

The Compactness Theorem

The compactness theorem is a fundamental theorem for the model theory of (classical) propositional and first-order logic. As well as having importance in several areas of mathematics, such as algebra and combinatorics, it also helps to pinpoint the strength of these logics, which are the standard ones used in mathematics and arguably the most important ones in philosophy.

The main focus of this article is the many different proofs of the compactness theorem, applying different Choice-like principles before later calibrating the strength of these and the compactness theorems themselves over the set theory ZF. Although our focus is mathematical, much of the discussion keeps an eye on philosophical applications and implications.

We first introduce some standard logics, detailing whether the compactness theorem holds or fails for these. We also broach the neglected question of whether natural language is compact. Besides algebra and combinatorics, the compactness theorem also has implications for topology and foundations of mathematics, via its interaction with the Axiom of Choice. We detail these results as well as those of a philosophical nature, such as apparent ‘paradoxes’ and non-standard models of arithmetic and analysis. We then provide several different proofs of the compactness theorem based on different Choice-like principles.

In later sections, we discuss several variations of compactness in logics that allow for infinite conjunctions / disjunctions or generalised quantifiers, and in higher-order logics. The article concludes with a history of the compactness theorem and its many proofs, starting from those that use syntactic proofs before moving to the semantic proofs model theorists are more accustomed to today.

Contents

1. Introduction

2. Compactness: Common Logics and Natural Language

- a. Common Logics
- b. Natural Language

3. Implications of Compactness

4. Some Non-topological Proofs

- a. Deductive: Proofs via Soundness and Completeness
- b. Syntactic: Henkin-Style Proofs
- c. Semantic: Ultraproduct Proofs

5. Connection to Topology

6. Extensions and Generalisations

7. Relative Strength

- a. ZF-Equivalents
- b. Principles Strictly ZF-Stronger Than the Compactness Theorem
- c. Principles Strictly ZF-Weaker Than the Compactness Theorem

8. History of the Compactness Theorem

9. References and Further Reading

1. Introduction

A *logic* consists of a language, grammar, semantics, and consequence relation \models . If Γ is a set of sentences of this logic and δ one of its sentences, $\Gamma \models \delta$ means that any model of Γ is a model of δ . (A model of Γ is a model of all sentences in Γ .) Informally, a logic is called *compact* if it is determined by its behaviour on finite sets of sentences; there may be infinitely many sentences in the language, but we can always reduce our considerations to finitely many in any given situation. More formally, a logic is compact just when:

- If every finite subset Γ^{fin} of Γ is satisfiable then Γ is also satisfiable.
- If Γ is an unsatisfiable set of sentences then so is Γ^{fin} for some finite subset Γ^{fin} of Γ .

Some authors take the compactness of a logic to be its satisfaction of these statements' biconditional versions. We have chosen to omit the reverse implication from the definition as it easily follows from the meaning of $\Gamma \models \delta$. These two characterisations of compactness are equivalent, since the second statement is effectively the contrapositive of the first. In a logic containing a classical negation connective (by which we mean for each sentence δ there is a sentence $\neg\delta$ such that $\mathfrak{M} \models \neg\delta$ if and only if $\mathfrak{M} \not\models \delta$), both statements are equivalent to:

- If $\Gamma \models \delta$ then $\Gamma^{\text{fin}} \models \delta$ for some finite subset Γ^{fin} of Γ .

This equivalence follows from

$$\Gamma \models \delta \text{ if and only if } \Gamma \cup \{\neg\delta\} \text{ is unsatisfiable.}$$

The compactness theorem is said to hold for a logic precisely when the logic is compact.

Alongside its close cousin, the completeness theorem for first-order logic, the compactness theorem for first-order logic is one of the most important theorems in contemporary logic. In this entry, we give a few examples of compact and incompact logics and briefly discuss whether natural languages such as English are compact (Section 2). We then mention some mathematical and philosophical implications of the compactness of first-order logic (Section 3). Following that, we give some non-topological proofs of the compactness of propositional and first-order logic (Section 4), followed by a topological proof of the propositional case, which gives the compactness theorem its name (Section 5). We continue with a sketch of some generalisations of the usual notion of compactness (Section 6), a calibration of the strength of the compactness theorems relative to the ZF axioms

(Section 7), and end with some notes on the history of the compactness theorems (Section 8). Our discussion concerns the logics philosophers are most familiar with.

2. Compactness: Common Logics and Natural Language

a. Common Logics

Propositional logic is usually taken to consist of a set of sentential atoms $\{p_1, \dots, p_n, \dots\}$ and some truth-functionally complete set of Boolean connectives, for instance $\{\neg, \vee, \wedge\}$. If we denote by PL_κ a propositional logic with κ sentential atoms, PL_ω is a propositional logic with a countable infinity of atoms. (κ will usually be taken to be an infinite cardinal in what follows.) Any propositional logic PL_κ is compact, whatever its set of truth-functional connectives may be. Notice that when $\kappa = n$ is finite, the compactness of any PL_n is a trivial consequence of the fact that any sentence is logically equivalent to a sentence drawn from a fixed set of size no greater than 2^{2^n} . (This set is of size exactly 2^{2^n} just when the set of connectives is truth-functionally complete.)

First-order logic with standard semantics is also compact. This fact is of tremendous importance for logic and its applications, since first-order logic remains the canonical logic to this day, the widespread interest in higher, supplementary and alternative logics notwithstanding. By ‘first-order logic’, we understand throughout first-order logic with identity. First-order logic without identity is of course also compact, since it is a sublogic of first-order logic with identity.

By second-order logic we mean second-order logic with standard or full semantics, in which second-order n -place predicate variables range over all the n -tuples from the domain of interpretation (and similarly for functional variables). In contrast to first-order logic, second-order logic is not compact. To see this, let $\exists_{\geq n}$ be a sentence of first-order logic satisfied in all and only models with domain of size $\geq n$; $\exists_{\geq 1}$ may thus be taken as $\exists x(x = x)$, $\exists_{\geq 2}$ as $\exists x \exists y \neg(x = y)$, and so on. Since first-order logic is a sublogic of second-order logic, $\exists_{\geq n}$ is a sentence of second-order logic too. Consider next the sentence

$$\exists R(R \text{ is functional} \wedge R \text{ is injective} \wedge \neg R \text{ is surjective})$$

where R is a binary predicate variable, ‘ R is functional’ abbreviates $\forall x \exists! y Rxy$, ‘ R is injective’ abbreviates $\forall x \forall y \forall z ((Rxz \wedge Ryz) \rightarrow x = y)$ and ‘ R is surjective’ abbreviates $\forall y \exists x Rxy$. Any interpretation of this sentence states that the domain is Dedekind-infinite. The following second-order argument is then valid:

$$\begin{array}{c} \exists_{\geq 1} \\ \exists_{\geq 2} \\ \vdots \\ \exists_{\geq n} \\ \vdots \end{array}$$

$$\exists R(R \text{ is functional} \wedge R \text{ is injective} \wedge \neg R \text{ is surjective})$$

However, no finite subset of the premisses entails the conclusion. For let the finite subset be $\{\exists_{\geq i_1}, \exists_{\geq i_2}, \dots, \exists_{\geq i_k}\}$ and take $m \geq \max\{i_1, i_2, \dots, i_k\}$. Then there is a model of size m in which the k premisses $\exists_{\geq i_1}, \exists_{\geq i_2}, \dots, \exists_{\geq i_k}$ are true but the argument's conclusion is false. Hence second-order logic is not compact.

The compactness theorem also typically, but not invariably, fails for infinitary logics. Any logic which allows infinite disjunctions, for example, is incompact, since the set of sentences $\{c \neq c_i : i \in \omega\} \cup \{\bigvee_{i \in \omega} c = c_i\}$ is *finitely-satisfiable* (every finite subset is satisfiable) but unsatisfiable. We return to infinitary logics and to generalisations of the notion of compactness in Section 6.

b. Natural Language

Natural languages are languages such as English, Mandarin, French, and Arabic. Formal languages in contrast are logical languages such as those of propositional, first-order and second-order logic. Although of great mathematical and philosophical importance, the latter are not 'natural' in the intended sense because they are not anyone's native language and are only ever 'spoken', if at all, in limited contexts.

Is natural language, say English, compact? We must first clarify what the question means. Assume there is such a thing as the relation of logical consequence in natural language. For example, consider these two natural-language arguments:

Hypatia is a woman. All women are mortal. <hr style="width: 80%; margin: 10px auto;"/> Hypatia is mortal.	Hypatia is mortal. All women are mortal. <hr style="width: 80%; margin: 10px auto;"/> Hypatia is a woman.
---	---

The argument on the left is logically valid, whereas the argument on the right is invalid. These two examples, of a logically valid English argument and a logically invalid one respectively, help us home in on the notion we are interested in, namely English logical validity, but they do not, of course, provide definitions of it. Next, let's say that a natural language is compact just when, for any logically valid argument in this language, there is a logically valid argument whose conclusion remains the same, yet the premiss set is a finite subset of the original argument's premiss set. This definition is the analogue of our last definition of compactness above for a formal language: if $\Gamma \models \delta$ then $\Gamma^{\text{fin}} \models \delta$ for some finite subset Γ^{fin} of Γ . An equivalent definition could be given based on the other definition of compactness for a formal language.

The overwhelming majority of linguists, philosophers, and other theorists of language take natural language to consist of infinitely many sentences. The idea is that since such sentences are of finite, but arbitrary, length, there must be infinitely many of them. The set of these sentences may be specified by a set of recursive procedures, which generate sentences of arbitrary length. For example, all the following are sentences of English:

My grandfather was tall;
 My great-grandfather was tall;
 My great-great-grandfather was tall;
 ...

Now as a matter of empirical fact, there is some finite number N such that I do not have a great ^{N} -grandfather (which N is the least such may be vague). That

does not affect the point at issue, which is that these infinitely many sentences are bona fide sentences of English.

Consider now an English argument roughly analogous to the following English analogue of the second-order-logic argument presented earlier:

There is at least one planet.
 There are at least two planets.
 ⋮
 There are at least n planets.
 ⋮

There are infinitely many planets.

This argument appears to be valid. Clearly, no finite subset of the premiss set entails the conclusion. If this is right, the English consequence relation is incompact. The moral carries over to any natural language into which the argument is translatable.

Resistance to the argument for the incompactness of English may take several forms. One line of resistance, for instance, would query whether any natural language has a *single* consequence relation. According to this objection, there are only various consequence relations that arise from looking at, say, English through a particular theoretical lens; none is the correct one. Logical pluralists would all take this view, as would some other philosophers of logic: Beall and Restall (2006) and Shapiro (2014) defend different versions of logical pluralism. The objection, then, is that the argument for the incompactness of English given above assumes that there is a determinate notion: *the* English consequence relation.

This objection, and others, must be addressed before we can conclude that English is incompact. Here we do not take sides but flag the issue as an important one. For more discussion, see Paseau and Griffiths (2021), and Griffiths and Paseau (2022, chap. 5).

3. Implications of Compactness

This section draws some implications of the compactness of first-order logic. The sample below is a small selection from a list that could fill volumes. Our choices are mostly guided by their philosophical implications, although there are a few examples of primarily mathematical interest. From this point on, we assume some knowledge of elementary model theory (Chang and Keisler 1990; Hodges 1997).

- (i) Any compact logic extending first-order logic cannot express the notions of finitude or infinitude (of a model). Suppose towards a contradiction that ϕ_F is satisfied by all and only finite models. Then $\{\phi_F\} \cup \{\exists_{\geq n} : n \in \omega\}$ is unsatisfiable, and hence by compactness it must have an unsatisfiable finite subset, which must be a subset of $\{\phi_F\} \cup \{\exists_{\geq i_1}, \dots, \exists_{\geq i_k}\}$ for some $i_1 < \dots < i_k$. But any finite model with domain of size $\geq i_k$ satisfies (any subset of) $\{\phi_F\} \cup \{\exists_{\geq i_1}, \dots, \exists_{\geq i_k}\}$, thereby contradicting our hypothesis.

And if there were a sentence ϕ_I satisfied by all and only infinite models then $\neg\phi_I$ would be satisfied by all and only finite models, a hypothesis we have just refuted.

This application of the compactness theorem is entirely typical. Schematically, one shows by contradiction that the class of models with some ω -property (expressible by the conjunction of an infinite set S of first-order sentences) is not definable by a single sentence ϕ . In these arguments, $\neg\phi$ together with any finite subset of S is satisfiable, but $\{\neg\phi\} \cup S$ is not, contradicting compactness. Informally speaking, in these applications the ω -property is the conjunction of the n -properties; in our example, the n -property was having size at least n and the ω -property was having infinite size.

However, it is possible that (depending on the language) there is a sentence which *implies* that any of these models must be infinite. As an example, take the first-order language with a single unary function symbol, and take the sentence

$$\phi := \forall x \forall y (f(x) = f(y) \rightarrow x = y) \wedge \exists x \forall y \neg (f(y) = x).$$

The models of this sentence are sets endowed with an injective, but non-surjective function. By the pigeonhole-principle, each such model must be infinite. Since this sentence cannot express infinitude, there must be an infinite model not satisfying ϕ . An example is the domain of natural numbers, over which f is interpreted as the identity function on the natural numbers.

As a further illustration of this technique, compactness implies that the class of torsion-free abelian groups is not finitely-axiomatisable in first-order logic. (A group is *torsion-free* if the identity is the only element with finite order. The prototypical example of a torsion-free abelian group is $(\mathbb{Z}, +)$.) For if it were, by the single sentence ϕ say, then the set

$$\{\text{abelian group axioms}\} \cup \{\neg\phi\} \cup \{\forall x \neq 0 \underbrace{(x + \dots + x)}_{n \text{ times}} \neq 0 : n = 1, 2, \dots\}$$

would be finitely-satisfiable (for finite sets where the sums are bounded by n in the sentences from the rightmost set, take the integers under addition modulo p for some prime $p > n$). But the set itself is unsatisfiable, since

$$\{\text{abelian group axioms}\} \cup \{\forall x \neq 0 \underbrace{(x + \dots + x)}_{n \text{ times}} \neq 0 : n = 1, 2, \dots\}$$

is satisfied by all and only torsion-free abelian groups. This contradicts compactness. Note that here the ω -property is that all non-identity elements of the model have infinite order, and the n -property that all non-identity elements have order $\neq n$ (all these properties incorporate the abelian group axioms too).

