Tools and Techniques in Modal Logic

Marcus Kracht
II. Mathematisches Institut
Freie Universität Berlin
Arnimallee 3
D – 14195 Berlin
kracht@math.fu-berlin.de

About this Book

This book is intended as a course in modal logic for students who have had prior contact with modal logic and wish to study it more deeply. It presupposes training in mathematics or logic. Very little specific knowledge is presupposed, most results which are needed are proved in this book. Knowledge of basic logic—propositional logic, predicate logic—as well as basic mathematics will of course be very helpful. The book treats modal logic as a theory, with several subtheories, such as *completeness theory*, *correspondence theory*, *duality theory* and *transfer theory*. Thus, the emphasis is on the inner structure of the theory and the connections between the subdisciplines and not on coverage of results. Moreover, we do not proceed by discussing one logic after the other; rather, we shall be interested in general properties of logics and calculi and how they interact. One will therefore not find sections devoted to special logics, such as **G**, **K4** or **S4**. We have compensated for this by a special index of logics, by which it should be possible to collect all major results on a specific system. Heavy use is made of algebraic techniques; moreover, rather than starting with the intuitively simpler Kripke–frames we begin with algebraic models. The reason is that in this way the ideas can be developed in a more direct and coherent way. Furthermore, this book is about modal logics with *any* number of modal operators. Although this may occasionally lead to cumbersome notation, it was felt necessary not to specialize on monomodal logics. For in many applications one operator is not enough, and so modal logic can only be really useful for other sciences if it provides substantial results about polymodal logics.

No book can treat a subject area exhaustively, and therefore a certain selection had to be made. The reader will probably miss a discussion of certain subjects such as modal predicate logic, provability logic, proof theory of modal logic, admissibility of rules, polyadic operators, intuitionistic logic, and arrow logic, to name the most important ones. The choice of material included is guided by two principles: first, I prefer to write about what I understand best; and second, about some subjects there already exist good books (see [**182**], [**43**], [**31**], [**157**], [**224**]), and there is no need to add another one (which might even not be as good as the existing ones).

I got acquainted with modal logic via Montague Semantics, but it was the book [**169**] by Wolfgang Rautenberg that really hooked me onto this subject. It is a pity that this book did not get much attention. Until very recently it was the only book which treated modal logic from a mathematical point of view. (Meanwhile, however, the book [**43**] has appeared in print, which is heartily recommended.) However, twenty years have passed from its publication and many strong and important results have been found, and this was the reason for writing this book.

My intellectual credits go not only to Wolfgang Rautenberg but also to Siegfried Breitsprecher — whose early death saddened me greatly — for teaching me algebra, Helmut Salzmann for his inspiring introduction to geometry and linear algebra, and to Walter Felscher for his introduction to logic and exact mathematics. Furthermore, I wish to thank Kit Fine for making an exception and taking me as his student in Edinburgh. He too taught me logic in his rather distinct way. More than anyone in the last years, Frank Wolter has been an inspiration and collaborator. Without him, this book would not have been written. Thanks to Carsten Grefe for his help both with some of the pictures as well as modal logic, and thanks also to Andreas Büll and Martin Mittelmaier. Thanks to Monika Budde, Sam Dorner, Kit Fine, Clemens Hendler, Carsten Ihlemann, Tomasz Kowalski and Timothy Surendonk for careful proofreading and Rajeev Goré and Misha Zakharyaschev for their advice in many matters. The final draft was carefully read by Hans Mielke and Birgit Nitzsche. Special thanks go to Armin Ecker for his never ending moral support.

No endeavour can succeed if it is not blessed by love and understanding. I am fortunate to have experienced both through my wife Johanna Domokos, my parents, my brother and my sister. This book is dedicated to all those to whom it gives pleasure. May it bring — in its own modest way — a deeper understanding of the human spirit.

Berlin, March 1999

Marcus Kracht

Added. A number of errors in the printed version have been brought to my attention by Guram Bezhanishvilij, Lloyd Humberstone, and Tomasz Kowalski.

Overview

The book is structured as follows. There are ten chapters, which are grouped into three parts. The first part contains the Chapters 1 – 3, the second part the Chapters 4 – 7 and the third part the Chapters 8 – 10. The first part contains roughly the equivalent of a four hour one semester course in modal logic. Chapter 1 presents the basics of algebra and general propositional logic inasmuch as they are essential for understanding modal logic. This chapter introduces the theory of consequence relations and matrix semantics. From it we deduce the basic completeness results in modal logic. The generality of the approach is justified by two facts. The first is that in modal logic there are several consequence relations that are associated with a given logic, so that acquaintance with the general theory of consequence relations is essential. Second, many results can be understood more readily in the abstract setting. After the first chapter follow the Chapters 2 and 3, in which we outline the basic terminology and techniques of modal logic, such as completeness, Kripke–frames, general frames, correspondence, canonical models, filtration, decidability, tableaux, normal forms and modal consequence relations. One of the main novelties is the method of *constructive reduction*. It serves a dual purpose. First of all, it is a totally constructive method, whence the name. It allows to give proofs of the finite model property for a large variety of logics without using infinite models. It is a little bit more complicated than the filtration method, but in order to understand proofs by constructive reduction one does not have to understand canonical models, which are rather abstract structures. Another advantage is that interpolation for the standard systems can be deduced immediately. New is also the systematic use of the distinction between local and global consequence relations and the introduction of the compound modalities, which allows for rather concise statements of the facts. The latter has largely been necessitated by the fact that we allow the use of any number of modal operators. Also, the fixed point theorem for **G** of DICK DE JONGH and GIOVANNI SAMBIN is proved. Here, we deduce it from the so–called Beth–property, which in turn follows from interpolation. This proof is originally due to CRAIG SMORYŃSKI [**200**].

The second part consists of chapters on duality theory, correspondence theory, transfer theory and lattice theory, which are an absolute necessity for understanding higher modal logic. In Chapter 4 we develop duality theory rather extensively, starting with universal algebra and Stone–representation. Birkhoff's theorems are proved in full generality. This will establish two important facts. One is that the lattice of normal modal logics is dually isomorphic to the lattice of subvarieties of the variety of modal algebras. Secondly, the characterization of modally definable classes of generalized frames in terms of closure properties is readily derived. After that we give an overview of the topological and categorial methods of GIOVANNI SAMBIN and VIRGINIA VACCARO, developed in [**186**] and [**187**]. Furthermore, we study the connection between properties of the underlying Kripke–frame and properties of the underlying algebra in a descriptive frame. We will show, for example, that subdirect irreducibility of an algebra and rootedness of dual descriptive frame are independent properties. (This has first been shown in [**185**].) We conclude this chapter with a discussion of the structure of canonical frames and some algebraic characterizations of interpolation, summarizing the work of LARISA MAKSIMOVA. An algebraic characterization of Halldén–completeness using coproducts is derived, which is slightly stronger than that of [**153**]. Chapter 5 develops the theme of first–order correspondence using the theory of internal descriptions, which was introduced in MARCUS KRACHT [**121**]. We will prove not only the theorem by HENDRIK SAHLQVIST [**183**] but also give a characterization of the elementary formulae which are definable by means of Sahlqvist formulae. This is done using a two–sided calculus by means of which correspondence statements can be systematically derived. Although this calculus is at the beginning somewhat cumbersome, it allows to compute elementary equivalents of Sahlqvist formulae with ease. Moreover, we will show many new corollaries; in particular, we show that there is a smaller class of modal formulae axiomatizing the Sahlqvist formulae. On the other hand, we also show that the class of formulae described by VAN BENTHEM in [**10**] which is larger than the class described by Sahlqvist does not axiomatize a larger class of logics. Next we turn to the classic result by KIT FINE [**65**] that a logic which is complete and elementary is canonical, but also the result that a modally definable first–order condition is equivalent to a positive restricted formula. This has been the result of a chain of theorems developed by SOLOMON FEFERMAN, ROBERT GOLDBLATT and mainly JOHAN VAN BENTHEM, see [**10**]. In Chapter 6 we discuss transfer theory, a relatively new topic, which has brought a lot of insights into modal logic. Its aim is to study how complex logics with several operators can be reduced to logics with less operators. The first method is that of a *fusion*. Given two modal logics, their fusion is the least logic in the common language which contains both logics as fragments. This construction has been studied by FRANK WOLTER in [**233**], by KIT FINE and GERHARD SCHURZ [**67**], and by FRANK WOLTER and MARCUS KRACHT in [**132**]. For many properties $\mathcal{P}$ it is shown that a fusion has $\mathcal{P}$ iff both fragments have $\mathcal{P}$. In the last section a rather different theorem is proved. It states that there is an isomorphism from the lattice of bimodal logics onto an interval of the lattice of monomodal logics such that many properties are left invariant. This isomorphism is based on the simulations defined by S. K. THOMASON in [**208, 210**]. Some use of simulations has been made in [**127**], but this theorem is new in this strong form. Only the simulations of THOMASON have these strong properties.

Extensive use of these results is made in subsequent chapters. Many problems in modal logic can be solved by constructing polymodal examples and then appealing to this simulation theorem. Chapter 7 discusses the global structure of the lattices of modal logics. This investigation has been initiated by Wim Blok and Wolfgang Rautenberg, whose splitting theorem [**170**] has been a great impulse in the research. We state it here in the general form of Frank Wolter [**234**], who built on [**120**]. The latter generalized the splitting theorem of [**170**] to non–weakly transitive logics and finitely presentable algebras. [**234**] has shown this use to be inessential; we show in Section 7.5 that there exist splitting algebras which are not finitely presentable. In the remaining part of this chapter we apply the duality theory of upper continuous lattices, which are also called frames or locales (see [**110**]) to modal logic. One result is a characterization of those lattices of logics which admit an axiomatization base. This question has been put and answered for **K4** by Alexander Chagrov and Michael Zakharyaschev [**42**]. The argument used here is rather simple and straightforward. We prove a number of beautiful theorems by Wim Blok about the degree of incompleteness of logics. The way these results are proved deserves attention. We do not make use of ultraproducts, only of the splitting theorem. This is rather advantageous, since the structure of ultraproducts of Kripke–frames is generally difficult to come to terms with. Finally, the basic structure of the lattice of tense logics is outlined. This is taken from [**123**].

The last part is a selection of issues from modal logic. Some topics are developed in great depth. Chapter 8 explores the lattice of transitive logics. It begins with the results of Kit Fine concerning the structure of finitely generated transitive frames and the selection procedure of Michael Zakharyaschev, leading to the *cofinal subframe logics* and the *canonical formulae*. The characterization of elementary subframe logics by Kit Fine is developed. After that we turn to the study of logics of finite width. These logics are complete with respect to *noetherian frames* so that the structure theory of Kit Fine [**66**] can be extended to the whole frame. This is the starting point for a rich theory of transitive logics of finite width. We will present some novel results such as the decidability of all finitely axiomatizable transitive logics of finite width and *finite tightness* and the result that there exist 13 logics of finite width which bound finite model property in the lattice of extensions of **S4**. The first result is a substantial generalization of [**247**], in which the same is shown for extensions of **K4.3**. In Chapter 9 we prove a series of undecidability results about modal logics using two main ingredients. The first is the simulation theorem of Chapter 6. And the other is the use of the logics $\mathbf{K.alt}_n$. The latter have been studied by Krister Segerberg [**197**] and Fabio Bellissima [**5**] and their polymodal fusions by Carsten Grefe [**91**]. The latter has shown among other that while the lattice of $\mathbf{K.alt}_1$ is countable, the lattice of the fusion of this logic with itself has $2^{\aleph_0}$ many coatoms. Moreover, the polymodal fusions of $\mathbf{K.alt}_1$ can be used to code word problems as decidability problems of logics. Using this method, a great variety of theorems on the undecidability of properties is obtained. This method is different from the one used by Lilia Chagrova [**44**], and Alexander Chagrov and Michael Zakharyaschev [**41**]. Their proofs establish undecidability for extensions of **K4**, but our proofs are essentially simpler. The proofs that global finite model property (global decidability) are undecidable even when the logic is known to have local finite model property (is locally decidable), are new.

We conclude the third part with Chapter 10 on propositional dynamic logic (**PDL**). This will be a good illustration of why it is useful to have a theory of arbitrarily many modal operators. Namely, we shall develop dynamic logic as a special kind of polymodal logic, one that has an additional component to specify modal operators. This viewpoint allows us to throw in the whole machinery of polymodal logic and deduce many interesting new and old results. In particular, we will show the finite model property of **PDL**, in the version of Rohit Parikh and Dexter Kozen [**118**], of **PDL** with converse, by Dimiter Vakarelov [**217**], and of deterministic **PDL** by Mordechai Ben–Ari, Joseph I. Halpern and Amir Pnueli, [**7**]. Again, constructive reduction is used, and this gives an additional benefit with respect to interpolation. We have not been able to determine whether **PDL** has interpolation, but some preliminary results have been obtained. Moreover, for the logic of finite computations we show that it fails to have interpolation and that it does not have a fixed point theorem. Largely, we feel that an answer to the question whether **PDL** has interpolation can be obtained by closely analysing the combinatorics of regular languages.

Contents

Part 1

The Fundamentals

CHAPTER 1

Algebra, Logic and Deduction

1.1. Basic Facts and Structures

In this section we will briefly explain the notation as well as the basic facts and structures which will more or less be presupposed throughout this book. We will assume that the reader is familiar with them or is at least willing to grant their truth.

Sets, Functions. We write $\{x : \varphi(x)\}$ for the set of all objects satisfying φ. Given a set S, $\wp(S)$ denotes the powerset of S, $\sharp S$ the *cardinality* of S. For functions we write $f : A \to B$ to say that f is a function from A to B, and $f : x \mapsto y$ to say that f maps (in particular) x to y. The image of x under f is denoted by $f(x)$. We write $f : A \rightarrowtail B$ if f is injective, that is, if $f(x) = f(y)$ implies $x = y$ for all $x, y \in A$; and we write $f : A \twoheadrightarrow B$ if f is surjective, that is, if for every $y \in B$ there is an $x \in A$ such that $y = f(x)$. For a set $S \subseteq A$, $f[S] := \{f(x) : x \in S\}$. We put $f^{-1}(y) := \{x : f(x) = y\}$. For a set $T \subseteq B$, $f^{-1}[T] := \{x : f(x) \in T\}$. If $f : A \to B$ and $g : B \to C$ then $g \circ f : A \to C$ is defined by $(g \circ f)(x) := g(f(x))$. The *image* of $f : A \to B$, denoted by $im[f]$, is defined by $im[f] := f[A]$. M^N denotes the set of functions from N to M. If $C \subseteq A$ then $f \restriction C$ denotes the restriction of f to the set C.

Cardinal and Ordinal Numbers. Finite ordinal numbers are constructed as follows. We start with the empty set, which is denoted by 0. The number n is the set $\{0, 1, \dots, n-1\}$. '$i < n$' is synonymous with '$i \in n$'. In general, an ordinal number is the set of ordinal numbers smaller than that number. So, in constructing ordinals, the next one is always the set of the previously constructed ordinals. There are two types of ordinal numbers distinct from 0, *successor ordinals* and *limit ordinals*. An ordinal λ is a *successor ordinal* if it is of the form $\kappa \cup \{\kappa\}$, and a *limit ordinal* if it is not 0 and not a successor ordinal. Finite numbers are successor ordinals, with the exception of 0. Ordinal numbers are *well–ordered* by the inclusion relation $\in$. A *well–order* $<$ on a set R is a linear ordering which is irreflexive and such that any nonempty subset $S \subseteq R$ has a least element with respect to $<$. Ordinal numbers can be characterized as sets well–ordered with respect to $\in$ such that every element of an ordinal κ is an ordinal. Any well–ordered set is isomorphic to a pair $\langle \kappa, \in \rangle$, κ an ordinal. '$\kappa < \lambda$' is synonymous with '$\kappa \in \lambda$'. The Axiom of Choice is equivalent to the statement that every set can be well–ordered. Throughout this book we will be working with the standard set–theory ZFC (Zermelo–Fraenkel set theory plus the Axiom of Choice). (See [**220**] and [**114, 115**] for an introduction to set theory.)

One can define the sum, product and exponentiation for ordinals. The sum and product of ordinals is generally not commutative. $1 + \omega = \omega$, but $\omega + 1 \neq \omega$. And $2 \cdot \omega = \omega$, but $\omega \cdot 2 = \omega + \omega \neq \omega$. Cardinal numbers are special kinds of ordinals, namely those ordinals λ for which no $\mu < \lambda$ can be mapped *onto* λ. Finite ordinals are cardinals. ω also is a cardinal here always denoted by $\aleph_0$. Cardinal arithmetic is different from ordinal arithmetic, except for finite numbers. If α or β is infinite, however, we have $\alpha + \beta = \alpha \cdot \beta = max\{\alpha, \beta\}$. A set M has *cardinality* α if there is a bijection $f : \alpha \to M$. The set 2^M of all functions from M to 2 has the same cardinality as the powerset $\wp(M)$. Therefore we will occasionally identify $\wp(M)$ with 2^M.

If α is a cardinal, α^+ denotes the least cardinal greater than α. This is always defined. For $\alpha = n$ we have $\alpha^+ = n + 1$. For $\alpha = \aleph_\kappa$, κ an ordinal, $\aleph_\kappa^+ := \aleph_{\kappa+1}$. We know that always $\alpha < 2^\alpha$. The *Generalized Continuum Hypothesis* (*GCH*) is the conjecture that $\alpha^+ = 2^\alpha$. For $\alpha = \aleph_0$, the cardinality of the set of natural numbers, this is the *Continuum Hypothesis* (*CH*). CH (and also GCH) is actually independent of ZFC. We will state our results so that they are independent of CH and GCH.

Lattices and Orderings. A *partial order* on a set S is a relation $\leq$ which is (1.) *reflexive*, that is, $x \leq x$ for all $x \in S$, (2.) *transitive*, that is, $x \leq y$ and $y \leq z$ implies $x \leq z$ for all $x, y, z \in S$, and (3.) *antisymmetric*, which means that for all $x, y \in S$ if $x \leq y$ and $y \leq x$ then $x = y$. A *chain* is a partial order in which for any two x, y we have $x \leq y$ or $y \leq x$. (If either $x \leq y$ or $y \leq x$ we say that x and y are *comparable*.) Let $X \subseteq S$. Then

$$\begin{aligned} \downarrow X &:= \{y : (\exists x \in X)(y \leq x)\} \\ \uparrow X &:= \{y : (\exists x \in X)(y \geq x)\} \end{aligned}$$

If $X = \{x\}$ then we write $\downarrow x$ and $\uparrow x$ rather than $\downarrow \{x\}$ and $\uparrow \{x\}$. A set of the form $\downarrow X$ ($\uparrow X$) for some X is called a *lower cone* (*upper cone*).

A *lattice* is a triple $\mathfrak{L} = \langle L, \sqcap, \sqcup \rangle$ satisfying the following laws for all $x, y, z \in L$

$$\begin{array}{lcllcl} x \sqcap (y \sqcap z) & = & (x \sqcap y) \sqcap z & x \sqcup (y \sqcup z) & = & (x \sqcup y) \sqcup z \\ x \sqcap y & = & y \sqcap x & x \sqcup y & = & y \sqcup x \\ x \sqcap x & = & x & x \sqcup x & = & x \\ x \sqcap (y \sqcup x) & = & x & x \sqcup (y \sqcap x) & = & x \end{array}$$

These laws are referred to as the laws of *associativity*, *commutativity*, *idempotence* and *absorption*. We call $x \sqcap y$ the *meet of* x *and* y and $x \sqcup y$ the *join of* x *and* y. In a lattice we can define a partial ordering $\leq$ by setting $x \leq y$ iff $x \sqcup y = y$. It turns out that $x \leq y$ iff $x \sqcap y = x$. From the laws above follows that $\leq$ is a partial order, and $x \sqcap y$ is the *greatest lower bound* (glb) of x and y, and $x \sqcup y$ the *least upper bound* (lub) of x and y. Conversely, if $\leq$ is a partial ordering on V such that the glb and the lub for any pair of elements exists, then $\langle V, glb, lub \rangle$ is a lattice.

There is a principle of *duality* in lattices which states that any law valid in all lattices is transformed into a valid law if $\sqcap$ and $\sqcup$ are exchanged. This is due to the fact that the laws postulated for lattices come in pairs, one the dual of the other. Let $\mathfrak{L} = \langle L, \sqcap, \sqcup \rangle$. We say that $\mathfrak{L}^{op} = \langle L, \sqcup, \sqcap \rangle$ is the *dual* lattice of $\mathfrak{L}$. The partial order obtained for $\mathfrak{L}^{op}$ is simply the converse ordering. (Usually, since $x \leq y$ means that in the graphical representation x is below y, $\mathfrak{L}^{op}$ is obtained from $\mathfrak{L}$ by putting everything upside down.)

We say that $\bot$ is a *bottom element* if $\bot \leq x$ for all x, and that $\top$ is a *top element* if $x \leq \top$ for all x. A structure $\langle L, \bot, \top, \sqcap, \sqcup \rangle$ whose reduct to $\sqcap$ and $\sqcup$ is a lattice with top element $\top$ and bottom element $\bot$ is called a *bounded lattice*. An element x is an *atom* if for no y, $\bot < y < x$, and a *coatom* if for no y, $x < y < \top$. Two lattices $\mathfrak{L} = \langle L, \sqcap, \sqcup \rangle$ and $\mathfrak{M} = \langle M, \sqcap, \sqcup \rangle$ are isomorphic iff the ordered sets $\langle L, \leq \rangle$ and $\langle M, \leq \rangle$ are isomorphic. If a lattice also has infinite meets and joins, that is, if the glb as well as the lub of infinite sets exists, then $\mathfrak{L}$ is called *complete*. These infinitary operations are denoted by $\bigsqcap$ and $\bigsqcup$. We write $\bigsqcap_{y \in Y} y$ or simply $\bigsqcap Y$ for the meet of Y. Similarly for the join. Notice that complete lattices always have a bottom and a top element. It can be shown that a lattice has infinitary glb's iff it has infinitary lub's. An element x of a lattice is *join compact* if from $x \leq \bigsqcup Y$ we may conclude that $x \leq \bigsqcup Y_0$ for a finite $Y_0 \subseteq Y$. A lattice is *algebraic* if every element is the least upper bound of a set of join compact elements. A lattice is *distributive* if it satisfies the identities

$$\begin{aligned} x \sqcap (y \sqcup z) &= (x \sqcap y) \sqcup (x \sqcap z) \\ x \sqcup (y \sqcap z) &= (x \sqcup y) \sqcap (x \sqcup z) \end{aligned}$$

It can be shown that either of the two equations implies the other.

A *filter* is a nonempty set $F \subseteq L$ such that $F = \uparrow F$ and such that if $x, y \in F$ then also $x \sqcap y \in F$. A filter is *principal* if it is of the form $\uparrow x$ for some $x \in L$. A subset $I \subseteq L$ is an *ideal* if $I = \downarrow I$ and for $x, y \in I$ also $x \sqcup y \in I$. An ideal I is *principal* if $I = \downarrow x$ for some $x \in L$.

A *boolean algebra* is a structure $\mathfrak{B} = \langle B, 0, 1, -, \cap, \cup \rangle$ such that the restriction $\mathfrak{B} \upharpoonright \{\cap, \cup, 0, 1\} := \langle B, 0, 1, \cap, \cup \rangle$ is a bounded distributive lattice, and $- : B \to B$ is a function satisfying for all x:

$$\begin{aligned} - - x &= x \\ x \cap -x &= 0 \\ x \cup -x &= 1 \end{aligned}$$

as well as the so–called *de Morgan Laws*

$$-(x \cup y) = (-x) \cap (-y)\ , \qquad -(x \cap y) = (-x) \cup (-y)\ .$$

We also use the notation $\bigcap_{x \in X} x$ or $\bigcap X$, and likewise for $\bigcup$, as for lattices. This is in general only defined if X is finite. A general reference for the kind of concepts introduced so far is [**37**], [**52**] and [**90**].

Closure Operators. Let S be a set. A map $C : \wp(S) \to \wp(S)$ is called a *closure operator* if it satisfies the following properties, referred to as *extensivity*, *monotonicity* and *idempotence*.

(ext)	$X \subseteq C(X)$
(mon)	$X \subseteq Y$ implies $C(X) \subseteq C(Y)$
(ide)	$C(C(X)) = C(X)$

A set of the form $C(X)$ is called a *closed set*. Obviously, $C(X)$ is the smallest closed set containing X. Given a closure operator C, any intersection of closed sets is closed again. Thus, the closure $C(X)$ can also be defined via

$$C(X) = \bigcap \langle Y : Y \supseteq X, Y = C(Y)\rangle$$

The closed sets form a lattice, with $\sqcap$ being standard set intersection, and $X \sqcup Y = C(X \cup Y)$. A closure operator is called *finitary* if it satisfies

(fin) $\qquad C(X) = \bigcup \langle C(E) : E \subseteq X, E \text{ finite} \rangle$

If a closure operator is finitary, the lattice of closed sets is algebraic, and conversely. The join compact elements coincide with the sets $C(E)$, E finite. For general reference see [**52**].

2–Valued Logic. We will assume familiarity with classical logic, and in some sections predicate logic and some model theory is required as well. Nevertheless, for reference, let us fix here what we mean by *classical logic*. First of all, to avoid a terminological clash, we talk of *2–valued logic* when we refer to classical logic in the usual sense. In fact, 2–valued only refers to the fact that we allow exactly two truth values, denoted by 0 and 1. Furthermore, in the languages of 2–valued logic we have primitive sentence letters functioning as propositional variables, and various logical symbols with fixed meaning throughout this book. These are the constants *verum* $\top$, and *falsum* $\bot$, the *negation* $\neg$, the *conjunction* $\wedge$, *disjunction* $\vee$, *implication* $\rightarrow$ and *biimplication* $\leftrightarrow$. We may identify the set of truth values with the set 2 (= $\{0, 1\}$, see above). When we speak of *boolean logic* we mean 2–valued logic for the language in which only $\top$, $\neg$ and $\wedge$ are basic, and all other symbols are defined in the usual way. A *valuation* is a function from the variables into the set 2. The truth value of a complex proposition is calculated using the truth tables of the symbols, which are standard. We say φ comes out *true* under a valuation, if it receives the value 1. A formula φ is a *tautology* if it is true under all valuations. Formally, boolean logic is the logic of the boolean algebra **2**, which is the algebra based on the set 2 with the usual interpretation of the symbols. Again, using the set interpretation of numbers, $\wedge$ will come out as *intersection*, $\vee$ as *union* and $\neg$ as relative *complement*. (Notice namely, that $0 = \varnothing$, $1 = \{\varnothing\}$ and $2 = \{\varnothing, \{\varnothing\}\}$.) A good reference for basic logical concepts is [**199**] or [**84**].

Binary Relations. The binary relations over a set M form a boolean algebra with respect to intersection, complement relative to $M \times M$, bottom element $\varnothing$ and top element $\nabla_M := M \times M$. Another special constant is the *diagonal*, $\Delta_M := \{\langle x, x\rangle : x \in M\}$. Moreover, the following operations can be defined. First, for two relations $R, S \subseteq M \times M$ we can define the *composition* $R \circ S := \{\langle x, z\rangle : (\exists y \in M)(x\, R\, y\, S\, z)\}$. From the composition we define the n–fold product R^n of a relation by $R^0 := \Delta_M$ and $R^{n+1} := R \circ R^n$. The *transitive closure* R^+ of R is the union $\bigcup_{0<n<\omega} R^n$. The *reflexive transitive closure* R^* of R is $\bigcup_{n<\omega} R^n$, or equivalently, $R^+ \cup \Delta_M$. Finally, for each relation R there is the *converse* $R^\smile := \{\langle y, x\rangle : x\, R\, y\}$. For the converse we have the following identities

$$\begin{array}{lcl} (R \cup S)^{\smile} & = & R^{\smile} \cup S^{\smile} \\ (R \circ S)^{\smile} & = & S^{\smile} \circ R^{\smile} \\ (R^*)^{\smile} & = & (R^{\smile})^* \end{array}$$

Groups and Semigroups. Given a set G and a binary operation $\cdot$ on G which is associative, $\langle G, \cdot \rangle$ is called a *semigroup*. If $1 \in G$ is such that $1 \cdot x = x \cdot 1 = x$ for all $x \in G$, then 1 is called a *unit*. Moreover, $\langle G, 1, \cdot \rangle$ is called a *monoid*. A particular example is provided by *strings*. Let A be a set. Then A^* denotes the set of finite strings over A. (For many purposes one may define strings as functions from a natural number to A; however, for us, strings are basic objects. They simply are sequences of symbols.) Strings are denoted by vector arrow, e. g. $\vec{x}$, $\vec{y}$, if it is necessary to distinguish them from simple symbols. Given $\vec{x}$ and $\vec{y}$, $\vec{x}^\frown \vec{y}$ or $\vec{x} \cdot \vec{y}$ denotes the concatenation of strings of $\vec{x}$ and $\vec{y}$. The empty string is denoted by ε. $\langle A^*, \varepsilon, ^\frown \rangle$ is a monoid. A string $\vec{x}$ is a *prefix* of $\vec{y}$ if there is a $\vec{u}$ such that $\vec{y} = \vec{x}^\frown \vec{u}$. $\vec{x}$ is a *postfix* of $\vec{y}$ if there exists a $\vec{u}$ such that $\vec{y} = \vec{u}^\frown \vec{x}$. $\vec{x}$ is a *substring* if there are $\vec{u}$ and $\vec{v}$ such that $\vec{y} = \vec{u}^\frown \vec{x}^\frown \vec{v}$. Every string $\vec{x}$ has a length, denoted by $|\vec{x}|$. It is defined inductively as follows:

$$\begin{array}{lcl} |\varepsilon| & := & 0, \\ |\vec{x}| & := & 1, \quad \text{if } \vec{x} \in A, \\ |\vec{x}^\frown \vec{y}| & := & |\vec{x}| + |\vec{y}| \ . \end{array}$$

Suppose we have an additional operation $^{-1} : G \to G$ such that the following laws hold for all $x, y \in G$:

$$\begin{array}{lcl} x \cdot x^{-1} & = & 1, \\ x^{-1} \cdot x & = & 1 \ . \end{array}$$

Then the structure $\langle G, 1, ^{-1}, \cdot \rangle$ is called a *group*.

1.2. Propositional Languages

A **propositional language** $\mathcal{L}$ consists of three things: (1) a set *var* of (propositional) variables, (2) a set F of (propositional) function symbols and (3) a function $\Omega : F \to \omega$. $\Omega(f)$ is the **arity** of f. Ω is called the **signature** of $\mathcal{L}$. The cardinality of *var* is a matter of choice. Unless otherwise stated we assume that $\sharp var = \aleph_0$. Hence we have countably and infinitely many variables. Sometimes we consider the case of languages with finitely many variables; these are called **weak languages**. Since the set *var* is usually fixed, $\mathcal{L}$ is uniquely identified by Ω alone.

To take an example, let us consider the language $\mathcal{B}$, consisting of the function symbols $\wedge$, $\vee$, $\neg$, $\bot$ and $\top$, where $\Omega(\wedge) = \Omega(\vee) = 2$, $\Omega(\neg) = 1$ and $\Omega(\top) = \Omega(\bot) = 0$. So, $\wedge$ and $\vee$ are *binary* function symbols, which is to say that their arity is 2; $\neg$ is a *unary* function symbol, in other words a function symbol of arity 1 and — finally — $\top$ and $\bot$ are *nullary* function symbols: their arity is zero. The propositional languages are a family of languages which are syntactically impoverished. There is only one type of well–formed expression, that of a proposition. The symbol f can also be understood as a function taking $\Omega(f)$ many propositions, returning a proposition (see below the definition of the term algebra). Predicate logic knows at least two types of meaningful expressions: terms and formulas. $x + y$ is a term, $x + y = 2$ is a formula. Natural languages have even more categories to classify meaningful expressions, for example noun phrases, verb phrases, adjectival phrases and so on. Thus, from this point of view, propositional languages are the poorest kind of language, those with a single type only. The freedom lies in the set of basic functions. There are quite meaningful n–ary functions for every n, for example **exactly one of** $A_0, A_1, \ldots$ **or** A_{n-1}. In boolean logic such functions are rarely studied because they can be produced by composing $\wedge$ and $\neg$; however, in other logics — e.g. intuitionistic logic — these definability results no longer hold.

Let X be a set. An **X–string** is a finite string consisting of symbols from $X \cup F$. For a string $\vec{x}$ we write $\vec{x} \subseteq S$ if all members from $\vec{x}$ are in S. The **set of Ω–terms over** X, $Tm_\Omega(X)$, is defined to be the smallest set of strings satisfying

(1) For all $x \in X$, $x \in Tm_\Omega(X)$.
(2) For all $f \in F$ and $t_k \in Tm_\Omega(X)$, $k < \Omega(f)$, also

$$f^\frown t_0^\frown \ldots^\frown t_{\Omega(f)-1} \in Tm_\Omega(X)$$

This way of writing a term will be called **prefix notation** otherwise also known as **Polish Notation**. The more conventional notation with brackets and binary function symbols in between their arguments will be referred to as **infix notation**. Typically, we will write $\vec{x}$ for a sequence of elements otherwise denoted by $x_0, \ldots, x_{k-1}$ for some k. Moreover, for an n–ary function f we will write $f(t_0, \ldots, t_{n-1})$ instead of $f^\frown t_0^\frown t_1^\frown \ldots^\frown t_{n-1}$. If we do not want to highlight the arity of f we will write $f(\vec{t})$ for $f(t_0, t_1, \ldots, t_{n-1})$. When the length of the sequence, n, is suppressed in the notation, in the context $f(\vec{x})$, the sequence is assumed to be of the required length, namely $\Omega(f)$. Finally, if f is a binary termsymbol we also write $t_0\ f\ t_1$ or $(t_0\ f\ t_1)$ (depending on readability) rather than $f^\frown t_0^\frown t_1$. (This is the infix notation.) The set X is taken in the context of propositional logic to be the set of **variables**, sometimes also denoted by var. We write $var(t)$ for the set of variables occurring in t and $sf(t)$ for the set of subterms or subformulae of t. We define them formally as follows. $var(t) := sf(t) \cap X$, and

$$\begin{array}{lcl} sf(x_i) & := & \{x_i\} \\ sf(f(t_0, \ldots, t_{\Omega(f)-1})) & := & \{f(t_0, \ldots, t_{\Omega(f)-1})\} \cup \bigcup_{i<\Omega(f)} sf(t_i) \end{array}$$

Nullary functions are also referred to as **constants**. The cardinality of the set of constants is usually considered independently of the cardinality of functions. The cardinality of $Tm_\Omega(X)$ is infinite except in some trivial cases.

Proposition 1.2.1. *$Tm_\Omega(X)$ is finite exactly when (i) X is empty, there are no constants and any number of n–ary functions for $n > 0$, or (ii) X is finite and there exist finitely many constants and no n–ary functions for any $n > 0$. If the cardinality of $Tm_\Omega(X)$ is infinite it is equal to the maximum of $\aleph_0$ and the cardinalities of the set of variables, the set of constants and the set of function symbols.*

Proof. Let C denote the set of constants and F the set of functions of arity > 0. Let $\alpha := \sharp X + \sharp C, \beta := \sharp F$ and $\gamma := \sharp Tm_\Omega(X)$. Obviously, if $\alpha = 0$, $\gamma = 0$ as well. If $\beta = 0$ then $\gamma = \alpha$. This is finite if α is. The theorem holds in all these cases. Now assume that $F \neq \varnothing$ and $X \cup C \neq \varnothing$. Define the sets S_i, $i < \omega$, as follows.

$$\begin{array}{rcl} S_0 & := & X \cup C \\ S_{i+1} & := & S_i \cup \{f(\vec{x}) : \vec{x} \subseteq S_i, f \in F\} \end{array}$$

Then

$$Tm_\Omega(X) = \bigcup_{i<\omega} S_i$$

If $F \neq \varnothing$ and $S_0 \neq \varnothing$, then $S_i \neq \varnothing$ and also $S_i \subsetneq S_{i+1}$. Hence γ is at least $\aleph_0$. If S_0 and F are finite, so is S_i for every $i < \omega$. Hence the theorem holds in these cases. Finally, let α be infinite. Then $\alpha = max\{\sharp X, \sharp C\}$. It can be shown that $\sharp S_{i+1} = \sharp S_i \cdot \beta = max\{\sharp S_i, \beta\}$. Hence for $i > 0$, $\sharp S_i = \sharp S_1$. So, $\gamma = \sharp S_1 = max\{\alpha, \beta\}$, showing the theorem. □

Typically, we think of the function symbols f as standing for functions, taking $\Omega(f)$ many inputs and yielding a value. This is codified in the notion of an Ω–algebra. (For basic concepts of universal algebra see [**37**] and [**89**].) An Ω–algebra is a pair $\mathfrak{A} = \langle A, I\rangle$, where A is a set, and I a function assigning to each $f \in F$ an $\Omega(f)$–ary function from A to A. This definition is somewhat cumbersome, and is replaced by the following, less rigorous definition. An **Ω–algebra** (or simply an **algebra**) is a pair $\mathfrak{A} = \langle A, \langle f^{\mathfrak{A}} : f \in F\rangle\rangle$, where A is a set, the **underlying set of** $\mathfrak{A}$, and for each $f \in F$, $f^{\mathfrak{A}}$ is an $\Omega(f)$–ary function on A, in symbols $f^{\mathfrak{A}} : A^{\Omega(f)} \to A$. One very important Ω–algebra is the algebra of Ω–terms over a given set X. Let us denote by f not only the function symbol f, but also the function $Tm_\Omega(X)^{\Omega(f)} \to Tm_\Omega(X) : \langle t_j : j < \Omega(f)\rangle \mapsto f(t_0, t_1, \ldots, t_{\Omega(f)-1})$. Now, since the set of Ω–terms is closed under (application of) the functions f for every $f \in F$, the following is well–defined.

$$\mathfrak{Tm}_\Omega(X) := \langle Tm_\Omega(X), F\rangle.$$

This is called the **algebra of Ω–terms over** X or simply the **termalgebra** (**over** X). In addition to the termalgebras we have the trivial algebra $\mathbf{1} := \langle\{0\}, \langle f^{\mathbf{1}} : f \in F\rangle\rangle$, where $f^{\mathbf{1}}(0, \ldots, 0) = 0$. In addition to the primitive functions $f^{\mathfrak{A}}$, $f \in F$, we can form complex functions by composition of functions; in algebra, for example, + and $\cdot$ are primitive functions, and $x\cdot(y+z)$, $x\cdot y+x\cdot z$ are complex functions (and the terms are distinct even though they represent identical functions from $\mathbb{R}^3$ to $\mathbb{R}$). A **term–function** of $\mathfrak{A}$ is now generally defined to be any primitive or complex function of $\mathfrak{A}$. Given a set A a **clone of term–functions** is a set **Cl** of functions $f : A^m \to A$ for some $m < \omega$ satisfying

(1) For all $n < \omega$ the projections $\mathbf{p}^n_i : A^n \to A$ of an n–sequence to the i–component are in **Cl**.
(2) For all $n < \omega$ if $f : A^k \to A$ is in **Cl** and $g_i : A^n \to A$ are in **Cl** for $i < k$, then the composition $f[g_0, \ldots, g_{k-1}] : A^n \to A$ is also in **Cl**. It is defined

by

$$f[g_0, \ldots, g_{k-1}](\vec{a}) := f(g_0(\vec{a}), \ldots, g_{k-1}(\vec{a}))$$

where $\vec{a} \in A^n$.

Let $\mathfrak{A}$ be an Ω–algebra. The clone generated by the functions $f^{\mathfrak{A}}$ is called the **clone of term functions** of $\mathfrak{A}$, and is denoted by $Clo(\mathfrak{A})$. $Clo_n(\mathfrak{A})$ denotes the set of n–ary term functions of $\mathfrak{A}$. A **polynomial** of the algebra $\mathfrak{A}$ is a termfunction of the algebra $\mathfrak{A}_A$, where $\mathfrak{A}_A$ denotes the algebra $\mathfrak{A}$ expanded by constants c_a with value a for each $a \in A$. We denote by $Pol_n(\mathfrak{A})$ the set of all n–ary polynomials of $\mathfrak{A}$, and by $Pol(\mathfrak{A})$ the set of polynomials of $\mathfrak{A}$. The elements of $Pol_1(\mathfrak{A})$ are called **translations**. The reader may verify the following fact. Given a polynomial $p(\vec{x}) \in Pol_n(\mathfrak{A})$ there is a term function $f(\vec{x}, \vec{y}) \in Clo_{n+m}(\mathfrak{A})$ for some $m \in \omega$ and some $\vec{b} \in A^m$ such that $p(\vec{a}) = f(\vec{a}, \vec{b})$ for all $\vec{a} \in A^n$. This means that every polynomial results from a term function by supplying constants for some of the arguments.

An Ω–**homomorphism** (or simply a **homomorphism**) from the algebra $\mathfrak{A} = \langle A, \langle f^{\mathfrak{A}} : f \in F\rangle\rangle$ to the algebra $\mathfrak{B} = \langle B, \langle f^{\mathfrak{B}} : f \in F\rangle\rangle$ is a map $h : A \to B$ such that for all $f \in F$ and all elements $a_j \in A$, $j < \Omega(f)$,

$$h(f^{\mathfrak{A}}(a_0, \ldots, a_{\Omega(f)-1})) = f^{\mathfrak{B}}(h(a_0), \ldots, h(a_{\Omega(f)-1})) \; .$$

In case $h : A \to B$ is a homomorphism we write $h : \mathfrak{A} \to \mathfrak{B}$. To rephrase the formal definition, homomorphisms are maps which preserve the structure of the source algebra; the source elements compose in the same way in the source algebra as their images do in the target algebra. A homomorphism $h : \mathfrak{A} \to \mathfrak{A}$ is called an **endomorphism** of $\mathfrak{A}$. A bijective endomorphism is called an **automorphism** of $\mathfrak{A}$. We write $End(\mathfrak{A})$ for the set of endomorphisms of $\mathfrak{A}$ and $Aut(\mathfrak{A})$ for the set of automorphisms of $\mathfrak{A}$. $End(\mathfrak{A})$ is closed under composition, and so the endomorphisms form a semigroup with id_A as unit. Moreover, if $h : A \to A$ is an automorphism, so is $h^{-1} : A \to A$.

Theorem 1.2.2. *Let Ω be a signature and $\mathfrak{A}$ an Ω–algebra. Put*

$$\begin{array}{rcl} \mathfrak{End}(\mathfrak{A}) & := & \langle End(\mathfrak{A}), id_A, \circ\rangle \; , \\ \mathfrak{Aut}(\mathfrak{A}) & := & \langle Aut(\mathfrak{A}), id_A, ^{-1}, \circ\rangle \; . \end{array}$$

Then $\mathfrak{End}(\mathfrak{A})$ is a semigroup and $\mathfrak{Aut}(\mathfrak{A})$ is a group.

A map $h : A \to B$ induces an equivalence relation $ker(h)$ on A via $\langle x, y\rangle \in ker(h)$ iff $h(x) = h(y)$. We call $ker(h)$ the **kernel** of h. Given an equivalence relation Θ, the sets $[x]\Theta := \{y : x \;\Theta\; y\}$ are called the **cosets** of Θ. For a set D, we write $[D]\Theta := \bigcup_{x \in D}[x]\Theta$. With $\Theta = ker(h)$ we call the cosets also **fibres** of h. If $h : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism, it induces a special equivalence relation on A, called a *congruence*. A **congruence** (**relation**) on $\mathfrak{A}$ is a set $\Theta \subseteq A \times A$ which is an equivalence relation and for all functions f and sequences $\vec{x} = \langle x_0, \ldots, x_{\Omega(f)-1}\rangle$ and $\vec{y} = \langle y_0, \ldots, y_{\Omega(f)-1}\rangle$ if $x_j \,\Theta\, y_j$ for all $j < \Omega(f)$ then also $f^{\mathfrak{A}}(\vec{x}) \,\Theta\, f^{\mathfrak{A}}(\vec{y})$. Each congruence relation on $\mathfrak{A}$

defines a so–called **factor algebra** $\mathfrak{A}/\Theta$ whose elements are the cosets $[x]\Theta$ of the equivalence relation, and the operations $[f^{\mathfrak{A}}]\Theta$ act blockwise in the following way.

$$[f^{\mathfrak{A}}]\Theta([x_0]\Theta, \ldots, [x_{\Omega(f)-1}]\Theta) := [f^{\mathfrak{A}}(x_0, \ldots, x_{\Omega(f)-1})]\Theta.$$

It is the fact that Θ is a congruence relation that makes this definition independent of the choice of the representatives. The reader may check this for known cases such as groups, lattices etc. The map $x \mapsto [x]\Theta$ is an Ω–homomorphism from $\mathfrak{A}$ onto $\mathfrak{A}/\Theta$. We write $\mathfrak{A} \twoheadrightarrow \mathfrak{A}/\Theta$ to highlight the fact that the map is surjective. The following is now straightforwardly proved.

Proposition 1.2.3. *Let $h : \mathfrak{A} \twoheadrightarrow \mathfrak{B}$ be a surjective homomorphism. Then $\Theta := ker(h)$ is a congruence, and $\mathfrak{B} \cong \mathfrak{A}/\Theta$. An isomorphism is given by the map $\iota : h(x) \mapsto [x]\Theta$.*

Since h is surjective, ι is defined on all elements of $\mathfrak{B}$ and is well–defined by construction of $[x]\Theta$.

