \in I : \mathcal{M}_i \models \perp\}$ is the empty set \emptyset . Since $\emptyset \notin \mathcal{U}$ by property (v) of the definition of an ultrafilter, the equivalence holds.

Case 3: The formula φ is $G \rightarrow H$. Assume the theorem holds for G and H . Recall that $\mathcal{M} \models G \rightarrow H$ is equivalent to saying that it is not the case that $(\mathcal{M} \models G$ and $\mathcal{M} \not\models H)$. Let $X_G = \{i \in I : \mathcal{M}_i \models G\}$ and $X_H = \{i \in I : \mathcal{M}_i \models H\}$. By the inductive hypothesis:

$$\mathcal{M} \models G \rightarrow H \iff (X_G \in \mathcal{U} \implies X_H \in \mathcal{U}).$$

Using the properties of an ultrafilter (specifically that \mathcal{U} is prime and closed under supersets), the condition $(X_G \in \mathcal{U} \implies X_H \in \mathcal{U})$ is equivalent to $(I \setminus X_G) \cup X_H \in \mathcal{U}$. This union is precisely the set $\{i \in I : \mathcal{M}_i \models G \rightarrow H\}$.

Case 4: The existential quantifier. For the case where φ is of the form $\exists v \psi(v, m_1, \dots, m_n)$, the argument involves picking representatives for the existential witness in each coordinate where the formula holds. This requires the Axiom of Choice and follows a standard construction; for the full details of this step, the reader is referred to [CL01, Theorem 8.31]. \square

Having established Łoś's Theorem, we now provide a purely semantic proof of the Compactness Theorem. The strategy is to “average out” a collection of models for finite subsets of Γ into a single structure that satisfies the entire set.

Outline of the proof: First, we define an index set I consisting of all finite subsets of Γ and associate each with a corresponding model. We then identify a collection of subsets of I that possess the finite intersection property, allowing us to extend them to an ultrafilter \mathcal{U} via the Ultrafilter Lemma. Finally, we construct an ultraproduct \mathcal{M} and apply Łoś's Theorem to show that \mathcal{M} is a model for Γ , completing the proof.

Proof. Let Γ be a set of first-order formulas such that every finite subset of Γ is satisfiable. We wish to show that Γ is satisfiable.

Let I be the collection of all finite subsets of Γ , i.e., $I = \{\Gamma_0 \subseteq \Gamma \mid \Gamma_0 \text{ is finite}\}$. By our hypothesis, for each $i \in I$, there exists a structure \mathcal{M}_i such that $\mathcal{M}_i \models i$. For each formula $\varphi \in \Gamma$, we define the set of indices $J_\varphi \subseteq I$ that “contain” φ :

$$J_\varphi = \{i \in I \mid \varphi \in i\}.$$

We observe that the collection $\{J_\varphi \mid \varphi \in \Gamma\}$ has the *finite intersection property*, (see 2.20). Indeed, for any finite selection of formulas $\varphi_1, \dots, \varphi_n \in \Gamma$, the index $i^* = \{\varphi_1, \dots, \varphi_n\}$ is an element of I , and clearly $i^* \in J_{\varphi_1} \cap \dots \cap J_{\varphi_n}$.

By the Ultrafilter Lemma, there exists an ultrafilter \mathcal{U} on I such that $\{J_\varphi \mid \varphi \in \Gamma\} \subseteq \mathcal{U}$. Now, consider the ultraproduct structure:

$$\mathcal{M} = \prod_{i \in I} \mathcal{M}_i / \mathcal{U}.$$

Claim: $\mathcal{M} \models \Gamma$.

Proof of claim: Let φ be any formula in Γ . By definition of our indices, if $i \in J_\varphi$, then $\varphi \in i$. Since $\mathcal{M}_i \models i$, it follows that for all $i \in J_\varphi$, we have $\mathcal{M}_i \models \varphi$. Thus:

$$J_\varphi \subseteq \{i \in I \mid \mathcal{M}_i \models \varphi\}.$$

Since $J_\varphi \in \mathcal{U}$ and \mathcal{U} is upward-closed, the set $\{i \in I \mid \mathcal{M}_i \models \varphi\}$ must also be an element of \mathcal{U} . By **Łoś’s Theorem**, $\mathcal{M} \models \varphi$ if and only if the set of indices where the formula holds is in the ultrafilter. Therefore, $\mathcal{M} \models \varphi$ for every $\varphi \in \Gamma$.