As a still further illustration, the same type of argument shows that the class of algebraically-closed fields is also not finitely-axiomatisable. The relevant ω -property is that the field is algebraically-closed, and the relevant n -property is

$$\forall y_1 \forall y_2 \dots \forall y_n \exists x (x^n + y_1 \cdot x^{n-1} + \dots + y_{n-1} \cdot x + y_n = 0)$$

This style of argument is easily applied to many other domains.

- (ii) Suppose Σ_1 and Σ_2 are two sets of sentences of a compact logic (in which conjunction and negation are definable) with the property that every model satisfies either Σ_1 or Σ_2 but not both. Then there is a sentence σ such that Σ_1 is logically equivalent to σ (meaning that $\{\sigma\}$ and Σ_1 have the same models) and Σ_2 is logically equivalent to $\neg\sigma$ (Chang and Keisler 1990, p. 12, cor. 1.2.15).
- (iii) The compactness theorem may be used to show that any first-order theory of arithmetic T_{AR} satisfied by the standard model has a non-standard model. (By the standard model of arithmetic, for the theory T_{AR} in question, we mean the structure of natural numbers with the standard interpretation of the non-logical symbols in the language of T_{AR} ; the constant $\bar{0}$ denotes 0, the two-place symbol $+$ denotes addition, \cdot denotes multiplication, and so forth. A nonstandard model can have ‘infinite’ elements: objects greater than those denoted by the terms \bar{n} for all n . As an ordered set, a nonstandard model looks like \mathbb{N} followed by blocks of \mathbb{Z} , which themselves form a dense linear order without endpoints.) Assuming that each numeral \bar{n} is definable in T_{AR} , consider

$$T_{AR}^+ = T_{AR} \cup \{c \neq \bar{n} : n \in \omega\}$$

where c is any constant not in the language of T_{AR} . Any finite subset of T_{AR}^+ is satisfied by the standard model, because we may interpret c as a number distinct from those n where $c \neq \bar{n}$ occurs in the given finite subset. Hence by compactness, T_{AR}^+ has a model \mathfrak{M} . The reduct of \mathfrak{M} to the language of T_{AR} is non-standard since it contains an element $c^{\mathfrak{M}}$ (the denotation of c in \mathfrak{M}) not identical to any natural number.

Supplementing the argument just given with an appeal to the *downward Löwenheim-Skolem Theorem* (Chang and Keisler 1990, p. 66, cor. 2.1.4; Hodges 1997, p. 72, cor. 3.1.4) shows that any such first-order theory of arithmetic T_{AR} is not even \aleph_0 -categorical, since it has a countably infinite non-standard model. (For an infinite cardinal κ , a theory is κ -categorical if it has exactly one model, up to isomorphism, of cardinality κ .) Observe that a theory may be complete even if it is not \aleph_0 -categorical. To see this, run the previous argument supplemented by an application of the downward Löwenheim-Skolem Theorem for a *complete* theory of arithmetic T_{AR} . More generally, a theory may fail to be λ -categorical for every infinite cardinal λ and still be complete—in other words the converse of the *Łoś-Vaught test* (Chang and Keisler 1990, p. 139, prop. 3.1.7) fails.

The same general idea can be used to demonstrate the existence of non-standard models of real analysis. Let T_{AN} be a first-order theory of analysis satisfied by the standard model (the ordered field of real numbers). As above, consider $T_{AN}^+ = T_{AN} \cup \{\bar{0} < c < \bar{n}^{-1} : n = 1, 2, \dots\}$ where c is any constant not in T_{AN} ’s language. Any finite subset of T_{AN}^+ is satisfied by the standard model, because we may interpret c as a positive real number smaller than $\frac{1}{n}$, where n is the largest number for which the sentence $\bar{0} < c < \bar{n}^{-1}$ is in the given finite subset. Hence by compactness, T_{AN}^+ has a model \mathfrak{M} . The reduct of \mathfrak{M} to the language of T_{AN} is non-standard since it contains an element $c^{\mathfrak{M}}$ not identical to any real number. Indeed, this element must be a positive infinitesimal, meaning that it is a number greater than $\bar{0}^{\mathfrak{M}}$ but smaller than every $(\bar{n}^{-1})^{\mathfrak{M}}$. As well as infinitesimals, our non-standard model also contains infinite elements, since the the model

satisfies $\forall x \neq 0 \exists y(x \cdot y = 1)$ and thus any non-zero element has an inverse. From these foundations, a consistent theory of the calculus that revives to a degree the use of infinitesimals in early modern mathematics may be constructed, called *non-standard analysis*: see Goldblatt (1998) or the original Robinson (1966). Furthermore, since \mathbb{R} and \mathfrak{M} have the same theory, any sentence in the language of T_{AN} that can be proven from T_{AN}^+ will hold in \mathfrak{M} and thus automatically holds in \mathbb{R} . This is the *transfer principle* of non-standard analysis.

Since second-order logic is incompact, the arguments just given fail for second-order theories of arithmetic and analysis. Indeed, there are categorical second-order axiomatisations of arithmetic (for example, Peano Arithmetic with second-order induction (Shapiro 1991, p. 82, thm. 4.8)) and real analysis (for example, the axioms for a complete ordered field (p. 84, thm. 4.10)). For more on second-order theories, consult (chap. 3–6).

- (iv) Assuming the downward Löwenheim-Skolem theorem, another corollary of the compactness of first-order logic is the *upward Löwenheim-Skolem theorem* (Chang and Keisler 1990, p. 67, cor. 2.1.6; Hodges 1997, p. 127, cor. 5.1.4). This upward version of the theorem states that if a first-order language \mathcal{L} has cardinality $\leq \lambda$ and \mathfrak{M} is an infinite model with domain of cardinality $\leq \lambda$ then \mathfrak{M} has an elementary extension of cardinality λ . For the proof, we consider the set of sentences consisting of the *elementary diagram* (the set $\{\phi \in \mathcal{L}(\mathfrak{M}) : \mathfrak{M} \models \phi\}$, where $\mathcal{L}(\mathfrak{M})$ is the extension of \mathcal{L} obtained by adding constant symbols c_m for each m in the domain of \mathfrak{M} , and the interpretation of c_m in \mathfrak{M} is simply m) together with the sentences $c_\alpha \neq c_\beta$ for all distinct $\alpha, \beta \in \lambda$, where the c_α are new constants. This set is finitely-satisfiable (because the infinite model \mathfrak{M} satisfies any finite subset), and hence by compactness it is satisfiable. Say that it is satisfied by a model \mathfrak{N} , which must be of size $\geq \lambda$ as it satisfies $\{c_\alpha \neq c_\beta : \alpha, \beta \in \lambda \text{ s.t. } \alpha \neq \beta\}$. Since \mathfrak{N} also satisfies the elementary diagram of \mathfrak{M} , an elementary embedding of \mathfrak{M} into \mathfrak{N} exists, and thus there is an elementary extension \mathfrak{D} of \mathfrak{M} with domain of size $\geq \lambda$ (\mathfrak{D} is an isomorph of \mathfrak{N} whose domain includes that of \mathfrak{M}). To find an elementary extension of \mathfrak{M} of size exactly λ , now apply the downward Löwenheim-Skolem theorem to \mathfrak{D} .

The upward Löwenheim-Skolem theorem may be applied to show not only that theories of arithmetic and analysis satisfied by their respective standard models have non-standard models, but also that they have non-standard models of every infinite cardinality. More generally, any first-order theory in a countable language satisfied by an infinite model has models of every cardinality. Applying this to ZF or ZFC, we obtain *Skolem's paradox*: if there exists a model of set theory, then there exists a countably-infinite model, which will nonetheless model the existence of uncountable sets (Skolem 1922, 295–96).

The upward and downward Löwenheim-Skolem theorems and the so-called Skolem paradox have generated much philosophical debate. The most famous philosophical use of these results is found in Putnam (1980). Wielding the theorems, Putnam argued that our mathematical-scientific theories do not admit a determinate interpretation. In particular, he claimed that set theory augmented with any theoretical principles from science and scientific data we care to add admits a countable model. This argument has given rise to an extensive literature. The responses to Putnam's argument, and

to other versions, include technical discussion of exactly what Putnam’s argument requires mathematically speaking, as well as philosophical commentary.

- (v) The compactness theorem may be used to prove the *Order-Extension Principle* (also known as *Szpilrajn’s Extension Theorem*): any partial order may be extended to a linear order. A partial order $(A, <)$ consists of a domain A with an irreflexive and transitive relation $<$. A linear order is a partial order satisfying the additional linearity axiom $\forall x \forall y (x < y \vee x = y \vee y < x)$. A linear order $(A, <_*)$ extends the partial order $(A, <)$ just when for all a_1, a_2 in A , if $a_1 < a_2$ then $a_1 <_* a_2$; in other words, the identity map from $(A, <)$ into $(A, <_*)$ is a homomorphism. The notions of partial and linear order are first-order definable in the language consisting of a single two-place non-logical predicate, which is what permits the application of the compactness theorem.

Unusually for applications of the compactness theorem, this result is of potential importance outside mathematics, logic, and philosophy. A standard assumption in economics is that a subject’s preferences over goods are linearly ordered. Empirically, however, we find that people’s preferences tend at best to be partially ordered: I may prefer going out for a Japanese meal rather than an Italian meal, but I may have no preference between going to the cinema and going out for a Japanese meal. In light of the Order-Extension Principle, one might try to argue that preferences being linearly ordered is a justifiable idealisation of the empirical data. See Szpilrajn (1930) for the first proof of this theorem and Richter (1966) for an early economic application, as well as Gonczarowski, Kominers, and Shorrer (2019) for more applications of the compactness theorem in economics.

How does one prove the Order-Extension Principle? A relatively straightforward proof by induction on the size of the domain shows that every partial order with a finite domain can be extended to a linear order. We may then use the compactness theorem to prove that every partial order can be extended to a linear order, whatever the size of the domain, be it finite or infinite. Given a partial order $(A, <)$, let Σ_A be a set of sentences consisting of the (*Robinson*) *diagram* of $(A, <)$ (the diagram of an \mathcal{L} -structure \mathfrak{A} is the set of literals—atomic sentences or negations of atomic sentences—in the language $\mathcal{L}' = \mathcal{L} \cup \{c_a : a \in A\}$, where c_a are new constant symbols, satisfied in the expansion \mathfrak{A}' of \mathfrak{A} where we interpret $c_a^{\mathfrak{A}'} := a$ for each $a \in A$), together with the axioms $c_a < c_b \vee c_b < c_a$ for all distinct $a, b \in A$. Since every partial order with a finite domain can be extended to a linear order, it follows that any finite subset of Σ_A is satisfiable. By compactness, Σ_A is therefore satisfiable. If \mathfrak{M} is a model of Σ_A then $(A, <)$ embeds into the $\mathcal{L}_{<}$ -reduct of \mathfrak{M} , which we call $(B, <_B)$, via $f : (A, <) \rightarrow (B, <_B)$ say. The required linear order extending $(A, <)$ is the inverse under f of the restriction of the order $<_B$ to $f(A)$, that is $a_1 <_* a_2$ if and only if $f(a_1) <_B f(a_2)$. Note in passing that the Order-Extension Principle implies that every set A can be linearly ordered: simply consider a linear order that extends the empty partial order on A .

The compactness theorem for first-order logic has a great many other applications to model theory—as Keisler (1965, 113) puts it, “the most useful theorem in model theory is probably the compactness theorem”—as well as to set theory, other parts of logic, combinatorics, algebra, algebraic geometry; see Hodges

(1997), especially chapter 5, for more.

4. Some Non-topological Proofs

In this section, we give three styles proofs of the compactness of propositional and first-order logic that do not draw on topology. These three proofs illustrate some important techniques, and could be broadly classified as deductive, syntactic, and semantic methods of proof.

a. Deductive: Proofs via Soundness and Completeness

We recall the definitions of *soundness* and *completeness* for a logic \mathcal{L} equipped with semantic consequence relation \models and deductive consequence relation \vdash . The logic is complete iff for any set of well-formed formulas Γ and any well-formed formula δ , if $\Gamma \models \delta$ then $\Gamma \vdash \delta$; and it is sound iff the converse obtains: if $\Gamma \vdash \delta$ then $\Gamma \models \delta$.

If a logic has a sound and complete proof procedure, it is compact. A simple argument demonstrates this:

- | | | |
|-----|--------------------------------------|--------------------------------------|
| (1) | $\Gamma \models \delta$ | Assumption |
| (2) | $\Gamma \vdash \delta$ | From (1) by Completeness |
| (3) | $\Gamma^{\text{fin}} \vdash \delta$ | From (2) by the finiteness of proofs |
| (4) | $\Gamma^{\text{fin}} \models \delta$ | From (3) by Soundness |

Here Γ^{fin} is some finite subset of Γ . Anything deserving of the name of ‘proof procedure’ usually satisfies a host of syntactic requirements. Given soundness and completeness the only such requirement needed for the validity of the inference above is that the step from (2) to (3) be valid, namely that proofs draw only on finitely many premisses. The argument just given therefore applies to any logic which has a sound and complete proof procedure in this liberal sense. Thus, no incompact logic can be completable by a sound proof procedure; second-order logic in particular cannot.

The proof of compactness via completeness is in an important respect unsatisfactory because it is based on properties incidental to the semantic property of interest. The proof derives compactness, a semantic property, from a property of the logic relating its syntax to its semantics. Thus Keisler (1965, 113): “unlike the completeness theorem, the compactness theorem does not involve the notion of a formal deduction, and so it is desirable to prove it directly without using that notion”. Indeed, from the perspective of a model theorist who sees talk of syntax as a heuristic for the study of certain relations between structures that happen to have syntactic correlates, proving compactness via completeness is tantamount to heresy (Poizat 2000, 53).

b. Syntactic: Henkin-Style Proofs

This proof is modelled on Henkin’s (1949b) proof of the completeness theorem for first-order logic, with its deductive core replaced by a semantic one. We begin with a relatively concrete argument for the simplest case of PL_ω before passing to more abstract versions. Unlike the usual versions of the proof, our argument does not assume any particular set of truth-functional connectives, only that the

set of sentences is countably-infinite (and thus the set of connectives is countable); the argument is usually simpler (but less general) when a particular set of connectives has been specified.