Let $E \subseteq A \times A$. The smallest congruence relation containing E is denoted by $\Theta(E)$. The map $E \mapsto \Theta(E)$ is a closure operator with closed sets being the congruences. The congruences on an algebra form a complete lattice, $\bigsqcap$ being intersection and $\bigsqcup$ the smallest congruence containing the members of the union. We denote this lattice by $Con(\mathfrak{A})$. The bottom element of this lattice is the **diagonal** $\Delta = \{\langle a, a\rangle : a \in A\}$, and the top element is $\nabla = A \times A$. An algebra is **simple** if it has only these congruences and they are distinct. The congruence $\Theta(E)$ can be computed explicitly. Let

$$E^{\tau} := \{\langle t(a), t(b)\rangle : \langle a, b\rangle \in E, t \in Pol_1(\mathfrak{A})\}.$$

This is the closure of E under translations. Then $\Theta(E)$ is the smallest equivalence relation containing E^{τ}. The rationale behind this is the following characterization of congruence relations.

Proposition 1.2.4. *Let $\mathfrak{A}$ be an algebra. A binary relation on A is a congruence on $\mathfrak{A}$ iff it is an equivalence relation closed under translations.*

Proof. Clearly, a congruence relation must be an equivalence relation. Conversely, let Θ be an equivalence relation and t a translation. Then $t(x) = f(x, \vec{a})$ for a term function $f(x, \vec{y})$. Assume $x \,\Theta\, y$. Then

$$t(x) = f(x, \vec{a}) \,\Theta\, f(y, \vec{a}) = t(y) \ .$$

Hence Θ is closed under translations. Conversely, let Θ be translation closed and $\vec{x} = \langle x_i : i < n\rangle$, $\vec{y} = \langle y_i : i < n\rangle$ be n–long sequences such that $x_i \,\Theta\, y_i$ for all $i < n$. Assume that f is a an n–ary term function. Then we have

$$\begin{array}{rcl} f(x_0, \ldots, x_{n-1}) & \Theta & f(y_0, x_1, \ldots, x_{n-1}) \\ & \Theta & f(y_0, y_1, x_2, \ldots, x_{n-1}) \\ & \ldots & \\ & \Theta & f(y_0, \ldots, y_{n-1}) \ . \end{array}$$

By transitivity, $f(\vec{x}) \,\Theta\, f(\vec{y})$. ∎

Proposition 1.2.5. *Let Θ_i, $i \in I$, be congruences of a given algebra $\mathfrak{A}$. Then let Ψ be the transitive closure of the equivalence relation $\bigcup_{i \in I} \Theta_i$. Then Ψ is a congruence on $\mathfrak{A}$ and identical to $\bigsqcup_{i \in I} \Theta_i$. Hence, $\langle x, y \rangle \in \bigsqcup_{i \in I} \Theta_i$ iff there is a number $n < \omega$, elements z_i, $i < n + 1$, and a sequence $j(i)$, $i < n$, of elements in I such that*

$$x = z_0 \, \Theta_{j(0)} \, z_1 \, \Theta_{j(1)} \, z_2 \ldots z_{n-1} \, \Theta_{j(n-1)} \, z_n = y$$

Proof. $\bigcup_{i \in I} \Theta_i$ is symmetric and reflexive; so is its transitive closure, Ψ. Therefore, we only need to verify that Ψ is closed under translations. To see that, let $\langle x, y \rangle \in \Psi$. Then there exist elements z_i, $i < n + 1$, and a sequence $j(i)$, $i < n$, of elements of I such that

$$x = z_0 \, \Theta_{j(0)} \, z_1 \, \Theta_{j(1)} \, z_2 \ldots z_{n-1} \, \Theta_{j(n-1)} \, z_n = y$$

Now let $f \in Pol_1(\mathfrak{A})$ be a translation. Then

$$f(x) = f(z_0) \, \Theta_{j(0)} \, f(z_1) \, \Theta_{j(1)} \, f(z_2) \ldots f(z_{n-1}) \, \Theta_{j(n-1)} \, f(z_n) = f(y)$$

Hence $\langle f(x), f(y) \rangle \in \Psi$. □

We derive the following useful consequence. Let $\langle a, b \rangle \in \bigsqcup \langle \Theta_i : i \in I \rangle$. Then for some finite set $I_0 \subseteq I$, $\langle a, b \rangle \in \bigsqcup \langle \Theta_i : i \in I_0 \rangle$. Moreover, for a set E of equations, $\Theta(E) = \bigcup \langle \Theta(E_0) : E_0 \subseteq E, \sharp E_0 < \aleph_0 \rangle$.

Proposition 1.2.6. *Let $\mathfrak{A}$ be an Ω–algebra. Then $Con(\mathfrak{A})$ is algebraic. The compact elements are of the form $\Theta(E)$, E a finite set of equations. Moreover,*

$$\Theta(E) = \bigcup \langle \Theta(E_0) : E_0 \subseteq E, \sharp E_0 < \aleph_0 \rangle$$

Term algebras have the important property that for any function $v : X \to A$ where A is the underlying set of $\mathfrak{A}$ there is exactly one Ω–homomorphism $\overline{v} : \mathfrak{Tm}_\Omega(X) \to \mathfrak{A}$ such that $\overline{v} \upharpoonright X = v$. $\overline{v}$ can be defined inductively as follows.

(1) For $x \in X$ we have $\overline{v}(x) = v(x)$.
(2) For every $f \in F$ and terms t_k ($k < \Omega(f)$):

$$\overline{v}(f(t_0, \ldots, t_{\Omega(f)-1})) = f^{\mathfrak{A}}(\overline{v}(t_0), \ldots, \overline{v}(t_{\Omega(f)-1})) \, .$$

We note the following useful fact. If $h : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism then $h \circ v : X \to B$, and $\overline{h \circ v} = h \circ \overline{v}$. Given $v : X \to A$ and a term $t = t(x_0, \ldots, x_{n-1})$ then $\overline{v}(t) \in A$. Hence each term t defines a term–function $t^{\mathfrak{A}} : A^n \to A$ on $\mathfrak{A}$ such that $t^{\mathfrak{A}}(a_0, \ldots, a_{n-1}) = \overline{v}(t(x_0, \ldots, x_{n-1}))$ where $v(x_i) = a_i$, $i < n$. A map $\sigma : X \to Tm_\Omega(X)$ is called a **substitution**. A substitution defines a unique homomorphism $\overline{\sigma} : \mathfrak{Tm}_\Omega(X) \to \mathfrak{Tm}_\Omega(X)$; conversely, any homomorphism of this type is determined by a substitution. So, substitutions are simply endomorphisms of the term algebra. We usually write t^σ for $\sigma(t)$. It is also customary to write $t[u/x]$ or $t[x \mapsto u]$ to denote the result of applying to t the substitution $\sigma : x \mapsto u, y \mapsto y$ (for all $y \neq x$). Given a set V of variables we denote by $t[u_x/x : x \in V]$ the result of simultaneously substituting u_x for x for all $x \in V$.

Exercise 1. Let A be a set, and $E \subseteq A \times A$. Show that the least equivalence relation over A containing E, $e(E)$, can be computed in three steps. Let $r(E) = E \cup \Delta_A$ denote the reflexive closure, $s(E) = E \cup E^{\smile}$ the symmetric closure, and $t(E) = E^+$ the transitive closure. Then $e(E) = t(s(r(E)))$.

Exercise 2. (Continuing the previous exercise.) Show that if E is reflexive, that is, $E = r(E)$ then so is E^τ; and that if E is symmetric, then so is E^τ. Hence show that $\Theta(E)$ can be computed in four steps as follows. (i) close E reflexively, (ii) close E symmetrically, (iii) close under translations and (iv) close E transitively.

Exercise 3. Show Theorem 1.2.2.

1.3. Algebraic Constructions

We have already encountered the notion of a homomorphism of Ω–algebras and congruences. Here we will state and prove some extremely useful theorems about these constructions and also introduce some (more) notation. If $h : \mathfrak{A} \to \mathfrak{B}$ is *surjective* we write $h : \mathfrak{A} \twoheadrightarrow \mathfrak{B}$ and call $\mathfrak{B}$ a **homomorphic image** of $\mathfrak{A}$. If h is injective we write $h : \mathfrak{A} \rightarrowtail \mathfrak{B}$. Furthermore, if $A \subseteq B$ and the natural inclusion $h : A \to B : x \mapsto x$ is a homomorphism we say that $\mathfrak{A}$ is a **subalgebra** of $\mathfrak{B}$ and write $\mathfrak{A} \leq \mathfrak{B}$. If h is both surjective and injective, it is called an **isomorphism**. $\mathfrak{A}$ and $\mathfrak{B}$ are **isomorphic** if there exists an isomorphism $h : \mathfrak{A} \to \mathfrak{B}$. We know that each homomorphism from $\mathfrak{A}$ to $\mathfrak{B}$ induces a congruence on $\mathfrak{A}$ and that each congruence on an algebra is associated with a surjective homomorphism. There is a one–to–one–connection between congruences and the natural factor algebras, where we compute with blocks rather than elements. Moreover, the following holds.

Proposition 1.3.1. *(1.) Let $\mathfrak{A}$ be an algebra and $\Theta_1 \subseteq \Theta_2$. Then Θ_2 induces a congruence on $\mathfrak{A}/\Theta_1$, denoted by Θ_2/Θ_1. Moreover,*

$$(\mathfrak{A}/\Theta_1)/(\Theta_2/\Theta_1) \cong \mathfrak{A}/\Theta_2 \ .$$

(2.) Let $h : \mathfrak{A} \twoheadrightarrow \mathfrak{B}$ be surjective and let Θ be a congruence on $\mathfrak{B}$. Then there exists a congruence Φ on $\mathfrak{A}$ such that $\mathfrak{A}/\Phi \cong \mathfrak{B}/\Theta$. (3.) Let $\mathfrak{A} \leq \mathfrak{B}$ be a subalgebra and Θ_B be a congruence on $\mathfrak{B}$. Then $\Theta_B \restriction A := \Theta_B \cap A \times A$ is a congruence on $\mathfrak{A}$.

Proof. **(1.)** Θ_2/Θ_1 can be defined as follows. $[[x]\Theta_1](\Theta_1/\Theta_2) := [x]\Theta_2$. This is independent of the choice of x as a representative of the class $[x]\Theta_1$, since Θ_2 includes Θ_1. The map $[x]\Theta_1 \mapsto [x]\Theta_2$ is a homomorphism with kernel Θ_2/Θ_1, as is immediately verified. **(2.)** Put $\Phi := \{\langle a, b\rangle : h(a)\,\Theta\,h(b)\}$. Let $h/\Phi : [x]\Phi \mapsto [h(x)]\Theta$. This is well defined because it does not depend on the choice of x as a representative of its class. By definition, h/Φ is injective; it is also surjective. It is also not hard to see that it is a homomorphism. **(3.)** Put $\Theta_A := \Theta_B \restriction A$. Θ_A is clearly an equivalence relation. Now let $t \in Pol_1(\mathfrak{B})$, $a, b \in A$. Then $t(a), t(b) \in A$, since $\mathfrak{A}$ is a subalgebra. Hence if $\langle a, b\rangle \in \Theta_A$ we have $\langle t(a), t(b)\rangle \in \Theta_A$ by the fact that Θ_B is a congruence and $\mathfrak{A}$ is closed under translations. ◻

Theorem 1.3.2. *The map* $\Phi \mapsto \Phi/\Theta$ *is an isomorphism from the interval* $[\Theta, \nabla]$ *in the lattice* $Con(\mathfrak{A})$ *onto the lattice* $Con(\mathfrak{A}/\Theta)$.

Proof. By the previous proposition this map is bijective. It is easy to see that it is an order isomorphism, thus it is a lattice isomorphism as well. □

Next we define the *product* of algebras. Let $\mathfrak{A}_j$, $j \in J$, be a family of algebras. Then the underlying set of the product is the cartesian product of the sets, $\mathsf{X}_{j\in J} A_j$, which is the set of functions s with domain J and value $s(j) \in A_j$. Denote this set by P. The operations are defined pointwise, that is, if f is an n–ary function symbol, and $s_0, \ldots, s_{n-1} \in P$, then

$$f^{\mathfrak{P}}(s_0, \ldots, s_{n-1})(j) := f^{\mathfrak{A}_j}(s_0(j), \ldots, s_{n-1}(j)) .$$

Then the **product** of the $\mathfrak{A}_j$ is

$$\prod_{j\in J} \mathfrak{A}_j := \langle P, \langle f^{\mathfrak{P}} : f \in F\rangle\rangle$$

The projection maps $p_i : P \to A_i : s \mapsto s(i)$ are homomorphisms. If Θ_j, $j \in J$, are congruences on the $\mathfrak{A}_j$ then there is a natural congruence $\mathsf{X}_{j\in J}\,\Theta_j$ on the product defined by $s\,(\mathsf{X}_{j\in J}\,\Theta_j)\,t$ iff for all $j \in J$ we have $s(j)\,\Theta_j\,t(j)$. For every family of maps $h_j : \mathfrak{A}_j \to \mathfrak{B}_j$ there exists a map

$$\prod_{j\in I} h_j : \prod_{j\in I} \mathfrak{A}_j \to \prod_{j\in I} \mathfrak{B}_j .$$

If all h_j are injective (surjective) then so is $\prod_{j\in I}$. The kernel of $\prod_{j\in I} h_j$ is exactly $\mathsf{X}_{j\in I} ker(h_j)$. However, not every congruence on the product can be obtained in this way. (The easiest example are *sets*, that is, algebras with no functions. There a congruence is just an equivalence relation. An equivalence on a product set is not necessarily decomposable.)

If $\mathcal{K}$ is a class of Ω–algebras for a fixed Ω, then by $\mathsf{H}(\mathcal{K})$ we denote the class of all algebras $\mathfrak{B}$ which are homomorphic images of algebras in $\mathcal{K}$, we denote by $\mathsf{S}(\mathcal{K})$ the class of subalgebras of algebras in $\mathcal{K}$ and by $\mathsf{P}(\mathcal{K})$ the class of algebras isomorphic to products of algebras in $\mathcal{K}$. A **variety** is a class closed under all three operators H, S and P.

Theorem 1.3.3. *Let* Ω *be a signature and* $\mathcal{K}$ *a class of* Ω*–algebras. The smallest variety containing* $\mathcal{K}$ *is the class* $\mathsf{HSP}(\mathcal{K})$.

Proof. All operators are individually closure operators. Namely, we have $\mathcal{K} \subseteq \mathsf{O}(\mathcal{K})$ for all $\mathsf{O} \in \{\mathsf{H}, \mathsf{S}, \mathsf{P}\}$. For if $\mathfrak{A}$ is in $\mathcal{K}$, then since the identity $1_{\mathfrak{A}} : \mathfrak{A} \to \mathfrak{A}$ is an isomorphism, $\mathfrak{A} \in \mathsf{S}(\mathcal{K})$ as well as $\mathfrak{A} \in \mathsf{H}(\mathcal{K})$. Moreover, $\mathfrak{A}$ is a product of $\mathfrak{A}$ (take a singleton index set). Secondly, if $\mathcal{K} \subseteq \mathcal{L}$ then $\mathsf{O}(\mathcal{K}) \subseteq \mathsf{O}(\mathcal{L})$ for $\mathsf{O} \in \{\mathsf{S}, \mathsf{H}, \mathsf{P}\}$, as is immediate from the definition. We will leave it as an exercise to show that $\mathsf{HH}(\mathcal{K}) \subseteq \mathsf{H}(\mathcal{K})$, $\mathsf{SS}(\mathcal{K}) \subseteq \mathsf{S}(\mathcal{K})$ and $\mathsf{PP}(\mathcal{K}) \subseteq \mathsf{P}(\mathcal{K})$. Furthermore, $\mathsf{SH}(\mathcal{K}) \subseteq \mathsf{HS}(\mathcal{K})$. For let $\mathfrak{C} \leq \mathfrak{B}$ and $\mathfrak{B} \cong \mathfrak{A}/\Theta$ for some congruence Θ. We may assume that

$\mathfrak{B} = \mathfrak{A}/\Theta$. Then C is a set of blocks of the form $[x]\Theta$. Put $D := \{x : [x]\Theta \in C\}$. Then $[D]\Theta = C$. Moreover, by the fact that C is closed under the operations $[f]\Theta$ it follows that D is closed under all operations of $\mathfrak{A}$. Hence $\mathfrak{D} \leq \mathfrak{A}$. Since we have $h_\Theta \restriction D : \mathfrak{D} \twoheadrightarrow \mathfrak{C}$, it follows that $\mathfrak{D} \in \mathsf{HS}(\mathcal{K})$. Also, we have noted that $\mathsf{PH}(\mathcal{K}) \subseteq \mathsf{HP}(\mathcal{K})$. Finally, $\mathsf{PS}(\mathcal{K}) \subseteq \mathsf{SP}(\mathcal{K})$ since the product of subalgebras is a subalgebra of the product of the $\mathfrak{B}_j$ as we have noted above. With these commutation laws we have $\mathsf{H}(\mathsf{HSP}(\mathcal{K})) \subseteq \mathsf{HSP}(\mathcal{K})$, $\mathsf{S}(\mathsf{HSP}(\mathcal{K})) \subseteq \mathsf{HSSP}(\mathcal{K}) = \mathsf{HSP}(\mathcal{K})$ and $\mathsf{P}(\mathsf{HSP}(\mathcal{K})) \subseteq \mathsf{HPSP}(\mathcal{K}) \subseteq \mathsf{HSPP}(\mathcal{K}) = \mathsf{HSP}(\mathcal{K})$. This shows the theorem. □

Let λ be a cardinal number. An algebra $\mathfrak{A}$ is λ**–generable** or λ**–generated** if there exist elements a_μ, $\mu < \lambda$, such that the smallest subalgebra containing all these elements is $\mathfrak{A}$ itself. Likewise, since the smallest subalgebra containing these elements is the set of all elements obtainable by applying the functions to these elements, $\mathfrak{A}$ is λ–generated iff there exists a surjective homomorphism $\mathfrak{Tm}_\Omega(X) \twoheadrightarrow \mathfrak{A}$, where X is a set of cardinality λ. An algebra $\mathfrak{A}$ is called **freely** λ**–generated** in a class $\mathcal{K}$ if if there is a set $X \subseteq A$ of cardinality λ such that for every map $v : X \to B$ there is exactly one homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ such that $h \restriction X = v$. We write $\overline{v}$ for h. We say also that $\mathfrak{A}$ is freely generated by X. In the class of all Ω–algebras, the term algebras $\mathfrak{Tm}_\Omega(X)$ are freely generated by X. (They are also called **absolutely free** Ω–algebras.)

Proposition 1.3.4. *Let $\mathcal{K}$ be a class of algebras, and let $\mathfrak{A}$ and $\mathfrak{B}$ be freely λ–generated in $\mathcal{K}$. Then $\mathfrak{A} \cong \mathfrak{B}$.*

Proof. Let $X \subseteq A$ and $Y \subseteq B$ be subsets of cardinality λ such that $\mathfrak{A}$ is freely generated by X and $\mathfrak{B}$ freely generated by Y. By assumption, there exist maps $p : X \to Y$ and $q : Y \to X$ such that $q \circ p = 1_X$ and $p \circ q = 1_Y$. Then $\overline{p} : \mathfrak{A} \to \mathfrak{B}$ and $\overline{q} : \mathfrak{B} \to \mathfrak{A}$ are (uniquely) defined extensions of p and q. Moreoever, $\overline{p} \circ \overline{q} \restriction Y = 1_Y$, and so $\overline{p} \circ \overline{q} = 1_{\mathfrak{B}}$, since there is exactly one homomorphism extending 1_Y, and the identity is a homomorphism. Likewise $\overline{q} \circ \overline{p} = 1_{\mathfrak{A}}$ is proved. Hence $\mathfrak{A}$ and $\mathfrak{B}$ are isomorphic. □

The previous theorem has established the uniqueness of free algebras. Their existence is not generally guaranteed. To have free algebras for all cardinalities of generating sets is a nontrivial property. A central theorem of universal algebra states that all nontrivial varieties have free algebras. The proof looks difficult, but the argument is simple. Take a cardinal λ and consider all pairs $\langle f, \mathfrak{A}\rangle$ where $\mathfrak{A} \in \mathcal{K}$ and $f : \lambda \to A$. Let S be the set of such pairs. It is used as an index set in the product

$$\mathfrak{P}_\lambda := \prod_{\langle f,\mathfrak{A}\rangle \in S} \mathfrak{A}$$

Let $\mathfrak{C}$ be the subalgebra generated by the functions s_μ, with $\mu \in \lambda$, where $s_\mu(\langle f, \mathfrak{A}\rangle) = f(\mu)$. Let $\iota : \mathfrak{C} \rightarrowtail \mathfrak{P}_\lambda$ be the inclusion map. We claim that $\mathfrak{C}$ is freely generated by the s_μ. To that end, let $v : s_\mu \mapsto a_\mu$ be any map into an algebra $\mathfrak{A} \in \mathcal{K}$. Then let g be defined by $g(\mu) := v(s_\mu)$. The pair $c = \langle g, \mathfrak{A}\rangle$ is in S. Hence there is a projection p_c :

$\mathfrak{P}_\lambda \to \mathfrak{A}$ such that $p_c(s_\mu) = s_\mu(c) = a_\mu$. The composition $p_c \circ \iota$ is a homomorphism from $\mathfrak{C}$ to $\mathfrak{A}$. By $\mathfrak{Fr}_{\mathcal{V}}(X)$ we denote the algebra freely generated by X in $\mathcal{V}$; and for a cardinal number λ we denote by $\mathfrak{Fr}_{\mathcal{V}}(\lambda)$ the freely λ–generated algebra in $\mathcal{V}$. Given a set X of generators, and a variety $\mathcal{V}$, there may be several terms denoting the same element in the free algebra $\mathfrak{Fr}_{\mathcal{V}}(X)$, since it is in general a homomorphic image of the term algebra $\mathfrak{Tm}_\Omega(X)$. Nevertheless, we will not distinguish between a term $t(\vec{x})$ and its equivalence class in $\mathfrak{Fr}_{\mathcal{V}}(X)$.

Theorem 1.3.5. *Let $\mathcal{V}$ be a variety of Ω–algebras for a given Ω. For every cardinal γ there exists in $\mathcal{V}$ a freely γ–generated algebra. Every algebra of $\mathcal{V}$ is the homomorphic image of a free algebra in $\mathcal{V}$.*

The reader may care to note that it may happen that the variety is trivial, containing up to isomorphism only the algebra **1**. In that case even though we start off with a set of larger cardinality, the free algebra will be isomorphic to **1**, so the generators turn out to be equal as elements of the algebra. If we insist on the generators as being elements of the free algebra then the previous theorem is false just in case we have a trivial variety. However, under a different reading the theorem makes sense, namely if we take the following definition of a free algebra. An algebra $\mathfrak{A}$ is *free over X* if there is a map $i : X \to A$ such that for any map $j : X \to B$ there is a homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ for which $j(x) = h \circ i(x)$ for all $x \in X$. In the case of the trivial variety, X can have any cardinality, and yet $A = \{0\}$. In the sequel we always assume to work with nontrivial varieties, whenever this should make a difference.

Theorem 1.3.6. *Let $\mathcal{V}$ be a variety, $i : X \rightarrowtail Y$ and $p : Y \twoheadrightarrow X$. Then $\overline{i} : \mathfrak{Fr}_{\mathcal{V}}(X) \rightarrowtail \mathfrak{Fr}_{\mathcal{V}}(Y)$ and $\overline{p} : \mathfrak{Fr}_{\mathcal{V}}(Y) \twoheadrightarrow \mathfrak{Fr}_{\mathcal{V}}(X)$.*

Proof. If $i : X \rightarrowtail Y$ there exists a q such that $q \circ i = 1_X$. It follows that $\overline{q \circ i}$ is the identity on $\mathfrak{Fr}_{\mathcal{V}}(X)$. Since $\overline{q \circ i} = \overline{q} \circ \overline{i}$, $\overline{q}$ is surjective and $\overline{i}$ is injective. Similarly it is shown that $\overline{p}$ is surjective. ☐

Proposition 1.3.7. *Let $\mathcal{V}$ be a variety and $\mathfrak{A}$ an algebra of $\mathcal{V}$. Then there exists a free algebra $\mathfrak{F}$ and a homomorphism $h : \mathfrak{F} \twoheadrightarrow \mathfrak{A}$.*

For a proof note that there exists a surjection $v : A \twoheadrightarrow A$. Hence we also have that $\overline{v} : \mathfrak{Fr}_{\mathcal{V}}(A) \twoheadrightarrow \mathfrak{A}$.

Exercise 4. Show that either a variety contains up to isomorphism only the trivial algebra **1** or it contains infinite algebras.

Exercise 5. Let $\mathcal{L}$ be a language with signature Ω. Let γ be the cardinality of $Tm_\Omega(\varnothing)$. Then show that any variety of Ω–algebras which is nontrivial contains an algebra of any infinite cardinality $\delta \geq \gamma$.

Exercise 6. Show with a specific example that the claim of the previous exercise need not hold for finite δ.

Exercise 7. Show that for any class of algebras, $\mathsf{HH}(\mathcal{K}) \subseteq \mathsf{H}(\mathcal{K})$ as well as $\mathsf{SS}(\mathcal{K}) \subseteq \mathsf{S}(\mathcal{K})$ and $\mathsf{PP}(\mathcal{K}) \subseteq \mathsf{P}(\mathcal{K})$. *Hint.* For the last claim take a collection $\langle I_j : j \in J\rangle$ of pairwise disjoint index sets, and assume that there is an algebra $\mathfrak{A}_{i_j}$ for every $i_j \in I_j$, $j \in J$. Then each $\mathfrak{B}_j := \prod_{k \in I_j} \mathfrak{A}_k$ is a product, and $\mathfrak{C} := \prod_{j \in J} \mathfrak{B}_j$ is a general element of $\mathsf{PP}(\mathcal{K})$. Now let $K := \bigcup_{j \in J} I_j$ and put $\mathfrak{D} := \prod_{k \in K} \mathfrak{A}_k$. Show that $\mathfrak{C} \cong \mathfrak{D}$.

1.4. General Logic

In our view, logic is the study of truth and consequence. In logic we study (among other things) whether a statement φ follows from some set Δ of other statements. We usually write $\Delta \vdash \varphi$ if this is the case. We interpret this as follows: if all $\chi \in \Delta$ are true then so is φ. Of course, we must specify what we mean by being true. However, already on these assumptions there are some nontrivial things that can be said about the relation $\vdash$. To write them down, we will — in accordance with our notation in connection with modal logic — use lower case Greek letters for terms of propositional logic, since these terms are thought of as formulae. We also will henceforth not distinguish between $\mathcal{L}$ as a set of function symbols and the terms of $\mathcal{L}$, namely the set $Tm_\Omega(var)$; given this convention $\vdash \subseteq \wp(\mathcal{L}) \times \mathcal{L}$. Moreover, we write $\Sigma \vdash \Gamma$ if for all $\varphi \in \Gamma$, $\Sigma \vdash \varphi$. It is also customary to use $\Sigma; \Delta$ for $\Sigma \cup \Delta$ and $\Sigma; \varphi$ instead of $\Sigma \cup \{\varphi\}$. This notation saves brackets and is almost exclusively used instead of the proper set notation.

(ext.)	If $\varphi \in \Sigma$ then $\Sigma \vdash \varphi$.
(mon.)	If $\Sigma \subseteq \Delta$ then $\Sigma \vdash \varphi$ implies $\Delta \vdash \varphi$.
(trs.)	If $\Sigma \vdash \Gamma$ and $\Gamma \vdash \varphi$ then $\Sigma \vdash \varphi$.

(Observe that (mon.) is derivable from (ext.) and (trs.).) For suppose that $\varphi \in \Sigma$. Then if all $\chi \in \Sigma$ are true, then φ is true as well. Thus (ext.) holds. Furthermore, if $\Sigma \vdash \varphi$ and $\Sigma \subseteq \Delta$ and if all $\chi \in \Delta$ are true, then all terms of Σ are true and so φ is true as well; this shows (mon.). The third rule is proved thus. If $\Sigma \vdash \Gamma$ and $\Delta \vdash \varphi$ and all terms of Σ are true then all formulae of Γ are true by the first assumption and so φ is true by the second.

In addition, there are two other postulates that do not follow directly from our intuitions about truth–preservation.

(sub.)	If $\Sigma \vdash \varphi$ and σ is a substitution then $\Sigma^\sigma \vdash \varphi^\sigma$.
(cmp.)	$\Sigma \vdash \varphi$ iff there exists a finite $\Sigma_0 \subseteq \Sigma$ such that $\Sigma_0 \vdash \varphi$.

The postulate (sub.) reflects our understanding of the notion of a *variable*. A variable is seen here as a name of an arbitrary (concrete) proposition and thus we may plug in all concrete things over which the variables range. Then the relation $\Sigma \vdash \varphi$ says that for any *concrete* instances of the occurring variables, the concretization of $\Sigma \vdash \varphi$ is valid. So, the rule $p \wedge q \vdash q \wedge p$ — being valid — should remain valid under all concretizations. For example, we should have *Aristotle was a philosopher and Socrates* *was a carpenter* $\vdash$ *Socrates was a carpenter and Aristotle was a philosopher*. Suppose now that we have a substitution σ. Then — contrary to what one might think — φ^σ is not a concretization of φ; however, *every concretization of* φ^σ *is a concretization of* φ. One may therefore think of σ as a *sharpening*, since given φ it returns a more specific proposition φ^σ. Thus, $\Sigma^\sigma \vdash \varphi^\sigma$ simply is a statement over a subset of concretizations of $\Sigma \vdash \varphi$. Since it holds for all of them $\Sigma^\sigma \vdash \varphi^\sigma$ holds as well. The algebraic argument is simple. Namely, a concretization is, formally speaking, just a homomorphism $\beta : \mathfrak{Tm}_\Omega(X) \to \mathfrak{A}$. A concretization of φ^σ is just the value $\beta(\sigma(\varphi))$ under a homomorphism. But $\beta(\varphi^\sigma) = \beta \circ \sigma(\varphi)$, and since $\beta \circ \sigma : \mathfrak{Tm}_\Omega(X) \to \mathfrak{A}$, it is a concretization as well. Hence any concretization of φ^σ is in fact a concretization of φ.

The last postulate (cmp.) is called **compactness**, a term borrowed from topology. Another term is **finitary**. (cmp.) is not at all justifiable from our notions of truth. In fact, it fails e. g. in logics with infinitary operations. We have $\{\varphi_i : i \in \omega\} \vdash \bigwedge\langle\varphi_i : i \in \omega\rangle$ even though no finite set $\Gamma \subseteq \{\varphi_i : i \in \omega\}$ exists such that $\Gamma \vdash \bigwedge\langle\varphi_i : i \in \omega\rangle$. One might expect that if we only have finitary operations, the truth of any term if entailed by Σ will already be entailed by some finite subset of Σ. But this is known to be false as well (see [**232**]). Indeed, the most plausible justification comes from the interpretation via deduction. For suppose we read $\Sigma \vdash \varphi$ as *there is a proof of* φ *from* Σ. Then, since such a proof is a finite object, we can use only finitely many terms of Σ in it.

Definition 1.4.1. *Let* $\mathcal{L}$ *be a propositional language and* $\vdash \subseteq \wp(\mathcal{L}) \times \mathcal{L}$. $\vdash$ *is called a* ***finitary consequence relation*** *over* $\mathcal{L}$ *if it satisfies (ext.), (mon.), (trs.), (sub.) and (cmp.). A pair* $\langle \mathcal{L}, \vdash \rangle$ *where* $\mathcal{L}$ *is a propositional language and* $\vdash$ *a finitary consequence relation over* $\mathcal{L}$ *is called a* ***(propositional) logic****.*

Notice the following. In contrast to the standard literature, we require both (sub.) and (cmp.), since we want to deal almost exclusively with such logics. If (cmp.) is not assumed, we explicitly say so. Hence, in sequel, by a *consequence relation* we actually understand a *finitary consequence relation*. Moreover, when there is no risk of confusion we will speak of the **logic** $\vdash$; this is justified especially when it is clear over which language $\vdash$ is defined. A general reference for consequence relations is [**232**]. A logic $\vdash$ is called **inconsistent** if $\vdash = \wp(Tm_\Omega(var)) \times Tm_\Omega(var)$. Alternatively, $\vdash$ is inconsistent if $\vdash p$ for some p. It follows that in an inconsistent logic $\Gamma \vdash \varphi$ for all Γ and φ. Likewise, a set Γ (a formula φ) is **consistent** if there is a formula χ such that $\Gamma \nvdash \chi$ ($\varphi \nvdash \chi$). The terms φ satisfying $\varnothing \vdash \varphi$ are called the **tautologies** of the logic. We denote the set of tautologies of $\vdash$ by $\mathsf{Taut}(\vdash)$.

A **rule** is a pair $\rho = \langle \Sigma, \varphi \rangle$ where Σ is a finite set. Σ is called the **set of premises of** ρ and φ the **conclusion**. If $\sharp\Sigma = n$, ρ is called an n–**ary** rule. Furthermore, ρ is a **proper** rule if $n > 0$, and an **axiom** otherwise. ρ is a **derived rule** of $\vdash$ if $\rho \in \vdash$.

Definition 1.4.2. *Let $\mathfrak{L}$ be a language and R a set of rules over $\mathfrak{L}$. Then by $\vdash^R$ we denote the least consequence relation $\vdash$ such that $R \subseteq \vdash$. R is a* **rule base** *of $\vdash$ if $\vdash = \vdash^R$.*

To characterize $\vdash^R$ define an **instance** of $\rho = \langle \Sigma, \varphi \rangle$ to be any pair $\langle \Delta, \psi \rangle$ such that there is a substitution σ with $\varphi^\sigma = \psi$ and $\Sigma^\sigma = \Delta$. Every rule justifies all instances of itself, by (sub.). Suppose that a set of rules R is given. Then $\Sigma \vdash^R \varphi$ iff there is a *proof tree* deriving φ from Σ. A **proof tree** of φ from Σ is a tree $\langle T, < \rangle$ (where the root is the largest element with respect to $<$ and $<$ is completely intransitive) labelled with formulae from $\mathfrak{L}$ in such a way that (i) the root has label φ, (ii) the leaves have labels from Σ and (iii) if $\{y_i : i < n\}$ is the set of elements $> x$, y_i has label φ_i and x has label ψ, then $\langle \{\varphi_i : i < n\}, \psi \rangle$ is an instance of some $\rho \in R$.

An alternative characterization is via sequences. Given R we define now an R–**proof from Σ to** φ to be any finite sequence $\langle \varphi_i : i \leq \lambda \rangle$ such that $\varphi_\lambda = \varphi$ and for any $\kappa \leq \lambda$ either (i) $\varphi_\kappa \in \Sigma$ or (ii) for some set Δ such that $\Delta \subseteq \{\varphi_\mu : \mu < \kappa\}$, $\langle \Delta, \varphi_\kappa \rangle$ is an instance of a rule in R. Now set $\Sigma \vdash_R \varphi$ iff there exists an R–proof from Σ to φ. The reader may verify that if we have a proof tree, there exists an enumeration of the nodes of a tree such that the corresponding labels, if written down according to that enumeration, form an R–proof; and that if we have an R–proof of φ from Σ, we can define a proof tree for φ from Σ. To see the correctness of this characterization of $\vdash^R$, we prove the following theorem.

Theorem 1.4.3. *Let R be a set of rules and put $\Sigma \Vdash \varphi$ iff there exists an R–proof of φ from Σ. Then $\Vdash = \vdash^R$.*

Proof. It is not hard to see that $\Vdash \subseteq \vdash^R$. To show that the two are equal it suffices to establish that $\Vdash$ as defined is a consequence relation. (ext.) If $\varphi \in \Sigma$ then the sequence consisting of φ alone is a R–proof of φ from Σ. (mon.) If $\vec{\alpha}$ is a R–proof of φ from Σ then it is also an R–proof of φ from any $\Delta \supseteq \Sigma$. (sub.) If $\vec{\alpha}$ is an R–proof of φ from Σ and σ is a substitution then $\vec{\alpha}^\sigma$ is an R–proof of φ^σ from Σ^σ. (cmp.) If $\vec{\alpha}$ is an R–proof of φ from Σ then let Σ_0 be the set of terms occurring both in Σ and $\vec{\alpha}$. Since $\vec{\alpha}$ is finite, so is Σ_0. Moreover, $\vec{\alpha}$ is a R–proof of φ from Σ_0. (trs.) Suppose that $\Sigma \Vdash \Gamma$ and that $\Gamma \Vdash \varphi$. By (cmp.) we can assume that Γ is finite, so without loss of generality let $\Gamma = \{\gamma_i : i < n\}$. Let $\vec{\alpha}$ be an R–proof of φ from Γ and $\vec{\beta}_i$ an R–proof of ψ_i from Σ, $i < n$. Now let

$$\vec{\zeta} := \vec{\beta}_0{}^\frown\vec{\beta}_1{}^\frown \ldots {}^\frown\vec{\beta}_{n-1}{}^\frown\vec{\alpha}$$

It is straightforward to check that $\vec{\zeta}$ is a proof of φ from Σ. ☐

Let R be a set of rules. We say that a rule ρ can be derived from R if $\rho \in \vdash^R$. Alternatively, $\rho = \langle \Delta, \varphi \rangle$ then $\rho \in \vdash^R$ if there exists an R–proof of φ from Δ. If ρ is a rule, let $\vdash^{+\rho}$ denote the least consequence relation containing both $\vdash$ and ρ. For a consistent $\vdash$ put

$$E(\vdash) := \{n : \text{ there is an } n\text{–ary rule } \rho \notin \vdash \text{ such that } \nvdash^{+\rho} p\}$$

$E(\vdash)$ contains the arity of all rules ρ such that $\vdash^{+\rho}$ is a proper, consistent extension of $\vdash$. We call $E(\vdash)$ the **extender set** of $\vdash$. Obviously, a consequence relation is maximal among the consistent consequences relations iff its extender set is empty.

Let Σ be a set of formulae. A rule $\langle\Delta, \varphi\rangle$ is called **admissible** for Σ if Σ is closed under the application of ρ. That is, if for some substitution σ, $\sigma[\Delta] \subseteq \Sigma$, then $\sigma(\varphi) \in \Sigma$. ρ is **admissible** for $\vdash$ if ρ is admissible for $\mathsf{Taut}(\vdash)$. Equivalently, ρ is admissible for $\vdash$ if for all substitutions σ, φ^σ is a tautology, that is, $\varnothing \vdash_R \varphi^\sigma$, whenever all members of Δ^σ are tautologies. $\vdash$ is called **structurally complete** if every admissible rule of $\vdash$ is derivable. $\vdash$ is called **Post–complete** if $0 \notin E(\vdash)$.

Proposition 1.4.4 (Tokarz). *(1) A consequence relation $\vdash$ is structurally complete iff $E(\vdash) \subseteq \{0\}$. (2) A consequence relation is maximally consistent iff it is both structurally complete and Post–complete.*

Proof. (2) follows immediately from (1). So, we show only (1). Let $\vdash$ be structurally complete. Let ρ be a rule such that $\vdash^{+\rho}$ is a proper and consistent extension of $\vdash$. Since $\vdash$ is structurally complete, the tautologies are closed under ρ. It follows that ρ must be a 0–ary rule. For the other direction assume that $\vdash$ is structurally incomplete. Then there exists some ρ which is admissible but not derivable. ρ is not an axiom. Hence $\vdash^{+\rho}$ properly extends $\vdash$. Since $\vdash$ is consistent, p is not a tautology of $\vdash$. Since ρ is admissible, p is not a tautology of $\vdash^{+\rho}$, and so the latter is also consistent. ∎

It will be proved later that 2–valued logics with $\top$, $\neg$ and $\wedge$ is both structurally complete and Post–complete. Now, given $\vdash$, let

$$T(\vdash) := \{\vdash' : \mathsf{Taut}(\vdash') = \mathsf{Taut}(\vdash)\}$$

Then $T(\vdash)$ is an interval with respect to set inclusion. Namely, the least element is the least consequence containing all tautologies of $\vdash$. The largest is the consequence $\vdash^{+R}$, where R is the set of all rules admissible for $\vdash$. As we will see in the context of modal logic, the cardinality of $T(\vdash)$ can be very large (up to $2^{\aleph_0}$ for countable languages).

Another characterization of logics is via *consequence operations* or via *theories*. This goes as follows. Let $\langle\mathcal{L}, \vdash\rangle$ be a logic. Write $\Sigma^\vdash = \{\varphi : \Sigma \vdash \varphi\}$. The map $\Sigma \mapsto \Sigma^\vdash$ is an operation satisfying the following postulates.

(ext.)	$\Sigma \subseteq \Sigma^\vdash$.
(mon.)	$\Sigma \subseteq \Delta$ implies $\Sigma^\vdash \subseteq \Delta^\vdash$.
(trs.)	$\Sigma^{\vdash\vdash} \subseteq \Sigma^\vdash$.
(sub.)	$\sigma[\Sigma^\vdash] \subseteq (\sigma[\Sigma])^\vdash$ for every substitution σ.
(cmp.)	$\Sigma^\vdash = \bigcup\langle\Sigma_0^\vdash : \Sigma_0 \subseteq \Sigma, \Sigma_0 \text{ finite}\rangle$.

Actually, given (cmp.), (mon.) can be derived. Thus the map $\Sigma \mapsto \Sigma^\vdash$ is a closure operator which is compact and satisfies (sub.). This correspondence is exact. Whenever an operation $Cn : \wp(\mathcal{L}) \to \wp(\mathcal{L})$ on sets of $\mathcal{L}$–terms satisfies these postulates, the relation $\Sigma \vdash_{Cn} \varphi$ defined by $\Sigma \vdash_{Cn} \varphi$ iff $\varphi \in Cn(\Sigma)$ is a consequence relation, that is to say, $\langle \mathcal{L}, \vdash_{Cn} \rangle$ is a logic. Moreover, two distinct consequence operations determine distinct consequence relations and distinct consequence relations give rise to distinct consequence operations.

Finally, call any set of the form $\Sigma^{\vdash}$ a **theory** of $\vdash$. By (trs.), theories are closed under consequence, that is, $\Sigma^{\vdash}$ not only contains all consequences of Σ but also all of its own consequences as well. We may therefore say that the theories of $\vdash$ coincide with the **deductively closed sets**. Given $\vdash$, the following can be said about $\vdash$–theories.

(top.) $\mathcal{L}$ is a $\vdash$–theory.
(int.) If $T_i, i \in I$, are $\vdash$–theories, so is $\bigcap\langle T_i : i \in I\rangle$.
(sub.) If T is a $\vdash$–theory, σ a substitution then
$\sigma^{-1}[T] = \{\varphi : \varphi^{\sigma} \in T\}$ is a $\vdash$–theory.
(cmp.) If T_i, $i \in I$, is an ascending chain of $\vdash$–theories then
$\bigcup\langle T_i : i \in I\rangle$ is a $\vdash$–theory.

Again, the correspondence is exact. Any collection $\mathfrak{T}$ of subsets of $\mathcal{L}$ satisfying (top.), (int.), (sub.) and (cmp.) defines a consequence operation $\Sigma \mapsto \Sigma^{\vdash}$ by $\Sigma^{\vdash} = \bigcap\langle T : T \in \mathfrak{T}, T \supseteq \Sigma\rangle$ which satisfies (ext.), (mon.), (trs.), (sub.) and (cmp.). The correspondence is biunique. Different consequence operations yield different sets of theories and different collections of theories yield different consequence operations. The theories of a logic form a lattice. This lattice is algebraic if the consequence relation is finitary. The converse does not hold; this has been shown by Burghard Herrmann and Frank Wolter [**103**].

Exercise 8. Give a detailed proof of Theorem 1.4.3.

Exercise 9. Let R be a set of axioms or 1–ary rules. Show that $\Delta \vdash^R \varphi$ iff there exists a $\delta \in \Delta$ such that $\delta \vdash^R \varphi$.

Exercise 10. Show that every consistent logic is contained in a Post–complete logic. *Hint.* You need Zorn's Lemma here. For readers unfamiliar with it, we will prove later Tukey's Lemma, which will give rise to a very short proof for finitary logics.

Exercise 11. Show that in 2–valued logic

$$\begin{array}{lll}
\varphi_1 \leftrightarrow \psi_1; \varphi_2 \leftrightarrow \psi_2 & \vdash & \varphi_1 \wedge \varphi_2 \leftrightarrow \psi_1 \wedge \psi_2 \\
\varphi_1 \leftrightarrow \psi_1; \varphi_2 \leftrightarrow \psi_2 & \vdash & \varphi_1 \vee \varphi_2 \leftrightarrow \psi_1 \vee \psi_2 \\
\varphi \leftrightarrow \psi & \vdash & \neg\varphi \leftrightarrow \neg\psi
\end{array}$$

Thus if $\varphi \equiv \psi$ is defined by $\vdash \varphi \leftrightarrow \psi$, then $\equiv$ is a congruence relation. What is the cardinality of a congruence class? *Hint.* We assume that we have $\aleph_0$ many propositional variables. Show that all congruence classes must have equal cardinality.