Since we have constructed a structure \mathcal{M} that satisfies every formula in Γ , the set Γ is satisfiable. □

4 Applications of Compactness

There are numerous ways to apply the Compactness Theorem to diverse areas of mathematics. In this section, we provide three distinct applications: the construction of non-standard models of PA , the topological interpretation of logical compactness, and the application of the theorem to mathematical tiling.

4.1 Non-standard models of PA

We let PA be the Peano axioms and let \mathcal{L}_A be the language of arithmetic. We consider the construction of models that are not isomorphic to the standard model $\mathcal{N} = \langle \mathbb{N}, 0, S, +, \times \rangle$.

Proposition 4.1. *There are 2^{\aleph_0} pairwise non-isomorphic countable models of PA .*

Proof. Let \mathbb{P} be the set of prime numbers. For any subset $S \subseteq \mathbb{P}$, we define a theory T_S in the expanded language $\mathcal{L}_c = \mathcal{L}_A \cup \{c\}$, where c is a new constant symbol:

$$T_S := PA \cup \{\bar{p} \mid c : p \in S\} \cup \{\bar{p} \nmid c : p \in (\mathbb{P} \setminus S)\} \cup \{c > \bar{n} : n \in \mathbb{N}\}.$$

To show T_S has a model, we apply the Compactness Theorem. Let $\Gamma_0 \subseteq T_S$ be a finite subset. Γ_0 can only contain finitely many divisibility requirements and finitely many inequalities of the form $c > \bar{n}$. We can satisfy Γ_0 in the standard model \mathcal{N} by choosing c to be a sufficiently large integer that is a multiple of all primes in $S \cap \Gamma_0$ and is not divisible by any primes in $(\mathbb{P} \setminus S) \cap \Gamma_0$. (For instance, one could take c to be the product of all primes in $S \cap \Gamma_0$ and, if necessary, add a large enough multiple of that product to satisfy the $c > \bar{n}$ requirements). Since every finite subset is satisfiable, T_S has a model.

Since the language \mathcal{L}_c is countable, the satisfiability of T_S implies the existence of a countable model \mathcal{M}_S . We now show that different subsets S lead to many non-isomorphic models.

Claim: For any fixed countable model $\mathcal{M} \models PA$, the set of $S \subseteq \mathbb{P}$ such that $\mathcal{M} \cong \mathcal{M}_S$ is at most countable.

Proof of Claim: Suppose $\mathcal{M} \cong \mathcal{M}_S$. Then there must exist an element a in the universe of \mathcal{M} that satisfies the same properties as the constant c (namely, it is greater than all standard integers and has the same prime divisors). The set S is

uniquely determined by this element a via:

$$S = \{p \in \mathbb{P} : \mathcal{M} \models \bar{p} \mid a\}.$$

Since the universe of \mathcal{M} is countable, there are only countably many candidates for the element a . Consequently, \mathcal{M} can represent at most countably many different sets S .

There are 2^{\aleph_0} possible subsets $S \subseteq \mathbb{P}$. Since each countable model accounts for at most a countable number of these sets, there must be 2^{\aleph_0} distinct isomorphism classes of models to represent all possible S . \square

4.2 The Topological Interpretation

The nomenclature of the “Compactness Theorem” is not accidental; it arises from a deep connection with point-set topology [Car13, Chapter 14.3].

Consider the set of all possible interpretations Ω of a language \mathcal{L} . We define an equivalence relation \sim_Ω on Ω by considering two interpretations \mathcal{M}_1 and \mathcal{M}_2 to be equivalent (specifically, *elementarily equivalent*) if they satisfy the same set of \mathcal{L} -formulas. That is:

$$\mathcal{M}_1 \sim_\Omega \mathcal{M}_2 \iff \forall \varphi \in \mathcal{L}, (\mathcal{M}_1 \models \varphi \iff \mathcal{M}_2 \models \varphi).$$

This relation partitions Ω into equivalence classes. Let $\mathcal{X} = \Omega / \sim_\Omega$ be the quotient set of these classes. We can then endow \mathcal{X} with a topology by defining a collection of *basic open sets*. For each formula φ in the language \mathcal{L} , let U_φ be the set of all equivalence classes of models that satisfy φ :

$$U_\varphi = \{[\mathcal{M}] \in \mathcal{X} : \mathcal{M} \models \varphi\}.$$

By taking the collection $\{U_\varphi : \varphi \in \mathcal{L}\}$ as a basis for the topology, the set of formulas true in a model corresponds to the “location” of that model in the space \mathcal{X} . Note that each U_φ is clopen (both closed and open), as its complement is the open set $U_{\neg\varphi}$.