Let $\{s_i : i \in \omega\}$ be an enumeration of the set of sentences of PL_ω . Given a finitely-satisfiable set Γ , define a denumerable sequence of sets Γ_n as follows:

$$\begin{aligned}\Gamma_0 &= \Gamma, \\ \Gamma_{n+1} &= \begin{cases} \Gamma_n \cup \{s_n\} & \text{if } \Gamma_n \cup \{s_n\} \text{ is finitely-satisfiable,} \\ \Gamma_n & \text{otherwise,} \end{cases} \\ \Gamma_\omega &= \bigcup_{n \in \omega} \Gamma_n.\end{aligned}$$

By definition, Γ_n is finitely-satisfiable for all n , hence Γ_ω is finitely-satisfiable since any finite subset of Γ_ω must be drawn from some Γ_n .

From Γ_ω , define the valuation v by letting $v(p) = 1$ if and only if $p \in \Gamma_\omega$, where p is a sentence letter. We now prove that v is a valuation in which all the sentences in Γ_ω (not just the sentence letters) are true. Suppose towards a contradiction that $v(\phi) = 0$ for some $\phi \in \Gamma_\omega$. Without loss of generality (renumbering sentence letters if necessary), let p_1, \dots, p_k be all the sentence letters in ϕ with truth-value 1 under v and p_{k+1}, \dots, p_n all the sentence letters in ϕ with truth-value 0 under v (one of these sets may be empty, in which case the argument to follow is easily modified). By the definition of v , none of p_{k+1}, \dots, p_n is in Γ_ω so for each of these, let Δ_i be a finite subset of Γ_ω such that $\Delta_i \cup \{p_i\}$ is unsatisfiable ($k+1 \leq i \leq n$); some such Δ_i must exist given that p_i was omitted in the construction of Γ_ω . Now consider the set

$$S := \{p_1, \dots, p_k, \phi\} \cup \left(\bigcup_{i=k+1}^n \Delta_i \right).$$

Any valuation in which all the elements of $\bigcup_{i=k+1}^n \Delta_i$ are true is one in which each of p_{k+1}, \dots, p_n is false, since $\Delta_i \cup \{p_i\}$ is unsatisfiable, and so any valuation in which p_1, \dots, p_k and all elements of $\bigcup_{i=k+1}^n \Delta_i$ are true is one in which the sentence letters p_{k+1}, \dots, p_n have truth-value 0. It follows that any such valuation is one in which ϕ is false, since ϕ contains no sentence letters other than $p_1, \dots, p_k, p_{k+1}, \dots, p_n$ and $v(\phi) = 0$. Hence S is unsatisfiable, contradicting the fact that this set is a finite subset of Γ_ω . Having proved by *reductio* that any sentence in Γ_ω is true under v , it follows that Γ , which is a subset of Γ_ω , is satisfiable. It is obvious from its construction that Γ_ω is a maximal finitely-satisfiable set, meaning that it is finitely-satisfiable and none of its proper extensions are finitely-satisfiable.

A more abstract version of the argument just presented runs as follows. Suppose Γ is finitely-satisfiable. Order by inclusion the set F_Γ of finitely-satisfiable sets of sentences of the language containing Γ . F_Γ is non-empty, since it contains at least Γ . Any chain in F_Γ has an upper bound, obtained by taking the union of the elements in the chain: this union contains Γ as a subset since all the members of the chain do, and it is finitely-satisfiable since any of its finite subsets must come from some element of the chain, which by hypothesis is finitely-satisfiable. *Zorn's Lemma*, provable in ZFC, states precisely that every partial order with the property that every chain has an upper bound has a maximal element. Since the conditions of Zorn's Lemma are satisfied, we deduce from it that F_Γ has a maximal element, that is to say, a maximal finitely-satisfiable set ex-

tending Γ . We then reason as in the previous paragraph to show that all the elements of Γ are true under the valuation v defined on sentence letters by $v(p) = 1$ if and only if p is a member of this maximal finitely-satisfiable set. Nowhere did we rely on the fact that the sentence letters are denumerably many, or on any assumption about the set of connectives. So this more general argument shows that PL_κ is compact for any cardinal κ whatsoever.

The abstract argument just given invoked Zorn's Lemma, well known to be equivalent to the *Axiom of Choice* in ZF. Is this use of a ZF-equivalent of Choice necessary? No. A weaker principle than Zorn's Lemma, namely the *Ultrafilter Lemma*, will do. (Section 7 has more on the relative strengths in ZF of the Ultrafilter Lemma and Zorn's Lemma. Consult Moore (1982) for more on the Axiom of Choice and its foundational significance.) To explain what the Ultrafilter Lemma is, we must first define the notion of a filter on a Boolean algebra. We denote the bottom, top, join, meet and complement in a Boolean algebra (with which we assume basic familiarity, see Givant and Halmos (2009) for an introduction) by the symbols $0, 1, \vee, \wedge, \neg$; the derived strong and weak inequality symbols have their customary meanings ($x \leq y$ means that $x \wedge y = x$ and so on). A *filter* on a Boolean algebra B is then a subset \mathcal{F} of B such that:

- (i) $1 \in \mathcal{F}; 0 \notin \mathcal{F}$;
- (ii) $\forall x, y \in \mathcal{F}, x \wedge y \in \mathcal{F}$;
- (iii) $\forall x, y \in B$, if $x \leq y$ and $x \in \mathcal{F}$ then $y \in \mathcal{F}$.

An *ultrafilter* \mathcal{F} is a maximal filter on B , or alternatively a filter on B that contains exactly one of b or $\neg b$ for each $b \in B$. The Ultrafilter Lemma, sometimes called the 'Ultrafilter Theorem' or 'Ultrafilter Principle', states that any filter on a Boolean algebra may be extended to an ultrafilter.

Armed with the notion of an ultrafilter, we may now modify the abstract proof of PL_κ 's compactness to rely not on Zorn's Lemma but on the ZF-weaker Ultrafilter Lemma instead. (Henceforth we assume that the set of propositional connectives is truth-functionally complete.) Consider the Boolean algebra B whose domain is the set of equivalence classes of sentences of PL_κ under logical equivalence, 0 and 1 being the equivalence class of a contradiction and tautology respectively, \vee, \wedge and \neg respectively denoting disjunction, conjunction and negation on the equivalence class representatives (easy check: this is well-defined); such a Boolean algebra is usually called a *Lindenbaum algebra*. To simplify notation, we denote the equivalence class of ϕ by ϕ itself; observe that the equivalence of $\psi \leq \phi$ and $\psi \models \phi$ follows from the fact that $\phi \wedge \psi$ and ψ are logically equivalent if and only if $\psi \models \phi$.

Given a finitely-satisfiable set Γ , form the set Γ^+ consisting of all the sentences entailed by some finite subset of Γ , that is,

$$\Gamma^+ = \{\phi : \exists \Gamma^{\text{fin}} \subseteq \Gamma \text{ finite s.t. } \Gamma^{\text{fin}} \models \phi\}.$$

Clearly, Γ is a subset of Γ^+ , and Γ^+ also has the following three properties:

- (i) $1 \in \Gamma^+$ but $0 \notin \Gamma^+$ (since Γ is finitely-satisfiable);
- (ii) $\forall x, y \in \Gamma^+, x \wedge y \in \Gamma^+$;
- (iii) $\forall x, y \in B$, if $x \leq y$ and $x \in \Gamma^+$ then $y \in \Gamma^+$.

In other words, Γ^+ is a filter on the Boolean algebra of sentences of PL_κ . It is the smallest filter containing Γ , that is to say, Γ is a filter-base for Γ^+ . By the Ultrafilter Lemma, Γ^+ may be extended to an ultrafilter. The alternative definition

of an ultrafilter shows that this ultrafilter is a maximal finitely-satisfiable set of sentences containing Γ . The rest of the proof proceeds as above. We have thus proved the compactness of PL_κ using the Ultrafilter Lemma rather than Zorn's Lemma.

The compactness of first-order logic may be demonstrated by a similar argument. We give two versions of the argument: the first and slightly more direct version uses the Axiom of Choice; the second uses only the weaker Ultrafilter Lemma. For the first version, let κ be the size of $\text{WFF}(\mathcal{L}) \times \text{Var}(\mathcal{L})$ where \mathcal{L} is the first-order language in which each of the sentences in our given finitely-satisfiable set Γ is expressed, $\text{WFF}(\mathcal{L})$ is the set of formulas of \mathcal{L} , and $\text{Var}(\mathcal{L})$ is \mathcal{L} 's set of variables. Add a distinct set of constants of size κ to the language (usually known as *nullary Skolem functions*), disjoint from \mathcal{L} and to be used as a source of *witnesses*. Then \mathcal{L}^+ also has size κ (since $\kappa^{<\aleph_0} = \kappa$ for infinite κ), and so we can well-order both the set of formulas \mathcal{L}^+ and the set of new constant symbols in order-type κ . We now augment the set Γ recursively by exploiting the well-ordering of \mathcal{L}^+ and the well-ordering of the set of new constants (each well-ordering being of type κ). Let $\Gamma_0 = \Gamma$ and $\Gamma_\lambda = \bigcup_{\alpha < \lambda} \Gamma_\alpha$ for λ a limit. For the successor case, let

$$\Gamma_{\alpha+1} := \Gamma_\alpha \cup \{\exists x\phi \rightarrow \phi[c \setminus x]\},$$

where $\phi[c \setminus x]$ denotes ϕ with c substituted for any free occurrences of x in ϕ , if $\exists x\phi$ is the α^{th} formula in the ordering of $\text{WFF}(\mathcal{L}^+)$ and c is the first constant in the ordering of the set of constants not to appear in any element of Γ_α nor in ϕ ; otherwise (if the α^{th} formula is not existential) let $\Gamma_{\alpha+1} := \Gamma_\alpha$. Finally, let $\Gamma_\kappa := \bigcup_{\alpha < \kappa} \Gamma_\alpha$ and define Γ^+ as the set of sentences entailed by any finite subset of Γ_κ , that is,

$$\Gamma^+ := \{\phi \in \mathcal{L}^+ : \exists \Gamma^{\text{fin}} \subseteq \Gamma_\kappa \text{ finite s.t. } \Gamma^{\text{fin}} \models \phi\}.$$

An identical argument to the one in the propositional case now shows that Γ^+ is a filter on the Boolean algebra of the first-order \mathcal{L}^+ -sentences quotiented by logical equivalence. The only difference lies in the verification that $0 \notin \Gamma^+$, in other words, that Γ^+ is finitely-satisfiable. For if

$$\{\gamma_1, \dots, \gamma_n, \exists x_{i_1} \phi_{j_1} \rightarrow \phi_{j_1}[c_{k_1} \setminus x_{i_1}], \dots, \exists x_{i_m} \phi_{j_m} \rightarrow \phi_{j_m}[c_{k_m} \setminus x_{i_m}]\}$$

were unsatisfiable, with $\gamma_1, \dots, \gamma_n \in \Gamma$ and c_{k_m} is the greatest constant subject to their well-order (m and n here are natural numbers, and the i, j and k -indices are ordinals smaller than κ), then

$$\{\gamma_1, \dots, \gamma_n\} \cup \{\exists x_{i_l} \phi_{j_l} \rightarrow \phi_{j_l}[c_{k_l} \setminus x_{i_l}] : l = 1, \dots, m-1\} \models \exists x_{i_m} \phi_{j_m} \wedge \neg \phi_{j_m}[c_{k_m} \setminus x_{i_m}].$$

But if the premisses

$$\gamma_1, \dots, \gamma_n, \exists x_{i_1} \phi_{j_1} \rightarrow \phi_{j_1}[c_{k_1} \setminus x_{i_1}], \dots, \exists x_{i_{m-1}} \phi_{j_{m-1}} \rightarrow \phi_{j_{m-1}}[c_{k_{m-1}} \setminus x_{i_{m-1}}]$$

are jointly satisfiable then they cannot entail both $\exists x_{i_m} \phi_{j_m}$ and $\neg \phi_{j_m}[c_{k_m} \setminus x_{i_m}]$. The reason is that if there is a model in which this premiss set and $\exists x_{i_m} \phi_{j_m}$ are satisfied, then there is a model in which the interpretation of c_{k_m} may be chosen to witness the truth of $\exists x_{i_m} \phi_{j_m}$, since c_{k_m} does not appear in any other sentence than $\phi_{j_m}[c_{k_m} \setminus x_{i_m}]$.

Having verified that Γ^+ is a filter, we invoke the Ultrafilter Lemma as before to extend Γ^+ to an ultrafilter Γ^{++} . Note that Γ^{++} contains Γ_κ and hence is *witness-complete*—every existential statement is satisfied by some constant.

To show that Γ^{++} (and hence Γ) is satisfiable, we construct a *term model* \mathcal{T} . The term model's domain consists of the closed terms of \mathcal{L}^+ , quotiented by the relation of appearing in an identity statement of Γ^{++} , that is to say $\tau_1 \sim \tau_2$ if and only if $\tau_1 = \tau_2 \in \Gamma^{++}$. The interpretation $a^{\mathcal{T}}$ of the constant a is a 's equivalence class, $[a]$; the interpretation $f^{\mathcal{T}}$ of the function symbol f applied to $[\tau_1], \dots, [\tau_n]$ is $[f(\tau_1, \dots, \tau_n)]$; and for any n -place relation symbol R and closed terms τ_1, \dots, τ_n , $R([\tau_1], \dots, [\tau_n])$ holds in the model if and only if $R(\tau_1, \dots, \tau_n)$ is an element of Γ^{++} . A routine argument shows that this interpretation is well-defined and that the term model \mathcal{T} is a model of Γ^{++} . It is an instructive exercise to determine where exactly the arguments just given break down in the case of second-order logic: see Paseau (2010a, 75–76) for details.