Exercise 12. Show that there is no term φ in $\top, \vee, \wedge$ such that $\varphi \vdash \neg p \vdash \varphi$ in classical logic. *Hint.* Show first that one can assume $var(\varphi) = \{p\}$.

Exercise 13. Show that if $\langle \mathcal{L}, \vdash \rangle$ is a logic, the map $\Sigma \mapsto \Sigma^{\vdash}$ is a finitary closure operator.

Exercise 14. Show that if $\Sigma \mapsto \Sigma^{\vdash}$ is a finitary closure operator, the system of deductively closed sets satisfies (top.), (int.), (sub.) and (cmp.).

Exercise 15. Show that if the system of $\vdash$–closed sets satisfies (top.), (int.), (sub.) and (cmp.), then $\langle \mathcal{L}, \vdash \rangle$ is a logic.

1.5. Completeness of Matrix Semantics

Fundamental for the study of algebraic logic is the notion of a logical matrix. While for algebraic purposes we need only an algebra in order to compute terms, *truth* is extraneous to the notion of an algebra. Boolean algebras as such are neutral with respect to the notion of truth, we must stipulate those elements which we consider as *true*. The link between 1 and *true* is conventionally laid. One must be aware, therefore, that this is *just a convention*. We might, for example, consider 0 rather than 1 as *true*, and it turns out, that the logic of the algebra $\langle 2, \wedge \rangle$ where 0 is considered *true* is the same as the logic of $\langle 2, \vee \rangle$ where 1 is considered *true*, if $\wedge$ is translated as $\vee$.

Definition 1.5.1. *An Ω–**matrix** for a signature Ω is a pair $\mathfrak{M} = \langle \mathfrak{A}, D \rangle$ where $\mathfrak{A}$ is an Ω–algebra and $D \subseteq A$ a subset. A is called the set of **truth values** and D the set of **designated truth values**. An **assignment** or a **valuation** into $\mathfrak{M}$ is a map v from the set of variables into A. We say that v makes φ **true** in $\mathfrak{M}$ if $\overline{v}(\varphi) \in D$; otherwise we say, it makes φ **false**.*

With respect to a matrix $\mathfrak{M}$ we can define a relation $\vdash_{\mathfrak{M}}$ by

$$\Delta \vdash_{\mathfrak{M}} \varphi \quad \Leftrightarrow \quad \text{for all assignments } v : \text{ If } \overline{v}[\Delta] \subseteq D \text{ then } \overline{v}(\varphi) \in D .$$

Given a class S of matrices (for the same signature) we define

$$\vdash_S \; := \bigcap \langle \vdash_{\mathfrak{M}} : \mathfrak{M} \in S \rangle .$$

Theorem 1.5.2. *Let Ω be a signature. For each class S of Ω–matrices, $\vdash_S$ is a (possibly nonfinitary) logic.*

Proof. We show this for a single matrix. The full theorem follows from the fact that the intersection of logics is a logic again. Let $\mathfrak{M} \in S$. (ext.) If $\varphi \in \Sigma$ and $\overline{v}[\Sigma] \subseteq D$ then $\overline{v}(\varphi) \in D$, by assumption. (mon.) Let $\Sigma \subseteq \Delta$ and $\Sigma \vdash_{\mathfrak{M}} \varphi$. Assume $\overline{v}[\Delta] \subseteq D$. Then $\overline{v}[\Sigma] \subseteq D$ as well, and so $\overline{v}(\varphi) \in D$, by assumption. (trs.) Let $\Sigma \vdash_{\mathfrak{M}} \Gamma$ and $\Gamma \vdash_{\mathfrak{M}} \varphi$. Assume $\overline{v}[\Sigma] \subseteq D$. Then we have $\overline{v}[\Gamma] \subseteq D$ and so $\overline{v}(\varphi) \in D$. (sub.) Assume $\Sigma \vdash_{\mathfrak{M}} \varphi$ and let σ be a substitution. Then $\overline{v} \circ \sigma$ is a homomorphism into the algebra underlying $\mathfrak{M}$ and $\overline{v} \circ \sigma[\Sigma] = \overline{v}[\Sigma^\sigma]$. Hence if $\overline{v}[\Sigma^\sigma] \subseteq D$ we also have $\overline{v}(\varphi^\sigma) \in D$, as required. □

Theorem 1.5.3 (Wójcicki). *For each logic $\langle \mathfrak{L}, \vdash \rangle$ there exists a class S of matrices such that $\vdash = \vdash_S$.*

Proof. Given the language, let S consist of all $\langle \mathfrak{Tm}_\Omega(var), T\rangle$ where T is a theory of $\vdash$. First we show that for each such matrix $\mathfrak{M}$, $\vdash \subseteq \vdash_{\mathfrak{M}}$. To that end, assume $\Sigma \vdash \varphi$ and that $\overline{v}[\Sigma] \subseteq T$. Now $\overline{v}$ is in fact a substitution, and T is deductively closed, and so $\overline{v}(\varphi) \in T$ as well, as required. Now assume $\Sigma \nvdash \varphi$. We have to find a single matrix $\mathfrak{M}$ of this form such that $\Sigma \nvdash_{\mathfrak{M}} \varphi$. For example, $\mathfrak{M} := \langle \mathfrak{Tm}_\Omega(var), \Sigma^\vdash\rangle$. Then with $\overline{v}$ the identity map, $\overline{v}[\Sigma] = \Sigma \subseteq \Sigma^\vdash$. However, $\overline{v}(\varphi) = \varphi \notin \Sigma^\vdash$ by definition of $\Sigma^\vdash$ and the fact that $\Sigma \nvdash \varphi$. □

We add the remark that if $\mathfrak{M}$ is a matrix for $\vdash$, then the set of truth values must be closed under the rules. The previous theorem can be refined somewhat. Let $\mathfrak{M} = \langle \mathfrak{A}, D\rangle$ be a logical matrix, and Θ a congruence on $\mathfrak{A}$. Θ is called a **matrix congruence** if D is a union of Θ–blocks, that is, if $x \in D$ then $[x]\Theta \subseteq D$, and likewise, if $x \notin D$ then $[x]\Theta \cap D = \varnothing$. Then we can reduce the whole matrix by Θ and define $\mathfrak{M}/\Theta := \langle \mathfrak{A}/\Theta, D/\Theta\rangle$.

Lemma 1.5.4. *Let $\mathfrak{M}$ be a matrix and Θ be a matrix congruence of $\mathfrak{M}$. Then $\vdash_{\mathfrak{M}} = \vdash_{\mathfrak{M}/\Theta}$.*

Proof. Let $\Sigma \vdash_{\mathfrak{M}} \varphi$. Let $v : var \to A/\Theta$ be a valuation such that $\overline{v}[\Sigma] \subseteq [D]\Theta$. Then let $w : var \to A$ be defined by taking $w(p_i) \in v(p_i)$. (Recall that $v(p_i)$ is a union of Θ–blocks.) By assumption, $[w[\Sigma]]\Theta \subseteq [D]\Theta$, since $[D]\Theta$ is a union of blocks, and Θ is a congruence. Hence $w(\varphi) \in D$, by which $\overline{v}(\varphi) = [\overline{w}(\varphi)]\Theta \in [D]\Theta$, as required. Now assume $\Sigma \vdash_{\mathfrak{M}/\Theta} \varphi$. Let w be a valuation such that $\overline{w}[\Sigma] \subseteq D$. Then define v to be the composition of $\overline{w}$ with the natural surjection $x \mapsto [x]\Theta$. Then $v[\Sigma] \subseteq [D]\Theta$. By assumption, $v(\varphi) \in [D]\Theta$, so that $v(\varphi) = [x]\Theta$ for some $x \in D$. Consider $\overline{w}(\varphi)$. We know that $v(\varphi) = [\overline{w}(\varphi)]\Theta = [x]\Theta$. Thus $\overline{w}(\varphi) = y$ for some $y \,\Theta\, x$. Since D consists of entire Θ–blocks, $y \in D$, as required. □

Call a matrix **reduced** if the diagonal, that is the relation $\Delta = \{\langle x, x\rangle : x \in A\}$, is the only matrix congruence. It follows that we can sharpen the Theorem 1.5.3 to the following

Theorem 1.5.5. *For each logic $\langle \mathfrak{L}, \vdash \rangle$ there exists a class S of reduced matrices such that $\vdash = \vdash_S$.*

Let $\mathcal{S}$ be a class of Ω–matrices. $\mathcal{S}$ is called a **unital semantics** for $\vdash$ if $\vdash = \vdash_{\mathcal{S}}$ and for all $\langle \mathfrak{A}, D\rangle \in \mathcal{S}$ we have $\sharp D \leq 1$. (See Janusz Czelakowski [**49, 50**]. A unital semantics is often called *algebraic*. This, however, is different from the notion of 'algebraic' discussed in Wim Blok and Don Pigozzi [**29**].) The following is a useful fact, which is not hard to verify.

PROPOSITION 1.5.6. *Let $\vdash$ have a unital semantics. Then in $\vdash$ the rules $p; q; \varphi(p) \vdash \varphi(q)$ are valid for all formulae φ.*

Notice that when a logic over a language $\mathcal{L}$ is given and an algebra $\mathfrak{A}$ with appropriate signature, the set of designated truth values must always be a deductively closed set, otherwise the resulting matrix is not a matrix for the logic. A theory is **consistent** if it is not the entire language, and **maximally consistent** if it is maximal in the set of consistent theories. One can show that each consistent theory is contained in a maximally consistent theory. A direct proof for modal logics will be given below. It is interesting to note that for classical logics the construction in the proof of Theorem 1.5.3 can be strengthened by taking as matrices in S those containing only maximally consistent theories. For if $\Sigma \nvdash \varphi$ then $\Sigma; \neg\varphi$ is consistent and so for some maximally consistent Δ containing Σ we have $\neg\varphi \in \Delta$. Taking v to be the identity, $\overline{v}[\Sigma] = \Sigma \subseteq \Delta$, but $\overline{v}(\varphi) \notin \Delta$, otherwise Δ is not consistent. Furthermore, there is a special matrix, $\mathfrak{Taut} = \langle \mathfrak{Tm}_\Omega(var), \varnothing^\vdash \rangle$. Recall that $\varnothing^\vdash$ are simply the tautologies of a logic.

THEOREM 1.5.7 (Wójcicki). *$\vdash$ is structurally complete iff $\vdash = \vdash_{\mathfrak{Taut}}$.*

Notes on this section. The concepts of logical consequence and logical matrix are said to date back to the work of JAN ŁUKASIEWICZ and ALFRED TARSKI [**141**]. Many results of this section are claimed to have been folklore in the 1930ies. Theorem 1.5.3 is due to RYSZARD WÓJCICKI [**229**]. The converse of the implication in Proposition 1.5.6 also holds on condition that the logic has tautologies. This is proved in [**50**], where it is attributed to unpublished work by ROMAN SUSZKO. The notion of structural completeness has been introduced by W. A. POGORZELSKI [**162**] who proved also that classical logic is structurally complete. For general reference on consequence relations and algebraic semantics see RYSZARD WÓJCICKI [**232**].

Exercise 16. Prove Proposition 1.5.7.

Exercise 17. Characterize $\vdash_{\mathfrak{M}}$ where $\mathfrak{M} = \langle \mathfrak{A}, D \rangle$, where $D = \varnothing$ or $D = A$.

1.6. Properties of Logics

Logics can have a number of highly desirable properties which one should always establish first whenever possible. The first is *decidability*. A logic $\langle \mathcal{L}, \vdash \rangle$ is said to be **decidable** if for all finite Σ and all terms φ we can decide whether or not $\Sigma \vdash \varphi$. In other words, we must have a procedure or an algorithm that yields a (correct) answer to the problem '$\Sigma \vdash \varphi$'. (To be exact, we should speak of the problem '$\Sigma \vdash \varphi$?', but the question mark generally will be omitted.) This comprises two things (i) the algorithm terminates, that is, does not run forever and (ii) the answer given is correct. Predicate logic is not decidable; the reason is that its expressive power is too strong. Many propositional logics, on the other hand, are decidable. 2–valued logics are an example. This follows from the next theorem, first shown by R. HARROP in [**100**].

Theorem 1.6.1 (Harrop). *Suppose that $\mathfrak{M} = \langle \mathfrak{A}, D\rangle$ is a finite logical matrix. Then $\langle \mathcal{L}, \vdash_{\mathfrak{M}}\rangle$ is decidable.*

Proof. Basically, the decision algorithm is that of building 'truth tables'. Given a finite Σ and a term φ, there exist only finitely many valuations $v : var(\Sigma) \cup var(\varphi) \to T$. It is a finite problem to compute all the values $\overline{v}(\psi)$ for $\psi \in \Sigma$ to check whether $\overline{v}[\Sigma] \subseteq D$ and then to see whether $\overline{v}(\varphi) \in D$ as well. ◻

This procedure is generally slower than tableaux–methods, however only mildly so (see [**51**]). Tableaux–methods allow for a guided search for falsifying assignments which in many cases (and certainly in many applications) reduces the search space rather drastically. However, the truth–table method is in certain cases also rather efficient. There is namely a certain tradeoff between the length of a formula and the number of variables it contains. The length of a computation for a given formula φ depends exponentially on the number of variables (so this is indeed expensive), but only quadratically on the length of φ. (See Section 1.8.) For if φ has n variables, and $\mathfrak{M}$ has k elements, then k^n assignments need to be checked. Given a particular assignment, the truth value of φ with respect to that assignment can be checked simply by induction on the constitution of φ. If φ contains ℓ many symbols different from variables, then ℓ many steps need to be performed. Each step takes time propotional to the length of φ. In total we get a bound on the time of $c \cdot \ell \cdot |\varphi| \cdot k^n$.

Secondly, we investigate the notion of implication. Of particular importance in logic are the *modus ponens* and the *deduction theorem*. To explain them, assume that we have a binary termfunction $\twoheadrightarrow (p, q)$, written $p \twoheadrightarrow q$. The rule of **modus ponens** for $\twoheadrightarrow$ — (mp$\twoheadrightarrow$.) for short — is the rule $\langle\{p, p \twoheadrightarrow q\}, q\rangle$. There are many connectives which fulfill modus ponens, for example $\wedge$ and $\to$. We write (mp.) for (mp$\to$.). $\twoheadrightarrow$ is said to satisfy the **deduction theorem** with respect to $\vdash$ if for all Σ, φ, ψ

$$(\dagger) \qquad \Sigma; \varphi \vdash \psi \quad \Leftrightarrow \quad \Sigma \vdash \varphi \twoheadrightarrow \psi .$$

A logic $\langle \mathcal{L}, \vdash\rangle$ is said to **admit a deduction theorem** if there exists a term $p \twoheadrightarrow q$ such that ($\dagger$) holds. Given the deduction theorem it is possible to transform any rule different from (mp.) into an axiom preserving the consequence relation. (To be precise, we can also rewrite (mp.) into an axiom, but we are not allowed to replace it by that axiom, while with any other rule this is possible in presence of the deduction theorem and (mp.).) For example, the rule $p; q \vdash p \wedge q$ can be transformed into $\vdash p \twoheadrightarrow (q \twoheadrightarrow (p \wedge q))$. Hence it is possible to replace the original rule calculus by a calculus where modus ponens is the only rule which is not an axiom. In such calculi, which are called **mp–calculi** or also **Hilbert–style calculi** for $\twoheadrightarrow$, validity of the deduction theorem is equivalent to the validity of certain rules.

Theorem 1.6.2. *An mp–calculus for $\twoheadrightarrow$, $\langle \mathcal{L}, \vdash\rangle$, has a deduction theorem for $\twoheadrightarrow$ iff $\twoheadrightarrow$ satisfies modus ponens and the following are axioms of $\vdash$:*

(wk.) $p \twoheadrightarrow (q \twoheadrightarrow p)$,

(fd.) $(p \twoheadrightarrow (q \twoheadrightarrow r)) \twoheadrightarrow ((p \twoheadrightarrow q) \twoheadrightarrow (p \twoheadrightarrow r))$.

Remark. (wk.) is the axiom of *weakening* and (fd.) is known as *Freges Dreierschluß*, named after Gottlob Frege. There is also a rather convenient notation of formulae using dots; it is used to save brackets. We write $\varphi. \twoheadrightarrow .\psi$ for $(\varphi) \twoheadrightarrow (\psi)$. For example, (fd.) can now be written as

$$(p. \twoheadrightarrow .q \twoheadrightarrow r) \twoheadrightarrow (p \twoheadrightarrow q. \twoheadrightarrow .p \twoheadrightarrow r).$$

Now we prove Theorem 1.6.2.

Proof. ($\Rightarrow$) Suppose both modus ponens and (†) hold for $\twoheadrightarrow$. Now since $\varphi \vdash \varphi$, also $\varphi; \psi \vdash \varphi$ and (by (†)) also $\varphi \vdash \psi \twoheadrightarrow \varphi$ and (again by (†)) $\vdash \varphi \twoheadrightarrow (\psi \twoheadrightarrow \varphi)$. For (fd.) note that the following sequence

$$\langle \varphi. \twoheadrightarrow .\psi \twoheadrightarrow \chi, \varphi \twoheadrightarrow \psi, \varphi, \psi \twoheadrightarrow \chi, \psi, \chi \rangle$$

proves $\psi. \twoheadrightarrow .\psi \twoheadrightarrow \chi; \varphi \twoheadrightarrow \psi; \varphi \vdash \chi$. Apply (†) three times and (fd.) is proved. ($\Leftarrow$) By induction on the length of an R–proof $\vec{\alpha}$ of ψ from $\Sigma \cup \{\varphi\}$ we show that $\Sigma \vdash \varphi \twoheadrightarrow \psi$. Suppose the length of $\vec{\alpha}$ is 1. Then $\psi \in \Sigma \cup \{\varphi\}$. There are two cases: (1) $\psi \in \Sigma$. Then observe that $\langle \psi \twoheadrightarrow (\varphi \twoheadrightarrow \psi), \psi, \varphi \twoheadrightarrow \psi \rangle$ is a proof of $\varphi \twoheadrightarrow \psi$ from Σ. (2) $\psi = \varphi$. Then we have to show that $\Sigma \vdash \varphi \twoheadrightarrow \varphi$. Now observe that the following is an instance of (fd.): $(\varphi. \twoheadrightarrow .(\psi \twoheadrightarrow \varphi) \twoheadrightarrow \varphi) \twoheadrightarrow (\varphi \twoheadrightarrow (\psi \twoheadrightarrow \varphi). \twoheadrightarrow .\varphi \twoheadrightarrow \varphi)$. But $\varphi. \twoheadrightarrow .(\psi \twoheadrightarrow \varphi) \twoheadrightarrow \varphi$ and $\varphi. \twoheadrightarrow .\psi \twoheadrightarrow \varphi$ are both instances of (wk.) and by applying modus ponens two times we prove $\varphi \twoheadrightarrow \varphi$. Now let $\vec{\alpha}$ be of length > 1. Then we may assume that ψ is obtained by an application of modus ponens from some formulae χ and $\chi \twoheadrightarrow \psi$. Thus the proof looks as follows:

$$\ldots, \chi, \ldots, \chi \twoheadrightarrow \psi, \ldots, \psi, \ldots$$

Now by induction hypothesis $\Sigma \vdash \varphi \twoheadrightarrow \chi$ and $\Sigma \vdash \varphi. \twoheadrightarrow .(\chi \twoheadrightarrow \psi)$. Then, as $\varphi \twoheadrightarrow (\chi \twoheadrightarrow \psi). \twoheadrightarrow .(\varphi \twoheadrightarrow \chi. \twoheadrightarrow .\varphi \twoheadrightarrow \psi)$ is a theorem we get that $\Sigma \vdash \varphi \twoheadrightarrow \psi$ with two applications of modus ponens. ∎

The significance of the deduction theorem lies among other in the fact that for a given set Σ there exists at most one consequence relation $\vdash$ with a deduction theorem such that Σ is the set of tautologies of $\vdash$. For assume $\Delta \vdash \varphi$ for a set $\Delta \subseteq \mathcal{L}$. Then by compactness there exists a finite set $\Delta_0 \subseteq \Delta$ such that $\Delta_0 \vdash \varphi$. Let $\Delta_0 := \{\delta_i : i < n\}$. Put

$$ded(\Delta_0, \varphi) := \delta_0 \twoheadrightarrow (\delta_1 \twoheadrightarrow \ldots (\delta_{n-1} \twoheadrightarrow \varphi) \ldots)$$

Then, by the deduction theorem for $\twoheadrightarrow$

$$\Delta \vdash \varphi \qquad \Leftrightarrow \qquad \varnothing \vdash ded(\Delta, \varphi)$$

Theorem 1.6.3. *Let $\vdash$ and $\vdash'$ be consequence relations with $\mathsf{Taut}(\vdash) = \mathsf{Taut}(\vdash')$. Suppose that there exists a binary termfunction $\twoheadrightarrow$ such that $\vdash$ and $\vdash'$ satisfy a deduction theorem for $\twoheadrightarrow$. Then $\vdash = \vdash'$.*

$\vdash$ has **interpolation** if whenever $\varphi \vdash \psi$ there exists a formula χ with $var(\chi) \subseteq var(\varphi) \cap var(\psi)$ such that both $\varphi \vdash \chi$ and $\chi \vdash \psi$. Interpolation is a rather strong property, and generally logics fail to have interpolation. There is a rather simple theorem which allows to prove interpolation for logics based on a finite matrix. Say that $\vdash$ has a **conjunction** if there is a term $p \wedge q$ such that the following are derivable rules: $\langle \{p, q\}, p \wedge q\rangle$ and both $\langle \{p \wedge q\}, p\rangle$ and $\langle \{p \wedge q\}, q\rangle$. In addition, if $\vdash\ =\ \vdash_{\mathfrak{M}}$ for some logical matrix we say that $\vdash$ has **all constants** if for each $\mathsf{t} \in T$ there exists a nullary term–function $\underline{\mathsf{t}}$ such that for all valuations v $\overline{v}(\underline{\mathsf{t}}) = \mathsf{t}$. (Note that since $var(\underline{\mathsf{t}}) = \varnothing$ the value of $\underline{\mathsf{t}}$ does not depend at all on v.) This rather complicated definition allows that we do not need to have a constant for each truth–value; it is enough if they are definable from the others. For example in classical logic we may have only $\top = \underline{1}$ as a primitive and then $\underline{0} = \neg\top$. Notice also the following. Say that an algebra is **functionally complete** if every function $A^n \to A$ is a term–function of $\mathfrak{A}$; and say that $\mathfrak{A}$ is **polynomially complete** if every function $A^n \to A$ is a polynomial function. Then any functionally complete algebra is polynomially complete; the converse need not hold. However, if $\mathfrak{A}$ has all constants, then it is functionally complete iff it is polynomially complete.

Theorem 1.6.4. *Suppose that $\mathfrak{M}$ is a finite logical matrix. Suppose that $\vdash_{\mathfrak{M}}$ has a conjunction $\wedge$ and all constants; then $\vdash_{\mathfrak{M}}$ has interpolation.*

Proof. Suppose that $\varphi(\vec{p}, \vec{q}) \vdash \psi(\vec{q}, \vec{r})$, where $\vec{r} = \langle r_i : i < n\rangle$. Clearly, $var(\psi) \not\subseteq var(\varphi)$, iff $n > 0$. We show that if $n \neq 0$ there exists a

$$\psi^1 = \psi^1(\vec{q}, r_0, \dots r_{n-2})$$

such that $\varphi \vdash \psi^1 \vdash \psi$. The claim is then proved by induction on n. We put

$$\psi^1(\vec{q}, r_0, \dots, r_{n-2}) := \bigwedge \langle \psi(\vec{q}, r_0, \dots, r_{n-2}, \underline{\mathsf{t}}) : \mathsf{t} \in T\rangle$$

Now observe that $\varphi \vdash \psi$ implies

$$\varphi(\vec{p}, \vec{q}) \vdash \psi(\vec{q}, r_0, \dots, r_{n-2}, \underline{\mathsf{t}})$$

for every t. Now we apply the rule for conjunction for each $\mathsf{t} \in T$, and obtain

$$\varphi(\vec{p}, \vec{q}) \vdash \bigwedge \langle \psi(\vec{q}, r_0, \dots, r_{n-2}, \underline{\mathsf{t}}) : \mathsf{t} \in T\rangle \quad (= \psi^1)\ .$$

Furthermore, $\psi^1 \vdash \psi$, that is,

$$\bigwedge \langle \psi(\vec{q}, r_0, \dots, r_{n-2}, \underline{\mathsf{t}})\rangle \vdash \psi(\vec{q}, r_0, \dots, r_{n-1}) \quad .$$

For if $\overline{v}(\psi^1)$ is true then for any extension v^+ with $r_{n-1} \in dom(v^+)$ we have $\overline{v^+}(\psi) \in D$. ∎

The following is an instructive example showing that we cannot have interpolation without constants. Consider the logic of the 2–valued matrix in the language $\mathcal{C}' = \{\wedge, \neg\}$, with 1 the distinguished element. This logic fails to have interpolation. For it holds that $p \wedge \neg p \vdash q$, but there is no interpolant if $p \neq q$. This is surprising because the algebra $\langle 2, -, \cap\rangle$ is polynomially complete. Hence, polynomial completeness is not enough, we must have functional completeness. In other words, we must have all constants. Let us also add that the theorem fails for intersections of logics with all constants and conjunction.

A property closely related to interpolation is *Halldén–completeness*. It is named after Sören Halldén, who discussed it first in [**93**]. (See also [**191**].) A logic is called **Halldén–complete** if for all formulae φ and ψ with $var(\varphi) \cap var(\psi) = \varnothing$ we have that if $\varphi \vdash \psi$ and φ is consistent then $\vdash \psi$. 2–valued logics are Halldén–complete. Namely, assume that φ is consistent. Let $v : var(\psi) \to 2$ be any valuation. Since φ is consistent there exists a $u : var(\varphi) \to 2$ such that $\overline{u}(\varphi) = 1$. Put $w := u \cup v$. Since u and v have disjoint domains, this is well–defined. Then $\overline{w}(\varphi) = 1$, and so $\overline{w}(\psi) = 1$. So, $\overline{v}(\psi) = 1$. This shows that $\vdash \psi$. The following generalization is now evident.

Theorem 1.6.5 (Łos & Suszko). *Let $\mathfrak{M}$ be a logical matrix. Then $\vdash_{\mathfrak{M}}$ is Halldén–complete.*

Thus, failure of Halldén–completeness can only arise in logics which are not determined by a single matrix. In classical logic, the property of Halldén–completeness can be reformulated into a somewhat more familiar form. Namely, the property says that for φ and ψ disjoint in variables, if $\varphi \vee \psi$ is a tautology then either φ or ψ is a tautology.

Finally we turn to structural completeness. Recall that structural completeness means that all admissible rules are derivable. The converse is always valid.

Theorem 1.6.6. *Suppose that $\mathfrak{M}$ is a logical matrix and $\vdash_{\mathfrak{M}}$ has all constants. Then it is structurally complete and Post–complete.*

Proof. Suppose that $\rho = \langle\{\varphi_0, \varphi_1, \dots, \varphi_{n-1}\}, \psi\rangle$. Assume $n > 0$. We show that if ρ is not derivable it is not admissible. So, assume $\rho \notin \vdash_{\mathfrak{M}}$. Then there is a $v : var \to T$ such that $\overline{v}(\varphi_0), \overline{v}(\varphi_1), \dots, \overline{v}(\varphi_{n-1}) \in D$ while $\overline{v}(\psi) \notin D$. Consider the substitution $\sigma(p) := \underline{v(p)}$, where $\underline{v(p)}$ is the constant with value $v(p)$. Then φ_0^σ, $\varphi_1^\sigma, \dots, \varphi_{n-1}^\sigma$ are in D under all homomorphisms, so they are theorems. But ψ^σ is not in D for any homomorphism. So ρ is not admissible in $\vdash_{\mathfrak{M}}$. If $n = 0$, the addition of ψ as an axiom makes the logic inconsistent. For ψ^σ is constant, with value $\notin D$. Hence, $\psi^\sigma \vdash_{\mathfrak{M}} p$. So, adding ψ^σ yields an inconsistent logic. Therefore, adding ψ makes the logic inconsistent. ☐

Notes on this section. In [**140**] a consequence relation $\vdash$ is called *uniform* if for sets Γ and Δ and a single formula φ such that $var(\Delta) \cap var(\Gamma; \varphi) = \varnothing$ we have: if $\Gamma; \Delta \vdash \varphi$ then $\Gamma \vdash \varphi$. Obviously a uniform consequence relation is Halldén–complete. It is shown in that paper that a consequence relation is uniform iff it is of the form $\vdash_{\mathfrak{M}}$ for a single logical matrix. It was noted by Ryszard Wójcicki [**230**] that for consequence relations that need not be finitary this works only on the assumption that $\vdash$ is *regular*. Finitary consequence relations are always regular.

Exercise 18. Let S be a finite set of finite matrices. Show that $\vdash_S$ is decidable.

Exercise 19. Let $\mathfrak{M} = \langle \mathfrak{T}, D\rangle$ be a logical matrix. Show that a connective (= termfunction) $\twoheadrightarrow$ satisfies the rule of modus ponens for $\vdash_{\mathfrak{M}}$ if whenever $a \in D$ and $a \twoheadrightarrow^{\mathfrak{T}} b \in D$ then also $b \in D$; in other words, this is the truth table for $\twoheadrightarrow^{\mathfrak{T}}$:

$\twoheadrightarrow$	D	$\overline{D}$
D	$\star$	$\overline{D}$
$\overline{D}$	$\star$	$\star$

Read this as follows. D stands for any element of D, $\overline{D}$ for any element of $T - D$. $\star$ stands for an arbitrary element. How many binary connectives of classical logic satisfy (mp.)?

Exercise 20. Let $\mathfrak{M} = \langle \mathfrak{T}, D \rangle$ be a logical matrix. Show that $\twoheadrightarrow$ satisfies the deduction theorem if it has the truth table below.

$\twoheadrightarrow$	D	$\overline{D}$
D	D	$\overline{D}$
$\overline{D}$	D	D

Thus the above truth table requires only that $a \twoheadrightarrow^{\mathfrak{T}} b \notin D$ if $a \in D$ but $b \notin D$.

Exercise 21. Let $\mathfrak{M} = \langle \mathfrak{T}, D \rangle$ be a logical matrix. Show that $\wedge$ is a conjunction if it has the following truth table.

$\wedge$	D	$\overline{D}$
D	D	$\overline{D}$
$\overline{D}$	$\overline{D}$	$\overline{D}$

1.7. Boolean Logic

The result of the previous sections will now be applied to the most fundamental logic, namely boolean logic. This chapter may be skipped by all those readers who are acquainted with the theory of boolean algebras. The main purpose is to repeat whatever will be essential knowledge for the rest of this book. Before we begin, let us agree that we will use the term *boolean logic* to denote what otherwise may also be called *classical logic*. The reason for not using the latter is a clash in terminology, because there are also classical modal logics. To distinguish them from the traditional classical logic we call the latter boolean logic.

We distinguish between boolean logic and 2–valued logic, which is a logic whose semantics consists of matrices with at most 2 elements. A set of term functions is **complete** or a **basis** if the smallest clone of functions containing this set is the clone of *all* term–functions. Examples of bases are $\{1, -, \cap\}$, $\{\rightarrow, \bot\}$, and $\{\downarrow, \bot\}$, where $p \downarrow q := \neg(p \cap q)$. Our set of primitive symbols is $\{1, -, \cap\}$. This set is a basis, as is well–known. Notice that if we need only a polynomially complete set of basic functions, $\top$ is redundant. (However, notice that by the theorems and exercises of the previous section, in the language of $\neg$ and $\wedge$ 2–valued logic does not have interpolation.)

Definition 1.7.1. *A* ***boolean algebra*** *is an algebra* $\mathfrak{B} = \langle B, 1, -, \cap\rangle$ *of the signature* $\langle 0, 1, 2\rangle$ *such that* $\cap$ *is commutative, associative and idempotent, with neutral element* 1*, and* $-$ *is a map satisfying the following conditions with* $0 := -1$ *and* $x \cup y := -(-x \cap -y)$.

$$\begin{array}{lcl} --x & = & x \\ x \cap -x & = & 0 \\ (x \cap y) \cup (x \cap -y) & = & x \end{array}$$

Recall that in Section 1.1 we defined a boolean algebra as a bounded distributive lattice with a negation satisfying the de Morgan laws and the identities above. We need to verify that if we define 0, $\cup$ as above, then $\langle B, 0, 1, \cap, \cup, -\rangle$ satisfies the description of Section 1.1. The most difficult part is the proof of the distributivity law.

Lemma 1.7.2. *In a boolean algebra,* $\cup$ *is associative, commutative, idempotent, and* $0 \cup x = x$.

The proof of this fact is left as an exercise. Put $x \leq y$ iff $x \cap y = x$. It follows that $x = y$ iff $x \leq y$ and $y \leq x$. For if the latter holds then $x = x \cap y = y \cap x = y$. Then $x \to y = -x \cup y$. Moreover, $x \leftrightarrow y := (x \to y) \cap (y \to x)$. For the proof of the next theorem observe that

$$x \cap 0 = 0$$

For $x \cap 0 = x \cap (x \cap -x) = (x \cap x) \cap -x = x \cap -x = 0$.

Lemma 1.7.3. *In a boolean algebra, the following holds.*

(1) $x \leq y$ *iff* $x \cap -y = 0$.
(2) $x \leq y$ *iff* $x \to y = 1$.
(3) $x = y$ *iff* $x \leftrightarrow y = 1$.
(4) $x \leq y$ *iff* $-y \leq -x$.

Proof. (1.) If $x \leq y$ then $x \cap -y = (x \cap y) \cap (-y) = x \cap (y \cap -y) = x \cap 0 = 0$. If on the other hand $x \cap -y = 0$ then $x = (x \cap y) \cup (x \cap -y) = (x \cap y) \cup 0 = x \cap y$, by the previous theorem. Hence $x \leq y$. (2.) $x \cap -y = -(x \to y)$, so $x \leq y$ iff $x \cap -y = 0$ iff $x \to y = 1$. (3.) $x = y$ iff $x \leq y$ and $y \leq x$ iff $x \to y = 1$ and $y \to x = 1$ iff $x \leftrightarrow y = 1$. (4.) $x \leq y$ iff $x \cap -y = 0$ (by (1.)) iff $-y \cap --x = 0$ iff $-y \leq -x$ (again by (1.)). □

Lemma 1.7.4. *In a boolean algebra, the following holds.*

(1) $x \to (y \to z) = (x \cap y) \to z$.
(2) $x \cap y \leq z \quad \Leftrightarrow \quad x \leq y \to z$

The law (2.) is known as the ***law of residuation***.

Proof. (1.) This is also easily proved from the other laws. For $x \to (y \to z) = -(x \cap -(-(y \cap -z))) = -(x \cap (y \cap -z)) = -((x \cap y) \cap -z) = (x \cap y) \to z$. (2.) $x \cap y \leq z$ iff $(x \cap y) \to z = 1$ iff $x \to (y \to z) = 1$ (by (1.)) iff $x \leq y \to z$ by (2.) of Lemma 1.7.3. □

Lemma 1.7.5. *In a boolean algebra,* $x \cap (x \to y) \leq y$.

Proof. From $-y \cap x \leq x \cap -y$ we deduce $-y \leq x \to (x \cap -y)$, by residuation. So, $-(x \to -(x \to y)) \leq y$, which is $x \cap (x \to y) \leq y$. □

Proposition 1.7.6. *In a boolean algebra,* $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$.

Proof. We have $x \cap y \leq x$ since $x \cap y \cap x = x \cap y$, and likewise $x \cap z \leq x$. Furthermore, $x \cap y \leq y \leq y \cup z$. Hence $x \cap y \leq y \cup z$, and likewise $x \cap z \leq y \cup z$. This shows one inequality. For the other, observe that

$$\begin{array}{rcl} x \cap (y \cup z) \cap -(x \cap y) \cap -(x \cap z) & = & x \cap (-y \to z) \cap (x \to -y) \\ & & \qquad\qquad \cap (x \to -z) \\ & \leq & (-y \to z) \cap -y \cap -z \\ & \leq & z \cap -z \\ & = & 0 \end{array}$$

So, by (1.) of Lemma 1.7.3, $x \cap (y \cup z) \leq (x \cap y) \cup (x \cap z)$. □

An alternative characterization of a boolean algebra is the following. $\langle B, 0, 1, -, \cap, \cup\rangle$ is a boolean algebra iff its reduct to $\{0, 1, \cap, \cup\}$ is a bounded distributive lattice, and $-$ a function assigning to each $x \in A$ its complement. Here, an element y is a **complement** of x if $y \cap x = 0$ and $y \cup x = 1$. In a distributive lattice, y is uniquely defined if it exists. For let y_1 and y_2 be complements of x. Then $y_1 = (x \cup y_2) \cap y_1 = (x \cap y_1) \cup (y_2 \cap y_1) = y_2 \cap y_1$, and so $y_1 \leq y_2$. By the same argument, $y_2 \leq y_1$, and so $y_1 = y_2$. In addition, if y is the complement of x, x is the complement of y. The second definition is easily seen to be equivalent (modulo the basic operations) to the one of Section 1.1. Call y a **complement of** x **relative to** z if $x \cap y = 0$ and $x \cup y = z$. The law $(x \cap y) \cup (x \cap -y) = x$ is (in presence of the other laws) equivalent to the requirement that $x \cap -y$ is the complement of $x \cap y$ relative to x. Namely, $(x \cap y) \cap (x \cap -y) = x \cap (y \cap -y) = 0$.

Given boolean logic, what are the deductively closed sets in $\mathfrak{B}$? To answer this note that we have $x = y$ iff $x \leq y$ and $y \leq x$ iff $x \leftrightarrow y = 1$. Now if S is deductively closed, it must contain all tautologies and be closed under the rule (mp.). So if $x \in S$ and $x \to y \in S$ then $y \in S$. We deduce first of all that S is not empty; namely, $1 \in S$. (Recall that 1 is assumed to be the value of $\top$, which is in the language; if it is not, then at least we have $x \cup -x \in S$ for an arbitrary x if S is not empty.) Furthermore, if $x \in S$ and $x \leq y$ then $x \to y = 1 \in S$ and so $y \in S$. Finally, if $x, y \in S$ then also $x \cap y \in S$ since $x \to (y \to x \cap y) = 1$ and by applying the previous rule twice we get the desired result. The following definition summarizes this.

Definition 1.7.7. *A* ***filter*** *in a boolean algebra* $\mathfrak{A}$ *is a subset* F *of* A *satisfying the following.*

(fi1.) $1 \in F$.
(fi≤.) *If* $x \in F$ *and* $x \leq y$ *then* $y \in F$.
(fi∩.) *If* $x, y \in F$ *then* $x \cap y \in F$.

A filter is ***trivial*** *or* ***improper*** *if it is the entire carrier set. A maximal proper filter is called an* ***ultrafilter****.*

Proposition 1.7.8. *A subset of a boolean algebra is deductively closed iff it is a filter.*

Proof. The direction from left to right has been shown already. Now let F be a filter. Consider $x \in F$ and $x \to y \in F$. Then $x \cap (x \to y) \in F$. But $x \cap (x \to y) = x \cap (-x \cup y) = 0. \cup .x \cap y = x \cap y$. So, $x \cap y \in F$. Now $x \cap y \leq y$ and so $y \in F$ as well. ◻

We have characterized the deductively closed sets; these turn out to be also the sets which are congruence classes of the top element. Furthermore, if F is a filter, then the relation $x \sim_F y$ defined by $x \leftrightarrow y \in F$ is a congruence. First of all, it is reflexive, since $x \leftrightarrow x = 1 \in F$. Second, it is symmetric since $x \leftrightarrow y = y \leftrightarrow x$. And it is transitive, for if $x \sim_F y$ and $y \sim_F z$ then $x \leftrightarrow y \in F$ and $y \leftrightarrow z \in F$ and

$$\begin{aligned}(x \leftrightarrow y) \cap (y \leftrightarrow z) &= (x \cap y. \cup . -x \cap -y) \cap (y \cap z. \cup . -y \cap -z)\\ &= (x \cap y \cap z. \cup . -x \cap -y \cap -z)\\ &\leq (x \cap z. \cup . -x \cap -z)\\ &= x \leftrightarrow z\end{aligned}$$

and so $x \leftrightarrow z \in F$ by (fi$\cap$.) and (fi$\leq$.). Finally, it has to be checked that it respects the operations. Now, if $x \leftrightarrow y \in F$, we have $-x \leftrightarrow -y = x \leftrightarrow y \in F$ as well. Secondly, if $x \leftrightarrow y \in F$ and $z \leftrightarrow u \in F$, then also $(x \cap z) \leftrightarrow (y \cap u) \leq (x \leftrightarrow z) \cap (y \leftrightarrow u) \in F$, as one can check. So, if we have a filter F, we also have a congruence, denoted by Θ_F. Moreover, $F = [1]\Theta_F$. The homomorphism with kernel Θ_F is denoted by h_F. Clearly, if F and G are filters and $F \subseteq G$ then $\Theta_F \subseteq \Theta_G$, and conversely. We conclude the following.

Proposition 1.7.9. *Let $\mathfrak{A}$ be a boolean algebra. The map $f : \Theta \mapsto F_\Theta := \{x : x \,\Theta\, 1\}$ is an isomorphism from the lattice of congruences of $\mathfrak{A}$ onto the lattice of filters of $\mathfrak{A}$. Furthermore, if F is a filter, then Θ_F defined by $x \,\Theta_F\, y$ iff $x \leftrightarrow y \in F$ is the inverse of F under f.*

The following are equivalent definitions of an ultrafilter.

Proposition 1.7.10. *Let $\mathfrak{B}$ be a boolean algebra not isomorphic to* $\mathbf{1}$*. A filter U is an ultrafilter iff either of (1.), (2.), (3.).*
(1.) For all x: $-x \in U$ iff $x \notin U$.
(2.) U is proper and for all x, y: if $x \cup y \in U$ then $x \in U$ or $y \in U$.
(3.) The algebra $\mathfrak{B}/U$ is isomorphic to $\mathbf{2}$*.*

Proof. First we show that U is an ultrafilter iff U satisfies (3.). Namely, U is maximal iff Θ_U is maximal in $Con(\mathfrak{A})$ iff the interval $[\Theta_U, \nabla] \cong \mathbf{2}$ iff $\mathfrak{A}/U$ is simple (by Proposition 1.3.2). A boolean algebra is simple iff it is isomorphic to $\mathbf{2}$. For if there exists an element $x \neq 0, 1$ then the filter $F = \{y : y \geq x\}$ is distinct from the trivial filters. Now we show that U is an ultrafilter iff U satifies (1.). Suppose that U

is an ultrafilter. Then, by (3.), $\mathfrak{A}/U \cong \mathbf{2}$. So for every $x \in A$, $[x]\Theta_U = 0$ or $[x]\Theta_U = 1$. Thus either $[x]\Theta_U = 1$, which means $x \in U$, or $[-x]\Theta_U = 1$, which means $-x \in U$. Both cannot hold simultaneously. Next suppose that (1.) holds. Then if $U \subsetneq V$, there exists an $x \in V - U$. By (1.), however, $-x \in U$ and so $-x \in V$. Hence $0 \in V$, and V is not proper. So, U is an ultrafilter. Thirdly, we show that U satisfies (2.) iff U satisfies (1.). Suppose U satisfies (1.). Then U is proper. For $0 \notin U$, since $1 \in U$ (by (1.)). Now suppose that $x \notin U$ and $y \notin U$. Then $-x, -y \in U$ and so $(-x \cap -y) \in U$. Therefore, by (1.), $x \cup y \notin U$. Hence U satisfies (2.). Conversely, assume that U satisfies (2.). Then since $(1 =) \, x \cup -x \in U$ we conclude that $x \in U$ or $-x \in U$. Not both can hold simultaneously, since U is proper. Hence $x \notin U$ iff $-x \in U$. So, U satisfies (1.). ◻

We know that boolean logic is complete with respect to matrices $\langle \mathfrak{B}, F\rangle$, where $\mathfrak{B}$ is a boolean algebra and F a deductively closed set. This is so because in the term algebra the relation $\equiv := \{\langle x, y\rangle : x \leftrightarrow y = 1\}$ is a matrix congruence. (See exercises of Section 1.4.) The deductively closed sets of a boolean algebra are the filters. The congruence associated with F, Θ_F, is a matrix congruence and so we may factor again by the congruence Θ_F. Hence boolean logic has a unital semantics. Now, suppose we can show that every nontrivial filter is contained in an ultrafilter. Then we may further deduce that boolean logic is complete with respect to matrices $\mathfrak{M} = \langle \mathfrak{B}, F\rangle$ where F is an ultrafilter. By the previous theorem, $\mathfrak{M}/\Theta_F = \langle \mathbf{2}, \{1\}\rangle$. So the matrix consisting of the algebra $\mathbf{2}$ with the distinguished element 1 characterizes boolean logic completely. This is less spectacular if we look at the fact that boolean logic is actually defined this way; rather it tells us that the set of equations that we have chosen to spell out boolean logic is complete — for it reduces an adequate set of matrices to just the algebra $\mathbf{2}$ with $D = \{1\}$, exactly as desired.