In this context, the Compactness Theorem of logic is precisely the statement that this topological space \mathcal{X} is compact. While we defined compactness via open covers, it is topologically equivalent to the statement that any collection of closed

sets with the finite intersection property has a non-empty total intersection (see [Con73, Chapter 2, proposition 4.4]). Logically, this translates perfectly: if every finite subset of a theory Γ is satisfiable (the FIP), then the total intersection of the corresponding closed sets U_φ for all $\varphi \in \Gamma$ is non-empty, meaning the entire theory Γ is satisfiable.

4.3 Tiling of the plane

For this section, we will concern ourselves with the integer plane \mathbb{Z}^2 . We can think of this plane as an infinitely extending chessboard where each square is represented by a coordinate $(x, y) \in \mathbb{Z}^2$. This grid serves as the underlying structure for our investigation into tilings.

$(-3, 3)$	$(-2, 3)$	$(-1, 3)$	$(0, 3)$	$(1, 3)$	$(2, 3)$	$(3, 3)$
$(-3, 2)$	$(-2, 2)$	$(-1, 2)$	$(0, 2)$	$(1, 2)$	$(2, 2)$	$(3, 2)$
$(-3, 1)$	$(-2, 1)$	$(-1, 1)$	$(0, 1)$	$(1, 1)$	$(2, 1)$	$(3, 1)$
$(-3, 0)$	$(-2, 0)$	$(-1, 0)$	$(0, 0)$	$(1, 0)$	$(2, 0)$	$(3, 0)$
$(-3, -1)$	$(-2, -1)$	$(-1, -1)$	$(0, -1)$	$(1, -1)$	$(2, -1)$	$(3, -1)$
$(-3, -2)$	$(-2, -2)$	$(-1, -2)$	$(0, -2)$	$(1, -2)$	$(2, -2)$	$(3, -2)$
$(-3, -3)$	$(-2, -3)$	$(-1, -3)$	$(0, -3)$	$(1, -3)$	$(2, -3)$	$(3, -3)$

The visual grid above illustrates the discrete nature of the space we are working with. To “tile” this space, we must decide on a set of allowed shapes, namely clusters of these coordinate points, and determine whether they can be arranged to cover the plane without gaps or overlaps. We formalize these shapes and their arrangements in the following definition.

Definition 4.2. A *tile* is a finite subset of \mathbb{Z}^2 . Note that a tile is not required to be “connected” in the topological sense. Let \mathcal{T} be a collection of tiles. A \mathcal{T} -*tiling* of a subset $S \subseteq \mathbb{Z}^2$ is a partition of S into disjoint, nonempty sets $\{C_i\}_{i \in I}$ such that

$$S \subseteq \bigcup_{i \in I} C_i$$

where each C_i is a translation/rotation of some tile $T \in \mathcal{T}$. That is, for each i , there exists $T \in \mathcal{T}$, a translation vector $\mathbf{v} \in \mathbb{Z}^2$, and a rotation matrix $M \in \mathcal{R}$ such that

$$C_i = \{M\tau + \mathbf{v} : \tau \in T\}$$

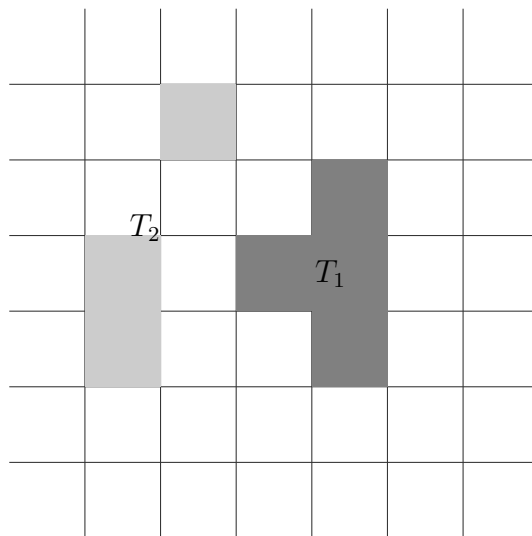
where

$$\mathcal{R} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right\}.$$

Example 4.3. We illustrate this definition with two simple tiles and a subset $S \subseteq \mathbb{Z}^2$. Let

$$T_1 = \{(0, 0), (1, 0), (1, 1), (1, -1)\}, \quad T_2 = \{(-2, 0), (-2, -1), (-1, 2)\}.$$