To see how to avoid the use of the Axiom of Choice, we show that turning Γ into a *Skolem set* (a set containing a witness for every existential statement) does not in fact require the *Well-Ordering Principle*. In this alternative argument for the compactness of first-order logic, we construct a set $\{c_{\langle\phi,x\rangle} : \phi \in \mathcal{L}, x \in \text{Var}(\mathcal{L})\}$ of constants disjoint from the set of constants of \mathcal{L} , one for each ordered pair in $\text{WFF}(\mathcal{L}) \times \text{Var}(\mathcal{L})$. (If we were to formalise this argument in our set-theoretic metatheory, say ZF , we may for instance code each constant $c_{\langle\phi,x\rangle}$ as $\langle\phi, x, \text{WFF}(\mathcal{L}) \times \text{Var}(\mathcal{L})\rangle$. The point is that no Choice principles are required for this recursive definition.) Then add Skolem sentences $\{\exists x\phi \rightarrow \phi[c_{\langle\phi,x\rangle}/x]\}$ for every such ordered pair $\langle\phi, x\rangle$. This gives us a new set of sentences $\Gamma_1 \supset \Gamma_0 = \Gamma$ in a language $\mathcal{L}_1 \supset \mathcal{L}_0 = \mathcal{L}$. Iterate this process ω -times, constructing a chain of sets of sentences $\Gamma_0 \subset \Gamma_1 \subset \dots \subset \Gamma_n \subset \dots$, and a corresponding chain of languages $\mathcal{L}_0 \subset \mathcal{L}_1 \subset \dots \subset \mathcal{L}_n \subset \dots$. By an inductive argument, each Γ_n is readily seen to be finitely-satisfiable (there is no need to assume the well-ordering of $\mathcal{L}_{n+1} \setminus \mathcal{L}_n$ at this point), hence so is $\bigcup_{n \in \omega} \Gamma_n$. Finally, define Γ^+ as the filter generated by Γ_ω . The rest of the argument, which defines a term model from the ultrafilter extending the filter Γ^+ proceeds as above. The version of the argument sketched in this paragraph invoked the Ultrafilter Lemma and at no point did it use the Axiom of Choice.

The set Γ^{++} is an example of a *Hintikka set*, which is a set of first-order sentences T that satisfies the following axioms. (For brevity, we use take \neg, \wedge, \forall as primitive symbols in our first-order logic.)

- For every atomic sentence ϕ , if $\phi \in T$ then $\neg\phi \notin T$,
- For every closed term t , $t = t \in T$,
- If $\phi(x)$ is an atomic formula with a single free variable x , and s, t are closed terms such that $\phi(s), s = t \in T$, then $\phi(t) \in T$,
- If ϕ is a sentence and $\neg\neg\phi \in T$, then $\phi \in T$,
- If Φ is a finite set of sentences, then
 - $\bigwedge \Phi \in T$ implies $\Phi \subseteq T$,
 - $\neg\bigwedge \Phi \in T$ implies $\neg\phi \in T$ for some $\phi \in \Phi$,
- If $\phi(x)$ is a formula with a single free variable x , then
 - $\forall x\phi(x) \in T$ implies $\phi(t) \in T$ for every closed term t ,
 - $\neg\forall x\phi(x) \in T$ implies $\neg\phi(t) \in T$ for some closed term t .

Given a Hintikka set T , we can construct a term model (using the same definition given above) that satisfies T (Hodges 1997, 40–42).

c. Semantic: Ultraproduct Proofs

In this subsection, we prove the compactness theorem for first-order logic from Łoś' Theorem. We recall the definition of an ultraproduct and state the theorem. For a proof, see Chang and Keisler (1990, p. 217, thm. 4.1.9) or Hodges (1997, p. 241–242, thm. 8.5.3).

Let $(\mathfrak{A}_i)_{i \in I}$ be a collection of first-order structures of the same signature, indexed by I , and U an ultrafilter over this indexing set I . In this context, the Boolean algebra is $\mathcal{P}(I)$ with the subset ordering \subset , so a filter $F \subset \mathcal{P}(I)$ is a set of subsets of I which does not contain the empty set and which is closed under finite intersection and upward containment. Denote the domain of each \mathfrak{A}_i by A_i . When U is an ultrafilter over I , defining $a \sim_U b$ by $\{i \in I : a(i) = b(i)\} \in U$ yields an equivalence relation over the product

$$\prod_{i \in I} A_i = \left\{ f : I \rightarrow \bigcup_{i \in I} A_i \mid \forall i \in I, f(i) \in A_i \right\}.$$

(We tend to view elements of the product as generalised sequences $(x_i)_{i \in I}$.) The characteristic function μ_U of U is a finitely-additive full measure on I and thus intuitively the elements of U are 'large subsets of I ': two functions in $\prod_{i \in I} A_i$ are identified by this relation precisely when they agree ' U -almost-everywhere'. We define a new structure $\mathfrak{B} = \prod_{i \in I} \mathfrak{A}_i / U$ as follows:

- Domain of $\mathfrak{B} = \prod_{i \in I} A_i$ quotiented by \sim_U ; $[a]$ denotes the equivalence class of a ;
- $c^{\mathfrak{B}} = [(\dots, c^{\mathfrak{A}_i}, \dots)]$;
- $f^{\mathfrak{B}}([a_1], \dots, [a_k]) = [(\dots, f^{\mathfrak{A}_i}(a_1(i), \dots, a_k(i)), \dots)]$;
- $R^{\mathfrak{B}}([a_1], \dots, [a_k])$ if and only if $\{i \in I : \mathfrak{A}_i \models R(a_1(i), \dots, a_k(i))\} \in U$.

for any k -place function and relation symbols f and R and any constant c of the first-order language in question. As may be checked, the fact that U is a filter means that the definitions just given do not depend on the choice of representatives from $\prod_{i \in I} A_i$. For ϕ atomic we have by definition:

$$\mathfrak{B} \models \phi(a_1, \dots, a_n) \text{ if and only if } \{i \in I : \mathfrak{A}_i \models \phi(a_1(i), \dots, a_n(i))\} \in U.$$

Łoś' theorem, which is proved by induction on the complexity of ϕ , extends this to all ϕ :

$$\text{for all } \phi, \mathfrak{B} \models \phi(a_1, \dots, a_n) \text{ if and only if } \{i \in I : \mathfrak{A}_i \models \phi(a_1(i), \dots, a_n(i))\} \in U.$$

The induction step for the universal / existential quantifier step does require some Choice. In fact, the Ultrafilter Lemma together with Łoś' theorem imply the full Axiom of Choice (Howard 1975).

To prove the compactness theorem for first-order logic from Łoś' Theorem, let Σ be a finitely-satisfiable set of first-order sentences. Let I be the set of finite subsets of Σ and suppose we are given for each $i \in I$ a model \mathfrak{A}_i for the sentences in i . For $i \in I$ let

$$J_i := \{j \in I : i \subset j\},$$

$$F := \{J \subset I : J_i \subset J \text{ for some } i \in I\}.$$

The collection of J_i is closed under finite intersection because $J_{i_1} \cap J_{i_2} = J_{i_1 \cup i_2}$, hence F is also closed under finite intersection; by definition, F is closed under

upward containment, and clearly F does not contain the empty set. Thus F satisfies the conditions for being a filter and may therefore be extended to an ultrafilter U . Now for any $\sigma \in \Sigma$, we have $\{\sigma\} \in I$ and $J_{\{\sigma\}} \subset \{i \in I : \mathfrak{A}_i \models \sigma\}$, from which it follows that $\{i \in I : \mathfrak{A}_i \models \sigma\} \in U$. By Łoś' theorem, $\mathfrak{B} \models \sigma$. This is true for every $\sigma \in \Sigma$, so $\mathfrak{B} \models \Sigma$, in other words if Σ is satisfiable. Intuitively, the model \mathfrak{B} was designed so as to agree with \mathfrak{A}_i on the 'large subset' J_i of I and in particular to agree with \mathfrak{A}_i on the truth-value of each element of i ; as this is true for all i , \mathfrak{B} satisfies Σ . This completes the ultraproduct proof of the compactness of first-order logic.

For *particular* applications, we may not need the full strength of the Axiom of Choice to apply then ultraproduct method. For example, suppose we wish to use the ultraproduct method to construct a non-standard model of arithmetic (see Section 2c) with language \mathcal{L} and standard model \mathfrak{A} with domain ω . We extend the language to \mathcal{L}' by introducing a new constant symbol c and for each $n \in \omega$, let \mathfrak{A}_n be the \mathcal{L}' -expansion of \mathfrak{A} given by interpreting $c^{\mathfrak{A}_n} := n$. Let F be the filter of cofinite subsets of ω and let U be an ultrafilter that extends F . Then the particular instance of Łoś' Theorem can be proven for $(\mathfrak{A}_n)_{n \in \omega}$ and U , utilising the well-foundedness of ω in place of the Axiom of Choice. Thus $\mathfrak{B} = \prod_{n \in \omega} \mathfrak{A}_n / U$ satisfies the same \mathcal{L} -sentences as \mathfrak{A} , and for all $m \in \omega$,

$$\{n \in \omega : \mathfrak{A}_n \models c \neq \underbrace{1 + \cdots + 1}_{m \text{ times}}\} = \omega \setminus \{m\} \in F \subseteq U,$$

so by Łoś' Theorem it follows that $\mathfrak{B} \models c \neq \underbrace{1 + \cdots + 1}_{m \text{ times}}$. Therefore \mathfrak{B} is a non-standard model of arithmetic.

5. Connection to Topology

Tarski (1952) article gives the compactness theorem its name (see theorems 13 and 17), observing its similarity with the finite-intersection-property definition of compactness in topologies. Here we will demonstrate the topological connection in two stages. First, we show that the compactness theorem for a propositional logic is equivalent to the claim that its associated valuation space is compact. (Recall that an *open cover* of a topological space X is a collection \mathcal{C} of open subsets of X whose union is X , and that a space X is *compact* if every open cover has a finite subcover—a finite subset that is also a cover. Equivalently, every collection of closed subsets with the *finite-intersection property*—every finite subset has non-empty intersection—has non-empty intersection. Intuitively, you cannot 'escape' a compact space; every collection of points will accumulate somewhere.) Second, we show that this space is indeed compact. We assume some basic knowledge of topology (Sutherland 2009; Willard 1970).

In this section, we spell out this reasoning for the case of propositional logic. We start with an argument that demonstrates the compactness of any PL_κ , initially assuming for the sake of simplicity that the version of PL_κ in question is equipped with a truth-functional set of connectives.

Let V_κ be the set of all valuations of PL_κ . For each sentence ϕ of PL_κ define $U(\phi) = \{v \in V_\kappa : v(\phi) = 1\}$, the set of valuations in which ϕ is true. Since $U(\phi_1) \cap U(\phi_2) = \{v \in V_\kappa : v(\phi_1) = 1\} \cap \{v \in V_\kappa : v(\phi_2) = 1\} = U(\phi_1 \wedge \phi_2)$, the sets $U(\phi)$ form a basis for a topology on V_κ ; call this basis \mathcal{B} . (A *basis* \mathcal{B} for a topology τ is a collection of open sets with the property that $U \in \tau$ if and only if $U = \bigcup \mathcal{C}$ for some $\mathcal{C} \subset \mathcal{B}$. Given a collection \mathcal{B} of subsets of X , if its

union is X and for all $B_1, B_2 \in \mathcal{B}$, $B_1 \cap B_2 = \bigcup \mathcal{C}$ for some $\mathcal{C} \subset \mathcal{B}$, then there exists a unique topology on X with basis \mathcal{B} (Willard 1970, p. 38, thm. 5.3.) The topological space we shall be interested in has domain V_κ and topology generated by the basis $\mathcal{B} = \{U(\phi) : \phi \text{ is a sentence of } \mathbf{PL}_\kappa\}$; with this topology, V_κ is the *dual space* of the Boolean algebra of sentences of \mathbf{PL}_κ . Thus properties of this space are determined only by the logical properties of valuations.

Now consider any set of sentences Σ of \mathbf{PL}_κ . Σ is unsatisfiable if and only if any valuation assigns truth-value 0 to at least one of its members, or equivalently any valuation assigns truth-value 1 to at least one sentence $\neg s$ such that $s \in \Sigma$. It follows that Σ is unsatisfiable if and only if each valuation is in some open set of the form $U(\neg s)$ where $s \in \Sigma$, or equivalently that $\{U(\neg s) : s \in \Sigma\}$ is an open cover of the topological space V_κ . The compactness of \mathbf{PL}_κ is therefore equivalent to the following claim:

if $\{U(\neg s) : s \in \Sigma\}$ is an open cover of V_κ then there is a finite subset Σ^{fin} of Σ such that $\{U(\neg s) : s \in \Sigma^{\text{fin}}\}$ is an open cover of V_κ .

Thus \mathbf{PL}_κ is compact if and only if V_κ is compact. (Note that any open set is the union of some basis sets and that any basis set $U(\phi)$ is identical to $U(\neg s)$ for $s = \neg\phi$. Also, a space is compact if every open cover drawn from a fixed basis has a finite subcover.) It suffices, then, to show that V_κ is compact in the usual topological sense.