The next result is of extreme importance in general logic. It deals with the existence of filters and ultrafilters — or, as is equivalent by Proposition 1.7.10, — with the existence of homomorphic images of boolean algebras. Let us take a subset $E \subseteq A$ and ask when it is possible to extend E to a proper filter on A.

Proposition 1.7.11. *The least filter containing an arbitrary subset E is the set $\langle E\rangle$ defined by*

$$\langle E\rangle = \{x : x \geq \bigcap X, X \text{ a finite subset of } E\}$$

$\langle E\rangle$ is proper iff every finite subset of E has a nonzero intersection. In this case we say that E has the ***finite intersection property****.*

Proof. First of all, $E \subseteq \langle E\rangle$. For $x \geq \bigcap\{x\}$. $\langle E\rangle$ is also a filter. For if $X = \varnothing$, then $\bigcap X = 1$, so (fi1.) is satisfied. Next let $x \in \langle E\rangle$ and $x \leq y$. We know that there is a finite X such that $x \geq \bigcap X$. Then also $y \geq \bigcap X$ and so $y \in \langle E\rangle$. This shows (fi≤.). Finally, if $x, y \in \langle E\rangle$ then $x \geq \bigcap X$, $y \geq \bigcap Y$ for some finite $X, Y \subseteq E$. Then $x \cap y \geq \bigcap(X \cup Y)$; and $X \cup Y$ is finite. To see that $\langle E\rangle$ is the least filter, observe that for every finite subset X we must put $\bigcap X \in \langle E\rangle$. For either $X = \varnothing$ and so $\bigcap X$ must be added to satisfy (fi1.), or $X = \{x\}$ for some x, and then $\bigcap X = x$ must be added to satisfy $E \subseteq \langle E\rangle$. Or $\sharp X > 1$ and then $\bigcap X$ must be added to satisfy (fi$\cap$.). Finally, to satisfy (fi$\leq$.) all elements of $\langle E\rangle$ must be taken as well. So no smaller set suffices. It is then clear that if for no finite set X we have $\bigcap X = 0$, $\langle E\rangle$ is not the full algebra, since $0 \notin \langle E\rangle$. But if $\bigcap X = 0$ for some finite X then by (fi$\leq$.), $y \in \langle E\rangle$ for all y. ◻

Now we show that every proper filter is contained in an ultrafilter. Equivalently, every set with the finite intersection property can be embedded in an ultrafilter. The proof will be of quite general character, by appealing to what is known as *Tukey's Lemma*. (See [**215**]. This lemma is actually the same as Oswald Teichmüller's Principle D, 3rd Version from [**205**].) To state the lemma, consider a set S and a property $\mathcal{P}$ of subsets. $\mathcal{P}$ is said to be of **finite character** if $\mathcal{P}$ holds of a set $X \subseteq S$ iff it holds of all finite subsets of X. If $\mathcal{P}$ is of finite character, then if T has $\mathcal{P}$ and $S \subseteq T$ then also S has $\mathcal{P}$. For every finite subset of S is a finite subset of T, and so if S fails to have $\mathcal{P}$, this is because some finite subset S_0 fails to have $\mathcal{P}$, which implies that T fails to have $\mathcal{P}$, since $S_0 \subseteq T$.

Theorem 1.7.12 (Tukey's Lemma). *Suppose $\mathcal{P}$ is a property of subsets of S and that $\mathcal{P}$ is of finite character. Then every set $X \subseteq S$ having $\mathcal{P}$ is contained in a set X^* which is maximal for $\mathcal{P}$.*

Proof. By the axiom of choice, S can be well–ordered by an ordinal λ; thus $S = \{s_\alpha : \alpha < \lambda\}$. By induction over λ we define X_α for $\alpha \leq \lambda$. We put $X_0 := X$. If X_α is defined, and $\alpha + 1 \neq \lambda$ then put $X_{\alpha+1} := X_\alpha \cup \{s_\alpha\}$ if this set has $\mathcal{P}$, and $X_{\alpha+1} := X_\alpha$ otherwise. For a limit ordinal α we put $X_\alpha := \bigcup_{\kappa<\alpha} X_\kappa$. Now let $X^* := X_\lambda$. We claim that X^* has the desired properties. X^* contains X. It has $\mathcal{P}$; for X_0 has $\mathcal{P}$ by assumption. In the successor step this remains true by construction and in the limit step by the finite character. For either the sequence $\langle X_\kappa : \kappa < \alpha\rangle$ is stationary from a certain ordinal β. Then $X_\alpha = X_\beta$, in which case X_α has $\mathcal{P}$ by induction hypothesis. Or the sequence is strictly ascending. Then any finite subset of X_α is contained in a X_β for some $\beta < \alpha$; and then X_α has $\mathcal{P}$ by the finite character of $\mathcal{P}$. Finally, we must verify that X^* is maximal. Suppose that $X^* \subseteq Y$ and Y has $\mathcal{P}$. Take an element $y \in Y$. $y = s_\gamma$ for some $\gamma < \lambda$. Consider the definition of $X_{\gamma+1}$. We know that $X_{\gamma+1} \subseteq X^* \subseteq Y$. Since Y has $\mathcal{P}$, $X^* \cup \{s_\gamma\}$ has $\mathcal{P}$ as well, since $\mathcal{P}$ is closed under subsets. Therefore $X_{\gamma+1} = X_\gamma \cup \{s_\gamma\}$, showing $y \in X^*$. Since y was arbitary, $X^* = Y$. Hence the set X^* is maximal. ◻

Corollary 1.7.13. *Every set with the finite intersection property is contained in an ultrafilter. In particular, every proper filter is contained in an ultrafilter.*

Proof. By Tukey's Lemma. The property $\mathcal{P}$ is taken to be *generates a proper filter*. It is of finite character. For X generates a proper filter iff X has the finite intersection property iff every finite subset of X has the finite intersection property iff every finite subset generates a proper filter. ◻

THEOREM 1.7.14. *Boolean logic in the language* $\top$, $\neg$ *and* $\wedge$ *is the logic of the matrix* $\mathbf{2} := \langle\{0, 1\}, 1, -, \cap, \{1\}\rangle$. *Boolean logic is structurally complete and Post–complete. It has interpolation and is Halldén–complete.*

For a proof note that we have established that boolean logic is the logic of the matrix **2**. Furthermore, it has all constants, since $\top$ has value 1 and $\bot$ has value 0. It follows that it has interpolation by Theorem 1.6.4 (since the matrix is finite, the logic has conjunction and all constants), and that it is structurally complete and Post–complete by Theorem 1.6.6. It is Halldén–complete by Theorem 1.6.5.

Exercise 22. Show that $\{-, \cap\}$ is polynomially complete and that $\{1, -, \cap\}$ is functionally complete.

Exercise 23. Let $-$ be an operation satisfying $- - x = x$. Let $\cap$ be a binary oeration, and put $x \cup y := -(-x \cap -y)$. Show that $\cap$ is associative (commutative, idempotent) iff $\cup$ is associative (commutative, idempotent). *Hint.* Show first that $x \cap y = -(-x \cup -y)$, so that only one direction needs to be established in each case. (This shows Lemma 1.7.2.)

Exercise 24. Show that each proper subspace of a vector space is contained in a hyperplane, i. e. a maximal proper subspace.

1.8. Some Notes on Computation and Complexity

In this section we will briefly explain the basic notions of computability and complexity. Although we shall prove only very few results on complexity of modal logics we shall nevertheless mention a fair number of them. This section provides the terminology and basic results so that the reader can at least understand the results and read the relevant literature. The reader is referred to MICHAEL R. GAREY and DAVID S. JOHNSON [**73**] for an introduction into complexity theory. For our purposes it is best to define computations by means of string handling devices. The most natural one is of course the Turing machine, but its definition is rather cumbersome. We shall therefore work with a slightly easier model, which is a mixture between a Turing machine and a so–called **Semi–Thue Process**. Let us now fix an alphabet A. Members of A shall be written in small caps to distinguish them from symbols that merely abbreviate or denote symbols or strings from A.

DEFINITION 1.8.1. *Let* A *be a finite set and* $\nabla \notin A$. *A* ***string handling machine*** *over* A *is a finite set* T *of pairs* $\langle\vec{x}, \vec{y}\rangle$ *such that* $\vec{x}$ *and* $\vec{y}$ *are strings over* $A \cup \{\nabla\}$ *in which* ∇ *occurs exactly once. We call members of* T ***instructions****.*

The symbol ∇ is used to denote the position of the read and write head of the Turing machine.

DEFINITION 1.8.2. *Let* T *be a string handling machine over* A *and* $\vec{x}$ *and* $\vec{y}$ *strings over* $A \cup \{\nabla\}$. *Then* $\vec{x} \Rightarrow_T \vec{y}$ *if there is a pair* $\langle\vec{u}, \vec{v}\rangle \in T$ *and strings* $\vec{w}_1$ *and* $\vec{w}_2$ *such that*

*$\vec{x} = \vec{w}_1{}^\frown\vec{u}{}^\frown\vec{w}_2$ and $\vec{y} = \vec{w}_1{}^\frown\vec{v}{}^\frown\vec{w}_2$. In that case we say that $\vec{y}$ is **1–step computable from** $\vec{x}$. $\vec{y}$ is **n+1–step computable from** $\vec{x}$, in symbols $\vec{x} \Rightarrow_T^{n+1} \vec{y}$, if there is a $\vec{z}$ such that $\vec{x} \Rightarrow_T^n \vec{z} \Rightarrow_T \vec{y}$. $\vec{y}$ is **computable from** $\vec{x}$ if $\vec{x} \Rightarrow_T^n \vec{y}$ for some $n \in \omega$. We write $\vec{x} \Rightarrow_T^* \vec{y}$.*

We also call a sequence $\langle \vec{x}_i : i < n\rangle$ a T–**computation** if for every $j < n - 1$ we have $\vec{x}_j \Rightarrow_T \vec{x}_{j+1}$. $\vec{y}$ is called a **halting string for** T if there is no string that is computable from $\vec{y}$. A **halting computation** is a computation whose last member is a halting string.

Definition 1.8.3. *Let $f : A^* \to A^*$ be a function and T a string handling machine over some alphabet $C = A \cup B$, where B is disjoint from A. We say that T **computes** f with respect to $\text{B}, \text{E}, \# \in B$ if for every $\vec{x} \in A^*$ the following holds:*

(1) *$\nabla{}^\frown\text{E}{}^\frown f(\vec{x}){}^\frown\#$ is computable from $\nabla{}^\frown\text{B}{}^\frown\vec{x}{}^\frown\#$.*
(2) *$\nabla{}^\frown\text{E}{}^\frown f(\vec{x}){}^\frown\#$ is a halting string for T.*
(3) *There is no halting string for T different from $\nabla{}^\frown\text{E}{}^\frown f(\vec{x}){}^\frown\#$ which is computable from $\nabla{}^\frown\text{B}{}^\frown\vec{x}{}^\frown\#$*

*f is called **nondeterministically computable** if there exists a string handling machine that computes f (with respect to some elements).*

The elements B and E are used to mark the begin and the end of the computation and # to mark the end of the input string. Notice that it is not possible to define a string handling machine that computes $f(\vec{x})$ from $\vec{x}$ simpliciter. For we would have no means to distinguish whether we just started the computation or whether we just ended it; nor could we tell where the input ended. Notice that the symbol B also marks the begin of the input string, whence a separate marker is not needed for that purpose. In the definition above we shall say that T computes $f(\vec{x})$ from $\vec{x}$ **in** n **steps** if

$$\nabla{}^\frown\text{B}{}^\frown\vec{x}{}^\frown\# \Rightarrow_T^n \nabla{}^\frown\text{E}{}^\frown f(\vec{x}){}^\frown\#$$

Note that it is not required that all computations terminate; but if a computation halts, it must halt in the just specified string. If f is only a partial function we require that if f is undefined on $\vec{x}$ then there is no halting computation starting with $\nabla{}^\frown\text{B}{}^\frown\vec{x}{}^\frown\#$. This allows us to define the notion of a computable function from A^* to B^*, which is namely also a partial function from $(A \cup B)^*$ to $(A \cup B)^*$. We shall present a simple example. Let $A = \{\text{R}, \text{S}\}$ and let $f : \vec{x} \mapsto \vec{x}{}^\frown\text{R}$. Here is a string handling machine that computes this function:

$$\begin{aligned} T_1 := \{&\langle \nabla\text{B}\#, \nabla\text{ER}\#\rangle, \\ &\langle \nabla\text{BR}, \text{BR}\nabla\text{C}\rangle, \quad \langle \nabla\text{BS}, \text{BS}\nabla\text{C}\rangle, \\ &\langle \nabla\text{CR}, \text{R}\nabla\text{C}\rangle, \quad \langle \nabla\text{CS}, \text{S}\nabla\text{C}\rangle, \\ &\langle \nabla\text{C}\#, \nabla\text{DR}\#\rangle, \\ &\langle \text{R}\nabla\text{D}, \nabla\text{DR}\rangle, \quad \langle \text{S}\nabla\text{D}, \nabla\text{DS}\rangle, \\ &\langle \text{B}\nabla\text{D}, \nabla\text{E}\rangle\} \end{aligned}$$

The verification that this machine indeed computes f is left to the reader.

The above definition of computability is technical. Yet, ALONZO CHURCH advanced the conjecture — also known as *Church's Thesis* — that any function that is computable in the intuitive sense is also computable in the technical sense. (See STEVEN C. KLEENE [**116**] for details.) The converse is of course clear. Notice that if A consists of a single letter, say a, then the set of natural numbers can be identified simply with the set of nonempty sequences over a, where n is represented by the sequence consisting of n consecutive as. (ALAN TURING used $n + 1$ as to code n, but this is unnecessary given the present setting.) It can be shown that under this coding any recursive function on the natural numbers is computable, and that any computable function is recursive. There are other codings of the natural numbers that work equally well (and are less space consuming), for example p–adic representations.

The above definitions can be generalized by admitting computation not on one string alone but on several strings at the same time. This model is more general but it can be shown that it does not generate a larger class of computable functions (modulo some coding). The benefit is that it is much easier to see that a certain function is computable. A string handling machine with k–strings is a finite set of k–tuples of pairs of strings over $A \cup \{\nabla\}$ in which ∇ occurs exactly once. Computations now run over k–tuples of strings. The definitions are exactly parallel. A replacement is done on all strings at once, unlike a multihead Turing machine, which is generally allowed to operate on a single tape only at each step. We can now define the notion of a computable k–ary function from A^* to A^* in much the same way. Notice that for T to compute f we shall simply require that the initial configuration is the sequence $\langle \nabla^\frown \mathrm{B}^\frown \vec{x}_i{}^\frown \# : i < k\rangle$, and that the the halting configuration is the sequence $\langle \nabla^\frown \mathrm{E}^\frown f(\vec{x})^\frown \#, \langle \nabla^\frown \mathrm{E}^\frown \# : i < k-1\rangle\rangle$. The fact that the starting state and the end state appear on each tape is harmless. Given $f : (A^*)^k \to A^*$, we shall define a unary function $f^\heartsuit : (A \cup \{\mathrm{c}\})^* \to A^*$, where $\mathrm{c} \notin A$, as follows. If $\vec{x} = \vec{y}_0{}^\frown \mathrm{c}^\frown \vec{y}_1{}^\frown \mathrm{c}^\frown \ldots {}^\frown \vec{y}_{k-1}{}^\frown \mathrm{c}^\frown \#$, then

$$f^\heartsuit(\vec{x}) := f(\langle \vec{y}_i : i < n\rangle)$$

Otherwise, $f^\heartsuit(\vec{x}) := \mathrm{c}$. The following now holds.

THEOREM 1.8.4. *f is computable on a string handling machine with k strings iff $f^\heartsuit$ is computable on a string handling machine with one string.*

The proof of this theorem is not hard but rather long winded.

With these definitions we can already introduce the basic measures of complexity. Let $f : A^* \to A^*$ and $h : \omega \to \omega$ be functions. We say that T **computes** f in h–**time** if for all $\vec{x}$, T computes $f(\vec{x})$ from $\vec{x}$ in at most $h(|\vec{x}|)$ steps. We say that T computes f in h–**space** if there is a computation of $\nabla^\frown \mathrm{E}^\frown f(\vec{x})^\frown \#$ from $\nabla^\frown \mathrm{B}^\frown \vec{x}^\frown \#$ in which every member has length $\leq |\vec{x}|$. Typically, one is not interested in the exact size of the complexity functions, so one introduces more rough measures. We say that T computes f in $O(h)$–**time** if there is a constant c such that T computes f is $c \cdot h$–time for almost all $\vec{x}$. Analogously with space. Now, before we can introduce the major classes of complexity, we shall have to define the notion of a deterministic machine.

Definition 1.8.5. *Let T be a string handling machine. T is called* ***deterministic*** *if for every string $\vec{x}$ there exists at most one string $\vec{y}$ which is 1–step computable from $\vec{x}$. A function $f : A^* \to A^*$ is* ***deterministically computable*** *if there is a deterministic string handling machine that computes f with respect to some elements.*

It is easy to see that the machine T_1 defined above is deterministic. Namely, a string handling machine T is deterministic if for any pair $\langle \vec{x}, \vec{y} \rangle \in T$ there exists no other pair $\langle \vec{u}, \vec{v} \rangle \in T$ such that $\vec{u}$ is a substring of $\vec{x}$. This is obviously satisfied. The following theorem shall be stated without proof.

Theorem 1.8.6. *A function $f : A^* \to A^*$ is nondeterministically computable iff it is deterministically computable.*

Definition 1.8.7. ***P*** *denotes the class of functions deterministically computable in polynomial time,* ***NP*** *the class of functions nondeterministically computable in polynomial time. Similarly,* ***EXPTIME*** *(**NEXPTIME**) denotes the class of function deterministically (nondeterministically) computable in exponential time.* ***PSPACE*** *denotes the class of functions deterministically computable in polynomial space.*

The reason why there is no class NPSPACE is the following result from [**189**].

Theorem 1.8.8 (Savitch). *If a function is computable by a nondeterministic machine using polynomial space then it is also computable by a deterministic machine using polynomial space.*

We have

$$\mathrm{P} \subseteq \mathrm{NP} \subseteq \mathrm{PSPACE} \subseteq \mathrm{EXPTIME} \subseteq \mathrm{NEXPTIME}$$

Most inclusions are by now obvious. Notice that if a deterministic machine runs in h–time it can write strings of length at most $O(h)$. When using a machine with several strings the complexity does not change. It can be shown, namely, that if T is a string handling machine with k strings computing f in $O(h)$ time, then there is a string handling machine $T^\heartsuit$ computing $f^\heartsuit$ in $O(h^2)$ time. Moreover, $T^\heartsuit$ can be chosen deterministic if T is. It is therefore clear that the above complexity classes do not depend on the number of strings on which we do the computation, a fact that is very useful.

Talk about computability often takes the form of problem solving. A problem can be viewed as a subset S of A^*. To be exact, the problem that is associated with S is the question to decide, given a string $\vec{x} \in A^*$, whether or not $\vec{x} \in S$.

Definition 1.8.9. *A* ***problem*** *is a function $f : A^* \to \{0, 1\}$. f is* ***trivial*** *if f is constant. We say that a problem f is $\mathcal{C}$ if $f \in \mathcal{C}$, we say that it is $\mathcal{C}$–**hard** if for any $g \in \mathcal{C}$ there exists a $p \in$ PSPACE such that $g = f \circ p$. Finally, f is $\mathcal{C}$–**complete** if it is both in $\mathcal{C}$ and $\mathcal{C}$–**hard**.*

A problem that is $\mathcal{C}$–complete is in a sense the hardest problem in $\mathcal{C}$. For any other problem is reducible to it. Given a set S, we shall denote by '$x \in S$?' the problem of deciding membership in S, which is simply defined as the function that gives 1 if the answer is 'yes' and 0 if the answer is 'no'. (This is also the characteristic function of S.) For amusement the reader may show that any nontrivial problem that is in P is also P–complete. In connection with these definitions we have

Definition 1.8.10. *A subset of A^* is* ***decidable*** *if its characteristic function is computable.*

We shall spend the rest of this section illustrating these concepts with some examples from logic. We have introduced languages in Section 1.2. To adapt the definitions of that section to the present context we need to make the alphabet finite. This is not entirely straightforward, since we have infinitely many variables. Therefore, let F be the set of function symbols, and $X := \{p_i : i \in \omega\}$ our set of variables. Let us first assume that F is finite. Take symbols P, 0 and 1 not occurring in F or X. So, we shall replace the variable p_i by the sequence $\mathrm{P}^\frown\vec{x}$, where $\vec{x} \in \{0, 1\}^*$ is a binary string representing the number i in the usual way. That is to say, we put $\mu(0) := 0$ and $\mu(1) := 1$, and if $\vec{x} = x_0{}^\frown x_1{}^\frown \dots x_{n-1}$ then $i = \mu(\vec{x})$, where

$$\mu(\vec{x}) := \sum_{j<n} 2^{n-1-j}\mu(x_j)$$

In the same way we can obviously also code a countable F by means of a single symbol F followed by a binary string. In this way, the entire logical language can be written using just the symbols F, P, 0 and 1. We will however refrain from using this coding whenever possible. We shall note here that the typical measure of length of a formula is the number of symbols occurring in it. We call this the **symbol count** of the formula. However, the actual string that we write to denote a formula can be longer. Since a variable counts as one symbol, the coding is not length preserving. Rather, a formula with n symbols is represented by a string of length at most $n \log_2 n$ if we allow renaming of variables. Given φ, $|\varphi|$ counts the length of a minimal representing string. This additional factor by which $|\varphi|$ is longer than the symbol count is usually (but not always!) negligeable. If the set of function symbols is infinite, there is no a priori upper bound on the length of the string in comparison to the symbol count! It is for these reasons that complexity is always measured in terms of the length of the string representing the formula. There is another way to represent a formula, which we refer to as the *packed representation*. It is described in the exercises below since it will only be relevant in Section 3.6.

We shall now go into the details of certain basic string properties and manipulations. First of all, let $\Omega : F \to \omega$ be a signature and $X = \{p_i : i \in \omega\}$. We shall provide a procedure to decide whether or not a given string is a term. Define the **weight** ρ_Ω of a symbol as follows.

$$\begin{array}{lcl}\rho_\Omega(p_i) & := & -1 \\ \rho_\Omega(f) & := & \Omega(f)-1\end{array}$$

For a string $\vec{x} = x_0{}^\frown x_1{}^\frown \ldots {}^\frown x_{n-1}$ we put

$$\rho_\Omega(\vec{x}) := \sum_{i<n} \rho_\Omega(x_i)$$

It is useful to observe the following

Lemma 1.8.11. *Let $\vec{x}$ be a string such that $\rho_\Omega(\vec{x}) < 0$. Then there exists a prefix $\vec{y}$ of $\vec{x}$ such that $\rho_\Omega(\vec{y}) = \rho_\Omega(\vec{x}) - 1$.*

This lemma is proved by induction on the length of $\vec{x}$. For the sake of precision define an **occurrence** of a string $\vec{x}$ in $\vec{y}$ to be a pair $\langle \vec{u}, \vec{x}\rangle$ such that $\vec{u}^\frown\vec{x}$ is a prefix of $\vec{y}$. Two string occurrences $\langle \vec{u}, \vec{x}\rangle$ and $\langle \vec{v}, \vec{z}\rangle$ **overlap** if either (a) $\vec{u}$ is a prefix of $\vec{v}$ and $\vec{y}$ is a proper prefix of $\vec{u}^\frown\vec{x}$ or (b) $\vec{v}$ is a prefix of $\vec{u}$ and $\vec{x}$ is a proper prefix of $\vec{v}^\frown\vec{y}$.

Proposition 1.8.12. *A string $\vec{x}$ is an Ω–term iff it has the property (P).*

(P). $\rho_\Omega(\vec{x}) = -1$ *and for every prefix $\vec{y}$ of $\vec{x}$:* $\rho_\Omega(\vec{y}) \geq 0$.

In particular, no proper prefix of a term is a term, and no two distinct occurrences of subterms overlap.

Proof. We begin by showing how the other claims follow from the first. If $\vec{x}$ is a term and $\vec{y}$ a proper substring, then $\rho_\Omega(\vec{y}) > -1$, and so $\vec{y}$ is not a term. Next, let $\vec{x} = \vec{u}_0{}^\frown\vec{u}_1{}^\frown\vec{u}_2{}^\frown\vec{u}_3{}^\frown\vec{u}_4$, and assume that $\vec{u}_1{}^\frown\vec{u}_2$ as well as $\vec{u}_2{}^\frown\vec{u}_3$ are terms and both $\vec{u}_1 \neq \varepsilon$ and $\vec{u}_3 \neq \varepsilon$. Then we have $\rho_\Omega(\vec{u}_2{}^\frown\vec{u}_3) = \rho_\Omega(\vec{u}_2)+\rho_\Omega(\vec{u}_3)$. Now, since $\vec{u}_2{}^\frown\vec{u}_3$ has Property (P), $\rho_\Omega(\vec{u}_2) \geq 0$. Now, likewise $-1 = \rho_\Omega(\vec{u}_1) + \rho_\Omega(\vec{u}_2)$, whence $\rho_\Omega(\vec{u}_1) < 0$. So, $\vec{u}_1{}^\frown\vec{u}_2$ does not have (P), and hence is not a term, contrary to our assumption.

We now show the first claim. This is done by induction on the length of $\vec{x}$. Clearly, for strings of length 0 and 1 the claim is true. (Notice that $\rho_\Omega(\varepsilon) = 0$.) So, let $|\vec{x}| > 1$. Assume that the claim is true for all strings of length $< |\vec{x}|$. Assume first that $\vec{x}$ is a term. Since $\vec{x}$ has length at least 2, the first symbol of $\vec{x}$ is some $f \in F$ of arity at least 1. So,

$$\vec{x} = f^\frown\vec{x}_0{}^\frown\vec{x}_1{}^\frown \ldots {}^\frown x_{\Omega(f)-1}$$

By inductive hypothesis, $\rho_\Omega(\vec{x}_i) = -1$ for all $i < \Omega(f)$. Hence we have $\rho_\Omega(\vec{x}) = \Omega(f) - 1 - \Omega(f) = -1$. Now it is easy to see that for no prefix $\vec{u}$ of $\vec{x}$ $\rho_\Omega(\vec{u}) \geq 0$. Hence $\vec{x}$ has Property (P). Conversely, assume that $\vec{x}$ has Property (P). Take the first symbol of $\vec{x}$. It is some $f \in F$ of arity at least 1. Let $\vec{y}$ be such that $\vec{x} = f^\frown\vec{y}$. Then $\rho_\Omega(\vec{y}) \geq -1$. If the arity of f is 1, then $\vec{x}$ is a term iff $\vec{y}$ is. Then $\vec{y}$ has (P) and so is a term, by inductive hypothesis. Therefore $\vec{x}$ is a term as well. Now suppose that $\Omega(f) > 1$. Then $\rho_\Omega(\vec{y}) = -\Omega(f)$. Using Lemma 1.8.11 it can be shown that $\vec{y}$ is the product of strings $\vec{u}_i$, $i < \Omega(f)$, which all have (P). By inductive hypothesis the $\vec{u}_i$ are terms. So is therefore $\vec{x}$. ◻

We shall note here the following corollary. If F is infinite, we code it as described above by strings of the form $\text{F}^\frown\vec{x}$. Ω is defined on these strings. We shall recode the signature as a function from binary strings to binary strings. Namely, we let

$$\Omega'(\vec{x}) := \mu^{-1}(\Omega(\text{F}^\frown\vec{x}))$$

We shall say that Ω is computable if Ω' is.

Proposition 1.8.13. *Assume that $\mathfrak{L}$ is a language with a computable signature. Then the set of terms over F is decidable.*

The proof is not difficult.

It is possible to convert a term in prefix notation into a string in the typical bracket notation and conversely. To that end, we assume that the signature consists of symbols of arity ≤ 2. The procedure is as follows. Take a string $\vec{x}$ in prefix notation. Start from the left end. Assume $\vec{x} = \vec{y}^\frown f^\frown\vec{z}$ and $n := |\vec{y}|$. Then let $\vec{u}$ be the smallest prefix of $\vec{z}$ such that $f^\frown\vec{u}$ is a term and let $\vec{z} = \vec{u}^\frown\vec{v}$. Then let $\vec{x}^n := \vec{y}^\frown(^\frown f^\frown\vec{u}^\frown)^\frown\vec{v}$. We call this the insertion of brackets at the place n. The procedure is now simply the following. Start at the left end of $\vec{x}$ and add brackets whenever the described situation occurs. Continue as long as possible. Call the resulting string $b(\vec{x})$. Now let $b(\vec{x})$ be given. Pick a symbol f of arity 2 following a symbol (. Then it is contained in a sequence of the form $(^\frown f^\frown\vec{u}_0{}^\frown\vec{u}_1{}^\frown)$ where $\vec{u}_0$ and $\vec{u}_1$ are terms. (This needs to be defined for sequences which contain brackets but is straightforward.) Replace this sequence by $(^\frown\vec{u}_0{}^\frown f^\frown\vec{u}_1{}^\frown)$. Continue whenever possible. Finally, some brackets are dropped (the outermost brackets, brackets enclosing variables). The resulting sequence is a term in usual bracket notation.

To close, let us describe a procedure that evaluates a term in a finite algebra given a particular valuation. It is commonly assumed that this can be done in time proportional to the length of the string. The procedure is as follows: identify a minimal subterm and evaluate it. Repeat this as often as necessary. This procedure if spelled out in detail is quadratic in the length of the string since identifying a minimal subterms also takes time linear in the length of a string. As this procedure is repeated as often as there are subterms, we get in total an algorithm quadratic in the length. We shall now describe the algorithm in detail. So, let $\mathfrak{A}$ be a finite Ω–algebra. For each element a of the algebra we assume to have a primitive symbol, which we denote by a. Let the term be given as a string, and the valuation as a sequence of pairs $\langle \text{P}^\frown\vec{x}, \beta(\text{P}^\frown\vec{x})\rangle$. It is not necessary to have all values, just all values for variables occurring in $\vec{x}$. We shall describe a procedure that rewrites $\vec{x}$ successively, until a particular value is determined. We start by inserting $\beta(\text{P}^\frown\vec{x})$ in place of $\text{P}^\frown\vec{x}$. We treat the elements as variables, assigning them the weight -1. Let $\langle\vec{u}, \vec{y}\rangle$ be an occurrence of a substring, where $\vec{y}$ has length > 1. We call $\rho_\Omega(\vec{u})$ its **embedding number**. An occurrence of a substring is with maximal embedding number is of the form $f^\frown a_0{}^\frown a_1{}^\frown \dots {}^\frown a_{\Omega(f)-1}$, where all $\vec{a}_i$ are (symbols denoting) elements of $\mathfrak{A}$. The procdure is therefore as follows. Look for an occurrence of a nontrivial substring with maximal embedding number and compute its value. Replace that string by its value. Continue as long as possible. The procedure ends when $\vec{x}$ is a single symbol.

Now let a formula φ of boolean logic be given. We can solve the problem whether φ is a theorem nondeterministically by guessing a valuation of the variables of φ and then evaluating φ. A valuation is linear in the length of φ. Hence, the problem whether a formula is a theorem of boolean logic is in NP. Alternatively, the problem whether a boolean formula is satisfiable is also in NP. Moreover, the following holds:

Theorem 1.8.14 (Cook). *Satisfiability of a boolean expression is NP–complete.*

This result may appear paradoxical since we have just proved that satisfiability is computable nondeterministically in $O(n^2)$ time. So, how come it is the hardest problem in NP? The answer is the following. If a problem S is in NP it is polynomially reducible to the satsifiability problem; the reduction function is itself a polynomial p and this polynomial can have any degree. Hence the harder S the higher the degree of p.

It has been shown by L. J. Stockmeyer and A. R. Meyer [**203**] that the problem of satisfiability of quantified boolean formulae is PSPACE–complete. (Here, quantifiers range over propositional variables.)

Exercise 25. We can represent the natural number n either as the sequence $\mu^{-1}(n)$ over $\{0, 1\}$ or as a sequence of n 1s. Show that the mappings converting one representation to the other are computable.

Exercise 26. Show that if $f, g : A^* \to A^*$ are computable then so is $g \circ f$. Show that if f and g are in $\mathcal{C}$ for any of the above complexity classes, then so is $g \circ f$.

Exercise 27. Let A be a fixed alphabet, and let c (for comma) $\langle$ and $\rangle$ be new symbols. With the help of these symbols we can code a string handling machine T by a string $T^\dagger$. (This is not unique.) Now let C be the enriched alphabet. Let $f : C^* \times C^* \to C^*$ be defined as follows. If $\vec{x} \in A^*$ and $\vec{y} = T^\dagger$ for some string handling machine using the alphabet A then $f(\vec{x}, \vec{y})$ is the result that T computes on $\vec{x}$ if it halts, and otherwise f is undefined. Show that f is computable. (This is analogous to the Universal Turing Machine.)

Exercise 28. Prove Lemma 1.8.11.

Exercise 29. Here is another way to represent a formula. Let φ be a formula, say $\wedge p_0 \vee p_1 \neg p_0$, which in infix notation is just $(p_0 \wedge (p_1 \vee \neg p_0))$. The string associated with it is $\wedge$p0 $\vee$ p1$\neg$p0. Now enumerate the subformulae of this formula. We shall use for the nth subformula the code $\mathrm{s}^\frown\vec{x}$, where $\vec{x}$ is the binary representation of n:

$$\begin{array}{lcl} \mathrm{s0} & : & \mathrm{P}0 \\ \mathrm{s1} & : & \mathrm{P}1 \\ \mathrm{s10} & : & \neg\mathrm{P}0 \\ \mathrm{s11} & : & \vee\mathrm{P}1\neg\mathrm{P}0 \\ \mathrm{s100} & : & \wedge\mathrm{P}0 \vee \mathrm{P}1\neg\mathrm{P}0 \end{array}$$

Now replace to the right of this list the immediate subformulae by their respective code, starting with the smallest subformulae that are not variables. (For example, the least line becomes $\mathrm{s100} \wedge \mathrm{s0s11}$.) Finally, write these lines in one continuous string:

$$\mathrm{s0P0s1P1s10}\neg\mathrm{s0s11} \vee \mathrm{s1s10s100} \wedge \mathrm{s0s11}$$

Denote the resulting string by $\varphi^\spadesuit$. Give upper and lower bounds for $|\varphi^\spadesuit|$ in comparison with $|\varphi|$. Show that given a sequence $\vec{x}$ one can compute in linear time a formula φ such that $\vec{x} = \varphi^\spadesuit$. This representation can be used to code a set Δ of formulae as follows. Each subformula φ of some member of Δ that is itself in Δ is denoted not by $\varphi^\spadesuit$ but simply by its code $\mathrm{s}^\frown\vec{x}$. How is $|\Delta^\spadesuit|$ related to $card(sf[\Delta])$?

CHAPTER 2

Fundamentals of Modal Logic I

2.1. Syntax of Modal Logics

The languages of propositional modal logic used in this book contain a set $var = \{p_i : i \in \gamma\}$ of variables, a set *cns* of propositional constants, the boolean connectives $\top, \neg, \wedge$ and a set $\{\Box_i : i \in \kappa\}$ of modal operators. $\Box_i\varphi$ is read *box i phi*. $\top$ is in *cns*. With each operator $\Box_i$ we also have its so–called **dual operator** $\Diamond_i$ defined by $\Diamond_i\varphi := \neg\Box_i\neg\varphi$. In what is to follow, we will assume that there are no basic constants except $\top$; and that there are countably many propositional variables. So, $\gamma = \aleph_0$ unless stated otherwise. The number of operators, denoted by κ throughout, is free to be any cardinal number except 0; usually, κ is finite or countably infinite, though little hinges on this. Theorems which require that κ has specific values will be explicitly marked. By $\mathcal{P}_\kappa$ we denote the language of κ–modal logic, with no constants and $\aleph_0$ many propositional variables. Also, $\mathcal{P}_\kappa$ denotes the set of terms also called well–formed formulae of that language. If $\kappa = 1$ we speak of (the language of) *monomodal* logic, if $\kappa = 2$ of (the language of) *bimodal logic*. Finally, $\neg$ and $\Box_j$ are assumed to bind stronger than all other operators, $\wedge$ stronger than $\vee$, $\vee$ stronger than $\rightarrow$ and $\leftrightarrow$. So, $\Box_j p \wedge q \rightarrow p$ is the same as $((\Box_j p) \wedge q) \rightarrow p$.

$\mathcal{P}_\kappa$ is the set of *terms* or *formulae* or *propositions* as defined earlier. Metavariables for propositional variables are lower case Roman letters, such as p, q, r, metavariables for formulae are lower case Greek letters such as φ, χ, ψ. In addition, rather than using indices to distinguish modal operators we will use symbols such as $\blacksquare$, $\boxminus$, $\Box$, $\Box$, $\Box$, $\Box$, and similar abbreviations for their duals. We denote by $\|\mathcal{P}_\kappa\|$ the cardinality of the set of terms over $\mathcal{P}_\kappa$. With the cardinality of the set of variables, of the set of constants and the set of operators the cardinality of $\mathcal{P}_\kappa$ is fixed. Namely, by Proposition 1.2.1 we have

$$\|\mathcal{P}_\kappa\| = max\{\aleph_0, \sharp var, \sharp cns, \kappa\} \ .$$

Since we standardly assume to have at most $\aleph_0$ variables and constants, the latter reduces generally to $\|\mathcal{P}_\kappa\| = max\{\aleph_0, \kappa\}$.

The **modal depth** or **(modal) degree** of a formula is the maximum number of nestings of modal operators. Formally, it is defined as follows.

$$\begin{array}{lll} dp(p) & := & 0 \qquad\qquad\qquad p \in var \cup cns \\ dp(\neg\varphi) & := & dp(\varphi) \\ dp(\varphi \wedge \psi) & := & max\{dp(\varphi), dp(\psi)\} \\ dp(\Box_j \varphi) & := & 1 + dp(\varphi) \end{array}$$

The boolean connectives will behave classically. As for the modal operators, the interest in modal logic lies in the infinitely many possibilities of defining their behaviour. First of all, according to the theory outlined in Chapter 1, a *modal logic* must be a relation $\vdash \subseteq \wp(\mathcal{P}_\kappa) \times \mathcal{P}_\kappa$ satisfying (ext.), (mon.), (cut.), (sub.) and (cmp.). Moreover, we generally assume that the boolean connectives behave as in boolean logic. There is a special set of consequence relations — by no means the only ones — which have a deduction theorem for $\to$. Such consequence relations are fully determined by their sets of tautologies. Indeed, it is standard practice to identify modal logics with their set of tautologies. We will stick to that tradition; however, we will see in Section 3.1 that for a given set of tautologies there exist other consequence relations with useful properties. Thus we call a set $\Lambda \subseteq \mathcal{P}_\kappa$ a **modal logic** if Λ contains all tautologies of boolean logic, is closed under substitution and modus ponens, that is, if $\varphi \in \Lambda$ then $\varphi^\sigma \in \Lambda$ for a substitution σ, and if $\varphi \in \Lambda$ and $\varphi \to \psi \in \Lambda$ then $\psi \in \Lambda$. The relation $\vdash_\Lambda$ is then defined via

(cmp.) $\Delta \vdash_\Lambda \varphi$ iff there is a finite set $\Delta_0 \subseteq \Delta$ such that $ded(\Delta_0, \varphi) \in \Lambda$.

Let a logic Λ be given. Fix an operator $\Box$ of the language for Λ. $\Box$ is called **classical** in Λ if the rule (cl$\Box$.) is admissible; if (mo$\Box$.) is admissible in Λ, $\Box$ is called **monotone** in Λ. Finally, if (mn.) is admissible, and if Λ contains the axiom of *box distribution*, which is denoted by (bd$\to$.) $\Box$ is called **normal** in Λ.

(cl$\Box$.) $\dfrac{\vdash \varphi \leftrightarrow \psi}{\vdash \Box\varphi. \leftrightarrow .\Box\psi}$ **(mo$\Box$.)** $\dfrac{\vdash \varphi \to \psi}{\vdash \Box\varphi. \to .\Box\psi}$ **(mn.)** $\dfrac{\vdash \varphi}{\vdash \Box\varphi}$

(bd$\to$.) $\vdash \Box(\varphi \to \psi). \to .\Box\varphi \to \Box\psi$

A normal operator $\Box$ of Λ is monotone in Λ; a monotone operator of Λ is classical in Λ. A *logic* is **classical** (**monotone**, **normal**) if all its operators are classical (monotone, normal). Two formulae φ, ψ are said to be **deductively** or (**locally**) **equivalent** in a logic Λ if $\varphi \leftrightarrow \psi \in \Lambda$. Classical logics have the property that ψ_1 and ψ_2 are deductively equivalent in Λ, if ψ_2 results from ψ_1 by replacing in ψ_1 an occurrence of a subformula by a deductively equivalent formula.

Proposition 2.1.1. *Let Λ be a classical modal logic and $\varphi_1 \leftrightarrow \varphi_2 \in \Lambda$. Let ψ_1 be any formula and let ψ_2 result from replacing an occurrence of φ_1 in ψ_1 by φ_2. Then $\psi_1 \leftrightarrow \psi_2 \in \Lambda$.*

Proof. Notice that the following rules are admissible in addition to **(cl$\Box$.)**, by the axioms and rules of boolean logic.

$$(\mathbf{cl}\wedge.)\ \frac{\vdash \alpha_1 \leftrightarrow \alpha_2 \quad \vdash \beta_1 \leftrightarrow \beta_2}{\vdash \alpha_1 \wedge \beta_1 \leftrightarrow \alpha_2 \wedge \beta_2} \qquad (\mathbf{cl}\neg.)\ \frac{\vdash \alpha_1 \leftrightarrow \alpha_2}{\vdash \neg\alpha_1 \leftrightarrow \neg\alpha_2}$$

By induction on the constitution of ψ_1 it is shown that $\psi_1 \leftrightarrow \psi_2 \in \Lambda$, starting with the fact that $\varphi_1 \leftrightarrow \varphi_2 \in \Lambda$ and $p \leftrightarrow p \in \Lambda$ the claim follows. ◻

Proposition 2.1.2. *Let Λ be a classical modal logic. Then for any formula φ there exists a formula ψ which is deductively equivalent to φ and is composed from variables and negated variables, $\bot$ and $\top$ using only $\wedge$, $\vee$, $\Box_j$ and $\Diamond_j$, $j < \kappa$.*

For this theorem it makes no difference whether the symbols $\bot$, $\vee$ and $\Diamond_j$, $j < \kappa$, are new symbols or merely abbreviations. The **dual** of a formula is defined as follows.

$$\begin{array}{lcl} p^d & := & p \\ (\top)^d & := & \bot \\ (\bot)^d & := & \top \\ (\varphi \wedge \psi)^d & := & \varphi^d \vee \psi^d \\ (\varphi \vee \psi)^d & := & \varphi^d \wedge \psi^d \\ (\Box_j\varphi)^d & := & \Diamond_j\varphi^d \\ (\Diamond_j\varphi)^d & := & \Box_j\varphi^d \end{array}$$

The dual is closely related to negation. Recall, namely, that in boolean algebra negation is an isomorphism between the algebras $\langle A, -, \cap, \cup\rangle$ and $\langle A, -, \cup, \cap\rangle$. The following theorem — whose proof is an exercise — states that for axioms there typically are two forms, one dual to the other.

Proposition 2.1.3. *Let Λ be a classical modal logic. Then $\varphi \to \psi \in \Lambda$ iff $\psi^d \to \varphi^d \in \Lambda$.*

In monotone logics we can prove that an alternative version of the (bd$\to$.) postulate, (bd$\wedge$.), is equivalent.