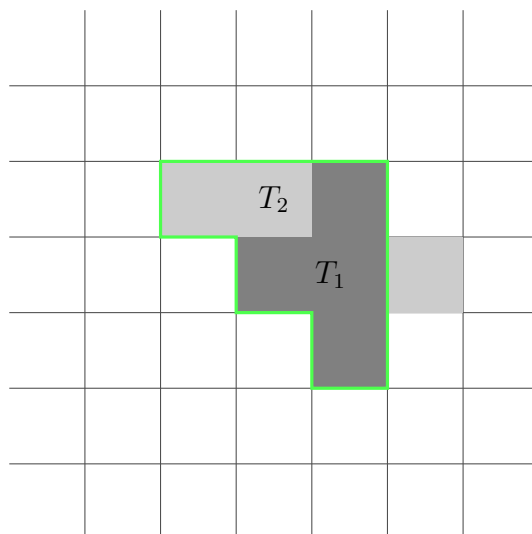
Both sets define valid tiles, whose shapes on the grid are shown below.



Now consider the subset

$$S = \{(-1, 1), (0, 1), (1, 1), (0, 0), (1, 0), (1, -1)\} \subseteq \mathbb{Z}^2.$$

The shape of S is outlined in the figure below, in green. Using one copy each of T_1 and T_2 , we obtain a \mathcal{T} -tiling of S . Note that, according to our definition, the tiling may extend beyond S ; in this example, the tile T_2 covers the additional cell $(2, 0)$ outside of S .



One may now see, at least intuitively, how the Compactness Theorem applies to this setting. Tiling the infinite grid \mathbb{Z}^2 can be viewed as a problem of global consistency determined by local constraints. Fix a set of tiles \mathcal{T} , and let $S \subset \mathbb{Z}^2$ be a finite region. The question of whether \mathcal{T} tiles S can be expressed in the language of first-order logic: we introduce propositions describing which tile is placed at each position, together with formulas enforcing the tiling rules (no gaps/overlaps). A tiling of S then corresponds precisely to a satisfying assignment, that is, a model of this finite set of formulas.

The Compactness Theorem states that if every finite subset of a set of formulas is satisfiable, then the entire set is satisfiable. Interpreted in this context, it implies that if a tile set \mathcal{T} can tile every finite region $S \subset \mathbb{Z}^2$, no matter how large, then there exists a valid \mathcal{T} -tiling of the entire infinite plane \mathbb{Z}^2 .

References

- [Car13] Jesper Carlström. *Logic*. Trans. by Christian Espíndola and Håkon Gylterud. Course notes compendium, Stockholm University. Translated from the Swedish. 2013.
- [CL00] René Cori and Daniel Lascar. *Mathematical logic. A course with exercises. Part I: Propositional calculus, Boolean algebras, predicate calculus, completeness theorems. Transl. from the French by Donald H. Pelletier*. English. Oxford: Oxford University Press, 2000. ISBN: 0-19-850049-1; 0-19-850048-3.
- [CL01] René Cori and Daniel Lascar. *Mathematical logic. A course with exercises. Part II. Recursion theory, Gödel's theorems, set theory, model theory. Translated from the 1993 French original by Donald H. Pelletier*. English. Oxford: Oxford University Press, 2001. ISBN: 0-19-850050-5; 0-19-850051-3.
- [Con73] John B. Conway. *Functions of one complex variable*. English. Vol. 11. Grad. Texts Math. Springer, Cham, 1973.
- [Daw93] John W. Dawson Jr. “The compactness of first-order logic: from Gödel to Lindström”. In: *Hist. Philos. Logic* 14.1 (1993), pp. 15–37. ISSN: 0144-5340,1464-5149. DOI: [10.1080/01445349308837208](https://doi.org/10.1080/01445349308837208). URL: <https://doi.org/10.1080/01445349308837208>.
- [Smi13] Peter Smith. *An introduction to Gödel's theorems*. English. 2nd ed. Camb. Introd. Philos. Cambridge: Cambridge University Press, 2013. ISBN: 978-1-107-60675-3; 978-1-107-02284-3. DOI: [10.1017/CB09780511800962.019](https://doi.org/10.1017/CB09780511800962.019).