The last claim is immediate from the fact that V_κ is homeomorphic to $\{0, 1\}^\kappa = \{f : \kappa \rightarrow \{0, 1\}\}$ with the (*Tychonoff*) *product topology*, when $\{0, 1\}$ is given the *discrete topology* $\{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$. Recall that the product topology on the product $\prod_{i \in I} X_i$ of a family of topological spaces (X_i, τ_i) indexed by I takes as its basic open sets $\prod_{i \in I} O_i$, where $O_i \in \tau_i$, and $O_i = X_i$ for all but finitely elements i in I . (For a finite collection $(X_1, \tau_1), \dots, (X_n, \tau_n)$, the product topology on $X_1 \times \dots \times X_n$ has basis $\{U_1 \times \dots \times U_n : \forall i, U_i \in \tau_i\}$. In general, the product topology is determined by the following *universal property*: for any space Y and function $f : Y \rightarrow \prod_{i \in I} X_i$, f is continuous if and only if $\pi_i \circ f$ is continuous for all $i \in I$, where $\pi_i : \prod_{j \in I} X_j \rightarrow X_i$ is the projection map onto the i -th coordinate. If we remove the restriction that all-but-finitely-many O_i 's are equal to X_i , we obtain the *box product* topology. This coincides with the product topology for finite indexes, but varies drastically for infinite products.) Since $\{0, 1\}^\kappa$ with the product topology is homeomorphic to the V_κ (if the sentence letters in \mathbf{PL}_κ are p_α for $\alpha \in \kappa$, then $f(v)(\alpha) := v(p_\alpha)$ defines a homeomorphism $f : V_\kappa \rightarrow \{0, 1\}^\kappa$) the latter is compact if and only if the former is. Now $\{0, 1\}$ with the discrete topology is compact because finite spaces are trivially compact. *Tychonoff's Theorem*, a ZF-equivalent of the Axiom of Choice, states that the product of compact spaces is compact. Putting the pieces together, it follows that $\{0, 1\}^\kappa$ with the product topology is compact. This proves the compactness of \mathbf{PL}_κ .

We do not in fact need the full power of Tychonoff's Theorem. The weaker principle that the product of compact *Hausdorff* spaces is compact will do, because $\{0, 1\}$ with the discrete topology is Hausdorff (any two of its elements—0 and 1—are contained in disjoint open sets $\{0\}$ and $\{1\}$). This weaker version of Tychonoff's Theorem is in fact equivalent to the compactness theorem for propositional logic (and first-order logic), as discussed in Section 7.

The topological proof of \mathbf{PL}_κ 's compactness just given assumed that \mathbf{PL}_κ 's set of connectives is truth-functionally complete. Without this assumption, there is no guarantee that the set \mathcal{B} is a basis for the product topology; for example, if a propositional logic cannot define negation, no element of \mathcal{B} may correspond to

the open set consisting of all elements of $\{0, 1\}^\kappa$ with first coordinate 0, and if it cannot define conjunction then \mathcal{B} will not be closed under intersection and thus cannot be a basis. The compactness of a propositional logic with κ sentence letters but without a truth-functional set of connectives is of course immediate from the fact that it is a sublogic of a truth-functionally complete propositional logic with the same κ sentence letters. The topological way of seeing this is to endow the space V_κ with the topology which has \mathcal{B} as a subbase and denote it by Y . (A *subbase* for a topological space (X, τ) is a collection of open sets $\mathcal{S} \subset \tau$ such that $\mathcal{B} := \{X\} \cup \{\bigcap \mathcal{F} : \mathcal{F} \subset \mathcal{S} \text{ finite}\}$ is a basis for (X, τ) . Given an arbitrary collection \mathcal{S} of subsets of a set X , there is a unique topology τ on X which has \mathcal{S} as a subbase. (Willard 1970, p. 39, thm. 5.6).) This topology is *coarser* than the dual space topology on V_κ : every open set in Y is open in the dual space topology. In general, if the topology τ on some set X is coarser than the topology τ^* on X and (X, τ^*) is compact, then so is (X, τ) . Intuitively, there are fewer means of ‘escaping’ (via the finite-intersection property) in τ than in τ^* , so if τ^* is compact then there are no such routes in τ . It follows that the topological space Y is compact, and thus any propositional logic is compact, whether or not its set of connectives is truth-functionally complete.

A ‘bare hands’ argument that avoids anything as strong as either the general Tychonoff Theorem or its compact Hausdorff version above may also be given for the compactness of V_ω . We present this argument, since PL_ω is the most common propositional logic.

Suppose that some open cover \mathcal{C} of the space $A = \{0, 1\}^\omega$ with the product topology lacks a finite subcover. We reduce this supposition to absurdity by constructing an element a of A with the following property: the set of all elements of A extending an arbitrary finite initial segment of a cannot be covered by any finite subset of \mathcal{C} . If f is a finite sequence of 0’s and 1’s, define A_f as the set of elements of A extending f . Let $a_0 = \langle 0 \rangle$ if $A_{\langle 0 \rangle}$ does not admit a finite subcover from \mathcal{C} , and $a_0 = \langle 1 \rangle$ otherwise. More generally, if a_m has been defined, let $a_{m+1} = a_m \hat{\ } \langle 0 \rangle$ if $A_{a_m \hat{\ } \langle 0 \rangle}$ does not admit a finite subcover from \mathcal{C} , and $a_{m+1} = a_m \hat{\ } \langle 1 \rangle$ otherwise. Since by the assumption to be reduced to absurdity $A = A_\emptyset$ does not admit a finite subcover from \mathcal{C} , an easy inductive argument shows that A_{a_m} does not admit a finite subcover from \mathcal{C} either, for any m . Letting a be the element of A whose restriction to m is a_m (more formally, $a = \bigcup_{m \in \omega} a_m$), it follows from the fact that \mathcal{C} is an open cover of A that a is an element of some basis set $B \subset O$, where O is an open set in \mathcal{C} . From $a \in B$ and the definition of the product topology, it further follows that for some m , $A_{a_m} \subset B$, so that A_{a_m} does in fact admit a finite subcover from \mathcal{C} , namely the open set $O \in \mathcal{C}$ which contains B as a subset. This contradicts our supposition and shows that $A = \{0, 1\}^\omega$ with the product topology is compact.

The argument just given, which constructs an infinite branch through a binary branching tree, did not require any version of Choice. The reason is that at each stage in the construction of a we used the ordering on $\{0, 1\}$ to extend a_m by 0 if possible and 1 if not. An alternative argument for the compactness of $\{0, 1\}^\omega$ uses the fact that a countable product $\prod_{n \in \omega} (X_n, d_n)$ of bounded (by 1) metric spaces (X_i, d_i) , is metrisable via the metric

$$d_\omega(x, y) = \sum_{n \in \omega} \frac{d_n(x_n, y_n)}{2^n}.$$

To say that $\prod_{n \in \omega} (X_n, d_n)$ is metrisable in this way is to say that the metric d_ω induces the product topology on $\prod_{n \in \omega} (X_n, d_n)$. The rest of the argument

turns on showing that the metric space $(\{0, 1\}^\omega, d_\omega)$ is compact, by exploiting the space's metric properties. This argument cannot be generalised to show that PL_κ 's valuation space is compact for any $\kappa \geq \omega_1$: the uncountable product of metric spaces, each having at least two points, is not metrisable, as no point has a countable neighbourhood base.

Moving to arbitrary logics \mathcal{L} , we construct the space $X_\mathcal{L}$ of theories $\text{Th}(\mathfrak{M})$ of \mathcal{L} -structures \mathfrak{M} , topologised by the subbasic open sets $U_\phi := \{T \in X_\mathcal{L} : \phi \notin T\}$ for each formula ϕ in \mathcal{L} . The compactness of \mathcal{L} is then easily seen to be equivalent to the compactness of $X_\mathcal{L}$ with this topology (see Figure 1). In the case of propositional logic, $X_{\text{PL}_\kappa} \cong V_\kappa$. (Note that unlike for bases, the implication that compactness of a space follows from every open cover drawn from a fixed *subbase* has a finite subcover requires the Axiom of Choice. This is the *Alexander subbase theorem* (Willard 1970, p. 129, problem 17S).) However, the argument for first-order logic is more complicated than that for propositional logic and not as elementary as exhibiting a simple homeomorphism.

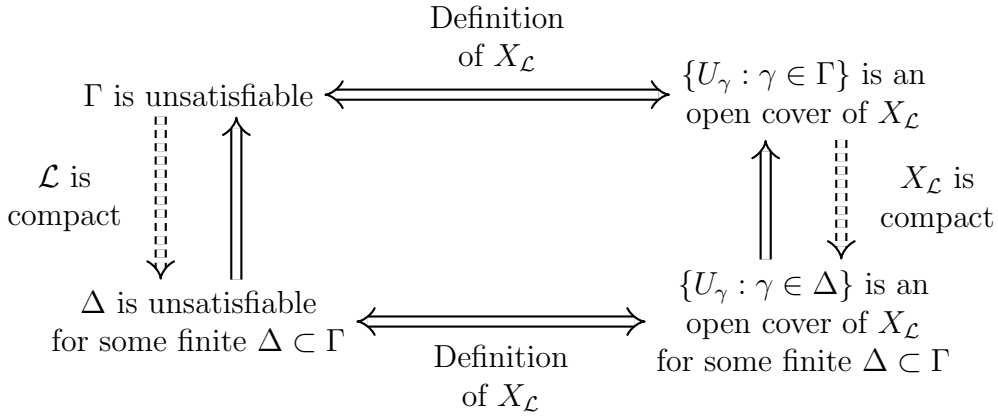


Figure 1: The connection between compactness of logics and its space of theories.

6. Extensions and Generalisations

The discussion of the compactness theorem in the most common logics, our concern here, could now ramify in several directions. There are two main dimensions of variation when considering a proof of compactness for a logic: we could vary the notion of compactness, or we could vary the logic. As an example of the latter, we might consider, for example, modal logics (K, T, S4, S5); infinitary logics $\mathcal{L}_{\kappa\lambda}$ ($\mathcal{L}_{\kappa\lambda}$ being the extension of first-order logic which allows for conjunctions and disjunctions of length less than κ and existential and universal quantifications of length less than λ . Usually $\kappa \geq \lambda$: we use ∞ to represent no bound on the connectives / quantifiers respectively.); other extensions of first-order logic, such as logics with cardinality quantifiers.

A rough rule of thumb is that infinitary logics are not compact. In particular, as demonstrated in Section 2 with the example of $\{c \neq c_i : i \in \omega\} \cup \{\bigvee_{i \in \omega} c = c_i\}$, $\mathcal{L}_{\kappa\lambda}$ is not compact whenever $\kappa > \omega$. In contrast, logics extending first-order logic with generalised quantifiers tend to be compact, but not if they contain cardinality quantifiers. Roughly speaking, the quantifiers Q_{\aleph_κ} , $Q_{>\aleph_\kappa}$, $Q_{<\aleph_\kappa}$, $Q_{\geq\aleph_\kappa}$, $Q_{\leq\aleph_\kappa}$

are interpretable as ‘there are exactly \aleph_κ ’, ‘there are more than \aleph_κ ’, ‘there are fewer than \aleph_κ ’, ‘there are at least \aleph_κ ’, and ‘there are no more than \aleph_κ ’. Consider the set

$$\{\neg Q_{\geq \aleph_\kappa} x(x = x)\} \cup \{c_\alpha \neq c_\beta : \alpha, \beta \in \kappa \text{ s.t. } \alpha \neq \beta\}$$

in an infinitary logic containing Q_{\aleph_κ} ; this set is finitely-satisfiable but unsatisfiable. (The arguments involving the other cardinality quantifiers are similar.) However, such logics tend to satisfy weaker notions of compactness. In particular, the logic that extends first-order logic with the quantifier $Q_{> \aleph_0}$, interpreted as ‘there are uncountably many’, is *countably-compact*: that is, a countable set of sentences is satisfiable if and only if it is finitely-satisfiable (Keisler 1970).

As an aside, it is worth observing that not all logics with higher-order (second-order or above) quantifiers are incompact. Second-order logic with non-standard (Henkin) semantics is compact (Enderton 2001, chap. 4), as is *existential second-order logic*, whose sentences are of the form $\exists S_1 \dots \exists S_n \phi$, where S_1, \dots, S_n are second-order functions / relations, and ϕ is a first-order sentence in a language with symbols for S_1, \dots, S_n . (Via embeddings, this logic is equivalent to many first-order logics extended with extra quantifiers and game-theoretic semantics, such as Henkin’s quantifier, Hintikka-Sandu’s independence-friendly logic, and Väänänen’s dependence logic. However, these logics do not have a classical negation and the alternative form of compactness given in the introduction of this article is false: see section 2a of this article. For further details on these logics, consult Väänänen (2007) and Mann, Sandu, and Sevenster (2011)). As another example, consider what has been called *pure second-order logic* with identity, that is, second-order logic with neither functional nor first-order variables but with both second-order and first-order identity (as well as predicate variables and quantifiers). Pure second-order logic may be thought of as the complement of first-order logic relative to second-order logic in the following sense: first-order logic has object but not predicate quantifiers, pure second-order logic has predicate but not object quantifiers, and second-order logic combines the two. In other words, second-order logic merges first-order and pure second-order logic. As Paseau (2010b) shows, pure second-order logic with identity is compact; Denyer (1992) gives an argument that applies to pure second-order logic without identity. The moral is that the incompactness of second-order logic is not owed solely to the presence of second-order quantifiers, but to the combination of both first- and second-order quantifiers. *Lindström’s Theorem* tells us that any regular logical system weakly extending first-order logic satisfying both the downward Löwenheim-Skolem property (if a sentence is satisfiable, it is satisfiable in an at most countable model) and the compactness theorem is in fact identical to first-order logic itself (Lindström 1969, p. 8, thm. 2).

The other natural dimension of generalisation is the notion of compactness. A fairly obvious generalisation is:

a logic is (κ, λ) -compact, where κ is an infinite cardinal or ∞ and λ is an infinite cardinal $\leq \kappa$, just when: whenever Γ is an unsatisfiable set of sentences of cardinality $\leq \kappa$ there is an unsatisfiable $\Gamma^\lambda \subset \Gamma$ of cardinality $< \lambda$. (By convention $\kappa < \infty$ for every cardinal κ .)

The ordinary notion of compactness is (∞, \aleph_0) -compactness: whenever Γ is any unsatisfiable set of sentences (of any cardinality) there is an unsatisfiable $\Gamma^{\aleph_0} \subset \Gamma$ of cardinality $< \aleph_0$, in other words of finite cardinality. A special case arises when $\kappa = \lambda^+$ and the logic is $\mathcal{L}_{\lambda\lambda}$. If this logic satisfies (λ, λ) -compactness, that is, if whenever Γ is an unsatisfiable set of sentences of cardinality $\leq \lambda$ then Γ has

an unsatisfiable subset of cardinality $< \lambda$, we say that it satisfies the *Weak Compactness Theorem*. One can show that for infinite cardinals, having the tree property is equivalent to satisfying the Weak Compactness Theorem, and either imply the weak inaccessibility of the cardinal. Cardinals satisfying the Weak Compactness Theorem are exactly those satisfying the combinatorial partition property $\kappa \rightarrow (\kappa)^2$, and indeed the latter is the usual definition of a *weakly-compact* cardinal. For the definitions of the tree property and the partition property, and more on weakly-compact cardinals, consult chapters 9 and 17 of Jech (2003).