(bd$\wedge$.) $\qquad \vdash \Box(\varphi \wedge \psi). \leftrightarrow .\Box\varphi \wedge \Box\psi$

Proposition 2.1.4. *Let Λ be a classical modal logic, and $\Box$ be a modal operator of the language of Λ. (1.) (mn.) is admissible for an operator $\Box$ in Λ if $\Box\top \in \Lambda$. (2.) If Λ contains (bd$\to$.) and $\Box\top$ then $\Box$ is normal in Λ. (3.) If $\Box$ is monotone in Λ, the postulates (bd$\wedge$.) and (bd$\to$.) are interderivable in Λ.*

Proof. **(1.)** By assumption, $\Box\top \in \Lambda$ and therefore $\vdash \Box\top \leftrightarrow \top$. Now assume $\vdash \varphi$. Then $\vdash \varphi \leftrightarrow \top$ and so, by (cl$\Box$.), $\vdash \Box\varphi \leftrightarrow \Box\top$. Using this equivalence we get $\vdash \Box\varphi \leftrightarrow \top$, that is, $\vdash \Box\varphi$, as required. **(2.)** By (1.), (mn.) is admissible. **(3.)** Assume (bd$\to$.) is in Λ. Then, since $\vdash \varphi \wedge \psi. \to .\varphi$ and $\vdash \varphi \wedge \psi. \to .\psi$ we have, by (mo$\Box$.), $\vdash \Box(\varphi \wedge \psi) \to \Box\varphi$ and $\vdash \Box(\varphi \wedge \psi) \to \Box\psi$. Now by boolean logic we get $\vdash \Box(\varphi \wedge \psi). \to .\Box\varphi \wedge \Box\psi$. Next, since it holds that $\vdash \varphi. \to .\psi \to (\varphi \wedge \psi)$, with (mo$\Box$.) we get $\vdash \Box\varphi. \to .\Box(\psi \to (\varphi \wedge \psi))$. Using (bd$\to$.) and some boolean equivalences we get $\vdash \Box\varphi. \to .\Box\psi \to \Box(\varphi \wedge \psi)$, from which the remaining implication of (bd∧.) follows. Now assume (bd∧.) is in Λ. Then $\vdash \Box(\varphi \to \psi) \wedge \Box\varphi. \to .\Box(\varphi \to \psi. \wedge .\varphi)$. Furthermore, we get $\vdash \Box(\varphi \to \psi.\wedge.\varphi) \to \Box\psi$, by (mo□.). Hence $\vdash \Box(\varphi \to \psi)\wedge\Box\varphi. \to .\Box\psi$, which is equivalent to (bd→.). □

The smallest classical modal logic will be denoted by $\mathbf{E}_\kappa$, the smallest monotone logic by $\mathbf{M}_\kappa$ and the smallest normal modal logic by $\mathbf{K}_\kappa$ (after Saul Kripke). The index κ is dropped when $\kappa = 1$, or whenever no confusion arises. Notice that since these logics are determined by their theorems, it is enough to axiomatize their theorems. This is different from an axiomatization of the proper rules (see Section 3.9). Notice, namely, that when our interest is only in axiomatizing the theorems, we can do this using the *admissible* rules for deriving theorems. A *classical logic* (*monotone logic*) can be identified with its set Λ of theorems, which is a set containing all boolean tautologies and which is closed under modus ponens, substitution and (cl□.) or, for monotone logics, (mo□.). A **quasi–normal logic** is a modal logic containing $\mathbf{K}_\kappa$. A quasi–normal logic is normal iff it is closed under (mn.). Notice that in a normal logic, the rule (mo□.) is in fact derivable. The smallest normal κ–modal logic containing a set X of formulae is denoted by $\mathbf{K}_\kappa(X)$, $\mathbf{K}_\kappa.X$ or $\mathbf{K}_\kappa \oplus X$, depending on the taste of authors and the circumstances. The smallest quasi–normal logic containing a set X is denoted by $\mathbf{K}_\kappa + X$. Similarly, if Λ is a (quasi–)normal logic then the result of adding the axioms X (quasi–)normally is denoted by $\Lambda \oplus X$ ($\Lambda + X$). In particular, there is a list of formulae that have acquired a name in the past, and modal logics are denoted by a system whereby the axioms are listed by their names, separated mostly by a dot. For example, there are the formulae $\mathbf{D} = \Diamond\top$, $\mathbf{4} = \Diamond\Diamond p \to \Diamond p$. The logic $\mathbf{K}(\Diamond\top)$ is denoted by **KD** or also **K**.**D**, the logic $\mathbf{K}(\Diamond\Diamond p \to \Diamond p)$ is denoted by **K4**, and so forth. We will return to special systems in Section 2.5.

Let us now turn to the calculi for deriving theorems in modal logic. Let a logic be axiomatized over the system $\mathbf{K}_\kappa$ by the set of formulae X, that is, consider the logic now denoted by $\mathbf{K}_\kappa \oplus X$. For deducing theorems we have the following calculus. (We write $\vdash \varphi$ for the fact $\varphi \in \mathbf{K}_\kappa \oplus X$. Also, $\vdash_{\mathbf{BC}} \varphi$ means that φ is a substitution instance of a formula derivable in the calculus of boolean logic.)

(bc.) $\vdash \varphi$ if $\vdash_{\mathbf{BC}} \varphi$

(bd→.) $\vdash \Box_i(\varphi \to \psi). \to .\Box_i\varphi \to \Box_i\psi \quad (i < \kappa)$

(ax.) $\vdash \varphi$ for all $\varphi \in X$

(mp.) $\dfrac{\vdash \varphi \to \psi \quad \vdash \varphi}{\vdash \psi}$

(sb.) $\dfrac{\vdash \varphi}{\vdash \varphi^\sigma}$

(mn.) $\dfrac{\vdash \varphi}{\vdash \Box_i\varphi}\ (i < \kappa)$

Thus, there are axioms (bc.), (ax.), (bd→.) and the rules (mp.), (sb.) and (mn.) (for all operators). Notice that the way in which we have stated the rules they are actually so called *rule schemata*. The difference is that while in an ordinary rule one uses propositional variables, here we use metavariables for formulae. Thus, with an ordinary rule the use of a rule schema like $\vdash \varphi / \vdash \Box_i\varphi$ is justified from the rule $\vdash p_0 / \vdash \Box_i p_0$ by uniform substitution of φ for p_0. In view of the fact proved below that any proof can be rearranged in such a way that the substitutions are placed at the beginning, we can eliminate the substitutions since we have rule schemata.

Although the ordering of the application of the deductive rules (mp.), (sb.) and (mn.) is quite arbitrary it is actually the case that any proof can be reorganized in quite a regular way by rearranging the rules. Consider an application of (mn.) after (mp.) as in the left hand side below. There is an alternative proof of $\vdash \Box\psi$ in which the order is reversed. This proof is shown to the right.

$$\dfrac{\dfrac{\vdash \varphi \quad \vdash \varphi \to \psi}{\vdash \psi}}{\vdash \Box\psi}$$

$$\dfrac{\dfrac{\vdash \varphi}{\vdash \Box\varphi} \qquad \dfrac{\dfrac{\vdash \varphi \to \psi}{\vdash \Box(\varphi \to \psi)} \qquad \vdash \Box(\varphi \to \psi). \to .\Box\varphi \to \Box\psi}{\vdash \Box\varphi \to \Box\psi}}{\vdash \Box\psi}$$

This shows that (mn.) can be moved above (mp.). However, this does not yet show that *all* applications of (mn.) can be moved from below applications of (mp.). A correct proof of this fact requires more than showing that the derivations can be permuted. One also needs to show that the procedure of swapping rule applications will eventually terminate after finitely many steps. In this case this is not hard to prove. Observe that the depth in the proof tree of the particular application of (mn.) decreases. Namely, if the depth of $\vdash \varphi$ is i and the depth of $\vdash \varphi \to \psi$ is j then the depth of $\vdash \psi$ is $max\{i, j\} + 1$ and so the depth of $\vdash \Box\psi$ is $max\{i, j\} + 2$. In the second tree, the sequent $\vdash \Box\varphi$ has depth $i + 1$ and the sequent $\vdash \Box(\varphi \to \psi)$ is $j + 1$. Both are smaller than $max\{i, j\} + 2$. Let the deepest applications of (mn.) be of depth δ. By starting with applications of depth δ we produce applications of depth $< \delta$, so the instances of (mn.) of depth δ can all be eliminated in favour of (twice as many) applications of depth $< \delta$. Now it is clear that the reduction will terminate.

Next we look at substitution. The place of (sb.) in the derivation can be changed quite arbitrarily. This is due to the fact that our rules are *schematic*, they are operative not only on the special formulae for which they are written down but for all substitution instances thereof. So, if we apply a certain rule, deriving a formula φ and apply (sb.) with the substitution σ, then in fact we could have derived φ^σ directly by applying (sb.) on all premises of the rule and then using the rule.

$$\dfrac{\dfrac{\vdash \varphi \quad \vdash \varphi \to \psi}{\vdash \psi}}{\vdash \psi^\sigma} \qquad\qquad \dfrac{\dfrac{\vdash \varphi}{\vdash \varphi^\sigma} \qquad \dfrac{\vdash \varphi \to \psi}{\vdash (\varphi \to \psi)^\sigma}}{\vdash \psi^\sigma}$$

Notice that $(\varphi \to \psi)^\sigma$ is the same as $\vdash \varphi^\sigma \to \psi^\sigma$. Finally, note that (mn.) can be permuted with (sb.), that is, the two derivations below are equivalent.

$$\dfrac{\dfrac{\vdash \varphi}{\vdash \varphi^\sigma}}{\vdash \Box_j \varphi^\sigma} \qquad\qquad \dfrac{\dfrac{\vdash \varphi}{\vdash \Box_j \varphi}}{\vdash (\Box_j \varphi)^\sigma}$$

Notice that $(\Box\varphi)^\sigma = \Box(\varphi^\sigma)$, so the second derivation is legitimate. We have now established that (sb.) can always be moved at the beginning of the proof. Hence it can be eliminated altogether, as we have remarked earlier. The proof that this commutation terminates is here simpler than in the first case.

We can now prove an important theorem about the relation between quasi–normal closure and normal closure. To state it properly, we introduce the following important bit of notation. Given a set Δ of formulae, put

$$\begin{array}{lcl} \boxtimes^0\Delta & := & \Delta \\ \boxtimes\Delta & := & \{\Box_i\delta : i < \kappa, \delta \in \Delta\} \\ \boxtimes^{k+1}\Delta & := & \boxtimes(\boxtimes^k\Delta) \\ \boxtimes^{\leq k}\Delta & := & \bigcup_{j \leq k} \boxtimes^j\Delta \\ \boxtimes^\omega\Delta & := & \bigcup_{k\in\omega} \boxtimes^k\Delta \end{array}$$

The notation $\boxtimes^{(k)}\Delta$ is also used for $\boxtimes^{\leq k}\Delta$. In all these definitions, Δ may also be replaced by a single formula. If, for example, $\kappa = 2$ then

$$\begin{array}{lcl} \boxtimes^1\varphi & = & \{\Box_0\varphi, \Box_1\varphi\} \\ \boxtimes^2\varphi & = & \{\Box_0\Box_0\varphi, \Box_0\Box_1\varphi, \Box_1\Box_0\varphi, \Box_1\Box_1\varphi\} \end{array}$$

$\boxtimes^\omega\Delta$ is effectively the closure of Δ under all rules (mn.) for each operator.

Theorem 2.1.5. *Let Λ be a normal logic. Then $\Lambda \oplus \Delta = \Lambda + \boxtimes^\omega\Delta$. Moreover, $\psi \in \Lambda \oplus \Delta$ iff ψ can be derived from $\Lambda \cup \boxtimes^\omega\Delta^s$ by using modus ponens alone, where Δ^s is the closure of Δ under substitution.*

Proof. We can arrange it that a derivation uses (mn.) only at the beginning of the proof. (mn.) is a unary rule, so it can under these circumstances only have been applied iteratively to an axiom. Likewise, the use of substitution can be limited to the beginning of the proof. ∎

In case $\boxtimes^k\Delta$ is finite, we also write $\boxtimes^k\Delta$ in place of $\bigwedge \boxtimes^k\Delta$ and $\boxtimes^{\leq k}\Delta$ in place of $\bigwedge \boxtimes^{\leq k}\Delta$. Here, for a finite set Γ of formulae, $\bigwedge \Gamma$ simply denotes the formula $\bigwedge\langle\gamma : \gamma \in \Gamma\rangle$. However, notice that the latter definition is not unambigous. First, we need to fix an enumeration of Γ, say, $\Gamma = \{\gamma_i : i < n\}$. Next, we let $\bigwedge \Gamma := \bigwedge_{i<n} \gamma_i$. The latter is defined inductively by

$$\begin{array}{lcl} \bigwedge_{i<0} \gamma_i & := & \top \\ \bigwedge_{i<1} \gamma_i & := & \gamma_0 \\ \bigwedge_{i<2} \gamma_i & := & \gamma_0 \wedge \gamma_1 \\ \bigwedge_{i<n+1} \gamma_i & := & (\bigwedge_{i<n} \gamma_i) \wedge \gamma_n \end{array}$$

In the last line, $n \geq 2$. Technically speaking, $\bigwedge \Gamma$ depends on the enumeration of Γ. This will not matter as long as we deal with formulae up to deductive equivalence.

However, in syntactic manipulations (for example, normal forms) the differences need to be dealt with. Here we fix beforehand a well–order on the set of formulae. This well–order defines a unique order on Γ.

Finally, we introduce the notion of a *compound modality*. Suppose that φ is a formula such that $\sharp var(\varphi) = 1$ (that is, φ contains occurrences of a single sentence letter) and which is built using $\wedge$ and $\Box_j$, $j < \kappa$. Then φ is called a **compound modality**. For example, $\Box_0\Box_1 p \wedge \Box_1\Box_0 p$ is a compound modality, and so is $\Box_0(p \wedge \Box_1 p)$. We say that φ is **normal** in Λ if from $\psi \in \Lambda$ we may infer $\varphi(\psi) \in \Lambda$, and if $\varphi(\psi_1 \to \psi_2). \to .\varphi(\psi_1) \to \varphi(\psi_2) \in \Lambda$. So, a normal compound modality satisfies the same axioms and rules as a normal operator, except that it is not necessarily a primitive symbol of the language.

Proposition 2.1.6. *Let $\varphi(p)$ be a compound modality. If all operators occurring in φ are normal, so is φ.*

Proof. We show by induction that the compound modalities admit (mn.) and satisfy (bd$\wedge$.). We leave it to the reader to verify that the operators are classical. Let $\varphi(p) = \psi_1(p) \wedge \psi_2(p)$. Assume that ψ_1 and ψ_2 are normal. Then

$$\begin{array}{rcl} \varphi(p \wedge q) & \vdash & \psi_1(p \wedge q) \wedge \psi_2(p \wedge q) \\ & \vdash & \psi_1(p) \wedge \psi_1(q) \wedge \psi_2(p) \wedge \psi_2(q) \\ & \vdash & \psi_1(p) \wedge \psi_2(p). \wedge .\psi_1(q) \wedge \psi_2(q) \\ & \vdash & \varphi(p) \wedge \varphi(q) \end{array}$$

and conversely. Thus φ satisfies (bd$\wedge$.). Now assume $\vdash \chi$. Then $\vdash \psi_1(\chi)$ and $\vdash \psi_2(\chi)$, by assumption. Thus $\vdash \psi_1(\chi) \wedge \psi_2(\chi)$, that is, $\vdash \varphi(\chi)$, as required. Now let $\varphi = \Box_j\psi$. Then we have

$$\varphi(p \wedge q) \vdash \Box_j\psi(p \wedge q) \vdash \Box_j(\psi(p) \wedge \psi(q)) \vdash \Box_j\psi(p) \wedge \Box_j\psi(q)\ ,$$

and conversely. Furthermore, from $\vdash \chi$ we may conclude $\vdash \psi(\chi)$, by normality of ψ, and $\vdash \Box_j\psi(\chi)$, by normality of $\Box_j$. ∎

It follows that a compound modality can be rewritten into the form $\bigwedge_{i<n} \psi_i(p)$, where each $\psi_i(p)$ consists just of a string of modal operators prefixed to the variable. We will use the following notation for such operators. If σ is a finite sequence in κ write $\Box^\sigma$ for the operator prefix obtained in the following way.

$$\Box^j := \Box_j \qquad \Box^{j^\frown\sigma} := \Box_j\Box^\sigma$$

Moreover, for the empty sequence ϵ, $\Box^\epsilon$ will be the empty prefix. So, $\Box^\epsilon\varphi = \varphi$ (the two are syntactically equal). Modulo deductive equivalence in $\mathbf{K}_\kappa$ any compound modality is of the form $\bigwedge_{i<k} \Box^{\sigma_i} p$. Let s be a finite set of finite sequences in κ. Then

$$\Box^s p := \bigwedge_{\tau \in s} \Box^\tau p$$

The following theorem is easily proved.

Proposition 2.1.7. *Any compound modality is deductively equivalent to a compound modality of the form $\Box^s p$, where s a finite set of finite sequences of indices.*

To make life easy, we use $\boxplus$ as a variable for an arbitrary compound modality. With Y a set we write $\boxplus Y$ for $\{\boxplus\delta : \delta \in Y\}$.

Proposition 2.1.8. *Let Λ be a normal modal logic, X a set. Then $\psi \in \Lambda \oplus X$ iff there is a compound modality $\boxplus$ and a finite set $Y \subseteq X^s$, X^s the closure of X under substitution, such that $\boxplus Y \vdash_\Lambda \psi$.*

Proof. Clearly, we know that if $\psi \in \Lambda \oplus X$ then ψ can be derived from Λ and a finite subset $Y \subseteq X^s$ using modus ponens and (mn.). Then ψ can be derived from Λ and a finite set Z of formulae of the form $\boxplus_i \chi_i$, $i < m$, $\chi_i \in Y$, using only (mp.). Put $\boxplus p := \bigwedge_{i<m} \boxplus_i p$. Then $\boxplus \chi_i \vdash_{\mathbf{K}_\kappa} \boxplus_i \chi_i$. Thus the compound modality $\boxplus$ and the set $\{\chi_i : i < m\}$ fulfill the requirements. ∎

Let $\boxplus_1$ and $\boxplus_2$ be two compound modalities. We write $\boxplus_1 \leq_\Lambda \boxplus_2$ if $\boxplus_2 p \to \boxplus_1 p \in \Lambda$. It is not hard to see that this ordering is transitive and reflexive. Put $\boxplus_1 \sim_\Lambda \boxplus_2$ if both $\boxplus_1 \leq_\Lambda \boxplus_2$ and $\boxplus_2 \leq_\Lambda \boxplus_1$. We define

$$\begin{array}{rcl}(\boxplus_1 \cup \boxplus_2)(p) & := & \boxplus_1 p \wedge \boxplus_2 p \\ (\boxplus_1 \circ \boxplus_2)(p) & := & \boxplus_1(\boxplus_2 p)\end{array}$$

Proposition 2.1.9. *The compound modalities, factored by the equivalence $\sim_\Lambda$, form a semilattice with respect to $\cup$ and a monoid with respect to $\circ$ and $\Box^\epsilon$. Moreover, the following distribution law holds.*

$$\boxplus_1 \circ (\boxplus_2 \cup \boxplus_3) \sim_\Lambda (\boxplus_1 \circ \boxplus_2) \cup (\boxplus_1 \circ \boxplus_3)$$

The proof of this theorem is straightforward.

Exercise 30. Show that in classical logic, the formula $ded(\Psi, \varphi)$ defined in Section 1.6 is equivalent to $\bigwedge_{\psi\in\Psi} \psi. \to .\varphi$.

Exercise 31. Show that *negation* is a classical modal operator but not monotone. Show also that $\Box_i\varphi \leftrightarrow \neg\Diamond_i\neg\varphi$ is a theorem in $\mathbf{E}_\kappa$.

Exercise 32. Show that in a classical logic, (mo$\Box$.) is interderivable with the rule (mo$\Diamond$.).

(mo$\Diamond$.) $\dfrac{\vdash \varphi \to \psi}{\vdash \Diamond\varphi \to \Diamond\psi}$

Exercise 33. Prove Propositions 2.1.2 and 2.1.3.

Exercise 34. The least modal logic is the logic which is just the closure of the tautologies of boolean logic under substitution and modus ponens. Show that in the minimal modal logic φ is a theorem iff there exists a substitution σ such that $\varphi = \psi^\sigma$

for some non–modal, boolean tautology ψ. Show further that $X \vdash \varphi$ iff there exists a substitution σ such that $X = Y^\sigma$ and $\varphi = \psi^\sigma$ for some non–modal Y and ψ such that $Y \vdash \psi$. Thus the minimal modal logic is decidable. *Hint.* Show that in each proof subformulae of the form $\Box_j\varphi$ can be replaced by a variable p_i for some i.

Exercise 35. Show that all compound modalities are classical if the basic operators are classical.

2.2. Modal Algebras

From the general completeness theorems for logics we conclude that for modal logics of any sort there is a semantics based on term algebras and deductively closed sets. Moreover, consequence relations for modal logics of any kind are determined by matrices of the form $\langle \mathfrak{Tm}(var), \Delta\rangle$, where Δ is deductively closed, or, to be more general, by matrices $\langle \mathfrak{A}, D\rangle$ where $\mathfrak{A} = \langle A, 1, -, \cap, \langle \blacksquare_i : i \in \kappa\rangle\rangle$ is an algebra of an appropriate signature and D a deductively closed set. We focus here on algebras whose reduct to the boolean operations is a boolean algebra. Call an **expanded boolean algebra** an algebra $\langle A, 1, -, \cap, F\rangle$, where $\langle A, 1, -, \cap\rangle$ is a boolean algebra and F a set of functions. We are interested in the question when a modal logic is complete with respect to matrices over expanded boolean algebras.

Theorem 2.2.1. *Let Λ be a classical modal logic. Then Λ is determined by a class of reduced matrices over expanded boolean algebras.*

Proof. Let $\mathfrak{Tm}(var)$ be the algebra of terms. Define an equivalence relation Θ on the set of formulae by $\varphi\ \Theta\ \psi$ iff $\varphi \leftrightarrow \psi \in \Lambda$. By Proposition 2.1.1, Θ is a congruence relation on $\mathfrak{Tm}(var)$. Moreover, if δ is a formula and Δ a deductively closed set then either $[\delta]\Theta \cap \Delta = \varnothing$ or $[\delta]\Theta \subseteq \Delta$. Hence Θ is admissible in any matrix $\mathfrak{M} := \langle \mathfrak{Tm}(var), \Delta\rangle$, where Δ is a deductively closed set with respect to $\vdash_\Lambda$. $\mathfrak{M}/\Theta$ is an expanded boolean algebra. Now, by the results of Section 1.5, $\vdash_\Lambda$ is the intersection of all $\vdash_{\mathfrak{M}}$, where $\mathfrak{M} = \langle \mathfrak{Tm}(var), \Delta\rangle$, Δ deductively closed in $\vdash_\Lambda$. Hence it is the intersection of $\vdash_{\mathfrak{M}/\Theta}$. ◻

The converse of this implication is not generally valid. This theorem explains the fundamental importance of classical logics in the general theory of modal logic. We can now proceed to stronger logics and translate the conditions that this logic imposes into the language of expanded boolean algebras. For example, the postulate of monotonicity is reflected in the following condition.

(mna.) If $a \leq b$ then $\blacksquare_i a \leq \blacksquare_i b$

In the general theory it has been customary to reserve the term *modal algebra* for expanded boolean algebras which correspond to normal modal logics.

Definition 2.2.2. *A* (**poly–**) **modal algebra** *is an algebra*

$$\mathfrak{A} = \langle A, 1, -, \cap, \langle \blacksquare_i : i < \kappa\rangle\rangle$$

where the following holds.

(1) $\langle A, 1, -, \cap\rangle$ *is a boolean algebra.*

(2) $\blacksquare_i 1 = 1$ *and* $\blacksquare_i(x \cap y) = \blacksquare_i x. \cap .\blacksquare_i y$ *for all* $i < \kappa$ *and* $x, y \in A$.

With κ *given,* $\mathfrak{A}$ *is called* κ**–modal**.

Let $\mathfrak{A}$ and $\mathfrak{B}$ be two boolean algebras. A map $h : A \to B$ with $h(1) = 1$ and $h(x \cap y) = h(x) \cap h(y)$ is called a **hemimorphism**. The name derives from Greek *hemi-* (*half*) and *morphē* (*shape*), just like *homomorphism* is from Greek *homo-* (*same*) and *morphē*. So, the name says that a hemimorphism preserves only half the shape. By definition, then, a modal algebra is a boolean algebra expanded by a set of endo–hemimorphisms. We will expand on this theme in Section 4.5.

The abstract machinery of general logic can provide us now with the canonical definition of a *model*. A model consists of a matrix and a valuation. A matrix is a pair consisting of an algebra and a deductively closed set. In classical logics we have seen that the algebras can be reduced to expanded boolean algebras and that we can choose maximally consistent sets, that is, ultrafilters. The general completeness theorem says that if in a logic Λ we have $X \nvdash_\Lambda \varphi$ then there is a *model* for Λ such that X holds in the model but φ does not.

Definition 2.2.3. *An* **algebraic model** *is a triple* $\mathfrak{M} = \langle\mathfrak{A}, \beta, U\rangle$ *where* $\mathfrak{A}$ *is a modal algebra,* β *a map from the set of variables into A and U an ultrafilter in* $\mathfrak{A}$*. We say that* φ **holds** *in* $\mathfrak{M}$*, in symbols* $\mathfrak{M} \vDash \varphi$*, if* $\overline{\beta}(\varphi) \in U$*. We write* $\langle\mathfrak{A}, U\rangle \vDash \varphi$ *if for all* β*,* $\langle\mathfrak{A}, \beta, U\rangle \vDash \varphi$ *and we write* $\mathfrak{A} \vDash \varphi$ *if for all ultrafilters U and all valuations* β *we have* $\langle\mathfrak{A}, \beta, U\rangle \vDash \varphi$.

Proposition 2.2.4. *Let* $\mathfrak{A}$ *be a modal algebra. Then* $\mathfrak{A} \vDash \varphi$ *iff for every* β*,* $\overline{\beta}(\varphi) = 1$.

Proof. Suppose that $\mathfrak{A} \vDash \varphi$. Then $\langle\mathfrak{A}, \beta, U\rangle \vDash \varphi$ for all valuations β and ultrafilters U. Hence, given β, $\overline{\beta}(\varphi) \in U$ for every ultrafilter. So, $\overline{\beta}(\varphi) = 1$ for all β. Now assume $\mathfrak{A} \nvDash \varphi$. Then there exists a valuation β and an ultrafilter U such that $\overline{\beta}(\varphi) \notin U$. Hence for this β, $\overline{\beta}(\varphi) \neq 1$. ◻

Proposition 2.2.5. *Let* $\mathfrak{M} = \langle\mathfrak{A}, \beta, U\rangle$ *be an algebraic model. Then*

(i.) $\mathfrak{M} \vDash \neg\varphi$ *iff* $\mathfrak{M} \nvDash \varphi$.

(ii.) $\mathfrak{M} \vDash \varphi \wedge \psi$ *iff* $\mathfrak{M} \vDash \varphi$ *and* $\mathfrak{M} \vDash \psi$.

This notion of *algebraic model* was chosen to contrast it with the *geometric models* based on Kripke–models. However, recall from Chapter 1.5 the notion of a unital semantics. A class of matrices based on modal algebras is called a unital semantics in the sense of that definition if it has at most one designated element. Since the set of designated elements is deductively closed, and so is a filter, in an algebra of the class of modal algebras there is always a designated element and it is the unit element, 1. By the Proposition 1.5.6, for a modal logic which has a unital semantics we must have $p; q; \varphi(p) \vdash \varphi(q)$. Putting $\top$ for p we get $q; \varphi(\top) \vdash \varphi(q)$. In particular, for $\varphi(q) := \Box_j q$ we deduce that $q \vdash \Box_j q$. Although this is not a rule of the standard consequence relation, we will show in Chapter 3.1 that for each logic there exists a consequence relation, denoted by $\Vdash_\Lambda$, with the same tautologies, in which the rules $\langle\{p\}, \Box_j p\rangle$, $j < \kappa$, are derived rules. Since we are mostly interested in tautologies only, it is justified to say that there is a unital semantics for modal logic.

For an algebra $\mathfrak{A}$ we write $\mathsf{Th}\,\mathfrak{A} = \{\varphi : \mathfrak{A} \vDash \varphi\}$ and for a class $\mathcal{K}$ of algebras we put $\mathsf{Th}\,\mathcal{K} := \bigcap\langle \mathsf{Th}\,\mathfrak{A} : \mathfrak{A} \in \mathcal{K}\rangle$. Furthermore, for a set Δ of formulae we write $\mathsf{Alg}(\Delta)$ to denote the class of algebras such that $\mathfrak{A} \vDash \Delta$. The operators Th and Alg are antitonic. That means, if $\mathcal{K} \subseteq \mathcal{L}$ then $\mathsf{Th}\,\mathcal{L} \supseteq \mathsf{Th}\,\mathcal{K}$, and if $\Delta \subseteq \Sigma$ then $\mathsf{Alg}\,\Delta \supseteq \mathsf{Alg}\,\Sigma$.

Proposition 2.2.6. *Let Δ be a set of κ–modal formulae, and $\mathcal{K}$ a class of κ–modal algebras. Then the following holds.*

(1) $\Delta \subseteq \mathsf{Th}\,\mathcal{K}$ *iff* $\mathsf{Alg}\,\Delta \supseteq \mathcal{K}$.
(2) $\Delta \subseteq \mathsf{Th}\,\mathsf{Alg}\,\Delta$.
(3) $\mathcal{K} \subseteq \mathsf{Alg}\,\mathsf{Th}\,\mathcal{K}$.
(4) $\mathsf{Alg}\,\Delta = \mathsf{Alg}\,\mathsf{Th}\,\mathsf{Alg}\,\Delta$.
(5) $\mathsf{Th}\,\mathcal{K} = \mathsf{Th}\,\mathsf{Alg}\,\mathsf{Th}\,\mathcal{K}$.

Proof. **(1.)** $\Delta \subseteq \mathsf{Th}\,\mathcal{K}$ iff for all $\mathfrak{A} \in \mathcal{K}$ we have $\mathfrak{A} \vDash \Delta$ iff for all $\mathfrak{A} \in \mathcal{K}$ we have $\mathfrak{A} \in \mathsf{Alg}\,\Delta$ iff $\mathcal{K} \subseteq \mathsf{Alg}\,\Delta$. **(2.)** From $\mathsf{Alg}\,\Delta \subseteq \mathsf{Alg}\,\Delta$ we deduce with (1.) that $\Delta \subseteq \mathsf{Th}\,\mathsf{Alg}\,\Delta$. **(3.)** From $\mathsf{Th}\,\mathcal{K} \subseteq \mathsf{Th}\,\mathcal{K}$ and (1.), $\mathcal{K} \subseteq \mathsf{Alg}\,\mathsf{Th}\,\mathcal{K}$. **(4.)** By (3.), $\mathsf{Alg}\,\Delta \subseteq \mathsf{Alg}\,\mathsf{Th}\,\mathsf{Alg}\,\Delta$. By (2.), $\Delta \subseteq \mathsf{Th}\,\mathsf{Alg}\,\Delta$, and so $\mathsf{Alg}\,\Delta \supseteq \mathsf{Alg}\,\mathsf{Th}\,\mathsf{Alg}\,\Delta$. The two together yield the claim. (5.) Analogous to (4.). ∎

Hence, the maps $\mathcal{K} \mapsto \mathsf{Alg}\,\mathsf{Th}\,\mathcal{K}$ and $\Delta \mapsto \mathsf{Th}\,\mathsf{Alg}\,\Delta$ are closure operators. The closed elements are of the form $\mathsf{Alg}\,\Delta$ and $\mathsf{Th}\,\mathcal{K}$, respectively. The next theorem asserts that the closed elements are *varieties* and *normal modal logics*, as expected.

Proposition 2.2.7. *For all Δ, $\mathsf{Alg}\,\Delta$ is a variety of κ–modal algebras. For all $\mathcal{K}$, $\mathsf{Th}\,\mathcal{K}$ is a normal κ–modal logic.*

Proof. For the first claim it will suffice to show that $\mathsf{Alg}\,\{\varphi\}$ is a variety. For in general, $\mathsf{Alg}\,\Delta = \bigcap_{\varphi\in\Delta} \mathsf{Alg}\,\{\varphi\}$. It is left as an exercise to verify that the intersection of varieties is a variety. So, let φ be given. We have to show that $\mathsf{Alg}\,\{\varphi\}$ is closed under products, subalgebras and homomorphic images. First, if $\mathsf{Th}\,\mathfrak{A}_i \supseteq \Phi$ then also $\mathsf{Th}\,\prod_{i\in I}\mathfrak{A}_i \supseteq \Phi$. For let $\mathfrak{B} := \prod_{i\in I}\mathfrak{A}_i$ and $q_i : \mathfrak{B} \twoheadrightarrow \mathfrak{A}_i$ be the projection onto the ith component. Let γ be a valuation on $\mathfrak{B}$. Then $\beta_i := q_i \circ \gamma$ is a valuation on $\mathfrak{A}_i$, and we have $\overline{\beta}_i(\varphi) = 1$. However, $\overline{\beta}_i(\varphi) = q_i \circ \overline{\gamma}(\varphi)$. Hence, for all $i \in I$, $q_i \circ \overline{\gamma}(\varphi) = 1$. Thus $\overline{\gamma}(\varphi) = 1$. So, $\mathfrak{B} \vDash \varphi$. This shows closure of $\mathsf{Alg}\,\Phi$ under products. Next let $i : \mathfrak{B} \rightarrowtail \mathfrak{A}$ and $\mathfrak{A} \vDash \varphi$. Suppose γ is a valuation into $\mathfrak{B}$. Then $\beta := i \circ \gamma$ is a valuation into $\mathfrak{A}$. By assumption, $\overline{\beta}(\varphi) = 1$. However, $\overline{\beta}(\varphi) = i \circ \overline{\gamma}(\varphi) = 1$; hence $\overline{\gamma}(\varphi) = 1$, since i is injective. Thus $\mathsf{Alg}\,\{\varphi\}$ is closed under subalgebras. Finally, we show that if $h : \mathfrak{A} \twoheadrightarrow \mathfrak{B}$ then $\mathsf{Th}\,\mathfrak{B} \supseteq \mathsf{Th}\,\mathfrak{A}$. Now suppose that γ is a valuation on $\mathfrak{B}$. Take a valuation β such that $h(\beta(p)) = \gamma(p)$. Then $1 = \overline{\beta}(\varphi)$ implies $1 = h(\overline{\beta}(\varphi)) = \overline{\gamma}(\varphi)$. Thus, $\mathsf{Alg}\,\{\varphi\}$ is also closed under homomorphic images; and so it is shown to be a variety. Now, take a class $\mathcal{K}$ of modal algebras. We want to show that its theory is a modal logic. We leave it to the reader to verify that the intersection of modal logics is again a modal logic. Hence we may specialize on the case $\mathcal{K} = \{\mathfrak{A}\}$. Clearly, since $\mathfrak{A}$ is an expanded boolean algebra, $\mathsf{Th}\,\mathfrak{A}$ contains all boolean tautologies. Moreover, by definition, it is a classical logic, contains $\Box_j\top$ and satisfies (bd$\wedge$.). Hence we have a monotonic logic by the first fact, and we have (bd$\to$.) by the equivalence of the latter with (bd$\wedge$.) in monotonic logics. ◻

Let us now note that we have shown that each set of formulae gives rise to a variety of algebras. They can be obtained rather directly by appeal to Theorem 1.5.5. It says that a logic is determined by its reduced matrices. Now define an equivalence $\equiv_\Lambda$ by $\varphi \equiv_\Lambda \psi$ iff $\varphi \leftrightarrow \psi \in \Lambda$. If Λ is classical, this is a congruence. Hence put

$$\mathfrak{Fr}_\Lambda(var) := \mathfrak{Tm}(var)/\equiv_\Lambda$$

The boolean reduct of $\mathfrak{Fr}_\Lambda(var)$ is a boolean algebra. So, the algebra is an expanded boolean algebra.

Lemma 2.2.8. *Let Λ be classical. Then $\mathsf{Th}\,\mathfrak{Fr}_\Lambda(var) = \Lambda$.*

Proof. Suppose $\psi \notin \Lambda$. Then $\psi \not\equiv_\Lambda \top$ and so because for the natural map $v : \psi \mapsto \psi/\equiv_\Lambda$ we have $v(\varphi) \neq 1$. Therefore we also have $\langle \mathfrak{Fr}_\Lambda(var), \{1\}\rangle \nvDash \varphi$. Hence there is an ultrafilter U not containing φ, and for that ultrafilter $\langle \mathfrak{Fr}_\Lambda(var), v, U\rangle \nvDash \varphi$, as required. On the other hand, if $\varphi \in \Lambda$ and $h : \mathfrak{Tm}_{\mathcal{P}}(var) \to \mathfrak{Fr}_\Lambda(var)$ is a homomorphism, then let σ be a substitution defined by $\sigma(p) = \psi_p$ for some ψ_p such that $h(p) = v(\psi_p)$. Then $h(p) = v \circ \sigma(p)$ and so $h = v \circ \sigma$. In particular $h(\varphi) = v \circ \sigma(\varphi)$. Since $\varphi \in \Lambda$ we also have $\sigma(\varphi) \in \Lambda$ by closure under substitutions. Thus $v(\sigma(\varphi)) = 1$, by definition. Consequently, $h(\varphi) = 1$ for all h, showing that $\mathfrak{Fr}_\Lambda(var) \vDash \varphi$. ◻

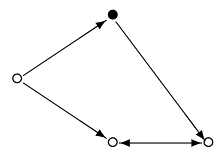
This last theorem is extremely important. It tells us not only that each logic has an adequate set of algebras, it also tells us the following.

Theorem 2.2.9. *The map Alg is a one–to–one map from normal κ–modal logics into the class of varieties of κ–modal algebras.*

Proof. Clearly, we have shown that each set of formulae defines a variety of κ–modal algebras, and each class of κ–modal algebras defines a normal κ–modal logic. Furthermore, for two logics $\Lambda \neq \Theta$ the varieties must be distinct, because either $\Lambda \not\supseteq \Theta$ or $\Theta \not\supseteq \Lambda$. In the first case $\mathfrak{Fr}_\Lambda(var) \notin \mathsf{Alg}\,\Theta$ and in the second case $\mathfrak{Fr}_\Theta(var) \notin \mathsf{Alg}\,\Lambda$. ◻

This shows that algebraic semantics provides enough classes to distinguish logics. We will see in Section 4.2 that distinct varieties give rise to distinct logics, so that the correspondence is actually exact. Notice also the following. The cardinality of the free algebra is at most $\|\mathcal{P}_\kappa\|$. Moreover, for a countermodel of φ we only need to consider finitely generated subalgebras of that algebra.

FIGURE 2.1.



THEOREM 2.2.10. *Let Λ be a κ–modal logic with at most countably many variables, $\kappa > 0$. If $\varphi \notin \Lambda$ then there exists an algebra $\mathfrak{A}$ of cardinality $\leq max\{\aleph_0, \kappa\}$ such that $\mathfrak{A} \nvDash \varphi$. In particular, if κ is at most countable, so is $\mathfrak{A}$.*

Exercise 36. Prove Proposition 2.2.5.

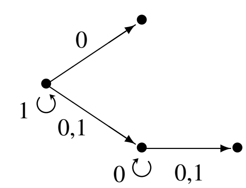
Exercise 37. Let $\mathfrak{B} \cong \prod_{i \in I} \mathfrak{A}_i$. Show that $\mathsf{Th}\, \mathfrak{B} = \bigcap_{i \in I} \mathsf{Th}\, \mathfrak{A}_i$.

Exercise 38. Show that the (possibly infinite) intersection of varieties is a variety again. Likewise, show that the (possibly infinite) intersection of normal modal logics is a normal modal logic again. *Hint.* Use closure operators.

2.3. Kripke–Frames and Frames

The most intuitive semantics for normal (and also quasi–normal) logics is based on relational structures, so–called *frames*. A frame is defined in two stages. First, a **Kripke–frame** for $\mathcal{P}_\kappa$ is a pair $\mathfrak{f} = \langle f, \langle \lhd_i : i < \kappa \rangle \rangle$ where f is a set, called the **set of worlds**, and each $\lhd_i$, $i < \kappa$, is a binary relation. $\lhd_i$ is called an **accessibility relation**. More precisely, $\lhd_i$ is the accessibility relation **associated with** $\Box_i$. It is not required that the set of worlds be nonempty. Frames can be pictured in much the same way as directed graphs. In fact, with only one accessibility relation, the two are one and the same thing. The worlds are denoted by some symbol, say $\bullet$, and the relation is just a collection of arrows pointing from a node x to a node y just in case $x \lhd y$. Some authors do not use arrows; instead they have an implicit convention that the arrows point from left to right or from bottom to top. (This is similar to the conventions for drawing lattices.) Especially when there is only one relation, several shorthand notations are used. First, $\circ$ standardly denotes a reflexive or self–accessible point, while $\bullet$ denotes an irreflexive point. Figure 2.1 illustrates this. There are four points, one is irreflexive. Another convention is to use $\bullet$ for reflexive and x for irreflexive points. In the case of more than one relation, Kripke–frames are the same as edge–coloured directed multigraphs. Technically, colours are realized as indices decorating the arrows. Notice that the shorthands for reflexive and irreflexive points are now insufficient, so it is generally better to use a subscripted turning arrow

FIGURE 2.2.



instead ($\circlearrowleft_1$ for example). In that case we will use $\bullet$ for worlds throughout. A substitute policy for a polymodal Kripke–frame $\mathfrak{f} = \langle f, \langle \lhd_i : i < \kappa \rangle\rangle$ is to present it as a set $\{\mathfrak{f}_i : i < \kappa\}$, $\mathfrak{f}_i = \langle f, \lhd_i \rangle$, of monomodal frames. This makes drawing of the frames simpler, although imagining such a frame is more difficult as compared to the coloured graphs. Each particular pair $\langle x, y \rangle$ such that $x \lhd_j y$ is also called a **j–transition** or simply **transition** of the frame. We also write $x \xrightarrow{j} y$ for the fact that there is a j–transition from x to y. This notation is generalized to sets of sequences s in the following way. Given a set of sequences s the symbols $x \lhd^s y$ and $x \xrightarrow{s} y$ are synonymous. For a sequence τ, $x \lhd^\tau y$ is defined by induction on the length. If $\tau = \epsilon$ (the empty string), $x \lhd^\tau y$ iff $x = y$; if $\tau = \tau_1^\frown j$, $j < \kappa$, then $x \lhd^\tau y$ iff there exists a z such that $x \lhd^{\tau_1} z$ and $z \lhd_j y$. Finally, for $s = \{\tau_i : i < n\}$, where each τ_i is a sequence, $x \lhd^s y$ iff $x \lhd^{\tau_i} y$ for some $i < n$.

Typically, the idea of these pictures is grasped rather quickly once the reader has played with them for a while. Readers who wish to have more examples of Kripke–frames might take a map of the bus connections of some area. The points are the bus stops, and each bus line defines its particular accessibility relation between these stops. A **pointed Kripke–frame** is a pair $\langle \mathfrak{f}, x \rangle$ where $x \in f$. A model based on $\mathfrak{f}$ consists of two more things, a valuation and a specification of a special reference world. A **valuation** is a function $\beta : var \to 2^f$. A **Kripke–model** is a triple $\mathfrak{M} := \langle \mathfrak{f}, \beta, x \rangle$, where $x \in f$ and β is a valuation on $\mathfrak{f}$. For every proposition in $\mathcal{P}_\kappa$ we can now say whether it is accepted or rejected by the model. This is defined formally as follows.

(md0.)	$\langle \mathfrak{f}, \beta, x \rangle \vDash p$	iff	$x \in \beta(p)$
(md¬.)	$\langle \mathfrak{f}, \beta, x \rangle \vDash \neg\varphi$	iff	$\langle \mathfrak{f}, \beta, x \rangle \nvDash \varphi$
(md∧.)	$\langle \mathfrak{f}, \beta, x \rangle \vDash \varphi \wedge \psi$	iff	$\langle \mathfrak{f}, \beta, x \rangle \vDash \varphi$ and $\langle \mathfrak{f}, \beta, x \rangle \vDash \psi$
(md□$_i$.)	$\langle \mathfrak{f}, \beta, x \rangle \vDash \Box_i\varphi$	iff	for all y with $x \lhd_i y$ $\quad \langle \mathfrak{f}, \beta, y \rangle \vDash \varphi$

Notice that when defining the model condition for a formula φ it is not necessary to assume that β is defined on all variables; all that is needed is that it is defined on all variables of φ. From a theoretical point of view it is mostly preferrable to assume β to be a total function. However, for practical computation and decidability proofs (see Section 2.6) we will prefer β to be a partial function, defined at least on the relevant variables. If β is a partial function we also say that it is a **partial valuation**.

Given a valuation β, we can extend β to a map $\overline{\beta}$ which assigns to each formula the set of worlds in which it is accepted.

(hex.) $\overline{\beta}(\varphi) = \{x : \langle \mathfrak{f}, \beta, x\rangle \vDash \varphi\}$.

It is not hard to see that $\overline{\beta}(\neg\varphi) = f - \overline{\beta}(\varphi)$ and that $\overline{\beta}(\varphi \wedge \psi) = \overline{\beta}(\varphi) \cap \overline{\beta}(\psi)$. Therefore, $\overline{\beta}$ is a boolean homomorphism from the boolean algebra of modal propositions into the powerset algebra of f. Moreover,

$$\overline{\beta}(\Box_i \varphi) = \{x : (\forall y)(x \lhd_i y. \Rightarrow .\langle \mathfrak{f}, \beta, y\rangle \vDash \varphi)\}.$$

Thus define the following operation on subsets of f

(alg□.) $\blacksquare_i a := \{x : (\forall y)(x \lhd_i y. \Rightarrow .y \in a)\}$.

Then $\overline{\beta}(\Box_j \varphi) = \blacksquare_j \overline{\beta}(\varphi)$.

Definition 2.3.1. *A (**polymodal**) **frame** is a pair $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ where $\mathfrak{f} = \langle f, \langle \lhd_i : i < \kappa\rangle\rangle$ is a Kripke–frame and $\mathbb{F}$ a set of subsets of f such that $\langle \mathbb{F}, f, -, \cap, \langle \blacksquare_i : i < \kappa\rangle\rangle$ is a polymodal algebra. Alternatively, $\mathbb{F}$ is a set of subsets closed under boolean operations and the operations $\blacksquare_i$ defined via (alg□.) on the basis of the relations $\lhd_i$ underlying $\mathfrak{f}$. A **pointed frame** is a pair $\langle \mathfrak{F}, x\rangle$ where $\mathfrak{F}$ is a frame, and $x \in f$ a world.*

Standardly, frames in our sense are called *generalized frames*. For the purpose of this book, however, we want to drop the qualifying phrase *generalized*. This has several reasons. First, the nongeneralized counterparts are called *Kripke–frames*, and so there will never be a risk of confusion. Second, from the standpoint of Duality Theory it is generalized frames and not Kripke–frames that we should expect as natural structures. (See Chapter 4.) And third, given that there exist numerous incomplete logics it is not a luxury but simply a necessity to use generalized frames in place of Kripke–frames.

In a frame $\langle \mathfrak{f}, \mathbb{F}\rangle$, a set $a \subseteq f$ is called **internal** or a **situation** if $a \in \mathbb{F}$. If $a \notin \mathbb{F}$, a is **external**. So, a frame combines two things in one structure: a Kripke–frame and an algebra. Due to the presence of the underlying relational structure, it is unnecessary to specify the operations of that algebra and so it shows up in an impoverished form only as a set of sets. From an ideological standpoint we may call frames also *realizations* of algebras. More on that in Section 4.6. A **valuation** into a frame is a valuation into the underlying Kripke frame which assigns only *internal sets* as values. Since the set $\mathbb{F}$ is closed under all relevant operations, the following is proved by induction on the structure of the formula.

Proposition 2.3.2. *Let $\langle \mathfrak{f}, \mathbb{F}\rangle$ be a frame and $\beta : var \to 2^f$ be a valuation on $\mathfrak{f}$. If $\beta(p)$ is internal for all $p \in var$ then $\overline{\beta}(\varphi)$ is internal for all φ. Moreover, $\overline{\beta}$ is a homomorphism from $\mathfrak{Tm}_{\mathcal{P}_\kappa}(var)$ to $\langle \mathbb{F}, 1, -, \cap, \langle \blacksquare_j : j < \kappa\rangle\rangle$.*

A (**geometric**) **model** based on the frame $\mathfrak{F}$ is a triple $\langle \mathfrak{F}, \beta, x \rangle$, where β is a valuation and x a world. Notice that models based on frames as well as Kripke–frames are called *geometric* because they use a world to evaluate propositions. Recall that in the *algebraic model* we had ultrafilters; an algebraic model is a triple $\langle \mathfrak{A}, \beta, U \rangle$, where U is an ultrafilter on $\mathfrak{A}$. We will see in Chapter 4 that every modal algebra is isomorphic to an algebra of internal sets over some Kripke–frame. Thus the principal difference between the algebraic and the geometric approach is the use of ultrafilters as opposed to worlds. In that respect an algebraic model is still more flexible. For if we have a world x, we also have a corresponding ultrafilter, $U_x := \{a : x \in a\}$. But not every ultrafilter is of this form. There exist, however, classes of frames where every model based on an ultrafilter has an equivalent model based on a world. These frames are called *descriptive*. More on that in Section 4.6.

Now we come to the interaction between models and logics. Consider a model $\mathfrak{M} = \langle \mathfrak{F}, \beta, x \rangle$. Define $\mathsf{Th}(\mathfrak{M}) := \{\varphi : \mathfrak{M} \vDash \varphi\}$. Then $\mathsf{Th}(\mathfrak{M})$ is closed under modus ponens. Moreover, for each φ we have either $\varphi \in \mathsf{Th}(\mathfrak{M})$ or $\neg\varphi \in \mathsf{Th}(\mathfrak{M})$ but not both, by (md¬.). So, the set $\mathsf{Th}(\mathfrak{M})$ is a theory (by mp–closure) and a *maximally consistent* theory. Now, suppose we abstract away from the valuation; that is, we take the pointed frame $\langle \mathfrak{F}, x \rangle$ and define

(tpf.) $\quad \langle \mathfrak{F}, x \rangle \vDash \varphi \quad \Leftrightarrow \quad$ for all valuations β we have $\langle \mathfrak{F}, \beta, x \rangle \vDash \varphi$

Then the theory of the pointed frame $\mathsf{Th}\,\langle \mathfrak{F}, x \rangle = \{\varphi : \langle \mathfrak{F}, x \rangle \vDash \varphi\}$ is still closed under modus ponens; however, we no longer have either $\varphi \in \mathsf{Th}\,\langle \mathfrak{F}, x \rangle$ or $\neg\varphi \in \mathsf{Th}\,\langle \mathfrak{F}, x \rangle$, even though $\varphi \vee \neg\varphi \in \mathsf{Th}\,\langle \mathfrak{F}, x \rangle$. Simply take $\varphi := p$, where p is a variable. However, $\mathsf{Th}\,\langle \mathfrak{F}, x \rangle$ is closed under substitution. Hence, it is a quasi–normal logic, since it contains all booelan tautologies and (bd→.). In a last step we abstract from the worlds and consider the theory of the frame alone.

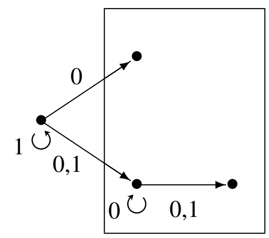
(tfr.) $\quad \mathfrak{F} \vDash \varphi \quad \Leftrightarrow \quad$ for all worlds x and all valuations β $\langle \mathfrak{F}, \beta, x \rangle \vDash \varphi$

$\mathsf{Th}\,\mathfrak{F} := \{\varphi : \mathfrak{F} \vDash \varphi\}$ is called the **theory of** $\mathfrak{F}$. This time we not only have closure under (mp.) and (sb.) but also under (mn.). Suppose, namely, that $\varphi \in \mathsf{Th}\,\mathfrak{F}$. Then for all points and all valuations $\langle \mathfrak{F}, \beta, y \rangle \vDash \varphi$. Take a valuation β and a point x. Since for all y such that $x \lhd_i y$ we already have $\langle \mathfrak{F}, \beta, y \rangle \vDash \varphi$, we now have $\langle \mathfrak{F}, \beta, x \rangle \vDash \Box_i \varphi$. Hence, since both x and β have been chosen arbitrarily, $\Box_i \varphi \in \mathsf{Th}\,\mathfrak{F}$.

THEOREM 2.3.3. *Let $\langle \mathfrak{F}, x \rangle$ be a pointed frame. Then $\mathsf{Th}\,\langle \mathfrak{F}, x \rangle$ is a quasi–normal logic. For all frames $\mathsf{Th}\,\mathfrak{F}$ is a normal logic.*

Now, given a quasi–normal logic Λ and a pointed frame $\langle \mathfrak{F}, x \rangle$, we say that $\langle \mathfrak{F}, x \rangle$ is a **pointed frame for** Λ if $\mathsf{Th}\,\langle \mathfrak{F}, x \rangle \supseteq \Lambda$; likewise, $\mathfrak{F}$ is a **frame for** Λ or a Λ**–frame** if $\mathsf{Th}\,\mathfrak{F} \supseteq \Lambda$. For a given logic Λ we denote the class of Λ–frames by $\mathsf{Frm}(\Lambda)$ and the class of Λ–Kripke–frames by $\mathsf{Krp}(\Lambda)$. We conclude this section with an important theorem concerning the generated substructures. Consider a generalized frame $\langle \mathfrak{f}, \mathbb{F} \rangle$ and an internal set $g \in \mathbb{F}$. g is called **open** if for all $x \in g$ and all y such that $x \lhd_j y$ for some j then also $y \in g$. So, g contains all successors of points contained in g.

FIGURE 2.3.



We put $\lhd_j^g := \lhd_j \cap (g \times g)$ and $\mathfrak{g} := \langle g, \langle \lhd_j^g : j < \kappa \rangle \rangle$. $\mathfrak{g}$ is a Kripke–frame. Finally, $\mathbb{G} := \{a \cap g : a \in \mathbb{F}\} = \{b \subseteq g : b \in \mathbb{F}\}$. $\mathbb{G}$ is closed under relative complements; for if $b \in \mathbb{G}$ then $g - b = g \cap (f - b) \in \mathbb{F}$; $\mathbb{G}$ is also closed under intersection. Furthermore, if $b \in \mathbb{G}$, then

$$\begin{array}{rcl} \blacksquare_j^g b & = & \{x \in g : (\forall y)(x \lhd_j^g y \Rightarrow y \in b)\} \\ & = & \{x \in g : (\forall y)(x \lhd_j y \Rightarrow y \in b)\} \\ & = & g \cap \blacksquare_j b\ , \end{array}$$

since g is successor closed. So the map $b \mapsto g \cap b$ is in fact a homomorphism of the modal algebras. Thus $\mathfrak{G} := \langle \mathfrak{g}, \mathbb{G} \rangle$ is a frame; we say, it is a **generated subframe of** $\mathfrak{F}$. We denote this by $\mathfrak{G} \leq \mathfrak{F}$. In the picture below the box encloses a generated subframe.

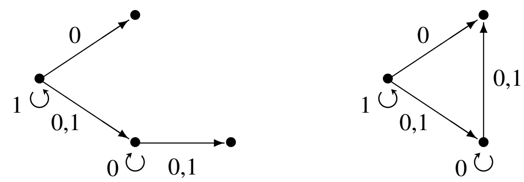
THEOREM 2.3.4. *Let* $\mathfrak{G}$ *be a generated subframe of* $\mathfrak{F}$ *and* $x \in g$. *Then* $\mathsf{Th}\,\langle \mathfrak{G}, x \rangle = \mathsf{Th}\,\langle \mathfrak{F}, x \rangle$. *Moreover,* $\mathsf{Th}\,\mathfrak{G} \supseteq \mathsf{Th}\,\mathfrak{F}$.

Notes on this section. The idea of a Kripke–frame is generally credited to SAUL KRIPKE ([**134**]), though it can already be found in works of RUDOLF CARNAP ([**38**]) and STIG KANGER. Also, BJARNI JÓNSSON and ALFRED TARSKI in [**113**] presented a fully fledged algebraic theory of algebras for modal logic. In it they also show that certain logics have the property that the algebraic structures of that logic are closed under completion (see Section 4.6), which they use to show that these logics are complete with respect to Kripke–frames with certain properties. The notion of a general frame first appeared with S. K. THOMASON in [**206**]. Before that it was customary to use the notion of a *model*, which was just a Kripke–frame together with a valuation. In the language of generalized frames, a model was equivalent to a generalized frame in which the internal sets were exactly the definable sets.

Exercise 39. Prove Theorem 2.3.4.

Exercise 40. Let $\varphi(p)$ be a compound modality. We say that $\varphi(p)$ is *based* on $\lhd$, if for all models: $\langle \mathfrak{F}, \beta, x \rangle \vDash \varphi(p)$ iff for all $y \rhd x$ we have $\langle \mathfrak{F}, \beta, y \rangle \vDash p$. For example,

FIGURE 2.4. A contraction



$\Box_j$ is based on $\lhd_j$. Show that if φ is a compound modality based on $\lhd$ and ψ is a compound modality based on $\blacktriangleleft$ then $\varphi \wedge \psi$ is based on $\lhd \cup \blacktriangleleft$, and $\varphi[\psi/p]$ is based on $\lhd \circ \blacktriangleleft$.

Exercise 41. (Continuing the previous exercise.) Show that $[p?]$ defined by $[p?]q := p \wedge q$ is a normal modal operator. On which relation is this operator based?

2.4. Frame Constructions I

In this section we will introduce a number of ways of creating frames from frames; in particular, these are the *subframes*, *p–morphic images* and *disjoint unions* or *direct sums*. These notions will first be introduced on Kripke–frames and then lifted to (general) frames. The most important notion is that of a *p–morphism*. Let $\pi : f \to g$ be a map. π is a **p–morphism** from $\mathfrak{f}$ to $\mathfrak{g}$, in symbols $\pi : \mathfrak{f} \to \mathfrak{g}$, if the following holds.

(pm1.) For $x, y \in f$: if $x \lhd_j y$ then $\pi(x) \lhd_j \pi(y)$.
(pm2.) For $x \in f$ and $u \in g$: if $\pi(x) \lhd_j u$ then there is a $y \in f$ such that $u = \pi(y)$ and $x \lhd_j y$.

We refer to (pm1.) as the *first condition* and to (pm2.) as the *second condition* on p–morphisms. If π is injective, we write $\pi : \mathfrak{f} \rightarrowtail \mathfrak{g}$. If in addition π is the identity on g, $\mathfrak{g}$ is a generated subframe of $\mathfrak{f}$. If π is surjective we call π a **contraction** and say that $\mathfrak{g}$ is a **p–morphic image** or **contractum** (plural: *contracta*) of $\mathfrak{f}$. We write $\pi : \mathfrak{f} \twoheadrightarrow \mathfrak{g}$. Figure 2.4 provides an example. The Kripke–frame on the right is a contractum of the Kripke–frame on the left. Another example is $\langle \omega, < \rangle$. The one–element Kripke–frame consisting of a reflexive point is a contractum of $\langle \omega, < \rangle$. Each contraction $\pi : \mathfrak{f} \twoheadrightarrow \mathfrak{g}$ induces an equivalence relation $\sim_\pi$ on f by $x \sim_\pi y$ iff $\pi(x) = \pi(y)$. The equivalence relations induced by p–morphisms can be characterized intrinsically as follows. A **net** on $\mathfrak{f}$ is an equivalence relation $\sim$ such that if $x \sim x'$ and $x \lhd_j y$ then there exists a $y' \sim y$ such that $x' \lhd_j y'$. This latter condition is called the **net condition**. Given a net $\sim$, define $[x] := \{x' : x \sim x'\}$, and put $[x] \lhd_j [y]$ iff there exist $x' \in [x]$ and $y' \in [y]$ such that $x' \lhd_j y'$. We denote by $\mathfrak{f}/{\sim}$ the Kripke–frame with worlds $[x]$, $x \in f$, and relations as just defined. We leave it to the reader to verify that the map $x \mapsto [x] : \mathfrak{f} \to \mathfrak{f}/{\sim}$ is a contraction if $\sim$ is a net, and that if π is a contraction, then the equivalence $\sim_\pi$ induced by π is a net. Nets are just a suitable way to picture contractions.

Proposition 2.4.1 (Net Extension I). *Suppose that* $\mathfrak{f} \rightarrowtail \mathfrak{g}$ *and that* $\sim$ *is a net on* $\mathfrak{f}$. *Define* $\approx \subseteq g \times g$ *by* $x \approx y$ *iff (i.)* $\{x, y\} \subseteq f$ *and* $x \sim y$ *or (ii.)* $\{x, y\} \not\subseteq f$ *and* $x = y$. *Then* $\approx$ *is a net on* $\mathfrak{g}$.

The proof is easy and omitted. This theorem is of extreme practical importance; it says that if $\mathfrak{c}$ is a contractum of a generated subframe $\mathfrak{f}$ of $\mathfrak{g}$, then it is a generated subframe of a suitably defined contraction image $\mathfrak{d}$ of $\mathfrak{g}$, see picture below. The maps denoted by dashed arrows are in some sense unique (we will come to that later), this is why we put an exclamation mark.

$$\begin{tikzcd} \mathfrak{f} \arrow[r] \arrow[d] & \mathfrak{g} \arrow[d, dashed, "!"] \\ \mathfrak{c} \arrow[r, dashed, "!"] & \mathfrak{d} \end{tikzcd}$$

If β is a valuation, and π a p–morphism, π is called **admissible** for β if for every $x \in g$ and every set $\beta(p)$ either $\pi^{-1}(x) \subseteq \beta(p)$ or $\pi^{-1}(x) \subseteq -\beta(p)$. In other words, the partition that π induces on f must be finer than the partition induced by the sets $\overline{\beta}(\varphi)$. In that case we can say that β **induces** a valuation γ on $\mathfrak{g}$ by taking $\gamma(p) := \{\pi(x) : x \in \beta(p)\}$. We say that γ is the **image** of β under π. Moreover, we will also write β for the valuation γ if no confusion arises. It should be clear that every valuation on $\mathfrak{g}$ can be seen as the image of a valuation on $\mathfrak{f}$ under a contraction. Now take an arbitrary p–morphism $\pi : \mathfrak{f} \to \mathfrak{g}$. Then the following important theorem holds.

Proposition 2.4.2. *Let* $\pi : \mathfrak{f} \to \mathfrak{g}$ *and let* π *be admissible for* β. *Then for every* $x \in f$

$$\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi \quad \Leftrightarrow \quad \langle \mathfrak{g}, \beta, \pi(x)\rangle \vDash \varphi$$

Proof. For variables this is true by construction; the steps for $\neg$ and $\wedge$ are easy. Now let $\varphi = \Diamond_j\psi$. Assume $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond_j\psi$. Then there is a y such that $x \lhd_j y$ and $\langle \mathfrak{f}, \beta, y\rangle \vDash \psi$. By (pm1.), $\pi(x) \lhd_j \pi(y)$, and by induction hypothesis $\langle \mathfrak{g}, \beta, \pi(y)\rangle \vDash \psi$. This gives $\langle \mathfrak{g}, \beta, \pi(x)\rangle \vDash \Diamond_j\psi$. Now assume that the latter holds. Then for some u with $\pi(x) \lhd_j u$ we have $\langle \mathfrak{g}, \beta, u\rangle \vDash \psi$. By (pm2.) there exists a $\widehat{y}$ such that $x \lhd_j \widehat{y}$ and $\pi(\widehat{y}) = u$. By induction hypothesis, $\langle \mathfrak{f}, \beta, \widehat{y}\rangle \vDash \psi$ and so $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond_j\psi$, as required. ∎

A remark on the proof. As is often the case, the induction is easier to perform using $\Diamond_j$ rather than $\Box_j$. Although the latter is a primitive symbol of the language, we allow ourselves for the purpose of proofs to take either as primitive and the other as composite, whichever is best suited for the inductive step.

Now let $\pi : \mathfrak{f} \to \mathfrak{g}$ be a p–morphism. Let $im[\pi] \subseteq g$ be the set of points $\pi(x)$ for $x \in f$. This is the so–called **image** of π. It is a subframe of $\mathfrak{g}$. Moreover, the following geometric analogue of the first Noether Isomorphism Theorem holds.

Proposition 2.4.3. *Let $\pi : \mathfrak{f} \to \mathfrak{g}$ be a p–morphism. Then there are maps $\rho : \mathfrak{f} \twoheadrightarrow im[\pi]$ and $\zeta : im[\pi] \rightarrowtail \mathfrak{g}$ such that $\pi = \zeta \circ \rho$.*

Proof. Put $\rho(x) := \pi(x)$ and $\zeta(u) := u$. First, we show that $im[\pi]$ is a generated subframe. This shows that ζ is a p–morphism. To that end let $u := \pi(x)$ and $u \lhd_i v$. Then by (pm2.) there is a y such that $x \lhd_i y$ and $\pi(y) = v$. Hence, $\zeta : im[\pi] \rightarrowtail \mathfrak{g}$. Now, consider the map ρ. Take x, y with $x \lhd_i y$. Then, by (pm1.) $\pi(x) \lhd_i \pi(y)$, whence $\rho(x) \lhd_i \rho(y)$. Finally, let $\rho(x) \lhd_i u$. Then, by definition, $\pi(x) \lhd_i u$ and by (pm2.) there is a y such that $x \lhd_i y$ and $\rho(y) = \pi(y) = u$. This shows $\rho : \mathfrak{f} \to im[\pi]$. That ρ is surjective follows immediately from the definition. ◻

Take now Kripke–frames, $\mathfrak{f}_i = \langle f_i, \langle \lhd^i_j : j < \kappa \rangle\rangle$, $i \in I$. Let

$$\bigoplus_{i \in I} f_i := \bigcup_{i \in I} \{i\} \times f_i$$

be the disjoint union of the sets f_i. (For simplicity, we standardly assume that the sets f_i are pairwise disjoint and then we put $\bigoplus_{i\in I} f_i := \bigcup_{i\in I} f_i$. In general, this is not without complications, however.) Based on this set we can define the frame $\bigoplus_{i\in I} \mathfrak{f}_i$, called the **direct sum** or **disjoint union**, via

$$\begin{array}{lcl} \lhd^{\oplus}_j & := & \{\langle\langle i, x\rangle, \langle i, y\rangle\rangle : i \in I; x, y \in f_i; x \lhd^i_j y\} \\ \bigoplus_{i\in I} \mathfrak{f}_i & := & \langle \bigoplus_{i\in I} f_i, \langle \lhd^{\oplus}_j : j < \kappa\rangle\rangle \end{array}$$

Intuitively, the direct sum consists of several components, and two points are j–related iff they are from the same component and are j–related in that component. The components are just placed next to each other, with no interaction. The direct sum has the following property.

Theorem 2.4.4. *Let $\mathfrak{f}_i$, $i \in I$, be Kripke–frames. Then there exist embeddings $\epsilon_i : \mathfrak{f}_i \rightarrowtail \bigoplus_{i\in I} \mathfrak{f}_i$. Moreover, for every Kripke–frame $\mathfrak{h}$ with p–morphisms $\kappa_i : \mathfrak{f}_i \to \mathfrak{h}$, $i \in I$, there exists a unique $\pi : \bigoplus_{i\in I} \mathfrak{f}_i \to \mathfrak{h}$, which is a p–morphism such that $\pi \circ \epsilon_i = \kappa_i$ for all $i \in I$.*

Proof. It is easy to check that the identity embeddings $\epsilon_i : \mathfrak{f}_i \to \bigoplus_{i\in I} \mathfrak{f}_i$ are injective p–morphisms. Now assume that $\mathfrak{h}$ is given and $\kappa_i : \mathfrak{f}_i \to \mathfrak{h}$, $i \in I$. We have to define π. Take an element $z \in \bigoplus_{i\in I} f_i$. There is an $i \in I$ such that $z = \langle i, x\rangle$ for some $x \in f_i$. Put $\pi(z) := \kappa_i(x)$. π as defined is a p–morphism. For if $z \lhd_i z'$ for some z and z', then there is an i and $x, x' \in f_i$ such that $z = \langle i, x\rangle$ and $z' = \langle i, x'\rangle$. By construction, $x \lhd_j x'$. Then $\pi(z) = \kappa_i(x) \lhd_j \kappa_i(x') = \pi(z')$, by the fact that κ_i satisfies (pm1.). Now assume $\pi(z) \lhd_i u$, $z = \langle i, x\rangle$. Then $\pi(z) = \kappa_i(x)$ and by the fact that κ_i satisfies (pm2.) we have a x' such that $x \lhd_j x'$ and $\kappa_i(x') = u$. Then $\pi(\langle i, x'\rangle) = \kappa_i(x') = u$. So, π is a p–morphism; moreover, by definition we get $(\pi \circ \epsilon_i)(x) = \pi(\langle i, x\rangle) = \kappa_i(x)$ for $x \in f_i$. Finally, let us see why π is unique. So let ζ be another map satisfying all requirements. Let $x \in f_i$. Then $\zeta(\langle i, x\rangle) = (\zeta \circ \epsilon_i)(x) = \kappa_i(x) = (\pi \circ \epsilon_i)(x) = \pi(\langle i, x\rangle)$. ◻

A valuation δ on $\bigoplus_{i\in I}\mathfrak{f}_i$ uniquely determines a valuation β_i on $\mathfrak{f}_i$ by $\{i\}\times\beta_i(p) = \delta(p)\cap(\{i\}\times f_i)$. Conversely, given a family β_i, $i \in I$, of valuations into $\mathfrak{f}_i$ there is a unique valuation $\bigoplus_{i\in I}\beta_i$ on $\bigoplus_{i\in I}\mathfrak{f}_i$ defined by

$$(\bigoplus_{i\in I}\beta_i)(p) := \bigcup_{i\in I}\{i\}\times\beta_i(p)$$

The following theorems are left as exercises.

Proposition 2.4.5. *Let* $\mathfrak{g}$, $\mathfrak{f}$, $\mathfrak{f}_i$, $i \in I$, *be Kripke–frames.*

(1) *If* $\mathfrak{g} \rightarrowtail \mathfrak{f}$ *then* $\mathsf{Th}(\mathfrak{f}) \subseteq \mathsf{Th}(\mathfrak{g})$.
(2) *If* $\mathfrak{g} \twoheadrightarrow \mathfrak{f}$ *then* $\mathsf{Th}(\mathfrak{g}) \subseteq \mathsf{Th}(\mathfrak{f})$.
(3) $\mathsf{Th}(\bigoplus_{i\in I}\mathfrak{f}_i) = \bigcap_{i\in I}\mathsf{Th}(\mathfrak{f}_i)$.

Theorem 2.4.6. *Let* Λ *be a normal polymodal logic. Then if* $\mathfrak{f}$ *is a Kripke–frame for* Λ*, so is any generated subframe and any p–morphic image. Moreover, any direct sum of Kripke–frames for* Λ *is again a Kripke–frame for* Λ*.*

To generalize these notions to frames we need to consider what happens to the internal sets. First, consider a subframe $\mathfrak{f} \rightarrowtail \mathfrak{g}$, and let $\mathfrak{F} = \langle\mathfrak{f}, \mathbb{F}\rangle$ as well as $\mathfrak{G} = \langle\mathfrak{g}, \mathbb{G}\rangle$. For simplicity we assume that $f \subseteq g$. The map $a \mapsto a \cap f$ is a boolean homomorphism from $\langle\mathbb{G}, g, -, \cap\rangle$ to $\langle 2^f, f, -, \cap\rangle$. Consider now the fact that any valuation β on $\mathfrak{g}$ defines a valuation γ on $\mathfrak{f}$, namely, $\gamma(p) := \beta(p) \cap f$. In order for theorems like Proposition 2.4.5 to hold we need that for every valuation β the corresponding valuation γ is a valuation on $\mathfrak{F}$. Hence, we must require that $a \mapsto a\cap f$ is a homomorphism from $\langle\mathbb{G}, g, -, \cap\rangle$ *onto* $\langle\mathbb{F}, f, -, \cap\rangle$.

Definition 2.4.7. *A map* $\pi : f \to g$ *is an* ***embedding*** *of the frame* $\mathfrak{F} = \langle\mathfrak{f}, \mathbb{F}\rangle$ *in the frame* $\mathfrak{G} = \langle\mathfrak{g}, \mathbb{G}\rangle$ *if (0.)* π *is injective, (1.)* $\pi(f) \in \mathbb{G}$*, (2.)* $\pi : \mathfrak{f} \to \mathfrak{g}$ *is a p–morphism and (3.) the map* $\pi^{-1} : a \mapsto \{x : \pi(x) \in a\}$ *is a surjective homomorphism from* $\langle\mathbb{G}, g, -, \cap, \langle\blacksquare_i : i \in \kappa\rangle\rangle$ *to the algebra* $\langle\mathbb{F}, f, -, \cap, \langle\blacksquare_i : i \in \kappa\rangle\rangle$*. If all that is the case we write* $p : \mathfrak{F} \rightarrowtail \mathfrak{G}$*. If in addition* $f \subseteq g$*, and* π *the natural inclusion map we call* $\mathfrak{F}$ *a* ***generated subframe*** *of* $\mathfrak{G}$ *and write* $\pi : \mathfrak{F} \leq \mathfrak{G}$ *or simply* $\mathfrak{F} \leq \mathfrak{G}$*.*

The surjectivity of the map π^{-1} for embeddings is actually quite important for generalizing the factorization theorem, Proposition 2.4.3. It would fail otherwise. Several remarks are in order. First, notice that while the map on the frames goes from $\mathfrak{f}$ to $\mathfrak{g}$, the corresponding map π^{-1} on the algebras goes from $\langle\mathbb{G}, g, -, \cap, \langle\blacksquare_i : i \in \kappa\rangle\rangle$ to $\langle\mathbb{F}, f, -, \cap, \langle\blacksquare_i : i \in \kappa\rangle\rangle$. Moreover, for generated subframes, rather than requiring that the map is a well–defined homomorphism we require that the map is onto, that is, all sets of $\mathbb{F}$ are restrictions of sets in $\mathbb{G}$. A last point is the requirement that the image of f under π is internal. This is added for theory internal reasons, since if this definition is generalized to subframes, this clause is needed. However, it can be shown to be unnecessary for generated subframes.

Next we turn to contractions. Again, we study the map $a \mapsto \pi^{-1}[a]$. This is a boolean homomorphism from 2^g to 2^f, and we need to make sure that it is also a homomorphism of the modal algebras on the frames. Hence, consider the set $c := \pi^{-1}[\blacksquare_j a]$. It contains all y such that for all u with $\pi(y) \lhd_j u$ we have $u \in a$. On the other hand, the set $d := \blacksquare_j(\pi^{-1}[a])$ is the set of all y such that for all z such that $y \lhd_j z$ we have $\pi(z) \in a$. If $y \in c$ and $y \lhd_j z$ then $\pi(y) \in \blacksquare_j a$. By (pm1.) $\pi(y) \lhd_j \pi(z)$, hence $\pi(z) \in a$, and so $z \in \pi^{-1}[a]$. Thus $y \in d$. Now let $y \in d$. To see that $y \in c$, we have to show that $\pi(y) \in \blacksquare_j a$. So, assume $\pi(y) \lhd_j u$. By (pm2.) there is a z such that $\pi(z) = u$ and $y \lhd_j z$. Then $z \in \pi^{-1}[a]$, by assumption on y, and so $u \in a$, as required. So, the map $a \mapsto \pi^{-1}[a]$ is indeed a homomorphism of modal algebras.

Definition 2.4.8. *Let $\mathfrak{F}$ and $\mathfrak{G}$ be frames and $\pi : f \to g$ a map. π is a* ***contraction*** *from $\mathfrak{F}$ to $\mathfrak{G}$ if (0.) π is surjective, (1.) π is a p–morphism and (2.) π^{-1} is an injective homomorphism from $\langle \mathbb{G}, g, -, \cap, \langle \blacksquare_j : j \in \kappa \rangle\rangle$ into $\langle \mathbb{F}, f, -, \cap, \langle \blacksquare_j : j \in \kappa \rangle\rangle$. If all that is the case, we write $\pi : \mathfrak{F} \twoheadrightarrow \mathfrak{G}$.*

A nontrivial example is $\Omega := \langle \omega, <, \mathbb{O} \rangle$ with $\mathbb{O}$ the set of finite and cofinite subsets of ω. Take the Kripke–frame $\mathfrak{f}_n = \langle \{0, \dots, n-1\}, \lhd \rangle$ where $i \lhd j$ iff $i < j$ or $i = j = n - 1$. Next, let $\mathfrak{g}_n := \langle \{0, 1, \dots, n-1\}, \blacktriangleleft \rangle$, where $\blacktriangleleft = \ n \times n$. It turns out that $\mathfrak{f}_n$ is a contractum of Ω, while $\mathfrak{g}_n$ is a contractum only for $n = 1$. However, $\mathfrak{g}_n$ is a contractum of $\langle \omega, < \rangle$, the underlying Kripke–frame of Ω. (It follows that also all $\mathfrak{f}_n$ are contracta of $\langle \omega, < \rangle$.) We can, finally, define the notion of a p–morphism for frames.

Definition 2.4.9. *Let $\mathfrak{F}$ and $\mathfrak{G}$ be frames and $\pi : f \to g$ a map. π is a* ***p–morphism*** *from $\mathfrak{F}$ to $\mathfrak{G}$, in symbols $\pi : \mathfrak{F} \to \mathfrak{G}$, if (1.) $\pi : \mathfrak{f} \to \mathfrak{g}$ is a p–morphism of the underlying Kripke–frames and (2.) π^{-1} is a homomorphism from the modal algebra $\langle \mathbb{G}, g, -, \cap, \langle \blacksquare_j : j \in \kappa \rangle\rangle$ to $\langle \mathbb{F}, f, -, \cap, \langle \blacksquare_j : j \in \kappa \rangle\rangle$.*

In fact, given (1.) it is not necessary to require π^{-1} to be a homomorphism. Rather, it is enough to require

(pm3.) If $a \in \mathbb{G}$ then $\pi^{-1}[a] \in \mathbb{F}$.

Proposition 2.4.10. *A map $h : f \to g$ is a p–morphism from $\mathfrak{F}$ to $\mathfrak{G}$ iff it satisfies (pm1.), (pm2.) and (pm3.).*

We refer to (pm3.) as the third p–morphism condition. Notice that with the definitions given an embedding is an injective p–morphism and a contraction a surjective p–morphism. A generalized frame is contractible to a Kripke–frame if the algebra of sets is finite. This is the content of the next theorem.

Theorem 2.4.11. *Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$ be a generalized frame and $\mathbb{F}$ finite. Define an equivalence relation $x \sim y$ on points by*

$$x \sim y \Leftrightarrow (\forall a \in \mathbb{F})(x \in a. \leftrightarrow .y \in a) .$$

Put $[x] := \{y : x \sim y\}$, $[f] := \{[x] : x \in f\}$ *and* $[x] \lhd_j [y]$ *iff there exist* $\widehat{x} \in [x]$ *and* $\widehat{y} \in [y]$ *such that* $\widehat{x} \lhd_j \widehat{y}$*. Then* $x \mapsto [x]$ *is a p–morphism from* $\mathfrak{F}$ *onto the frame* $\langle [f], \langle \lhd_j : j < \kappa \rangle, 2^{[f]} \rangle$.

Proof. We have to check the second clause of the p–morphism condition. (The first is satisfied by definition of $\lhd_j$ on $[f]$.) Let $[x] \lhd_j u$. Then there is a y such that $u = [y]$. We have to show that there is a $\widehat{y}$ such that $x \lhd_j \widehat{y}$ and $\widehat{y} \sim y$. To see this, let a_y be the intersection of all elements in $\mathbb{F}$ containing y (equivalently, let a_y be the unique atom containing y). Since $\mathbb{F}$ is finite, $a_y \in \mathbb{F}$. Then $\widehat{y} \sim y$ iff $\widehat{y} \in a_y$ iff $a_{\widehat{y}} = a_y$. It is checked that $x \lhd_j y$ iff $a_x \leq \blacklozenge_j a_y$. Now since $[x] \lhd_j [y]$ there are $\widehat{x} \in [x]$ and $\widehat{y} \in [y]$ such that $\widehat{x} \lhd_j \widehat{y}$. It follows that $a_{\widehat{x}} \leq a_{\widehat{y}}$. Since $a_{\widehat{x}} = a_x$ and $a_{\widehat{y}} = a_y$ we conclude $a_x \leq \blacklozenge_j a_y$, from which $x \in \blacklozenge_j a_y$. And so there is a $y' \sim y$ such that $x \lhd_j y'$. So, the map is a p–morphism. We have seen that the classes $p^{-1}([x])$ are internal sets of the form a_x defined earlier. Hence, each set has a preimage. ◻

Definition 2.4.12. *Let* $\mathfrak{F}_i$*,* $i \in I$*, be frames. The* ***disjoint union*** *of the* $\mathfrak{F}_i$*, denoted by* $\bigoplus_{i \in I} \mathfrak{F}_i$*, is defined as follows. Then underlying Kripke–frame is* $\bigoplus_{i \in I} \mathfrak{f}_i$*. A set is* ***internal*** *if it is a union* $\bigcup_{i \in I} \{i\} \times a_i$*, where* $a_i \in \mathbb{F}_i$ *for each* $i \in I$*.*

Proposition 2.4.13. *Let* $\mathfrak{F}_i$*,* $i \in I$*, be frames. There exist embeddings* $e_i : \mathfrak{F}_i \rightarrowtail \bigoplus_{i \in I} \mathfrak{F}_i$ *such that for all* $\mathfrak{G}$ *and embeddings* $d_i : \mathfrak{F}_i \rightarrowtail \mathfrak{G}$ *there exists a p–morphism* $\pi : \bigoplus_{i \in I} \mathfrak{F}_i \to \mathfrak{G}$ *such that* $d_i = \pi \circ e_i$ *for all* $i \in I$*.*

Proof. Follow the proof of Theorem 2.4.4. The construction is completely analogous. We only have to check that the map π satisfies the third condition on p–morphisms. To this end consider an internal set a of $\mathfrak{G}$. Let $a_i := d_i^{-1}[a]$. By the fact that d_i is a p–morphism, this is an internal set of $\mathfrak{F}_i$. Now, $\pi^{-1}[a] = \bigcup_{i \in I} \{i\} \times d_i^{-1}[a]$. By definition of $\bigoplus_{i \in I} \mathfrak{F}_i$, this is an internal set. ◻

A **net** on a generalized frame $\mathfrak{F}$ is a net $\sim$ on $\mathfrak{f}$ such that for each $a \in \mathbb{F}$ the set $[a]_\sim := \{y : (\exists x \in a)(x \sim y)\}$ is a member of $\mathbb{F}$. We write $\mathfrak{F}/{\sim}$ for the quotient, which is defined by

$$\mathfrak{F}/{\sim} := \langle \mathfrak{f}/{\sim}, \langle \{[x] : x \in a\} : a \in \mathbb{F} \rangle \rangle$$

Proposition 2.4.14 (Net Extension II). *Let* $\mathfrak{F}$ *be a frame, and* $\mathfrak{G} \leq \mathfrak{F}$*. Let* $\sim$ *be a net on* $\mathfrak{G}$*. Let* $x \approx y$ *iff (i.)* $x, y \in g$ *and* $x \sim y$ *or (ii.)* $x, y \in f - g$ *and* $x = y$*. Then* $\approx$ *is a net on* $\mathfrak{F}$*.*

Proof. In view of Proposition 2.4.1 we only have to check that for each $a \in \mathbb{F}$ we have $[a]_\approx \in \mathbb{F}$. To see this, let $b := a \cap g$ and $c := a \cap (f - g)$. Since $g \in \mathbb{F}$ we have $b, c \in \mathbb{F}$. Then $[a]_\approx = [b]_\sim \cup c$. By assumption on $\sim$, $[b]_\sim \in \mathbb{F}$. Hence the claim is proved. ◻

Exercise 42. Let $\mathfrak{G}$ be a frame and $f \subseteq g$. Show that the map $a \mapsto a \cap f$ is a boolean homomorphism from 2^g to 2^f. Hence, the condition (2.) of Definition 2.4.7 is necessary only to ensure that this map is a homomorphism of the modal algebras, that is, that it commutes with the modal operators. Show that it is sufficient.

Exercise 43. Show that if in Definition 2.4.7 we did not require π^{-1} to be surjective, then there were injective p–morphisms which are not embeddings.

Exercise 44. Let $\pi : \mathfrak{F} \twoheadrightarrow \mathfrak{G}$ be a p–morphism and $x \in g$. Show that if there is no chain of length k $x = x_0 \lhd_j x_1 \lhd_j x_2 \ldots \lhd_j x_k$ then the same holds for $\pi(x)$ as well. Show that if $\pi(x) \not\lhd_j \pi(x)$ then $\pi^{-1}(\pi(x))$ is a j–antichain, that is, for all $y, z \in \pi^{-1}(\pi(x))$ $y \not\lhd_j z$.

Exercise 45. Prove Proposition 2.4.5 and Theorem 2.4.6.

Exercise 46. Formulate and prove Proposition 2.4.5 and Theorem 2.4.6 for (generalized) frames instead of Kripke–frames.

Exercise 47. Let $\mathfrak{f} = \langle f, \langle \lhd_j : j < \kappa \rangle \rangle$ be a Kripke–frame and G a subgroup of the group $\mathfrak{Aut}(\mathfrak{f})$ of automorphisms of $\mathfrak{f}$. Put

$$[x] := \{y : \text{ there exists } g \in G : g(x) = y\} \ .$$

Also, put $[x] \lhd_j [y]$ if there exist $\widehat{x} \in [x]$ and $\widehat{y} \in [y]$ such that $\widehat{x} \lhd_j \widehat{y}$. Show that $x \mapsto [x]$ is a p–morphism. (An automorphism of $\mathfrak{f}$ is a bijective p–morphism from $\mathfrak{f}$ to $\mathfrak{f}$. The automorphisms of a structure generally form a group.)

2.5. Some Important Modal Logics

Among the infinitely many logics that can be considered there are a number of logics that are of fundamental importance. Their importance is not only historical but has as we will see also intrinsic reasons. We begin with logics of a single operator. Here is a list of axioms together with their standard names. (In some cases we have given alternate forms of the axioms. The first is the one standardly known, the second a somewhat more user friendly variant.)

Inc	$\bot$
4	$\Diamond\Diamond p \to \Diamond p$
D	$\Diamond\top$
B	$p \to \Box\Diamond p$
T	$p \to \Diamond p$
5	$\Diamond p \to \Box\Diamond p$
alt$_1$	$\Diamond p \to \Box p$
	$\Diamond p \wedge \Diamond q. \to .\Diamond(p \wedge q)$
1	$\Box\Diamond p \to \Diamond\Box p$
2	$\Diamond\Box p \to \Box\Diamond p$
3	$\Diamond p \wedge \Diamond q. \to .\Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p) \vee \Diamond(p \wedge q)$
G	$\Box(\Box p \to p) \to \Box p$
	$\Diamond p \to \Diamond(p \wedge \neg\Diamond p)$
Grz	$\Box(\Box(p \to \Box p) \to p) \to p$
	$p \to \Diamond(p \wedge \Box(\neg p \to \Box\neg p))$

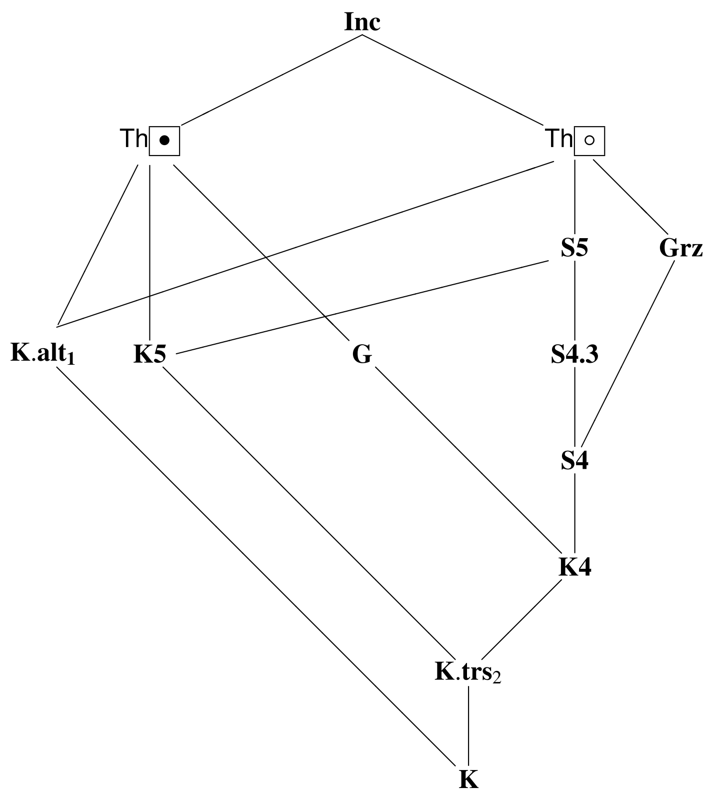
In addition, there are also logics with special names. For example, **S4** is **K4**.**T**, **S5** is **K4**.**BT** or, equivalently, **S4**.**B**. (The dot has no meaning; it is inserted for readability. Occasionally, several dots will be inserted.) The letter **S** stems from a classification by C. I. Lewis, who originally introduced modal logic as a tool to analyse conditionals. He considered five systems, called **S1** to **S5**, among which only the last two — namely **S4** and **S5** — were based on normal modal logics. **D** originally comes from *deontic*, since this postulate was most prominent in deontic logic, the logic of obligations. Nowadays, **D** is associated with *definal*, lit. meaning *without end*, because in frames satisfying this postulate every world must see at least one world. **G** is named after Kurt Gödel, because this axiom is related to the logic derived by interpreting $\Box$ as *provable in Peano Arithmetic (= PA)*; the logic **K4.G** is called the **provability logic**. Often **G** is also called **GL**, where **L** stands for M. H. Löb, who contributed the actual axiomatization. In [**201**], Robert Solovay showed that if $\Box$ is read as *it is provable in* ***PA*** *that* then the logic of this modal operator is exactly **K4.G**. For the history of this logic see the entertaining survey by George Boolos and Giovanni Sambin, [**32**]. Finally, the axiom **Grz** is named after the Polish logician Grzegorczyk. Usually, the logic **S4**.**Grz** is called Grzegorczyks logic. Some authors use **G** for **K4.G** and **Grz** for **S4.Grz**. The reason will be provided in the exercises; namely, it turns out that **K**.**G** as well as **K**.**Grz** contain the axiom **4**. We will use the same convention here, if no confusion arises. **1** is known as McKinsey's axiom, therefore also denoted by **M**. **2** is called the Geach axiom.