Another generalisation of the notion of compactness is strong compactness. An uncountable cardinal κ is *strongly-compact* if for any set S , any κ -complete filter on S can be extended to a κ -complete ultrafilter on S . (A filter \mathcal{F} on a set is κ -complete if $\bigcap \mathcal{A} \in \mathcal{F}$ for every $\mathcal{A} \subset \mathcal{F}$ with $|\mathcal{A}| < \kappa$.) One may then show that κ is strongly-compact if and only if the logic $\mathcal{L}_{\kappa\omega}$ (or $\mathcal{L}_{\kappa\kappa}$) has the property of (∞, κ) -compactness, that is, if Γ is an unsatisfiable set of sentences then it has an unsatisfiable subset of cardinality $< \kappa$. (p. 366, lemma 20.2)

Both weakly- and strongly-compact cardinals are *large cardinal* properties. (A sentence $\phi(\kappa)$ is a large cardinal property if it satisfies the following: $\text{ZFC} \cup \{\neg\exists\kappa\phi(\kappa)\}$ is equiconsistent with ZFC , $\text{ZFC} \vdash \forall\kappa(\phi(\kappa) \rightarrow \kappa \text{ is a cardinal})$, and the consistency strength of $\text{ZFC} \cup \{\exists\kappa\phi(\kappa)\}$ is strictly larger than ZFC , by which we mean the consistency of the former implies that of the latter, but there is no finitistic argument for the reverse implication, for example, if $\text{ZFC} \cup \{\exists\kappa\phi(\kappa)\}$ implies the arithmetised consistency of ZFC , by Gödel's second incompleteness theorem.) Another large cardinal property is the *extendible cardinal* property, which is defined in terms of elementary embeddings of levels in the von Neumann cumulative hierarchy. However, it was later discovered that extendibility of a cardinal κ is equivalent to (∞, κ) -compactness of $\mathcal{L}_{\kappa\omega}^2$, or the (∞, κ) -compactness of $\mathcal{L}_{\kappa\kappa}^n$ for each positive integer n (Kanamori 2009, p. 315, thm. 23.4). $\mathcal{L}_{\kappa\lambda}^n$ is the n th-order extension of $\mathcal{L}_{\kappa\lambda}$.

Generalising the argument in Section 5 above, it is easily shown that a logic \mathcal{L} is (κ, λ) -compact precisely when every subbasic open cover $\{U_\phi : \phi \in \Gamma\}$ of $X_{\mathcal{L}}$, where Γ is a set of \mathcal{L} -sentences of cardinality at most κ , has a subcover of cardinality less than λ . This is weaker than stating that $X_{\mathcal{L}}$ is (κ, λ) -compact, meaning *every* cover of cardinality at most κ has a subcover of cardinality less than λ . For example, if \mathcal{L} is the extension of first-order logic with the quantifier $Q_{>\aleph_0}$, then \mathcal{L} is (\aleph_0, \aleph_0) -compact, but $X_{\mathcal{L}}$ is not (\aleph_0, \aleph_0) -compact. See Ebbinghaus (1985, sec. 5.1), Makowsky (1985, sec. 1.1), Mannila (1983), and Stephenson (1984) for more on (κ, λ) -compact logics and spaces.

These remarks do no more than gesture at an extensive literature generalising the notion of compactness and connecting it to topics of set-theoretic interest. We would be remiss if we did not at least mention in passing the *Barwise compactness theorem*, which is as an important theorem in generalised recursion theory (Keisler and Knight 2004). The interested reader is invited to explore the references contained within this section.

7. Relative Strength

As we saw, the compactness theorem for both propositional and first-order logic may be proven in ZFC ; but as also indicated, all that is required for its proof, in the case of both propositional and first-order logic, is the Ultrafilter Lemma, which is in fact weaker than Choice relative to ZF . Indeed, the Ultrafilter Lemma turns out to be equivalent to the compactness theorem in ZF . By the capitalised

name ‘Compactness Theorem’ we hereafter understand the compactness theorem for both propositional and first-order logic, since these are of equivalent strength relative to ZF. The relative strength of these and related principles has been well-studied; in this section we report several of these results.

a. ZF-Equivalents

Tychonoff’s Theorem states that every product of compact spaces (with the product topology) is compact. In proving compactness earlier, we only required Tychonoff’s Theorem in the case when the product consisted of Hausdorff spaces. This weakened version is in fact equivalent to the Compactness Theorem. Some other equivalents are:

- The Boolean Prime Ideal Theorem: every Boolean algebra has a prime ideal.
- The Ultrafilter Lemma: every filter on a Boolean algebra can be extended to an ultrafilter.
- Stone’s Representation Theorem: every Boolean algebra is isomorphic to a field of sets.
- Alexander subbase theorem: if X is a topological space with a subbase \mathcal{S} and every open cover $\mathcal{U} \subseteq \mathcal{S}$ has a finite subcover, then X is compact (Rubin and Scott 1954).
- For every graph G , if every finite subgraph of G is 3-colourable then G is also 3-colourable (Cowen 1990).
- If Σ is a set of propositional sentences consisting of a disjunction of at most three literals and every finite subset is satisfiable, then Σ is also satisfiable (Cowen 1990).

The equivalences of the first three statements with the Compactness Theorem can be found in Jech (1973, chap. 2). For more information, consult Howard and Rubin (1998, form 14).

b. Principles Strictly ZF-Stronger Than the Compactness Theorem

It is known that the full Tychonoff Theorem is equivalent to the Axiom of Choice in ZF (Kunen 2011, 72). To show that the Compactness Theorem does not imply Choice, we will briefly sketch the construction of a model where Choice fails but the Compactness Theorem holds. Suppose we are working in ZFA + AC with a countably infinite set of atoms A (objects which do not contain any elements and are different from the empty set); ZFA is a variant of ZF with atoms (Jech 1973, sec. 4.1). We can develop all of our standard theory in this system, with some minor modifications (for example, the Axiom of Extensionality states that two sets are equal if they have the same elements). Let \triangleleft be a dense linear ordering on A without endpoints. Every permutation π of A has an extension $\tilde{\pi}$ to the entire universe, define recursively by $\tilde{\pi}(X) := \{\tilde{\pi}(x) : x \in X\}$ for each set x . Let G be the group of order-preserving permutations of A and for every set x define $\text{fix}(x) := \{\pi \in G : \forall y \in x, \tilde{\pi}(y) = y\}$ and $\text{sym}(x) := \{\pi \in G : \tilde{\pi}(x) = x\}$. We shall say that an object x is *symmetric* if there is a finite set of atoms E such that $\text{fix}(E) \subset \text{sym}(x)$. Denote the class of hereditarily symmetric objects by M .

If V denotes the standard cumulative hierarchy in this theory, then it is easily seen that $V \subset M$. We may also show that ZFA holds in M . As the Axiom of Choice holds in V , a set $X \in M$ is well-orderable, according to M , precisely when there is a symmetric function mapping X to some set in V . This turns out to be the case precisely when there is a finite set of atoms E such that $\text{fix}(E) \subset \text{fix}(X)$. Since for every finite set of atoms E , there is a non-identity order-preserving permutation of A fixing E , it follows that A is not well-orderable according to M . However, the Compactness Theorem still holds in M .

Although we worked in a theory with atoms, it is possible to translate this model to a model without atoms where Choice fails but the Compactness Theorem holds. The interested reader should consult (Jech 1973, pp. 44–54, chap. 4) for details of this proof. It now follows that, working in ZF, Tychonoff's Theorem is not implied from the Compactness Theorem.

Here are some more examples of theorems equivalent to Choice (a principle we restate for the sake of completeness), and thus unprovable from the Compactness Theorem:

- Axiom of Choice: every family of non-empty sets has a choice function.
- Axiom of Multiple Choice: every collection of non-empty sets has a *multiple choice function*—a function that picks out a non-empty, finite subset from each set in the collection (Chapter 9).
- Well-Ordering Principle: every set can be well-ordered (Kunen 2011, 68).
- Zorn's Lemma: every partial order in which every chain has an upper bound has a maximal element (68).
- Every vector space has a basis (Blass 1984).
- Every non-empty set can be endowed with a group operation (Hajnal and Kertész 1972).

See Howard and Rubin (1998, form 1) for more statements equivalent to the Axiom of Choice.

c. Principles Strictly ZF-Weaker Than the Compactness Theorem

The following is a small selection of Choice-like principles that are weaker than the Compactness Theorem. See Jech (1973) for the construction of models satisfying one of these statements but not the Compactness Theorem.

- Order-Extension Principle, and consequently the statement that *every set can be linearly-ordered*.
- Hahn-Banach theorem: *If p is a sublinear functional on a set X and ϕ is a linear functional defined on a subspace $V \leq X$ such that $\phi(v) \leq p(v)$ for all $v \in V$, then there exists a linear extension ψ of ϕ to all of X such that $\psi(x) \leq p(x)$ for all $x \in X$.* In fact, the Hahn-Banach theorem is equivalent to the existence of a finitely-additive probability measure on every Boolean algebra (28). It is easily seen that a prime ideal gives rise to a finitely-additive, $\{0, 1\}$ -valued probability measure on every Boolean algebra, and thus the Hahn-Banach theorem follows from the Boolean Prime Ideal Theorem.
- Every infinite set has a non-principal ultrafilter.

8. History of the Compactness Theorem

The first proof of a compactness theorem was published by Gödel:

Theorem X. *For a denumerably infinite system of formulas to be satisfiable it is necessary and sufficient that every finite subsystem be satisfiable. (Gödel 1930, 118–19)*

The formulas of the logic, called *restricted (functional) calculus* (103n3) or *first-order predicate logic* (Mal'cev 1936, 1), are those from first-order logic that do not contain any function, constant, nor equality symbols, but do contain *propositional variables*; these variables can only be interpreted as 1 (true) or 0 (false) in any model. Thus, this system extends that of propositional logic PL_ω .

The central idea behind Gödel's proof is to create an equivalent sequence (ϕ_n) of satisfiable formulas, where each ϕ_{n+1} is a conjunction, with one of the conjuncts being ϕ_n , for each n . The models for each ϕ_n (in the language restricted to only those non-logical symbols that occur in ϕ_n) form a finitely-branching tree of height ω , ordered by extension, which by *König's lemma* must contain a branch. (Whilst König's lemma in general requires some Choice to prove, if the tree is countable then it can be proven in ZF. A finitely-branching tree of height ω is a countable union of finite sets, which is not necessarily countable in ZF, even if the finite sets are pairs (Pincus 1974, 224).) The interpretations along this branch are consistent and form a model of the whole sequence (ϕ_n) , which in turn models the original sequence. We have not included this approach to the compactness theorem in this entry—for details, see Paseau (2010a, 84).

This result also proves the corresponding result for first-order logic with a countable infinity of symbols. Indeed, Gödel does extend his result to allow for equality symbols (Gödel 1930, p. 117, thm. VIII). To allow for constant and function symbols, we introduce new predicate symbols R_c and R_f of arity 1 and $n + 1$ for each constant symbol c and n -ary function symbol f , and make the following substitutions:

- For a constant symbol c occurring in a sentence ϕ , replace each occurrence of c with a variable y that does not occur in ϕ to obtain the formula. If $\hat{\phi}$ denotes this new sentence, we then replace ϕ with $\exists y(\hat{\phi} \wedge R_c(y))$.
- For an n -ary function symbol f in a sentence ϕ , replace each occurrence of

$$R(\sigma_1, \dots, \sigma_{k-1}, f(\tau_1, \dots, \tau_n), \sigma_{k+1}, \dots, \sigma_m)$$

with

$$\exists y(R(\sigma_1, \dots, \sigma_{k-1}, y, \sigma_{k+1}, \dots, \sigma_m) \wedge R_f(\tau_1, \dots, \tau_n, y)),$$

where R is an m -ary predicate symbol, $\sigma_1, \dots, \sigma_{k-1}, \sigma_{k+1}, \dots, \sigma_m, \tau_1, \dots, \tau_n$ are terms and y is a variable that does not occur in ϕ . Viewing equality as a binary predicate, we can perform the same procedure on subformulas of this form.

Repeating this procedure until there are no occurrences of constant nor function symbols, we obtain a new set of sentences $\hat{\Gamma}$. Appending to this set sentences of the form $\exists!zR_c(z)$ for each constant symbol c and $\forall x_1, \dots, x_n \exists!zR_f(x_1, \dots, x_n, z)$ for each n -ary function symbol f , we find that the satisfiability of this new set is equivalent to the satisfiability of Γ , thus extending Gödel's result to encompass languages with constant and function symbols.

Note that Gödel's result is only applicable to the countable versions of propositional and first-order logic. In Gödel's (1932) short paper, he proves the compactness theorem for propositional logic for arbitrary languages. Given a deductively-closed, consistent set Γ of propositional formulas and a well-order on the set of propositional formulas, Gödel defines by transfinite recursion:

- $\Gamma_0 := \Gamma$,
- for all ordinals α , if there is a formula ϕ such that neither ϕ nor $\neg\phi$ belong to Γ_α , let ϕ_α be the least such ϕ and define

$$\Gamma_{\alpha+1} := \{\psi : (\phi_\alpha \rightarrow \psi) \in \Gamma_\alpha\}.$$

Otherwise, define $\Gamma_{\alpha+1} := \Gamma_\alpha$.

- for all limit ordinals λ , $\Gamma_\lambda := \bigcup_{\alpha < \lambda} \Gamma_\alpha$.

For an ordinal α where $\Gamma_{\alpha+1} = \Gamma_\alpha$, it follows that the valuation v defined as follows satisfies every formula in Γ : for all propositional letters p ,

$$v(p) = 1 \iff p \in \Gamma_\alpha.$$

Independently, Mal'cev (1936, thm. 1) also proved the compactness theorem for propositional logic, again using the full strength of the Axiom of Choice. His proof relies on transfinite induction, letting κ be an uncountable cardinal and assuming that the compactness theorem holds for sets of propositional sentences with cardinality strictly less than κ . Let Γ be a finitely-satisfiable set of sentences with cardinality κ and well-ordered in this order type. At each non-limit stage α , he substitutes the α th sentence ϕ_α for a conjunction of literals. Supposing this has been done up to, but not including, stage $\alpha < \kappa$, so we have a set $\Gamma_\alpha = \{\widehat{\phi}_\beta : \beta < \alpha\} \cup \{\phi_\beta : \alpha \leq \beta < \kappa\}$. The α th substitution is made according to the following rule: as the initial segments of Γ_α are satisfiable, we have a collection of κ -many models. One of the truth-value assignments to the variables that occur in ϕ_α must coincide with κ -many of the assignments for these models. Replace ϕ_α with $\widehat{\phi}_\alpha$, which is the conjunction of the letters in ϕ_α that are assigned true, together with the negation of the letters that are assigned false. By construction, the final set $\Gamma_\kappa = \{\widehat{\phi}_\alpha : \alpha < \kappa\}$ is satisfiable, and any satisfying model is also a model of the original Γ .

Mal'cev claimed to have extended this result to first-order logic in the same paper. However, his terminology is slightly ambiguous. What seems to be at issue is the claim that every sentence is 'equivalent' to a $\forall\exists$ -sentence; one of the form $\forall x_1 \dots \forall x_m \exists y_1 \dots \exists y_n \phi$, where ϕ is an atomic formula. This is patently false: for instance, the arithmetical hierarchy does not collapse, so there exists a sentence that is not equivalent to a Π_2 -sentence, even over Peano Arithmetic. However, by 'equivalent' Mal'cev appears to mean *equisatisfiable*:

as is well known, every formula of FOPL can be replaced with an equivalent formula in the Skolem normal form for satisfiability. (4)

This is also noted in a review of Mal'cev's later articles (Henkin and Mostowski 1959, 57). Under this interpretation, the statement is true. Skolem's normal form for satisfiability can be found in (Skolem 1920):

Theorem 1. *If U is an arbitrary first-order proposition, there exists a first-order proposition U' in normal form with the property that U is satisfiable in a given domain whenever U' is, and conversely. (255)*

Note that the languages of U and U' are not in general the same. This result readily extends to sets of sentences and is a well-known tool in model theory. (Nowadays, *Skolemisation* of a theory usually refers to the introduction of function symbols directly in order to prove, amongst other things, the downward Löwenheim-Skolem theorem (Skolem 1920, 257–59). As we do not require this theorem, the predicate approach has the added advantage of not requiring Choice.)

The first explicit, published proofs of a compactness theorem from completeness were given independently by Henkin (1948, 1949b, 1950) and Robinson (1949, 1951). The logics considered in these dissertations are *first-order logic / theory of simple types* for Henkin, and *restricted functional calculus* for Robinson. However, the proof relies only upon the finitary nature of the deductive system and completeness of the logic. (The completeness theorem for propositional logic is independently proven by Bernays (1918, sec. 3)—see Zach (1999, 340–48) for a discussion on authorship—and Post (1921, 169). For the restricted functional calculus, it was first asked, for an empty set of premisses, in Hilbert and Ackermann (1928, 68). It was proven in the form of the model existence theorem—*consistency implies satisfiability*, which is also the form Henkin uses to prove compactness—for countable sets in Gödel (1929, 96–101) and for arbitrary sets, using the Axiom of Choice, in Henkin (1949b, p. 164, cor. 2) and Robinson (1951, chap. 3). As noted above, this easily extends to completeness of first order logic.)

The compactness theorem finally receives its name in Tarski (1952) in two forms. The second form (p. 467, thm. 17) is the form we are familiar with. By **AC**, Tarski denotes the collection of *arithmetical classes*, which are classes of the form

$$\mathcal{CL}(\phi) := \{\mathfrak{M} : \mathfrak{M} \models \phi\}$$

where ϕ is a formula over a fixed first-order language \mathcal{L} , and the \mathfrak{M} range over \mathcal{L} -structures. In his terminology, the compactness theorem is the statement that if Γ is a collection of \mathcal{L} -formulas and $\bigcap\{\mathcal{CL}(\phi) : \phi \in \Gamma\} = \emptyset$ (which means Γ is unsatisfiable) then there exists a finite set $\Delta \subset \Gamma$ such that $\bigcap\{\mathcal{CL}(\phi) : \phi \in \Delta\} = \emptyset$ (which means Γ is not finitely-satisfiable). The first form (p. 466, thm. 13) is a generalisation that applies to *types* (Chang and Keisler 1990, 78–79; Hodges 1997, 130–31).

Whilst Tarski does not provide a topological argument for these theorems, relying upon completeness instead, he does draw some connection between a topological space based on **AC** and first-order logic. By defining a *closure operator* and then quotienting by equivalence of models, he obtains a space homeomorphic to the one constructed at the end of section 5 of this article. (There are some foundational issues at play here, since each equivalence class is a *proper class*, as are the open sets. Tarski takes only a *set* of models as opposed to the whole class. We avoided this issue in our construction by considering the space of theories. As the class of theories is a set, by taking enough representatives (using, for example, the Axiom of Collection) we find that Tarski’s space will be homeomorphic to our construction in section 5.) As he observes, the resulting space is a Stone space, noting that compactness of the space follows from the compactness of the logic. (Tarski uses the term *bicomact* for what we call *compact*. Compact used to refer to the *countable compactness* property (Willard 1970, 304).)

The ultraproduct proof is given in Frayne, Morel, and Scott (1962, p. 216, thm. 2.10), which is based on Łoś’ Theorem (p. 213, lemma 2.1)—see Łoś (1971, 105, “(2.6)”) for the original statement (without proof). The argument is similar to the one given in Section 4c. Frayne, Morel, and Scott (1962, 195) acknowledge

that Alfred Tarski suggested the use of ultraproducts, which had several applications in the literature already, in proving the compactness theorem. David Gale gave a topological proof of the compactness theorem for propositional logic (as credited by Henkin (1949a, 48n4)), whilst Beth (1951) gives a topological proof of the completeness theorem before deriving compactness as a corollary. A direct topological proof is given by Frayne, Morel, and Scott (1962, p. 225, ex. 2), using a combination of ultraproducts and Stone spaces.

More details on this history can be found in Dawson (1993) and the references therein.

9. References and Further Reading

Barwise, J., and Solomon Feferman, eds. 1985. *Model-Theoretic Logics*. Perspectives in Mathematical Logic. New York: Springer.

A book on the model-theory of many logics that extend first-order logic.

Beall, J. C., and Greg Restall. 2006. *Logical Pluralism*. Oxford: Clarendon Press.

The book that started the contemporary debate on whether there is one correct foundational logic (logical monism) or more than one (logical pluralism). The authors defend logical pluralism.

Bernays, Paul. 1918. “Beiträge zur axiomatischen Behandlung des Logik-Kalküls.” In Ewald and Sieg 2013, 231–68.

Bernays’ habilitation thesis, in which he proves the completeness theorem for propositional logic, amongst other results related to the chosen propositional calculus.

Beth, E. W. 1951. “A Topological Proof of the Theorem of Löwenheim-Skolem-Gödel.” *Indagationes Mathematicae (Proceedings)* 54:436–44. [https://doi.org/10.1016/S1385-7258\(51\)50062-8](https://doi.org/10.1016/S1385-7258(51)50062-8).

Beth presents a proof of the Löwenheim-Skolem-Gödel theorem: a set of first-order sentences Γ has a countable model if and only if Γ is consistent.

Blass, Andreas. 1984. “Existence of Bases Implies the Axiom of Choice.” In *Axiomatic Set Theory*, edited by James E. Baumgartner, Donald A. Martin, and Saharon Shelah, 31–33. Contemporary Mathematics 31. Providence, RI: American Mathematical Society.

Blass proves that if every vector space has a basis, then the Axiom of Choice holds, via the Axiom of Multiple Choice.

Chang, C. C., and H. Jerome Keisler. 1990. *Model Theory*. 3rd ed. Studies in Logic and the Foundations of Mathematics 73. Amsterdam: North-Holland.

A classic model theory text.

Cowen, Robert H. 1990. “Two Hypergraph Theorems Equivalent to BPI.” *Notre Dame Journal of Formal Logic* 31 (2): 232–40. <https://doi.org/10.1305/ndjfl/1093635418>.

Cowen proves that the Boolean Prime Ideal theorem (and thus the Compactness Theorem) is equivalent to the Compactness Theorem restricted to propositional formulas formed from a disjunct of at most 3 literals, and is also equivalent to the statement “a graph G is 3-colourable if every finite subgraph is 3-colourable”.

Dawson, John W., Jr. 1993. “The Compactness of First-Order Logic: From Gödel to Lindström.” *History and Philosophy of Logic* 14 (1): 15–37. <https://doi.org/10.1080/01445349308837208>.

Another article detailing the development and history of the compactness theorems.

Denyer, Nicholas. 1992. “Pure Second-Order Logic.” *Notre Dame Journal of Formal Logic* 33 (2): 220–24. <https://doi.org/10.1305/ndjfl/1093636099>.

Denyer proves that pure second-order logic (where formulas only have predicate variables) without identity is compact, whilst the corresponding logic with only functional variables is not.

Ebbinghaus, H.-D. 1985. “Extended Logics: The General Framework.” In Barwise and Feferman 1985, chap. 2.

Ebbinghaus presents a general framework for logics extending first-order logic and discusses various model-theoretic properties of these logics, in particular compactness and variations thereof.

Enderton, Herbert B. 2001. *A Mathematical Introduction to Logic*. 2nd ed. San Diego: Harcourt/Academic Press.

An introductory textbook on propositional, first-order, and second-order logic, as well as Gödel’s incompleteness theorems.

Ewald, William, and Wilfried Sieg, eds. 2013. *David Hilbert’s Lectures on the Foundations of Arithmetic and Logic: 1917–1933*. Vol. 3 of *David Hilbert’s Lectures on the Foundations of Mathematics and Physics: 1891–1933*. Heidelberg: Springer.

A collection of Hilbert’s works in the foundations of mathematics, with additional commentary and historical background, as well as a reproduction of Bernays’ habilitation thesis.

Feferman, Solomon, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort, eds. 1986. *Publications 1929–1936*. Vol. 1 of *Kurt Gödel Collected Works*. Oxford: Oxford University Press.

A collection of many of Gödel’s important works, including his famous results on completeness, incompleteness, and compactness.

Frayne, T., A. C. Morel, and Dana S. Scott. 1962. “Reduced Direct Products.” *Fundamenta Mathematicae* 51:195–228. <https://doi.org/10.4064/fm-51-3-195-228>.

This paper presents several results on properties reduced products and ultraproducts.

Givant, Steven, and Paul Halmos. 2009. *Introduction to Boolean Algebras*. Undergraduate Texts in Mathematics. New York: Springer.

An introductory textbook on algebraic, order-theoretic, and topological aspects of Boolean algebras.

Gödel, Kurt. 1929. “Über die Vollständigkeit des Logikkalküls.” Translated by Stefan Bauer-Mengelberg and Jean van Heijenoort. In Feferman, Dawson, Kleene, Moore, Solovay, and Van Heijenoort 1986, 60–101.

Gödel’s doctoral dissertation, in which he proves the completeness of the restricted functional calculus.

Gödel, Kurt. 1930. “Die Vollständigkeit der Axiome des logischen Funktionenkalküls.” Translated by Stefan Bauer-Mengelberg. In Feferman, Dawson, Kleene, Moore, Solovay, and Van Heijenoort 1986, 102–23.

Based on his 1929 doctoral dissertation, Gödel proves the compactness (for countable sets of sentences) of restricted functional calculus from the completeness theorem.

Gödel, Kurt. 1932. “Eine Eigenschaft der Realisierungen des Aussagenkalküls.” Translated by John W. Dawson Jr. In Feferman, Dawson, Kleene, Moore, Solovay, and Van Heijenoort 1986, 238–41.

Gödel presents his proof of the compactness theorem in general for propositional logic.

Goldblatt, Robert. 1998. *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*. Graduate Texts in Mathematics 188. New York: Springer.

A graduate-level textbook on the subject of non-standard analysis, including the foundational and model-theoretic justification for its methods.

Gonczarowski, Yannai A., Scott Duke Kominers, and Ran I. Shorrer. 2019. “To Infinity and Beyond: Scaling Economic Theories via Logical Compactness.” *Harvard Business School Entrepreneurial Management Working Paper*, no. 19–127, revised November 9, 2020. <https://doi.org/10.2139/ssrn.3409828>.

This substantial working paper demonstrates several applications of the compactness theorem in economics.

Griffiths, O., and A. C. Paseau. 2022. *One True Logic*. Oxford: Oxford University Press.

The authors argue that there is one correct foundational logic, and that it is highly infinitary.

Hajnal, A., and A. Kertész. 1972. “Some New Algebraic Equivalents of the Axiom of Choice.” *Publicationes Mathematicae Debrecen* 19:339–40.

The equivalence, under ZF, of the Axiom of Choice with the existence of groups structures (or other algebraic structures) on any set is proven in this short article.

Henkin, Leon A. 1948. “The Completeness of Formal Systems.” PhD diss., Princeton University.

Henkin’s doctoral dissertation which contains his proof of the completeness and compactness theorems for first-order functional calculus and the simple theory of types, as well as applications of the compactness theorem to algebra among other results. These results were published in Henkin (1949b, 1950).

Henkin, Leon A. 1949a. “Fragments of the Propositional Calculus.” *Journal of Symbolic Logic* 14:42–48. <https://doi.org/10.2307/2268976>.

Henkin demonstrates how to obtain complete axiomatisations for fragments of propositional logic.

Henkin, Leon A. 1949b. “The Completeness of the First-Order Functional Calculus.” *Journal of Symbolic Logic* 14:159–66. <https://doi.org/10.2307/2267044>.