For theoretical purposes, the following two infinite series of axioms are important.

alt$_n$	$\bigwedge_{i<n+1} \Diamond p_i. \to . \bigvee_{i<j<n+1} \Diamond(p_i \wedge p_j)$
trs$_m$	$\Diamond^{\leq m} p. \to .\Diamond^{\leq m+1} p$

The first holds in a Kripke–frame iff any point has at most n successors, the second

Figure 2.5. Important Extensions of **K**



holds in a Kripke–frame iff any point that can be reached at all can be reached in m steps. There are not so many polymodal logics which have acquired fame. However, the most useful logics occur as extensions of a logic with several operators in which the logic of a single operator on its own belongs to one of the systems above, and the axioms specifying the interaction of these operators are of a rather simple type. Therefore, the polymodal logics in which there are no axioms mixing the operators, are an important basic case. The following notation is used here. Given two monomodal logics, Λ_1 and Λ_2, the symbol $\Lambda_1 \otimes \Lambda_2$ denotes the smallest normal bimodal logic in which the first operator satisfies the axioms of Λ_1 and the second operator the axioms of Λ_2. $\Lambda_1 \otimes \Lambda_2$ is called the **fusion** or **independent join** of the two logics. Similarly, the notation $\bigotimes_{i<\mu} \Lambda_i$ denotes the fusion of μ many modal logics; in general each Λ_i can be polymodal as well. In this logic the operator $\Box_i$

satisfies exactly the postulates of Λ_i. Given a logic Λ, the operator $\Box$ is called **m–transitive** if $\Box^{\leq m} p \to \Box^{\leq m+1} p \in \Lambda$. $\Box$ is **weakly transitive** if it is m–transitive for some m. A polymodal logic Λ is called **weakly transitive** if there is a compound modality $\boxplus$ such that for every compound modality $\boxplus'$, $\boxplus p \to \boxplus' p \in \Lambda$. If κ is finite, we call Λ *m*–**transitive** if it contains the axiom $\mathbf{trs}_m$ below and **weakly transitive** if it is m–transitive for some m. This notion of weak transitivity coincides with the one defined earlier for arbitrary κ, as can easily be demonstrated.

$\mathbf{trs}_m \qquad \boxtimes^{\leq m} p. \to . \boxtimes^{\leq m+1}\, p$

(Recall the definition of $\boxtimes$ from Section 2.1.) If each basic operator is weakly transitive, Λ is said to be **weakly operator transitive**. $\mathbf{K4} \otimes \mathbf{K4}$ is operator transitive, but not transitive (as can be shown). A logic is of **bounded operator alternativity** if for each operator there is a d such that that operator satisfies $\mathbf{alt}_d$. Λ is of **bounded alternativity** if it has finitely many operators and is of bounded operator alternativity. A logic Λ is called **cyclic** if for every compound modality $\boxplus$ there exists a compound modality $\boxminus$ such that $p \to \boxplus \Diamond p \in \Lambda$.

Furthermore, the postulates $\Diamond_i p \to \Diamond_j p$ are considered. On Kripke–frames they force the relation $\lhd_i$ to be included in $\lhd_j$. Also, the postulates $\Diamond_i \Diamond_j p \leftrightarrow \Diamond_j \Diamond_i p$ say that any point reachable following first $\lhd_i$ and then $\lhd_j$ is reachable following $\lhd_j$ and then $\lhd_i$ — and vice versa. This is a kind of Church–Rosser Property with respect to i and j. Sometimes, only one implication is considered. Quite interesting is the following construction. Consider a (polymodal) logic Λ with operators $\Box_i$, $i < \kappa$. Add a new operator $\blacksquare = \Box_\kappa$. Then call $\blacksquare$ a **master modality** if it satisfies the postulates of **K4** and the interaction postulates $\Diamond_i p \to \blacklozenge p$. If $\blacksquare$ satisfies **S5** it is called a **universal modality**.

Important bimodal logics are *tense logics*. A **tense logic** is a normal extension of

$$\mathbf{K}.t := \mathbf{K}_2 \oplus \{p \to \overrightarrow{\Box} \Diamond p, p \to \overleftarrow{\Box} \Diamond p\}$$

Here, we have used $\overrightarrow{\Box}$ for $\Box_0$ and $\overleftarrow{\Box}$ for $\Box_1$. If Λ is a monomodal logic, then $\Lambda.t$ is obtained by interpreting $\Box$ as $\overrightarrow{\Box}$. There is also the possibility to interpret $\Box$ as $\overleftarrow{\Box}$. Tense logical axioms can be derived from monomodal axioms by choosing either interpretation for the operator. Thus, if φ is an axiom and we want to interpret the operator as $\overrightarrow{\Box}$, we write φ^+, and if we want to interpret the operator as $\overleftarrow{\Box}$ then we write φ^-. So, we have logics like $\mathbf{S4}.t$, $\mathbf{S5}.t$ and $\mathbf{K4}.t.\mathbf{D}^+.\mathbf{D}^-$. The latter is actually the same as $\mathbf{K4D}.t.\mathbf{D}^-$. A logic Λ with $2k$ operators is called **connected** if there exists a permutation $\pi : 2k \to 2k$ such that $\pi^2 = id$ and for each $i < 2k$, $p \to \Box_i \Diamond_{\pi(i)}, p \to \Box_{\pi(i)} \Diamond_i \in \Lambda$. A connected logic is cyclic.

Exercise 48. Recall from Proposition 2.1.3 that $\mathbf{K} \oplus \varphi \to \psi = \mathbf{K} \oplus \psi^d \to \varphi^d$. Write down the axioms you obtain for the axioms presented in this section if you apply this operation.

Exercise 49. Show that **K**.**G** $\supseteq$ **K4**. Proceed as follows. Suppose we have an intransitive Kripke–frame $\mathfrak{f}$ and we want to show that it is also not a Kripke–frame for **K.G**. Then there must be points $x \lhd y \lhd z$ such that $x \ntriangleleft z$. Now put as a valuation $\beta(p) := \{y, z\}$. Then $x \vDash \Diamond p; \neg\Diamond(p \wedge \neg\Diamond p)$. But this trick works only for Kripke–frames. Nevertheless, it gives a clue to a solution which is completely syntactical, and therefore completely general. Assume that $\{\Diamond\Diamond p, \neg\Diamond p\}$ is consistent in **K.G**. Then put $\varphi := p \vee \Diamond p$. Now $\{\Diamond\Diamond p, \neg\Diamond p\} \vdash_{\mathbf{K.G}} \Diamond\varphi$. Show that this leads to a contradiction. The relation with the Kripke–frame is the following. A violation of transitivity can be documented by taking $\gamma(p) := \{z\}$. Now we have $\overline{\gamma}(\varphi) = \{y, z\} = \overline{\beta}(p)$, the desired set documenting the failure of **G**. (The proof that transitivity is deducible in **K.G** is attributed to Dick de Jongh and Giovanni Sambin in [**9**].)

Exercise 50. Show as in the previous exercise that **K**.**Grz** $\supseteq$ **K4**. Hence, **4** is dispensable in the axiomatization.

Exercise 51. Show that **S4**.**Grz** = **K** $\oplus\, p \to \Diamond(p \wedge \Box(\neg p \to \Box\neg p))$. (See [**14**].)

Exercise 52. Show that **K4** $\otimes$ **K4** is operator transitive but not weakly transitive. *Hint.* Consider the frame $\mathfrak{Z} = \langle \omega, \lhd, \blacktriangleleft \rangle$ with $x \lhd y$ iff $y = x + 1$ and x is even, $x \blacktriangleleft y$ iff $y = x + 1$ and x is odd.

2.6. Decidability and Finite Model Property

Recall that a logic Λ is called *decidable* if one can decide for every finite set Δ and a formula φ whether or not $\Delta \vdash_\Lambda \varphi$. It follows from the deduction theorem that a logic is decidable iff for every formula φ we can decide whether or not $\varphi \in \Lambda$. In other words, Λ is decidable iff the problem '$\varphi \in \Lambda$?' is computable iff the set Λ is decidable. We shall also say that a modal logic Λ is $\mathcal{C}$–**computable** or **in** $\mathcal{C}$, where $\mathcal{C}$ is a complexity class, if '$\varphi \in \Lambda$?' is in $\mathcal{C}$. Likewise, $\mathcal{C}$–**hardness** and $\mathcal{C}$–**completeness** of Λ are defined. Now let A be a finite set. A set $M \subseteq A^*$ is called **recursively enumerable** if it is either empty or there exists a computable function $f : \omega \to A^*$ such that the set $\mathit{range} f = f[\omega] = M$. (So, $M \neq \varnothing$ is recursively enumerable if we can, so to speak, make an infinite list of M.) A set is **co–recursively enumerable** or **co–r. e.** if its complement is recursively enumerable. M is called **recursive** if it is recursively and co–recursively enumerable.

Proposition 2.6.1 (Post). *Let A be finite and $M \subseteq A^*$. The problem '$x \in M$?' is decidable iff M is both recursively enumerable and co–recursively enumerable.*

First, let $h : \omega \to A^*$ be a computable bijection. For a proof note that if '$x \in M$?' is decidable, define f as follows. If M is not empty, pick $\vec{x} \in M$. Then put $f(n) := h(n)$ if $h(n) \in M$, otherwise $f(n) := \vec{x}$. This function enumerates M. So, M is recursively enumerable. Likewise we show that it is co–recursively enumerable. Now assume that M is both recursively and co–recursively enumerable,

$\varnothing \neq M$ and $A^* \neq M$, and let f and g enumerate M and $A^* - M$. Define h as follows. $h(\vec{x}) := 1$ if there is an $n \in \omega$ such that $f(n) = \vec{x}$; $h(\vec{x}) := 0$ if there is an $n \in \omega$ such that $g(n) = \vec{x}$. It is easy to see that this is a computable function. It follows from Proposition 1.8.13 that if $\kappa \leq \omega$ then the set of well–formed formulae is decidable. A particular consequence of Proposition 2.6.1 is that a logic Λ is decidable iff Λ is recursive. Unfortunately, this does not apply to the particular axiomatizations. There are recursively axiomatized (even finitely axiomatized) logics which are undecidable. However, if Λ is recursive then it is recursively axiomatizable. Moreover we have the following.

Proposition 2.6.2. ($\kappa < \aleph_1$.) *Let Δ be a recursively enumerable set of modal formulae. Then $\mathbf{K}_\kappa \oplus \Delta$ is recursively enumerable.*

Proof. We show how to make an infinite list of the tautologies. Classical logic is decidable, hence the tautologies of **PC** are recursively enumerable. The primitive instances of (bd→.), which are of the form

$$\Box_j(p_0 \to p_1). \to .\Box_j p_0 \to \Box_j p_1$$

are enumerable, since κ is. Fix an enumeration g of Δ, an enumeration h of all classical tautologies and an enumeration ℓ of the instances of (bd→.). Put $f(3i) := g(i)$, $f(3i+1) := h(i)$, and $f(3i+2) := \ell(i)$. This gives an enumeration of the axioms. We show how to enumerate the theorems of $\mathbf{K}_\kappa \oplus \Delta$. The problem is that we have to calculate the consequences of Δ with respect to the rules, namely, *modus ponens*, the *necessitation rule* and the *substitution rule*. The reader is asked to think about the fact that it is enough to use *finitary substitutions* rather than substitutions. A substitution σ is called **finitary** if $\sigma(p) \neq p$ only for finitely many p. The finitary substitutions can be enumerated. We leave this to the reader. (Basically, it amounts to showing that the finite sequences of natural numbers are enumerable. In effect, this is what we will be showing here as well, though in disguise.) Thus, assume that the substitutions are somehow enumerated. Now begin the enumeration of the theorems simply as a list. The list is produced in cycles. The nth cycle consists of $f(n)$ and all one–step consequences of theorems of the previous cycles, but with substitution restricted to the first n substitutions and (mn.) restricted to the first n boxes according to the enumeration. If the list has k entries up to the nth cycle, then there are $k \times n$ consequences with respect to (mn.), at most $k \times k$ consequences with respect to (mp.) and $k \times n$ consequences with respect to (sb.). So in each cycle the list is finite. Let us show that this list contains all theorems. The proof is by induction on the length of the derivation of φ from Δ. **Case 1.** φ is a classical tautology or a member of Δ. Then for some i $\varphi = f(i)$, and so φ is in the ith cycle. **Case 2.** $\varphi = \Box_j\psi$. By inductive hypothesis, ψ occurs in the list, say in the kth cycle. Let $\Box_j$ be the jth modality according to the enumeration ℓ. Then φ occurs in the cycle $max\{k, j\} + 1$. **Case 3.** φ is the result of applying modus ponens to $\psi \to \varphi$ and ψ. By induction hypothesis, the latter are in the list. Then φ is the next cycle. **Case 4.** $\varphi = \sigma_k(\psi)$. Let

ψ be in the mth cycle and $\ell := max\{k, m\}$. Then φ occurs in the $\ell + 1$st cycle. This concludes the proof. □

The last proof applies as well to all other logics, classical, monotone etc. Typically, since logics are mostly given in such a way that one can deduce that they are enumerable, it is mostly the enumerability of the nontheorems which is problematic. Only very recently, Valentin Goranko in [**85**], has given some proof procedures for enumerating the nontheorems directly. (The proof of their correctness has interestingly been given by means of semantic arguments.) Let us now say that a logic Λ is **recursively axiomatizable** (**finitely axiomatizable**) if there exists a recursively enumerable (finite) set Δ such that $\Lambda = \mathbf{K}_\kappa \oplus \Delta$. And let us say that Λ is **strongly recursively axiomatizable** if a recursive set axiomatizing Λ can be given. It is a priori possible that all modal logics are finitely axiomatizable; it may, namely, very well be that although a logic can be axiomatized by an infinite set of formulae, a finite set would have been enough. We will show below that this is false. The following is a consequence of the compactness theorem.

Proposition 2.6.3. *Let Λ be finitely axiomatizable and $\Lambda = \boldsymbol{K}_\kappa \oplus \Delta$. Then there exists a finite $\Delta_0 \subseteq \Delta$ such that $\Lambda = \boldsymbol{K}_\kappa \oplus \Delta_0$.*

Proof. Let $\Lambda = \mathbf{K}_\kappa \oplus \Delta$. Since Λ is finitely axiomatizable, there is a finite set Γ such that $\Lambda = \mathbf{K}_\kappa \oplus \Gamma$. Then there exists a proof of $\bigwedge \Gamma$ from Δ and the classical tautologies. This proof is finite, so it uses only a finite number of formulas in Δ. Let them be collected in Δ_0. Then we have $\Lambda \supseteq \mathbf{K}_\kappa \oplus \Delta_0 \supseteq \mathbf{K}_\kappa \oplus \Gamma = \Lambda$, and so $\Lambda = \mathbf{K}_\kappa \oplus \Delta_0$. □

So if Δ is a set of axioms for Λ such that no finite subset axiomatizes Λ, then Λ cannot be finitely axiomatized. Decidability is usually brought into correspondence with the finite model property defined below (see next section). However, rather than with finite model property it is primarily connected with constructibility of models, at least if the language is countable, which we will now assume. Recall that we have shown that any logic Λ is complete with respect to some class of algebras; in particular Λ is the theory of the algebra $\mathfrak{Fr}_\Lambda(var)$. This being so it is nevertheless not at all clear that we can always produce these algebras. In particular, if Λ is undecidable, then even though we can enumerate all formulae, we are not able to construct $\mathfrak{Fr}_\Lambda(var)$ from the definition, since we have

$$\mathfrak{Fr}_\Lambda(var) = \mathfrak{Tm}(var)/{\equiv_\Lambda}$$

where $\varphi \equiv_\Lambda \psi$ iff $\varphi \leftrightarrow \psi \in \Lambda$. The problem is simply that we cannot even decide whether or not a formula φ is in the class of $\top$. On the other hand, suppose that Λ is decidable. Then choose an enumeration *for* of the formulae; for simplicity $for(0) := \top$. Furthermore, we assume to have an inverse $\ulcorner\varphi\urcorner$, yielding for each formula φ a $k \in \omega$ such that $for(k) = \varphi$. We are going to produce an enumeration γ of the equivalence classes as follows. We start with $\gamma(0) := for(0) = \top$. Then $\gamma(i + 1) := for(k)$ where k is the smallest number such that for no $k' < k$, $for(k') \equiv_\Lambda$

$for(k)$. In other words, we choose a subsequence s_i of ω such that $for(s_{i+1})$ is the first formula in the enumeration that does not belong to one of the already established equivalence classes. Then $\gamma(i) := for(s_i)$. Since the logic is decidable, this is indeed an (algorithmic) enumeration of the equivalence classes. Furthermore, for each φ we can decide whether or not it belongs to a given class, and we can compute the number $\mu(\varphi)$ of the class of φ. Namely, take $k := \ulcorner\varphi\urcorner$, and enumerate all $for(k')$ for all $k' < k$. Then calculate the $\gamma(i)$ up to (at most) k and see which is the first i such that $\gamma(i) \leftrightarrow \varphi \in \Lambda$. Now, the algebra $\mathfrak{Fr}_\Lambda(var)/\equiv_\Lambda$ formed by calculating with numbers instead of formulae. For example, let m and n be given. The conjunction is a function $conj : \omega \times \omega \to \omega$ defined by $conj(i, j) := \mu(\gamma(i) \wedge \gamma(j))$. Likewise, all other functions of $\mathfrak{Fr}_\Lambda(var)$ can be reproduced as functions over ω. Instead, we could also use the representatives $\gamma(i)$ as the underlying set of the algebra.

Definition 2.6.4. ($\kappa < \aleph_1$.) *A modal algebra is called* ***effective*** *if its underlying set is* ω*, and the functions* 1, $-$, $\cap$, *and* $\blacksquare_j$, $j < \kappa$, *are computable. In general, an algebra is called* ***effective*** *if its underlying set is* ω *and all basic term–functions are computable.*

Definition 2.6.5. *Let* $\mathcal{V}$ *be a variety of* κ*–modal algebras.* $\mathcal{V}$ *is said to have* ***constructible free algebras*** *if for any finite set of generators the congruence* $\equiv_\Lambda$*, defined by* $\varphi \equiv_\Lambda \psi$ *iff* $\varphi \leftrightarrow \psi \in \Lambda$*, is a decidable subset of the set of pairs of terms.*

What we have shown is that if Λ is decidable, its variety has constructible free algebras, which are also effective algebras.

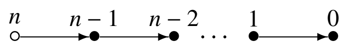
Proposition 2.6.6. *Suppose that* $\mathfrak{A}$ *is an effective algebra. Then* $\mathsf{Th}\,\mathfrak{A}$ *is co–recursively enumerable.*

Proof. Enumerate all partial valuations into $\mathfrak{A}$, and enumerate the formulae. Given a partial valuation and a formula with variables in the domain of that valuation we can compute the value of the formula under the given valuation, since the algebra is effective. It is therefore possible to enumerate all pairs $\langle\varphi, a\rangle$ where φ is a formula and a a value of φ under some valuation. Consequently, choosing among this set only the pairs for which $a \neq 1$ we obtain an enumeration of the nontheorems. □

This proof needs some explanations. If an algebra based on the set ω is not effective, there is a formula φ and a valuation β such that $\overline{\beta}(\varphi)$ cannot be determined even when $\beta(p)$ is known for all relevant variables. For by definition of effectiveness the primitive functions are not all computable. So, let f_i be a primitive function of $\mathfrak{A}$ which is not computable. Then $\varphi(\vec{p}) := f_i(\vec{p})$ is a formula such that $\overline{\beta}(\varphi)$ cannot be computed for any given β.

Corollary 2.6.7. *A logic* Λ *over a countable language is decidable iff* Λ *is recursively axiomatizable and* Λ *is the logic of an effective algebra.*

Figure 2.6.



(It is enough to have completeness with respect to a recursively enumerable class of effective algebras in addition to recursive axiomatizability. This has been pointed out to me by Michael Zakharyaschev.) Above we have seen that it is actually enough to have the free algebra be effective, with no assumptions on axiomatizability. Why is it then that the constructibility of $\mathfrak{Fr}_\Lambda(var)$ is enough to guarantee the decidability, while otherwise effectiveness is apparently not enough? The reason for that is that the theory of the free algebra is the theory of $\langle \mathfrak{Fr}_\Lambda(var), \nu \rangle$, where ν is the natural valuation. If we have completeness with respect to such a pair then theoremhood is easy to decide. In fact, then the assumption that the underlying algebra is effective is sufficient.

Definition 2.6.8. *A logic Λ has the* **finite model property** *(fmp) if for all $\varphi \notin \Lambda$ there exists a finite frame $\mathfrak{F}$ such that $\mathfrak{F} \nvDash \varphi$. Λ is* **tabular** *if there is a finite Kripke–frame $\mathfrak{f}$ such that $\Lambda = \mathsf{Th}\,\mathfrak{f}$.*

By Theorem 2.4.11 we know that Λ has the finite model property iff for each non–theorem φ there exists a finite Kripke–frame $\mathfrak{f}$ for Λ with $\mathfrak{f} \nvDash \varphi$.

Theorem 2.6.9 (Harrop). ($\kappa < \aleph_0$.) *Suppose that Λ has the finite model property. If Λ is finitely axiomatizable, Λ is decidable.*

Proof. Suppose that $\Lambda = \mathbf{K}_\kappa \oplus \Delta$ for a finite Δ. Then Λ is recursively enumerable; we need to show that it is co–recursively enumerable. Let us first show that it is possible to enumerate the frames for the logic Λ. To see that, observe that in order to decide for $\mathfrak{F}$ whether or not $\mathfrak{F} \vDash \Lambda$ we just have to check whether or not $\mathfrak{F} \vDash \Delta$. Since Δ is finite, this can be decided in finite time. Hence, since we can enumerate all frames, we can also enumerate the Λ–frames. Furthermore, we can enumerate all models $\langle \mathfrak{F}, \beta, x \rangle$ where β assigns values only for finitely many variables. For each model we can enumerate easily all formulas which are false. Hence we have an enumeration $n : \omega \times \omega \to wff$ which returns for a pair $\langle i, j \rangle$ the j^{th} formula refuted by model number i. Since $\omega \times \omega$ can be enumerated, say by $p : \omega \to \omega \times \omega$, we can finally enumerate all nontheorems by $n \circ p$. ◻

The use of the finite axiomatizability is essential. For recursive axiomatizability this theorem is actually false, see Alasdair Urquhart [**216**] and also [**122**].

Finally we give the proof that there are uncountably many logics. We work here in monomodal logic, that is, there is just one operator. Let us take the following frames. $\mathfrak{c}_n := \langle \{0, 1, 2, \ldots, n\}, \lhd \rangle$ with $i \lhd j$ iff (a) $j = i - 1$ or (b) $i = j = n$. Consider the following formulae

$$\kappa_n := \Diamond^{n+1}\top \wedge \Diamond(\Box^n\bot \wedge \neg\Box^{n-1}\bot)$$

LEMMA 2.6.10. *$\mathfrak{c}_n \vDash \neg\kappa_m$ iff $n \neq m$. $\langle \mathfrak{c}_n, j\rangle \vDash \kappa_n$ iff $j = n$.*

PROOF. Consider $\langle \mathfrak{c}_m, j\rangle \vDash \kappa_n$. If $j > n$ then $\Diamond\Box^n\bot$ is not satisfied; if $j < n$ then $\Diamond^{n+1}\top$ is not satisfied unless j is reflexive; but if it is, the formula $\Diamond(\Box^n\bot \wedge \neg\Box^{n-1}\bot)$ is not satisfied. So we must have $j = n$. Now assume $m > n$. Then $\Diamond\Box^n\bot$ is not satisfied at $j = n$. If, however, $m < n$ then $\Diamond(\Box^n\bot \wedge \neg\Box^{n-1}\top)$ is false at $j = m$. Thus $m = n = j$, as required. ☐

This innocent example has a number of consequences. Take any subset $M \subseteq \omega$ and let $\iota(M) = \mathbf{K}\oplus\{\neg\kappa_n : n \in M\}$. Then $\iota : 2^\omega \to \mathcal{E}(\mathbf{K})$ is injective. For if $M \neq N$ then there is a $m \in M$ but $m \notin N$ (or the other way around). Then all axioms of $\iota(N)$ are satisfied on $\mathfrak{c}_m$, but not all axioms of $\iota(M)$. We conclude that $\mathsf{Frm}(\iota(M)) \neq \mathsf{Frm}(\iota(N))$. Thus the two logics are different.

THEOREM 2.6.11. *There are $2^{\aleph_0}$ κ–modal logics, for all $\kappa > 0$. Moreover, there exist $2^{\aleph_0}$ many non–recursively axiomatizable logics and there exist recursively axiomatizable, undecidable logics.*

PROOF. We have seen that the map ι is injective, so there are at least $2^{\aleph_0}$ logics. However, our language has countably many formulae, and a logic is a set of formulae, so there are at most $2^{\aleph_0}$. Since there can be only countably many algorithms, there are $2^{\aleph_0}$ non–recursively enumerable subsets of ω and so there are $2^{\aleph_0}$ many non–recursively axiomatizable logics. Finally, take a recursively enumerable, but non–recursive set M (such sets exist). The logic $\iota(M)$ is recursively enumerable, by definition. But it cannot be decidable, since that would mean that we can decide '$\neg\kappa_j \in \iota(M)$', or, equivalently, '$j \in M$'. ☐

There exist also finitely axiomatizable undecidable logics. The first was established by STEPHEN ISARD [**107**], basically through coding the action of a machine in modal logic. Subsequently, many alternative ideas have been used, for example undecidable problems of group theory by VALENTIN SHEHTMAN ([**198**]), the tiling problem in EDITH SPAAN [**202**] and Thue–problems (see among others MARCUS KRACHT [**127**]). We will return to this subject in Section 9.4.

Exercise 53. Show the following variant of Proposition 2.6.3. *Let Θ be a logic and Λ be finitely axiomatizable over Θ. Suppose $\Theta = \Lambda \oplus \Delta$. Then there exists a finite set $\Delta_0 \subseteq \Delta$ such that $\Lambda = \Theta \oplus \Delta_0$.*

Exercise 54. Give an example of a logic Θ which is finitely axiomatizable as a normal extension of $\mathbf{K}_1$ but not as a quasi–normal extension.

Exercise 55. Let $\mathcal{K}$ be a recursively enumerable set of effective κ–modal algebras.

Show that $\mathsf{Th}\,\mathcal{K}$ is co–recursively enumerable.

Exercise 56. Show that Theorem 2.6.9 holds also for infinite κ. *Hint.* A finite axiom system uses only finitely many basic operators.

Exercise 57. Let T be a theory of classical predicate logic, in any given signature. Show that T is recursively axiomatizable iff it is strongly recursively axiomatizable.

2.7. Normal Forms

This chapter introduces a very basic method for proving that the logic $\mathbf{K}_\kappa$ is decidable, using the fact that it has the finite model property. This proof was first given by Kit Fine [**64**]. The finite model property is among the best–studied properties of logics. We will show a fair number of strong results on the finite model property later but shall be content in this section to show only a single result, namely the finite model property of the base logic $\mathbf{K}_\kappa$. There are many proofs of this fact but only very few proofs are constructive and do not presuppose heavy theory. For example, the proof by filtration — which we will present later — presupposes that we can show the existence of at least one model, from which we then obtain a finite model. The basic method here is syntactic in nature. We will start by proving that formulae can be rewritten into a somewhat more user–friendly form.

Definition 2.7.1. *A formula φ is called* ***strictly simple of degree 0*** *if it is of modal degree 0 and of the form $\top$ or $\bigwedge_{j<n} q_j$, $n > 0$, where each q_j is either a variable or a negated variable; moreover, no conjunct may occur twice. φ is called* ***simple of degree 0*** *if it is $\bot$ or a nonempty disjunction of pairwise distinct strictly simple formulae of degree 0. φ is called* ***strictly simple of degree d+1*** *if it is of modal degree $d + 1$ and of the form*

$$\mu \wedge \bigwedge_{j<p} \Box_{s(j)}\chi_j \wedge \bigwedge_{j<q} \Diamond_{t(j)}\omega_j \ ,$$

where μ is strictly simple of degree 0, no conjunct of φ occurs twice, $s : p \to \kappa$, $t : q \to \kappa$ are functions, χ_j are simple of degree $\leq d$ and all ω_j are strictly simple of degree $\leq d$. φ is ***simple of degree d+1*** *if it is of modal degree $d + 1$ and a disjunction of pairwise distinct strictly simple formulae of degree $\leq d + 1$. Moreover, a modal formula of degree 0 is called* ***standard of degree 0*** *if it is simple of degree 0; a modal formula of degree $d + 1$ is called* ***standard of degree d+1*** *if it is a disjunction of strictly simple formulae in which the functions s are injective and in which for a subformula of the form $\Diamond_j\psi$ ψ is standard of degree $dp(\psi)$.*

It is important to get used to simple formulae, so the reader is asked to prove some easy properties of them.

Lemma 2.7.2. *(i) Any subformula of a simple formula is simple. (ii) If a formula is simple of degree $d+1$ it is composed from variables (constants) and their negations, and formulae $\Box_j\psi$, where ψ is simple of degree d, using only $\wedge$, $\vee$ and $\Diamond_j$.*

Proposition 2.7.3. *Every formula φ can effectively be tranformed into a simple formula deductively equivalent to φ in $\boldsymbol{K}_\kappa$.*

Proof. We prove in a rather detailed way how to obtain a form that contains negation only directly in front of variables. The method to convert φ into full simple form is similar. Take a formula φ and apply the following reductions from left to right as often as possible.

$$\begin{array}{lcl} \neg(\psi \wedge \chi) & \leadsto & \neg\psi \vee \neg\chi \\ \neg(\psi \vee \chi) & \leadsto & \neg\psi \wedge \neg\chi \\ \neg\Diamond_j\psi & \leadsto & \Box_j\neg\psi \\ \neg\Box_j\psi & \leadsto & \Diamond_j\neg\psi \\ \neg\neg\psi & \leadsto & \psi \end{array}$$

It is clear that some reduction will apply as long as some operator is in the scope of $\neg$. Moreover, each step is an equivalence. (This follows from Proposition 2.1.1.) Thus all we have to show is that there is a terminating reduction series. To see this, we have to monitor two parameters, namely ℓ, the length of a longest subformula of the form $\neg\psi$, and k, the number of the subformulae of length ℓ of the form $\neg\psi$. As long as there is a subformula $\neg\psi$ of length ℓ, we apply the reduction algorithm to that subformula. The result is always a formula in which k is decreased by 1. If $k = 1$, then in the next step ℓ decreases. Proceeding this way, we will eventually reach $\ell \leq 2$, which means that the subformulae are of the form $\neg p_i$, p_i a variable, or $\neg c_i$, c_i a constant. This shows how to throw negation in front of variables and constants. To obtain simple form use the following reductions

$$\begin{array}{lcl} \varphi \wedge (\psi \vee \chi) & \leadsto & (\varphi \wedge \psi) \vee (\varphi \wedge \chi) \\ \Diamond_j(\psi \vee \chi) & \leadsto & \Diamond_j\psi \vee \Diamond_j\chi \\ \varphi \wedge \varphi & \leadsto & \varphi \\ \varphi \vee \varphi & \leadsto & \varphi \end{array}$$

It is proved analogously that these reduction terminate and that after their termination the formula is in the desired form. ◻

Proposition 2.7.4. *Every formula φ can be effectively transformed into a deductively equivalent standard formula.*

Proof. First, transform φ into simple form. Let χ be a strictly simple subformula of minimal degree that is not standard. Then it contains two conjuncts of the form $\Box_j\sigma_1$, $\Box_j\sigma_2$. φ_1 is defined by eliminating that occurrence of $\Box_j\sigma_2$ and replacing the occurrence of $\Box_j\sigma_1$ by $\Box_j(\sigma_1 \wedge \sigma_2)$. φ_1 is deductively equivalent to φ. Iterate this as often as possible. Now repeat this construction for other nonstandard subformulae of minimal degree. Each time the construction is performed it either reduces the number of nonstandard subformulae of least degree, or it increases the minimal degree of a nonstandard subformula. The procedure terminates and yields a standard formula. ◻

The next definition is crucial for the definition of the finite model. In order to understand it, we explain first the notion of *in conjunction with*. First, say that an occurrence of ψ in χ is a **conjunct** of χ if this occurrence of ψ in χ is only in the scope of $\wedge$. An occurrence of ψ_1 is **in conjunction with** an occurrence of ψ_2 in the formula φ if ψ_1 and ψ_2 are conjuncts of some subformula containing both occurrences of ψ_1 and ψ_2.

Definition 2.7.5. *Let φ be a standard formula. φ is called **explicit** if for every strictly simple subformula $\mu \wedge \bigwedge_{i<p} \Box_{s(i)}\chi_i \wedge \bigwedge_{j<q} \Diamond_{t(j)}\omega_j$ and every $j < q$ there exists an i such that $t(j) = s(i)$ and a disjunct α of χ_i, such that every conjunct of α is a conjunct of ω_j.*

Theorem 2.7.6. *For each φ there exists a standard and explicit ψ such that $\varphi \leftrightarrow \psi \in \boldsymbol{K}_\kappa$.*

Proof. First, turn φ into standard form. Call a subformula $\Box_j(\bigvee_i \psi_i)$ **unleashed** if for every $\Diamond_j\chi$ it occurs in conjunction with, one ψ_i is such that all conjuncts of ψ_i are conjuncts of χ. Let δ be the largest number such that there is a subformula $\Box_j(\bigvee_i \psi_i)$ of degree δ which is not unleashed. Now take the subformulae which are of degree δ and not unleashed. Let $\Box_j(\bigvee_i \psi_i)$ be one of them. Suppose it occurs in conjunction with a subformula $\Diamond_j\tau$. Then add $\bigvee_i \psi_i$ as a conjunct to τ; perform this for formulae of degree δ which are not unleashed. Now distribute $\wedge$ over $\vee$, and then $\Diamond_j$ over $\vee$. We will then end up with formulae of the form $\Diamond_j(\psi_i \wedge \tau)$ in place of $\Diamond_j\tau$. ψ_i is not of the form $\chi_1 \vee \chi_2$, and standard. Thus the resulting subformulae are simple. Finally, to convert the formula into standard form we only have to drive $\vee$ outside, and so all subformulae $\Box_j(\bigvee_i \psi_i)$ of degree δ are now unleashed. Thus, we may proceed to smaller subformulae. Since we never change the modal degree of the modal formulae involved, this procedure ends. ∎

We work through a particular example to give the reader a feeling for these definitions. Take the language $\kappa = 2$, the operators being $\Box$ and $\Box$. Then let φ be the formula

$$p. \wedge .\Box\neg(\Diamond\Diamond p \wedge \Diamond\Box\neg p). \wedge .\Diamond(\neg p \vee \Diamond p)$$

We can push negation inside in the second conjunct, distribute $\Diamond$ over $\vee$ in the third, and kill double negation.

$$p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p). \wedge .(\Diamond\neg p) \vee (\Diamond\Diamond p)$$

Next, we can distribute $\vee$ and get

$$\begin{array}{ll} & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p). \wedge .\Diamond\neg p] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p). \wedge .\Diamond\Diamond p] \end{array}$$

Now the formula is in standard form. However, it is not explicit. Namely, in both disjuncts, the second conjunct is of the form $\Box\psi$ while it is in conjunction with a formula of the form $\Diamond\chi$. Thus, we must add ψ as a conjunct to χ. Subsequently, we can distribute $\Diamond$ and $\vee$.

$$\begin{array}{ll} & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Box\neg p \wedge \neg p)] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Diamond p \wedge \neg p)] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Box\neg p \wedge \Diamond p)] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Diamond p \wedge \Diamond p)] \end{array}$$

We have unleashed formulae of degree 3, but there have appeared new formulae of lower depth with must also be unleashed. After distribution etc. the formula is in standard and explicit form.

$$\begin{array}{ll} & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Box\neg p \wedge \neg p)] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Diamond p \wedge \neg p)] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Box\neg p \wedge \Diamond(p \wedge \Box\neg p))] \\ \vee & [p. \wedge .\Box(\Box\Box\neg p \vee \Box\Diamond p) \wedge \Diamond(\Box\Diamond p \wedge \Diamond(p \wedge \Diamond p))] \end{array}$$

Call φ **clash–free** if there do not exist occurrences of subformulae of the form p_i and $\neg p_i$, for some i, in conjunction with each other.

Lemma 2.7.7. *Let φ be standard and explicit and not of the form $\chi_1 \vee \chi_2$. Suppose that it contains an occurrence of a formula of the form $p_i \wedge \neg p_i \wedge \omega$ which is not in the scope of a box. Then φ is inconsistent in $\boldsymbol{K}_\kappa$.*

Proof. By assumption, p_i and $\neg p_i$ are not in the scope of $\vee$ and $\Box_j$, for any $j < \kappa$. Clearly, from the assumptions, φ is composed from $p_i \wedge \neg p_i \wedge \omega$ and other formulae using only $\wedge$ and $\Diamond_j$. Then φ is inconsistent. ∎

There is an algorithm which converts a formula into a clash–free formula. Moreover, there is an algorithm which in addition preserves simplicity, explicitness and being standard. Namely, suppose that p_i and $\neg p_i$ are in conjunction with each other. Then remove from φ all subformula occurrences in conjunction with these occurrences, and replace p_i and $\neg p_i$ together by $\bot$. If necessary, remove $\bot$ in disjunction with some formula. This converts φ into clash–free and standard form. It is somewhat cumbersome but not difficult to verify that the resulting formula is also explicit if φ has been explicit.

Given a standard, explicit and clash–free formula φ we build a set of models as follows. Let us assume $\varphi = \bigvee_{i<n} \varphi_i$, φ_i strictly simple. Then for each φ_i we build a separate model; the collection of the models is the model–set of φ. We will see that φ is consistent iff the model set is non–empty iff $n > 0$. Thus assume $n = 1$, that is, φ is now strictly simple. Take a node x_φ as the root, and for each subformula $\Diamond_j\chi$ not in the scope of a box take a point x_χ. Then, as it is directly verified, χ is strictly simple, standard, explicit and clash–free. For two subformulae $\Diamond_p\psi$ and $\Diamond_q\chi$ put $x_\psi \lhd_j x_\chi$ iff $\Diamond_j\chi$ is a conjunct of ψ. (In that case, $j = q$.) A valuation is defined as follows. Let x_χ be a point constructed for the formula χ. If p_i is a conjunct of χ then $x_\chi \in \beta(p_i)$, and if $\neg p_i$ is a conjunct of χ then $x_\chi \notin \beta(p_i)$. In case where neither p_i nor $\neg p_i$ is a conjunct $\beta(p_i)$ can be fixed arbitrarily. Since χ is clash–free, a model can always be defined. A model for φ thus obtained is called a **direct model**.

Lemma 2.7.8. *Let φ be strictly simple, standard, explicit, and clash–free. Let $\langle \mathfrak{f}, \beta, x_\varphi \rangle$ be a direct model of φ. Let $\Diamond_j \chi$ not occur in the scope of a box. Then if ψ is a conjunct of χ, $x_\chi \in \overline{\beta}(\psi)$.*

Proof. By induction on the constitution of ψ (and χ). For ψ a variable, this is true by construction. Also, the definition is sound, by assumption. Now assume $\psi = \omega_1 \wedge \omega_2$. Then if ψ is a conjunct of χ so are ω_1 and ω_2. By induction hypothesis, $x_\chi \in \overline{\beta}(\omega_1)$ and $x_\chi \in \overline{\beta}(\omega_2)$. Then $x_\chi \in \overline{\beta}(\omega_1 \wedge \omega_2)$, as required. The case $\psi = \psi_1 \vee \psi_2$ does not arise. Next assume $\psi = \Diamond_j \omega$. Then there exists a j–successor x_ω of x_χ. By induction hypothesis $x_\omega \in \overline{\beta}(\omega)$, and so $x_\chi \in \overline{\beta}(\psi)$. Finally, let $\psi = \Box_j \omega$. Let y be a j–successor of x_χ. Then $y = x_\tau$ for some τ such that $\Diamond_j \tau$ is a conjunct of χ. By explicitness, some disjunct ω_i of ω (or χ itself) is a conjunct of τ. By hypothesis, $x_\tau \in \overline{\beta}(\omega_i)$. Hence $x_\chi \in \overline{\beta}(\psi)$. ◻

Theorem 2.7.9. ***K**$_\kappa$ has the finite model property.*

Proof. Start with φ and convert it into standard and explicit form. Let φ be a disjunction of φ_i, $i < n$. If φ_i contains a clash we have $\varphi_i \vdash \bot$ by Lemma 2.7.7. If φ_i does not contain a clash then by Lemma 2.7.8 there exists a finite model for φ_i. Hence, either all φ_i contain a clash, in which case $\varphi \vdash \bot$, or there is a model for some φ_i and hence for φ. ◻

Related to standard formulae are the *normal forms* which are also in use in boolean logic. The difference with standard formulae is that normal forms give rise to unambiguous direct models on a given set of variables. One can define normal forms with respect to standard formulae by the following fact. In addition to being standard (i) a normal form is consistent, (ii) a normal form is always reduced; it contains no occurrences of φ in conjunction with a different occurrence of φ and no occurrence of φ in disjunction with another occurrence of φ, and (iii) a normal form is complete for given modal depth δ; if φ is in normal form, χ of depth less than δ and $\varphi \wedge \chi$ is consistent, then $\varphi \vdash \chi$. (i) is easy to achieve. We can start from a simple formula and just throw away all disjuncts containing a clash. (ii) is likewise easy to get. Simply drop multiple ocurrences of the same formula. (iii) is possible only on one condition, namely that we work over a finite vocabulary. By *finite vocabulary* we mean both that there are finitely many modal operators and that there only finitely many propositional variables and constants. Thus, let us assume that we have a finite set $J \subseteq \kappa$ and a finite set $P = \{p_i : i < n\}$. Then the **normal forms of degree** k over

J and P are defined inductively as follows.

$$\begin{array}{lcll}
\chi^0_C & := & \bigwedge_{i\in C} p_i \wedge \bigwedge_{i\notin C} \neg p_i & C \subseteq n \\
nf(J,P,0) & := & \{\chi^0_C : C \subseteq n\} & \\
\chi^{k+1}_{D(j)} & := & \bigwedge_{i\in D(j)} \Diamond_j \chi^k_i \wedge \bigwedge_{i\notin D(j)} \neg\Diamond_j\chi^k_i & D(j) \subseteq nf(J,P,k) \\
\chi^{k+1}_{C,\overline{D}} & := & \chi^0_C \wedge \bigwedge_{j\in J} \chi^{k+1}_{D(j)} & C \subseteq n, D(j) \subseteq nf(J,P,k) \\
nf(J,P,k+1) & := & \{\chi^{k+1}_{C,\overline{D}} : C \subseteq n, \overline{D} : J \to \wp(nf(J,P,k))\} &
\end{array}$$

Proposition 2.7.10. *Let φ be a modal formula of depth k based on the variables of P and the operators of J. Then there is a set $\Psi \subseteq nf(J,P,k)$ such that*

$$\varphi \leftrightarrow \bigvee \Psi \in \boldsymbol{K}$$

Proof. By induction on k. For $k = 0$ this is the familiar disjunctive normal form for boolean logic. Now let $k > 0$. Then φ is a boolean combination of variables and formulae $\Diamond_j\psi$ with $dp(\psi) < k$. By inductive hypothesis each ψ is equivalent to a disjunction of a set N_ψ of normal forms of degree $k-1$. If $N_\psi = \varnothing$, then $\Diamond_j\psi$ is equivalent to $\Diamond_j\bot$, hence to $\bot$. If $N_\psi \neq \varnothing$ then $\Diamond_j\psi \equiv \Diamond_j \bigvee N_\psi \equiv \bigvee\langle\Diamond_j\chi : \chi \in N_\psi\rangle$. After this rewriting, bring φ into disjunctive normal form. Each conjunct of φ is now of the form

$$\mu = \nu \wedge \bigwedge_{C\in G} \Diamond_j\chi^{k-1}_C \wedge \bigwedge_{C\in H} \neg\Diamond_j\chi^{k-1}_C$$

where ν is nonmodal and $G \cap H = \varnothing$, G and H sequences of subsets of $nf(J,P,k-1)$. This is not necessarily in normal form, since we may have $G \cup H \subsetneq nf(J,P,k-1)$. But if there is a $\Diamond_j\chi$ which has not yet been included in $G \cup H$ we expand μ by the disjunction $\Diamond_j\chi \vee \neg\Diamond_j\chi$, and we get

$$\mu \equiv \mu \wedge \Diamond_j\chi. \vee .\mu \wedge \neg\Diamond_j\chi$$

In this way we can expand μ so as to include all $\Diamond_j\chi$, $\chi \in nf(J,P,k-1)$. The same with ν. Repeat this procedure as often as necessary. Finally, we reach normal form. □

Lemma 2.7.11. *Any two distinct normal forms of $nf(J,P,k)$ are jointly inconsistent.*

Proof. By induction on k. If $k = 0$, then let $C, C' \subseteq n$ two distinct subsets. Without loss of generality we may assume $C - C' \neq \varnothing$. Then $\chi^0_C \wedge \chi^0_{C'} \equiv \bot$. For there is an $i \in C$ such that $i \notin C'$. Then $\chi^0_C \vdash p_i$ but $\chi^0_{C'} \vdash \neg p_i$. Now let $k > 0$. If χ and χ' are distinct forms, then either they have distinct nonmodal components, or there is a $j \in J$ such that $D(j) \neq D'(j)$. In the first case we already have seen that there arises a contradiction. In the second case, assume without loss of generality, $\omega \in D(j) - D'(j)$. Then $\chi \vdash \Diamond_j\omega$ but $\chi' \vdash \neg\Diamond_j\omega$, a contradiction. □

Proposition 2.7.12. *The formulae of degree k with variables from the set $P = \{p_i : i < n\}$ and modal operators from J form a boolean algebra. The number of* *atoms is bounded by the number $b(j, n, k)$ defined by*

$$\begin{array}{lcl} b(j, n, 0) & := & 2^n \\ b(j, n, k+1) & := & b(j, n, 0) \cdot 2^{j \cdot b(j,n,k)} \end{array}$$

Proof. We have seen that each formula is equivalent to a disjunction of normal forms, so there are at most $2^{b(j,n,k+1)}$ formulae, where $b(j, n, k+1)$ is the cardinality of $nf(J, P, k+1)$. Moreover, since the normal forms are mutually inconsistent, they form the atoms of this algebra. The number of atoms is obtained by multiplying the choices for a normal form of degree 0 with the choices of j–long sequences of sets of normal forms of degree k. The latter is nothing but $2^{b(j,n,k)^j} = 2^{j \cdot b(j,n,k)}$. ◻

Thus, each set of normal forms individually presents a representative for a class of equivalent propositions. With a normal form we can associate a direct model as before. However, this time there are no clashes, and the valuation β is uniquely defined. For if x_φ is given, then φ is equivalent to $\varphi \wedge p_i \vee \varphi \wedge \neg p_i$, so that it must be equivalent to either of them, showing that p_i or $\neg p_i$ must be a conjunct of φ.