This article contains Henkin’s proof of the completeness and compactness theorems of first-order functional calculus, based on his doctoral dissertation.

Henkin, Leon A. 1950. “Completeness in the Theory of Types.” *Journal of Symbolic Logic* 15:81–91. <https://doi.org/10.2307/2266967>.

This article contains Henkin proves the completeness theorem for the theory of simple types, based on his doctoral dissertation.

Henkin, Leon A., and Andrzej Mostowski. 1959. Review of Anatoliĭ Ivanovič Mal’cev. 1941. “Ob odnom obščem metodé polučeniá lokal’nyh teorém teorii grupp.” *Ivanovskij Gosudarstvennyj Pédagogičeskij Institut, Učényé zapiski, Fiziko-matématiceskíe nauki* 1 (1): 3–9 and “O predstavleniáh modéléj.” 1956 by Anatoliĭ Ivanovič Mal’cev. *Doklady Akademii Nauk SSSR* 108:27–29. *Journal of Symbolic Logic* 24 (1): 55–57. <https://doi.org/10.2307/2964581>.

Henkin and Mostowski’s review of two of Mal’cev’s important papers on applications of the compactness theorems in algebra. English translations of these article are found in Mal’cev (1971, 15–26).

Hilbert, D., and W. Ackermann. 1928. “Grundzüge der theoretischen Logik.” In Ewald and Sieg 2013, 809–915.

A collection of Hilbert’s works in the foundations of mathematics, with additional commentary and historical background, as well as a reproduction of Bernays’ habilitation thesis.

Hodges, Wilfrid. 1997. *A Shorter Model Theory*. Cambridge: Cambridge University Press.

A comprehensive textbook on model theory.

Howard, Paul E. 1975. “Łoś’ Theorem and the Boolean Prime Ideal Theorem Imply the Axiom of Choice.” *Proceedings of the American Mathematical Society* 49:426–28. <https://doi.org/10.2307/2040659>.

In this paper, Howard proves the title statement and discusses the possibility of assumption of the Boolean Prime Ideal Theorem.

Howard, Paul E., and Jean E. Rubin. 1998. *Consequences of the Axiom of Choice*. Mathematical Surveys and Monographs 59. Providence, RI: American Mathematical Society.

A comprehensive source of the myriad Choice principles that occur throughout mathematics, organised by equivalence as well as topic. It also includes descriptions of the many models of set theory that demonstrate non-provable implications between the different forms of Choice.

Jech, Thomas J. 1973. *The Axiom of Choice*. Studies in Logic and the Foundations of Mathematics 75. Amsterdam: North-Holland.

A textbook on the Axiom of Choice and other Choice-like principles, which covers the standard permutation models used for proving the independence of one statement from another.

Jech, Thomas J. 2003. *Set Theory*. 3rd ed. Springer Monographs in Mathematics. Berlin: Springer.

A graduate-level textbook on axiomatic set theory, covering infinitary combinatorics, large cardinals, inner models, and forcing.

Kanamori, Akihiro. 2009. *The Higher Infinite: Large Cardinals in Set Theory from Their Beginnings*. 2nd ed. Springer Monographs in Mathematics. Berlin: Springer.

A graduate-level textbook on the large cardinal hierarchy.

Keisler, H. Jerome. 1965. "A Survey of Ultraproducts." In *Logic, Methodology and Philosophy of Science II: Proceedings of the 1964 International Congress*, edited by Yehoshua Bar-Hillel, 112–26. Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland.

A survey on the use of ultraproducts in mathematics, including its relationship with the compactness theorem for first-order logic via Łoś' theorem.

Keisler, H. Jerome. 1970. "Logic with the Quantifier 'There Exist Uncountably Many'." *Annals of Pure and Applied Logic* 1:1–93. [https://doi.org/10.1016/S0003-4843\(70\)80005-5](https://doi.org/10.1016/S0003-4843(70)80005-5).

An extensive paper on the model-theoretic properties of logics that include the quantifier 'there exist uncountably many'.

Keisler, H. Jerome, and Julia F. Knight. 2004. "Barwise: Infinitary Logic and Admissible sets." *The Bulletin of Symbolic Logic* 10 (1): 4–36. <http://doi.org/10.2178/bsl/1080330272>.

A survey of the Barwise compactness theorem of infinitary logic.

Kunen, Kenneth. 2011. *Set Theory*. Studies in Logic 34. London: College Publications.

A textbook on axiomatic set theory, covering infinitary combinatorics and forcing.

Lindström, Per. 1969. "On Extensions of Elementary Logic." *Theoria* 35:1–11. <https://doi.org/10.1111/j.1755-2567.1969.tb00356.x>.

Lindström proves his famous theorem characterising first-order logic via the (countable) compactness and Löwenheim-Skolem theorems.

Łoś, Jerzy. 1971. "Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres." In *Mathematical Interpretation of Formal Systems*, 2nd ed., 16:98–113. Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland.

In this article (originally published in 1955), Łoś states without proof the theorem which bears his name.

Makowsky, J. A. 1985. "Compactness, Embeddings and Definability." In Barwise and Feferman 1985, chap. 18.

Makowsky presents results on several model-theoretic properties of abstract logics, including generalised compactness properties.

Mal'cev, Anatolii Ivanovič. 1936. "Investigations in the realm of mathematical logic." In Mal'cev 1971, 1–14.

In this article, Mal'cev proves the compactness theorems for propositional and first-order logic.

Mal'cev, Anatoliĭ Ivanovič. 1971. *The Metamathematics of Algebraic Systems: Collected papers: 1936–1967*. Edited and translated by Benjamin Franklin Wells III. Amsterdam: North-Holland.

Translations of most of Mal'cev's works on algebraic applications of mathematical logic.

Mann, Allen L., Gabriel Sandu, and Merlijn Sevenster. 2011. *Independence-Friendly Logic: A Game-Theoretic Approach*. London Mathematical Society Lecture Note Series 386. Cambridge: Cambridge University Press.

This monograph provides an introduction to independence-friendly logic, a logic with game-theoretic semantics.

Mannila, Heikki. 1983. "A Topological Characterization of $(\lambda, \mu)^*$ -Compactness." *Annals of Pure and Applied Logic* 25 (3): 301–5. [https://doi.org/10.1016/0168-0072\(83\)90022-2](https://doi.org/10.1016/0168-0072(83)90022-2).

Mannila gives a model-theoretic characterisation of the (λ, μ) -compactness of the model space of a logic.

Moore, Gregory H. 1982. *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*. Studies in the History of Mathematics and Physical Sciences 8. New York: Springer-Verlag.

A comprehensive book detailing the history of Zermelo's explication of the Axiom of Choice and the consequent controversy surrounding it in the early twentieth century.

Paseau, A. C. 2010a. "Proofs of the Compactness Theorem." *History and Philosophy of Logic* 31 (1): 73–98. <https://doi.org/10.1080/01445340903495340>. Corrected in "Proofs of the Compactness Theorem." *History and Philosophy of Logic* 32, no. 4 (2011): 407. <https://doi.org/https://doi.org/10.1080/01445340.2011.618261>.

This article examines different proofs of the compactness theorem, including some of the ones included in the present entry, and draws some philosophical conclusions.

Paseau, A. C. 2010b. "Pure Second-Order Logic with Second-Order Identity." *Notre Dame Journal of Formal Logic* 51 (3): 351–60. <https://doi.org/10.1215/00294527-2010-021>.

Pure second-order logic is second-order logic without first-order or functional variables. In the context of second-order logic, it is therefore, in a sense, the complement of first-order logic. The article states and proves some metalogical results for this logic.

Paseau, A. C., and Owen Griffiths. 2021. “Is English Consequence Compact?” *Thought: A Journal of Philosophy* 10 (3): 188–98. <https://doi.org/10.1002/tht3.492>.

A detailed investigation of the validity of the ‘planets’ argument discussed in section 2b of the present entry.

Pincus, David. 1974. “The Strength of the Hahn-Banach Theorem.” In *Victoria Symposium on Nonstandard Analysis: University of Victoria 1972*, edited by Albert Hurd and Peter Loeb, 203–48. Lecture Notes in Mathematics 369. Berlin: Springer-Verlag.

Pincus shows the consistency (relative to ZF) of the Hahn-Banach theorem holding whilst the Boolean Prime Ideal Theorem failing.

Poizat, Bruno. 2000. *A Course in Model Theory*. Translated by Moses Klein. Universitext. An Introduction to Contemporary Mathematical Logic. New York: Springer.

A graduate-level textbook on model theory.

Post, Emil L. 1921. “Introduction to a General Theory of Elementary Propositions.” *American Journal of Mathematics* 43 (3): 163–85. <https://doi.org/10.2307/2370324>.

Based on his doctoral dissertation, Post proves the completeness theorem for propositional calculus, as well as some results on many-valued logics.

Putnam, Hilary. 1980. “Models and Reality.” *Journal of Symbolic Logic* 45 (3): 464–82. <https://doi.org/10.2307/2273415>.

In this article, Putnam sets out the argument mentioned in section 3 of the present entry.

Richter, Marcel K. 1966. “Revealed Preference Theory.” *Econometrica* 34 (3): 635–45. <https://doi.org/10.2307/1909773>.

Richter characterises the representability and rationality of consumers, where the latter proof uses the Order-Extension Principle.

Robinson, Abraham. 1949. “On the Metamathematics of Algebraic Systems.” PhD diss., Birkbeck College, University of London.

Robinson’s doctoral dissertation, which includes his proof of the completeness and compactness theorems for first-order functional calculus. These results were later published in Robinson (1951).

Robinson, Abraham. 1951. *On the Metamathematics of Algebra*. Studies in Logic and the Foundations of Mathematics 4. Amsterdam: North-Holland.

A monograph largely based on Robinson’s doctoral dissertation.

Robinson, Abraham. 1966. *Non-Standard Analysis*. Studies in Logic and the Foundations of Mathematics 42. Amsterdam: North-Holland.

A foundational text in non-standard analysis, providing a rigorous justification of its methods.

Rubin, Herman, and Dana S. Scott. 1954. *Some Topological Theorems Equivalent to the Boolean Prime Ideal Theorem*. Presented at 503rd meeting of the American Mathematical Society, May 1, 1954. Abstract in J. W. Green. "The May Meeting in Yosemite." *Bulletin of the American Mathematical Society* 60 (1954): 386–99. <https://doi.org/10.1090/S0002-9904-1954-09827-0>.

At this talk, several topological theorems are proven to be equivalent to the Boolean Prime Ideal Theorem.

Shapiro, Stewart. 1991. *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford Logic Guides 17. Oxford: Clarendon Press.

A detailed and extensive account of second-order logic, taking in its historical, mathematical and philosophical aspects.

Shapiro, Stewart. 2014. *Varieties of Logic*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199696529.001.0001>.

The author defends a more radical form of logical pluralism (the thesis that there is more than one correct foundational logic) than that of Beall and Restall (2006).

Skolem, Thoralf. 1920. "Logico-Combinatorial Investigations in the Satisfiability or Provability of Mathematical Propositions: A Simplified Proof of a Theorem by L. Löwenheim and Generalizations of the Theorem." Translated by Stefan Bauer-Mengelberg. In *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, edited by Jean van Heijenoort, 254–63. Cambridge, MA: Harvard University Press.

A translated extract from Skolem's 1920 paper, in which he presents his normal form for formulas in first-order logic as well as presenting a fully-correct proof of the downward Löwenheim-Skolem theorem for countable sets of sentences, which uses his Axiom of Choice.

Skolem, Thoralf. 1922. "Some Remarks on Axiomatized Set Theory." Translated by Stefan Bauer-Mengelberg, 290–301.

A translated transcript of Skolem's address at the Fifth Congress of Scandinavian Address held in Helsinki 4th–7th July 1920, in which he discusses several foundational aspects of set theory, and in particular demonstrates his 'paradox'.

Stephenson, R. M., Jr. 1984. “Initially κ -Compact and Related Spaces.” In *Handbook of Set-Theoretic Topology*, edited by Kenneth Kunen and Jerry E. Vaughan, 603–32. Amsterdam: North-Holland.

This article presents a survey of results on (κ, λ) -compactness of topological spaces.

Sutherland, Wilson A. 2009. *Introduction to Metric and Topological Spaces*. 2nd ed. Oxford: Oxford University Press.

An introductory textbook to topology. Chapter 13 on compactness is particularly relevant to this article.

Szpilrajn, Edward. 1930. “Sur l’extension de l’ordre partiel.” *Fundamenta Mathematicae* 16:386–89. <https://doi.org/10.4064/fm-16-1-386-389>.

In this short paper, Szpilrajn (also known as Marczewski) proves his Order Extension Principle.

Tarski, Alfred. 1952. “Some Notions and Methods on the Borderline of Algebra and Metamathematics.” In *Proceedings of the International Congress of Mathematicians: Cambridge, Massachusetts, U.S.A., August 30–September 6, 1950*, 1:705–20. Providence, RI: American Mathematical Society.

Tarski presents several results in the language of arithmetical classes (the classes of structures that satisfy a particular first-order sentence). Theorem 13 gives the compactness theorem its name.

Väänänen, Jouko. 2007. *Dependence Logic: A New Approach to Independence Friendly Logic*. 70. Cambridge: Cambridge University Press.

A monograph on dependence logic, and other logics with game-theoretic semantics.

Willard, Stephen. 1970. *General Topology*. Reading, MA: Addison-Wesley.

An introductory textbook on topology. Chapter 6 on compactness is particularly relevant to this article.

Zach, Richard. 1999. “Completeness Before Post: Bernays, Hilbert, and the Development of Propositional Logic.” *Bulletin of Symbolic Logic* 5 (3): 331–66. <https://doi.org/10.2307/421184>.

A historical article on Bernays and Hilbert’s role in development of mathematical logic, particularly focused on propositional logic.

Authors

A. C. Paseau

Email: alexander.paseau@philosophy.ox.ac.uk

University of Oxford

United Kingdom

Robert Leek

Email: r.leek@bham.ac.uk

University of Birmingham

United Kingdom