Notice that there are nontrivial propositions with no variables as we have seen earlier. Their number is bounded by the modal degree and the set of occurring operators. In many logics, however, there are up to equivalence only two distinct constant propositions, $\bot$ and $\top$. We say that such logics have **trivial constants**.

Theorem 2.7.13. *A modal logic Λ has trivial constants iff for every operator $\Box_j$ either $\Box_j \bot \in \Lambda$ or $\Diamond_j \top \in \Lambda$.*

The proof is simple and is omitted. Notice that the postulate $\Box_j \bot$ says that no world has a j–successor, while $\Diamond_j \top$ says that every world has a j–successor.

Normal forms are closely connected with a technique called *unravelling* introduced in Section 3.3. The method of unravelling can be used to show that $\mathbf{K}_\kappa$ is complete with respect to completely intransitive trees, by first showing that it is complete and then using unravelling to get a totally intransitive tree from a model. This is somewhat better than the proof via normal forms, which established completeness with respect to acyclic frames only. Furthermore, we can show now the following rather important fact. (Recall the definition of $\boxtimes$ from Section 2.1.)

Theorem 2.7.14. *Let κ be finite and $\mathfrak{A}$ an n–generated κ–modal algebra. If $\mathfrak{A} \vDash \boxtimes^k \bot$, $k > 0$, then $\mathfrak{A}$ is finite with at most $2^{b(\kappa,n,k-1)}$ elements.*

Proof. Let $a_0, \ldots, a_{n-1}$ be the generators of $\mathfrak{A}$. An arbitrary element of $\mathfrak{A}$ is of the form $\varphi[a_0, \ldots, a_{n-1}]$, where φ is a formula in the variables $p_0, \ldots, p_{n-1}$. We will show that for any formula φ there exists a formula $[\varphi]_k$ of degree $\leq k$ such that $\boxtimes^k \bot \wedge \varphi$ is deductively equivalent to $\boxtimes^k \bot \wedge [\varphi]_k$ in $\mathbf{K}_\kappa$. Since the number of such formulae is at most $2^{b(\kappa,n,k-1)}$, we are done. So, let φ be of depth $> k$. We assume that negation is in front of variables, double negations killed. Let $[\varphi]_k$ denote the formula obtained as follows.

$$\begin{array}{lcllcl}
[p]_0 & := & \bot & [\neg p]_0 & := & \bot \\
[p]_{k+1} & := & p & [\neg p]_{k+1} & := & \neg p \\
[\varphi \wedge \psi]_k & := & [\varphi]_k \wedge [\psi]_k & [\varphi \vee \psi]_k & := & [\varphi]_k \vee [\psi]_k \\
[\Diamond_j \varphi]_{k+1} & := & \Diamond_j [\varphi]_k & [\Box_j \varphi]_{k+1} & := & \Box_j [\varphi]_k
\end{array}$$

Roughly speaking, $[\varphi]_k$ is the result of replacing all occurrences of subformulas embedded exactly k times by modal operators by $\bot$. Then φ is the result of replacing some occurrences of $\bot$ by some formulae χ we have $[\varphi]_k \vdash \varphi$. (Notice, namely, that the occurrences are not embedded in any negation.) It remains to be shown that $\boxtimes^k \bot; \varphi \vdash [\varphi]_k$. This is done by induction on k and φ. The case $k = 0$ is straightforward. Now let $k > 0$. Let $\varphi = \psi \wedge \omega$. By inductive hypothesis, $[\varphi]_k = [\psi]_k \wedge [\omega]_k$. By inductive hypothesis $\boxtimes^k \bot; \psi \vdash [\psi]_k$ and $\boxtimes^k \bot; \omega \vdash [\omega]_k$. Hence $\boxtimes^k \bot; \psi \wedge \omega \vdash [\psi]_k \wedge [\omega]_k$. Analogously the case $\varphi = \psi \vee \omega$ is treated. Now assume $\varphi = \Box_j \psi$ and $k > 0$. Then $\boxtimes^{k-1} \bot; \psi \vdash [\psi]_{k-1}$. From this we obtain $\boxtimes^k \bot; \Box_j \psi \vdash \Box_j [\psi]_{k-1}$. Since $[\Box_j \psi]_k = \Box_j [\psi]_{k-1}$ we obtain the claim. Analogously for $\varphi = \Diamond_j \psi$. ∎

Exercise 58. Prove Lemma 2.7.2.

Exercise 59. Let $\langle \mathfrak{f}, \beta, x\rangle$ be a model. Show that there exists exactly one normal form χ of degree n for given J and P such that $\langle \mathfrak{f}, \beta, x\rangle \vDash \chi$. This form will be denoted by $\chi^n(x)$.

Exercise 60. With notation as in the previous exercise, let $w \sim_n x$ if $\chi^n(w) = \chi^n(x)$. Let $w \approx x$ if $w \sim_n x$ for all n. Show that the map $x \mapsto x/\approx$ is a p–morphism, and compatible with the valuation.

Exercise 61. Call $\langle \mathfrak{f}, \beta\rangle$ *condensed* if $\approx$ is the identity. Let $\langle \mathfrak{f}, \beta, w_0\rangle$ be condensed and $\mathfrak{f}$ generated by w_0. Show that there exists a ℓ–localic map from $\chi^\ell(w_0)$ into $\mathfrak{f}$. (See Section 3.3 for a definition.)

Exercise 62. Let χ and χ^+ be normal forms. Call χ^+ an *elaboration* if it is of depth at least that of χ, and if $\chi^+ \vdash \chi$. Characterize this notion syntactically.

Exercise 63. Show that in $\mathbf{K.alt}_1$ every formula is equivalent to a formula of the following $\mathbf{alt}_1$*–form*. For degree 0, a formula is in $\mathbf{alt}_1$–form iff it is in disjunctive normal form. A formula of degree $d+1$ is in $\mathbf{alt}_1$–normal form iff it is of the form either $\mu \wedge \Box\bot$ and $d = 1$ or $\mu \wedge \Diamond\varphi$ where μ is a $\mathbf{alt}_1$–form degree 0 and φ an $\mathbf{alt}_1$–form of degree d.

2.8. The Lindenbaum–Tarski Construction

An important question from the model theoretic point of view is the problem whether for a given logic Λ there is a frame $\mathfrak{F}$ such that $\Lambda = \mathsf{Th}(\mathfrak{F})$. Notice that in Section 2.2 we have shown that there exist algebras of this sort. We will now show that there are also frames with this property. The solution is due to Lindenbaum and Alfred Tarski. The basic idea is that a *world* is a maximally consistent set of formulae. Given a logic Λ this is a set $W \subseteq$ *wff* such that W is consistent but no proper superset is. The intuition behind this terminology is that a world is an existing thing and everything we say about it should therefore be either true or false. Clearly, one is more inclined to say that there can be only *one* world, the world we are living in, so speaking of collections of worlds can then only be metaphorical. There are ways to get around this apparent problem. For now we are interested in the connection between our logic and the worlds that might exist. The basic result of this section is that if we define a certain natural frame from Λ over a given set of propositional variables then the theory of that frame will be exactly Λ. This shows that every modal logic is the theory of a single frame and so frame semantics is as good as algebraic semantics. In Chapter 4 we will see that this is no accident. The construction is a specialization of a general technique to form geometric models from algebraic models. We proceed as follows. First, we show that there are enough maximally consistent sets (or worlds). This proof is completely analogous to that of Corollary 1.7.13. Second, we establish the frame based on these worlds. And thirdly we show that the logic of this frame is Λ.

Let us begin with the question of the existence of worlds. With respect to a logic Λ a **world** is a maximally Λ–consistent set of formulas. The next lemma asserts that for any consistent collection of facts there is a world in which it is realized.

Lemma 2.8.1. *Every Λ–consistent set is contained in a maximally Λ–consistent set.*

The proof is immediate from Tukey's Lemma. A maximally consistent set is also deductively closed, as can easily be shown. We note the following properties, of which we will make tacit use later on. These are easy to prove (cf. Section 1.7).

Lemma 2.8.2. *Let W be a deductively closed set of formulae.*

(1) *W is consistent iff for no φ: $\varphi \in W$ and $\neg\varphi \in W$.*
(2) *$\varphi \wedge \chi \in W$ iff $\varphi \in W$ and $\chi \in W$.*
(3) *If $\varphi \in W$ or $\chi \in W$ then $\varphi \vee \chi \in W$.*
(4) *W is maximally consistent iff it is consistent and $\varphi \vee \chi \in W$ implies $\varphi \in W$ or $\chi \in W$, for all φ and χ.*
(5) *W is maximally consistent iff for all φ: $\neg\varphi \in W$ exactly if $\varphi \notin W$.*

We now need to specify the relations in which these worlds stand to each other. For that we have a plausible definition. Given a modal operator $\Box_j$ (and its dual $\Diamond_j$) and two worlds W and X we want to say that X is *j–possible* at W if the total collection of facts in X is j–possible at W, that is, we would like to say that

$$W \lhd_j X \quad \Leftrightarrow \quad \Diamond_j \bigwedge X \in W$$

However, X is infinite, so we cannot define things in this way. We have no infinite conjunction, only a finitary one. So our best approximation is

$$W \lhd_j X \quad \Leftrightarrow \quad \text{for all finite subsets } X_0 \text{ of } X\text{: } \Diamond_j \bigwedge X_0 \in W$$

This is actually how we define the accessibility relation. However, this definition can be phrased more elegantly. Notice, namely, that if X_0 is a finite set then $\bigwedge X_0 \in X$. For from $\bigwedge X_0 \notin X$ follows $\bigvee\langle\neg\varphi : \varphi \in X_0\rangle \in X$, so for one $\varphi \in X_0$ also $\neg\varphi \in X$, which cannot be.

To introduce the definition in its final form let us agree on the abbreviation $\Diamond_j S := \{\Diamond_j\varphi : \varphi \in S\}$. Then we define

(acc.) $\quad W \lhd_j X \quad \Leftrightarrow \quad \Diamond_j X \subseteq W$

There is an alternative characterization as follows.

LEMMA 2.8.3. *Let W and X be worlds. Then the following are equivalent.*

(1) *For all $\varphi \in X$: $\Diamond_j\varphi \in W$.*
(2) *For all $\Box_j\varphi \in W$: $\varphi \in X$.*

PROOF. Suppose that the first holds and assume $\varphi \notin X$. Then $\neg\varphi \in X$ and so $\Diamond_j\neg\varphi \in W$. Since $\vdash_{\mathbf{K}} \Diamond_j\neg\varphi. \leftrightarrow .\neg\Box_j\varphi$ we also have $\neg\Box_j\varphi \in W$. Thus $\Box_j\varphi \notin W$, by consistency of W. So, the second holds. Now suppose that the second holds and assume $\Diamond_j\varphi \notin W$. Then we have $\neg\Diamond_j\varphi \in W$, thus $\Box_j\neg\varphi \in W$, which by assumption implies $\neg\varphi \in X$. Hence $\varphi \notin X$. So, the first holds. ∎

The construction also yields enough worlds in the following sense. If $\Diamond_j\varphi \in W$ then there is an X such that $W \lhd_j X$ and $\varphi \in X$. For consider the set $S := \{\varphi\} \cup \{\chi : \Box_j\chi \in W\}$. If it is consistent, there is a world $X \supseteq S$ and we must have $W \lhd_j X$ by construction. So we have to show that S is consistent. Suppose it is not. Then there is a finite set $S_0 \subseteq S$ which is inconsistent. Without loss of generality $S_0 = \{\varphi\} \cup T_0$. Now $T_0 \vdash_\Lambda \neg\varphi$, and so $\Box_j T_0 \vdash_\Lambda \Box_j\neg\varphi$. Since $\Box_j T_0 \subseteq W$ we must have $\Box_j\neg\varphi \in W$, which is to say $\neg\Diamond_j\varphi \in W$, and so by consistency of W, $\Diamond_j\varphi \notin W$. This is contrary to our assumptions, however. Thus, S is consistent and successor worlds containing φ exist.

PROPOSITION 2.8.4. *Whenever W is a world and $\Diamond_j\varphi \in W$ there exists a world X such that $W \lhd_j X$ and $\varphi \in X$.*

Finally, the internal sets must be specified. This is most straightforward. Internal sets are those which are specifiable by a proposition. Define $\widehat{\varphi}$ by

$$\widehat{\varphi} := \{X : \varphi \in X\}$$

We call $v : p \mapsto \widehat{p}$ the **natural valuation**.

Lemma 2.8.5. *Let* $\widehat{\varphi} := \{X : \varphi \in X\}$ *and* $\nu : p \mapsto \widehat{p}$. *Then for all* φ, $\overline{\nu}(\varphi) = \widehat{\varphi}$.

Proof. By induction on φ. Suppose $\varphi = \neg\chi$. Then $\widehat{\neg\chi} = \{W : \neg\chi \in W\}$ and by induction hypothesis $\overline{\nu}(\chi) = \widehat{\chi} = \{W : \chi \in W\}$. Since the worlds are maximally consistent, $\chi \in W$ is the same as $\neg\chi \notin W$, so $\widehat{\neg\chi} = -\widehat{\chi} = -\overline{\nu}(\chi)$, from which the claim follows. Now assume $\varphi = \chi_1 \wedge \chi_2$. We have $\widehat{\chi_1 \wedge \chi_2} = \widehat{\chi_1} \cap \widehat{\chi_2}$, as is easily checked. Therefore $\overline{\nu}(\chi_1 \wedge \chi_2) = \overline{\nu}(\chi_1) \cap \overline{\nu}(\chi_2) = \widehat{\chi_1} \cap \widehat{\chi_2} = \widehat{\chi_1 \wedge \chi_2}$. Finally, assume that $\varphi = \Diamond_j\chi$. We have $W \in \widehat{\Diamond_j\chi}$ iff $\Diamond_j\chi \in W$ iff there is a j–successor X such that $\chi \in X$ (by Proposition 2.8.4) iff there is a j–successor X such that $X \in \widehat{\chi}$ iff $W \in \blacklozenge_j\widehat{\chi}$. Hence $\overline{\nu}(\varphi) = \overline{\nu}(\Diamond_j\chi) = \blacklozenge_j\overline{\nu}(\chi) = \blacklozenge_j\widehat{\chi} = \widehat{\Diamond_j\chi} = \widehat{\varphi}$, as required. ☐

Under the supposition that the valuation is the natural valuation the internal sets are exactly those which are values of the formulae of our modal language.

Definition 2.8.6. *Let* Λ *be a normal logic. Denote the set of worlds by* W_Λ *and let* $\mathbb{W}_\Lambda = \{\widehat{\varphi} : \varphi \in wff\}$. *Define* $\lhd_j$ *by* $X \lhd_j Y$ *iff for all* $\Box_j\varphi \in X$, $\varphi \in Y$ *(= (acc.)). Then the* ***canonical frame*** *for* Λ *is the frame* $\mathfrak{Can}_\Lambda(var) = \langle W_\Lambda, \langle \lhd_j : j < \kappa\rangle, \mathbb{W}_\Lambda\rangle$. *The underlying Kripke–frame is denoted by* $\mathfrak{can}_\Lambda(var)$. *The* ***global canonical model*** *for* Λ *is the pair* $\langle \mathfrak{Can}_\Lambda(var), \nu\rangle$ *where* $\nu(p) = \widehat{p} = \{W : p \in W\}$. *A* ***local canonical model*** *is a triple* $\langle \mathfrak{Can}_\Lambda(var), \nu, X\rangle$, *where* X *is a world.*

What we know now is that if a formula φ is not in Λ, then there exists a W such that $\neg\varphi \in W$, since $\neg\varphi$ is consistent with Λ. Then $\langle \mathfrak{Can}_\Lambda, \nu, W\rangle \vDash \neg\varphi$. But can we be sure that Λ is the logic of the frame? Is it possible that there are countermodels for axioms of Λ? We will show that this is not the case. The reason is the choice of the internal sets. Notice that the internal sets can be 'named' in the canonical model by a formula. Namely, for every $a \in \mathbb{W}_\Lambda$ there exists a φ such that $a = \widehat{\varphi} = \overline{\nu}(\varphi)$. So, let β be an arbitrary valuation. Then for each p there exists a formula φ_p such that $\beta(p) = \overline{\nu}(\varphi_p)$. Then for an arbitrary formula ψ,

$$\overline{\beta}(\psi) = \overline{\nu}(\psi[\varphi_p/p : p \in var(\psi)])$$

Thus, if a model exists for ψ based on the valuation β, then for some substitution σ, a model for ψ^σ exists based on the valuation ν. Another way of seeing this is using the Theorem 2.8.8 below. Suppose namely that an axiom φ of Λ is violated. Then, by our arguments, a substitution instance φ^σ (which is also an axiom) is violated on the canonical model. This means, however, that there is a world W such that $\neg\varphi^\sigma \in W$. Now, since $\neg\varphi^\sigma$ is inconsistent with Λ this simply cannot be. Hence we have shown that no axiom can be refuted on any model based on the canonical frame.

Theorem 2.8.7 (Canonical Frames). *Let* Λ *be a normal polymodal logic. Then* $\Lambda = \mathsf{Th}\,\mathfrak{Can}_\Lambda(var)$.

In writing $\mathfrak{Can}_\Lambda(var)$ we have indicated that the canonical frame also depends on the set of variables that we have available. In fact, the structure of the canonical frame is up to isomorphism determined only by the cardinality of the set of variables.

(We will see that even the logic really depends on the latter.) We have assumed that this set has countably many elements. Thus we also write $\mathfrak{Can}_\Lambda(\aleph_0)$.

A different proof of Theorem 2.8.7 consists in showing that the algebra of internal sets is isomorphic to $\mathfrak{Fr}_\Lambda(var)$. This follows from the fact that the algebra of sets is the direct image of the map $\varphi \mapsto \widehat{\varphi}$. We have verified above that this map is a homomorphism and that $\widehat{\varphi} = \widehat{\psi}$ iff $\vdash_\Lambda \varphi \leftrightarrow \psi$.

Theorem 2.8.8. *The algebra of internal sets of* $\mathfrak{Can}_\Lambda(var)$ *is isomorphic to* $\mathfrak{Fr}_\Lambda(var)$. *An isomorphism is given by the map* $\varphi \mapsto \widehat{\varphi}$.

Analogous techniques can be used for quasi–normal logics. Suppose that Ψ is a quasi–normal logic extending Λ. Then any Ψ–consistent set is contained in a maximally Ψ–consistent set. For this set there is a model of the form $\langle \mathfrak{Can}_\Lambda(var), v, W\rangle$. Again, since a quasi–normal logic is closed under substitution, we get that $\langle \mathfrak{Can}_\Lambda(var), W\rangle \vDash \Psi$.

Theorem 2.8.9 (Canonical Frames for Quasi–Normal Logics). *Any quasi–normal logic* $\Psi \supseteq \Lambda$ *can be obtained by*

$$\Psi = \bigcap_{W \in S} \mathsf{Th}\,\langle \mathfrak{Can}_\Lambda(var), W\rangle$$

for a set $S \subseteq W_\Lambda$.

In a normal logic we have $S = W_\Lambda$ by closure under (mn.).

As we have noted earlier in connection with free algebras, the structure of $\mathfrak{Can}_\Lambda(var)$ only depends on the cardinality of the set *var*. Let it be α. Then we also write $\mathfrak{Can}_\Lambda(\alpha)$ and call it the α–**canonical frame for** Λ. As it will turn out, the structure of a canonical frame depends nontrivially on the cardinality of the set of variables. The question is whether the canonical frames for Λ for different cardinalities of the variables are related by certain p–morphisms. We will show that a function between two cardinal numbers induces a p–morphism of the associated canonical frames. So, take two cardinals α and β and a function $f : \alpha \to \beta$. f induces a homomorphism $h_f : \mathfrak{Tm}(\alpha) \to \mathfrak{Tm}(\beta)$ defined by $h_f(p_i) := p_{f(i)}$. h_f is uniquely determined by f. Moreover, the Theorem 2.8.11 below mirrors Theorem 1.3.6, showing that the maps between the algebras have a correlate for the canonical frames.

Lemma 2.8.10. *Let X be a world in the language over* $\{p_i : i < \beta\}$ *and* $f : \alpha \to \beta$. *Then* $h_f^{-1}[X]$ *is a world in the language with variables* $\{p_i : i < \alpha\}$.

Proof. (1.) $h_f^{-1}[X]$ is deductively closed. For let $\varphi; \varphi \to \psi \in h_f^{-1}[X]$. Then $h_f(\varphi) \in X$ and $h_f(\varphi) \to h_f(\psi) \in X$. Since X is deductively closed, $h_f(\psi) \in X$. So, $\psi \in h_f^{-1}[X]$. (2.) $h_f^{-1}[X]$ is consistent. For if $\varphi; \neg\varphi \in h_f^{-1}[X]$ then $h_f(\varphi) \in X$ and $\neg h_f(\varphi) \in X$. But X is consistent. So, either $h_f(\varphi) \notin X$ or $\neg h_f(\varphi) \notin X$. Hence either $\varphi \notin h_f^{-1}(X)$ or $\neg\varphi \notin h_f^{-1}[X]$. (3.) $h_f^{-1}[X]$ is maximally consistent. For let

$\varphi \vee \psi \in h_f^{-1}[X]$. Then $h_f(\varphi) \vee h_f(\psi) \in X$. By maximality of X, $h_f(\varphi) \in X$ or $h_f(\psi) \in X$. Thus we have $\varphi \in h_f^{-1}[X]$ or $\psi \in h_f^{-1}[X]$. ∎

Theorem 2.8.11. *Let α and β be cardinal numbers and $f : \alpha \to \beta$ a map. Let X_f denote the map $X \mapsto h_f^{-1}[X]$. Then $X_f : \mathfrak{Can}_\Lambda(\beta) \to \mathfrak{Can}_\Lambda(\alpha)$. Moreover, if f is injective, X_f is surjective and if f is surjective then X_f is injective.*

Proof. By the previous lemma, X_f is a map between the sets of worlds. Now assume that Y and Z are worlds over $\{p_i : i < \beta\}$ and $Y \lhd_j Z$. We claim that $X_f(Y) \lhd_j X_f(Z)$. For let $\Box_j\varphi \in X_f(Y) = h_f^{-1}[Y]$. Then $\Box_j h_f(\varphi) \in Y$. Hence $h_f(\varphi) \in Z$, since $Y \lhd_j Z$. This shows $\varphi \in h_f^{-1}[Z] = X_f(Z)$. The first condition for p–morphisms is proved. Now assume that Y is a world over $\{p_i : i < \beta\}$ and U a world over $\{p_i : i < \alpha\}$; and let $X_f(Y) \lhd_j U$. Then for every $\Box_j\varphi$ such that $\Box_j h_f(\varphi) \in Y$ we have $\varphi \in U$. Let $Y_\Box := \{\varphi : \Box_j\varphi \in Y\}$. The set $h_f[U] \cup Y_\Box$ is consistent. For take finite sets $\Delta_0 \subseteq h_f[U]$ and $\Delta_1 \subseteq Y_\Box$. We have $h_f^{-1}[\Delta_0] \subseteq U$ and $h_f^{-1}[\Box_j\Delta_1] \subseteq h_f^{-1}[Y]$. By assumption on Y and U, $h_f^{-1}[\Delta_1] \subseteq U$. Therefore, the set $h_f^{-1}[\Delta_0; \Delta_1]$ is Λ–consistent. Then $\Delta_0; \Delta_1$ is Λ–consistent as well. So, every finite subset of $h_f[U] \cup Y_\Box$ is Λ–consistent. This set is therefore Λ–consistent and has a maximally consistent extension. Call it V. Then for every $\Box_j\varphi \in Y$ we have $\varphi \in V$, and so $Y \lhd_j V$. Furthermore, $h_f^{-1}[V] \supseteq h_f^{-1}[h_f[U]] \supseteq U$. Since $h_f^{-1}[V]$ is consistent and U maximally consistent, $h_f^{-1}[V] = U$, and that had to be shown. This proves the second p–morphism condition. Finally, let $\widehat{\varphi}$ be an internal set of $\mathfrak{Can}_\Lambda(\alpha)$. Then $X_f^{-1}[\widehat{\varphi}] = \{X_f(Y) : \varphi \in Y\} = \{Z : h_f(\varphi) \in Z\} = \widehat{h_f(\varphi)}$. Therefore, $X_f^{-1}[\widehat{\varphi}]$ is an internal set of $\mathfrak{Can}_\Lambda(\beta)$ showing that X_f is indeed a p–morphism.

Now assume that $f : \alpha \to \beta$ is injective and let U be a world over the set $\{p_i : i < \alpha\}$. Then $h_f[U]$ is a world over $\{p_{f(i)} : i < \alpha\}$. Hence it is a consistent set over $\{p_i : i < \beta\}$. Let $V \supseteq h_f[U]$ be a world. Then $h_f^{-1}[V] \supseteq h_f^{-1}[h_f[U]]$. Hence, $X_f(V) = h_f^{-1}[V] = U$. So X_f is onto. Now assume that f is surjective and let Y and Z be two different worlds over $\{p_i : i < \beta\}$. Then there exists a formula φ such that $\varphi \in Y$ but $\varphi \notin Z$. There exists a formula ψ over $\{p_i : i < \alpha\}$ such that $h_f(\psi) = \varphi$. (Simply choose a function $g : \beta \to \alpha$ such that $g \circ f(i) = i$, for all $i < \alpha$. Then put $\psi = h_g(\varphi)$.) It follows that $\psi \in X_f(Y)$ but $\psi \notin X_f(Z)$. ∎

An application of these theorems can be found in so–called *weak canonical frames*. Suppose we restrict ourselves to a finite subset, say $\{p_0, \ldots, p_{n-1}\}$. Languages and logics based on finitely many variables are called **weak** in [**66**]. If Λ is a logic then $\Lambda \restriction n$ denotes the set of theorems of Λ in the variables $\{p_i : i < n\}$. It is clear that $\Lambda = \bigcup_{n\in\omega} \Lambda \restriction n$. So a logic is determined already by its weak fragments. As it turns out, $\mathfrak{Can}_{\Lambda\restriction n}(n) = \mathfrak{Can}_\Lambda(n)$. (This is straightforward to verify.) We call a frame of the form $\mathfrak{Can}_\Lambda(n)$, $n \in \omega$, a **weak canonical frame**. It follows from Theorem 2.8.11 that if $n < m$, $\mathfrak{Can}_\Lambda(n) \rightarrowtail \mathfrak{Can}_\Lambda(m)$ and so $\mathsf{Th}\,\mathfrak{Can}_\Lambda(n) \supseteq \mathsf{Th}\,\mathfrak{Can}_\Lambda(m)$.

The situation is therefore as follows.

$$\Lambda = \mathsf{Th}\,\mathfrak{Can}_\Lambda(\aleph_0) \subseteq \ldots \subseteq \mathsf{Th}\,\mathfrak{Can}_\Lambda(2) \subseteq \mathsf{Th}\,\mathfrak{Can}_\Lambda(1) \subseteq \mathsf{Th}\,\mathfrak{Can}_\Lambda(0)$$

Equality need not hold. Furthermore, the following is easy to establish.

Theorem 2.8.12 (Weak Canonical Frames). *Let Λ be a normal logic. Then* $\Lambda = \bigcap_{n\in\omega} \mathsf{Th}\,\mathfrak{Can}_\Lambda(n)$.

The question arises what happens if a logic is equal to the theory of one of its weak frames.

Definition 2.8.13. *Let Λ be a modal logic. We call Λ n–**characterized** if*

$$\Lambda = \mathsf{Th}\,\mathfrak{Can}_\Lambda(n)\ .$$

Θ is n–**axiomatizable over** Λ if $\Theta = \Lambda \oplus \Delta$ for some set Δ with $var[\Delta] \subseteq \{p_i : i < n\}$. Θ is n–**axiomatizable** if Θ is n–axiomatizable over $\mathbf{K}_\kappa$. If Θ is n–axiomatizable over Λ it need not be n–axiomatizable simpliciter, for Λ itself may not be n–axiomatizable. Clearly, if Θ is finitely axiomatizable, it is also n–axiomatizable for some n, but the converse need not hold. Furthermore, for every n one can find a logic which is $n+1$–axiomatizable, but not n–axiomatizable.

Exercise 64. Let $\mathfrak{M} = \langle \mathfrak{F}, \beta, x\rangle$ be a local model. Show that $\mathsf{Th}\,\mathfrak{M}$, the theory of the point x in the model, is a maximally consistent set.

Exercise 65. Show that two worlds X, Y in the canonical frame are different iff there is a formula φ such that $\varphi \in X$ and $\varphi \notin Y$. Show that $X \ntriangleleft_j Y$ iff there is a formula φ such that $\Box_j\varphi \in X$ but $\varphi \notin Y$. Show that for every ultrafilter U in $\mathbb{W}_\Lambda$, $\bigcap U \neq \varnothing$. (The first property is called *differentiatedness*, the second *tightness* and the third *compactness*. See Section 4.6.)

Exercise 66. Let Λ_1, Λ_2 be two normal logics. Show that $\Lambda_1 \subseteq \Lambda_2$ iff $\mathfrak{Can}_{\Lambda_2}(\aleph_0)$ is a generated subframe of $\mathfrak{Can}_{\Lambda_1}(\aleph_0)$. *Hint.* Every Λ_2–consistent set is also Λ_1–consistent.

Exercise 67. Show that $\mathfrak{Can}_{\mathbf{K}}(\aleph_0)$ contains $2^{\aleph_0}$ many worlds, and $2^{\aleph_0}$ worlds without successor. Show that if Λ is a normal logic and T, U worlds such that $T \triangleleft_i U$ then there are $2^{\aleph_0}$ i–predecessors of U. What about the weak canonical frames?

Exercise 68. Let Λ be a logic and $\mathfrak{g}$ a finite Kripke–frame for Λ. Show that $\langle \mathfrak{g}, 2^g\rangle$ is a generated subframe of $\mathfrak{Can}_\Lambda(\aleph_0)$.

Exercise 69. Show that $\mathsf{Th}\,\langle n, >\rangle$ is 0–axiomatizable. *Hint.* The axiom $\Box^n\bot$ is obviously not sufficient, but a good start. Now choose additional constant axioms carefully so that only the intended frames remain as frames for the logic.

2.9. The Lattices of Normal and Quasi–Normal Logics

Recall that a lattice is complete if any set S has a least upper bound, denoted by $\bigsqcup_{x \in S} x$ or simply by $\bigsqcup S$, and a greatest lower bound, denoted by $\bigsqcap_{x \in S} x$ or simply by $\bigsqcap S$. A lattice has greatest lower bounds iff it has least upper bounds. Hence, a lattice is complete iff it has least upper bounds. We can rephrase completeness by using *limits* of directed systems. Let $\mathfrak{I} = \langle I, \leq \rangle$ be a partially ordered set. This set is **directed** if for $i, j \in I$ there exists a $k \in I$ such that $i, j \leq k$. An indexed family $\langle x(i) : i \in I \rangle$ is called an **upward directed system over** $\mathfrak{I}$ if $x(i) \leq x(j)$ whenever $i \leq j$. For example, let S be a subset of L. Take $I := S^{<\aleph_0}$, the set of all finite subsets of S, ordered by inclusion. For a $d \in I$ put $x(d) := \bigsqcup d$. Let $S^+ := \langle x(d) : d \in S^{<\aleph_0} \rangle$. S^+ is an upward directed system. For an upward directed system $X = \{x_i : i \in I\}$ we write $\lim_{\mathfrak{I}} X$ to denote the least upper bound. It is clear that we have $\lim_{\mathfrak{I}} S^+ = \bigsqcup S$. (With S given, the set I is uniquely defined, and may be dropped.) Analogously a **downward directed system** is defined. For a downward directed system we write $\lim_{\mathfrak{I}} S$ for the intersection. If S is upward (downward) directed, so is $x \sqcup S = \{x \sqcup y : y \in S\}$ and $x \sqcap S = \{x \sqcap y : y \in S\}$. A lattice is called **upper continuous** if intersection commutes with upward limits, that is, if $x \sqcap \lim S = \lim (x \sqcap S)$ for all upward directed sets S. It is called **lower continuous** if join commutes with downward limits, that is, if $x \sqcup \lim S = \lim (x \sqcup S)$. A complete lattice is **continuous** if it is both upper and lower continuous.

Theorem 2.9.1. *A complete, distributive lattice is upper continuous iff it satisfies the law (jdi.) and lower continuous iff it satisfies the law (mdi.).*

$$\begin{array}{lrcl} (jdi.) & a \sqcap \bigsqcup B & = & \bigsqcup (a \sqcap B) \\ (mdi.) & a \sqcup \bigsqcap B & = & \bigsqcap (a \sqcup B) \end{array}$$

Proof. We show only the first claim, the second is dual. Let B be a set and B^+ be the family of finite joins of elements of B. This is an upward directed system. Then, by distributivity, the family $a \sqcap B^+$ defined by $a \sqcap B^+ := \langle a \sqcap \bigsqcup d : d \in B^{<\aleph_0} \rangle$ is identical to the family

$$(a \sqcap B)^+ := \langle \textstyle\bigsqcup_{x \in d} a \sqcap x : d \in B^{<\aleph_0} \rangle .$$

Thus we have the following identities.

$$\begin{array}{rcl} a \sqcap \bigsqcup B & = & a \sqcap \lim B^+ \\ & = & \lim a \sqcap B^+ \\ & = & \lim (a \sqcap B)^+ \\ & = & \bigsqcup (a \sqcap B) \end{array}$$

∎

Definition 2.9.2. *A* ***locale*** *is a complete, distributive and upper continuous lattice. A homomorphism from a locale $\mathfrak{L}$ to a locale $\mathfrak{M}$ is a map $h : L \to M$ commuting with finite intersections and arbitrary joins.*

We need to comment on this definition. Locales are also called *frames* in the literature and what we have called a homomorphism of locales is actually a homomorphism of frames (see [**110**]); a homomorphism between locales goes in the opposite direction. (The open sets of a topological space form a locale; the maps between them are continuous maps. More about this in Chapter 7.) However, due to a clash in terminology we have departed from this convention.

Logics in general are ordered by set inclusion. Moreover, they form a lattice with respect to this ordering. Normal modal logics are identified with their tautologies, so $\Lambda_1 \leq \Lambda_2$ is equivalent with the fact that all tautologies of Λ_1 are tautologies of Λ_2. It is possible to spell out exactly how to compute the join and meet of two logics if their axiomatization is known. Moreover, we will show that the lattice of normal logics is distributive and upper–continuous. First of all, however, note that if Λ_i, $i \in I$, is an indexed family of normal logics, then the intersection $\bigcap_{i \in I} \Lambda_i$ is a normal modal logic as well. The reason is simply that if each Λ_i is individually contains the classical tautologies and the (bd→.) postulates and is closed under the rules (sub.), (mp.) and (mn.), so is the intersection. (The reader might also recall that logics are defined as closed sets of a closure operator; an intersection of any number of closed sets is always closed.) The union, however, is generally not closed under these operations. On the other hand, if the logics Λ_i are axiomatized as $\mathbf{K}_\kappa \oplus X_i$, then the least logic containing all Λ_i must be $\mathbf{K}_\kappa \oplus \bigcup_{i \in I} X_i$. So, an axiomatization of the union is quite easily obtained. We are interested in an axiomatization of the meet as well. To this end, let us concentrate on the case of a finite intersection, say of $\mathbf{K}_\kappa \oplus X_1$ and $\mathbf{K}_\kappa \oplus X_2$. The next theorem tells us how the intersection can be axiomatized. The notation $\varphi \dot{\vee} \psi$ is used to denote the disjunction of φ^σ and ψ for some suitable renaming σ of variables such that $var(\varphi^\sigma) \cap var(\psi) = \varnothing$.

Theorem 2.9.3. *(i) Let $\Lambda_1 = \boldsymbol{K}_\kappa + X_1$ and $\Lambda_2 = \boldsymbol{K}_\kappa + X_2$ be quasi–normal logics. Then $\Lambda_1 \cap \Lambda_2 = \boldsymbol{K}_\kappa + Y$ where*

$$Y = \{\varphi \dot{\vee} \psi : \varphi \in X_1, \psi \in X_2\} .$$

(ii) Let $\Lambda_1 = \boldsymbol{K} \oplus X_1$ and $\Lambda_2 = \boldsymbol{K} \oplus X_2$ be two normal modal logics. Then $\Lambda_1 \cap \Lambda_2 = \boldsymbol{K} \oplus Y$ where

$$Y = \{\boxplus\varphi \dot{\vee} \boxplus\psi : \varphi \in X_1, \psi \in X_2, \boxplus \textit{ a compound modality}\} .$$

Remark. The logic $\Lambda_1 \cap \Lambda_2$ defined above does not depend on the choice of the renaming σ of variables in $\varphi \dot{\vee} \psi$.

Proof. The second claim follows from the first as follows. Assume that If $\Lambda_1 = \mathbf{K}_\kappa \oplus X_1$. Then $\Lambda_1 = \mathbf{K}_\kappa + \{\boxplus\varphi : \varphi \in X_1, \boxplus \text{ compound}\}$, and similarly for Λ_2. Then by (i), $\Lambda_1 \cap \Lambda_2 = \mathbf{K}_\kappa + Z$ for the set

$$Z := \{\boxplus_1\varphi \dot{\vee} \boxplus_2\psi : \varphi \in X_1, \psi \in X_2, \boxplus_1, \boxplus_2 \text{ compound}\}$$

However, the set Y as given in (ii) is actually sufficient. For notice that if $\boxplus(p) := \boxplus_1(p) \wedge \boxplus_2(p)$ then in $\mathbf{K}_\kappa$ we have $\boxplus\alpha \vee \boxplus\beta \vdash \boxplus_1\alpha \vee \boxplus_2\beta$; and so every member of Z can be deduced from a member of Y. Since the intersection is a normal logic, we actually have $\Lambda_1 \cap \Lambda_2 = \mathbf{K}_\kappa \oplus Y$. So let us prove the first claim. Let $\varphi \dot{\vee} \psi \in Y$. Then $\Lambda_1 \vdash \varphi \dot{\vee} \psi$ and $\Lambda_2 \vdash \varphi \dot{\vee} \psi$ so that $\mathbf{K}_\kappa + Y \subseteq \mathbf{K}_\kappa + X_1$ as well as $\mathbf{K}_\kappa + Y \subseteq \mathbf{K}_\kappa + X_2$. For the converse implication let $\chi \in \mathbf{K} \oplus X_1$ as well as $\chi \in \mathbf{K} \oplus X_2$. Then χ can be derived from substitution instances φ^τ of some $\varphi \in X_1$ using modus ponens, and from instances ψ^υ of some $\psi \in X_2$ using modus ponens. Now, in boolean logic, a formula χ follows both from a set V of formulae and from a set W iff it follows from $\{\alpha \vee \beta : \alpha \in V, \beta \in W\}$. (Namely, we have $\bigwedge V \to \chi$ as well as $\bigwedge W \to \chi$ and so $\bigwedge V \vee \bigwedge W. \to .\chi$, from which the claim follows by the distributivity laws.) So we can deduce χ from formulae of the form $\varphi^\tau \vee \psi^\upsilon$ with $\varphi \in X_1$, $\psi \in X_2$ for some substitutions τ, υ. Generally, it is not possible to find a single substitution ρ such that $\varphi^\tau \vee \psi^\upsilon = (\varphi \vee \psi)^\rho$, since τ and υ might disagree on a common variable. This is why we take instead of φ the formula φ^σ, which is disjoint in variables from φ. Then there is a substitution ρ such that $\varphi^\tau \vee \psi^\upsilon = (\varphi^\sigma \vee \psi)^\rho$. ∎

We remark here that the infinite meet of logics cannot be given a canonical axiomatization in this way. This should be at least plausible from the construction of the set Y above. This construction breaks down in the infinite case. This is the reason why lattices of modal logics are not continuous, that is, do not satisfy all infinite distributive laws. At the moment it is not proven that the construction does not work. We will provide an explicit counterexample in the exercises.

Theorem 2.9.4. *The set of quasi–normal κ–modal logics operators is a locale.*

Proof. Let $\Theta = \mathbf{K}_\kappa + X$, $\Lambda_i = \mathbf{K}_\kappa + Y_i$. Then

$$\begin{aligned}
\Theta \sqcap \textstyle\bigsqcup_{i\in I}\Lambda_i &= (\mathbf{K}_\kappa + X) \sqcap (\mathbf{K}_\kappa + \textstyle\bigcup_{i\in I} Y_i) \\
&= \mathbf{K}_\kappa + \{\varphi \dot{\vee} \psi : \varphi \in X, \psi \in \textstyle\bigcup_{i\in I} Y_i\} \\
&= \mathbf{K}_\kappa + \textstyle\bigcup_{i\in I}\{\varphi \dot{\vee} \psi : \varphi \in X, \psi \in Y_i\} \\
&= \textstyle\bigsqcup_{i\in I} \mathbf{K}_\kappa + \{\varphi \dot{\vee} \psi : \varphi \in X, \psi \in Y_i\} \\
&= \textstyle\bigsqcup_{i\in I} (\mathbf{K}_\kappa + X) \sqcap (\mathbf{K} + Y_i) \\
&= \textstyle\bigsqcup_{i\in I} \Theta \sqcap \Lambda_i
\end{aligned}$$

Moreover,

$$\begin{aligned}
\Theta \sqcup (\Lambda_1 \sqcap \Lambda_2) &= \mathbf{K}_\kappa + X \cup \{\varphi \dot{\vee} \psi : \varphi \in Y_1, \psi \in Y_2\} \\
&= \mathbf{K}_\kappa + \{\varphi \dot{\vee} \psi : \varphi \in X \cup Y_1, \psi \in X \cup Y_2\} \\
&= (\Theta \sqcup \Lambda_1) \sqcap (\Theta \sqcup \Lambda_2)
\end{aligned}$$

The step from the first to the second line needs justification. Put $\Delta := X \cup \{\varphi \dot{\vee} \psi : \varphi \in Y_1, \psi \in Y_2\}$, and $\Sigma := \{\varphi \dot{\vee} \psi : \varphi \in X \cup Y_1, \psi \in X \cup Y_2\}$. Assume $\varphi \in \Delta$. Then either $\varphi \in X$ or φ is of the form $\psi \dot{\vee} \chi$ where $\psi \in Y_1$ and $\chi \in Y_2$. Assume the first. Observe that $\varphi \dot{\vee} \varphi \in \Sigma$. Clearly, $\varphi \in \mathbf{K}_\kappa + \varphi \dot{\vee} \varphi$, and hence $\varphi \in \mathbf{K}_\kappa + \Sigma$. Now assume $\varphi \notin X$. Then it is of the form $\psi \dot{\vee} \chi$, with $\psi \in Y_1$ and $\chi \in Y_2$. Then also

$\varphi \in \Sigma$. This shows $\mathbf{K}_\kappa + \Delta \subseteq \mathbf{K}_\kappa + \Sigma$. For the converse inclusion, assume $\varphi \in \Sigma$. Then $\varphi = \psi \dot{\vee} \chi$ where $\psi \in X \cup Y_1$ and $\chi \in X \cup Y_2$. If either $\psi \in X$ or $\chi \in X$, then $\varphi \in \mathbf{K}_\kappa + X$ and so $\varphi \in \mathbf{K}_\kappa + \Delta$. However, if $\psi \in Y_1$ and $\chi \in Y_2$, then $\varphi \in \mathbf{K}_\kappa + \Delta$, since it is in Δ modulo renaming of some variables. □

Theorem 2.9.5. *The set of normal κ–modal logics forms a locale. Moreover, the natural embedding into the locale of quasi–normal logics is continuous, that is, it is a homomorphism with respect to the infinitary operations.*

Proof. Clearly, since the infinite intersection of normal logics is a normal logic, the embedding is faithful to arbitrary intersections. What we have to show is that the quasi–normal join of normal logics is also normal. To see this, let Θ_i, $i \in I$, be logics and let φ be deducible via (mp.) from $X \subseteq \bigcup_{i \in I} \Theta_i$. Then we know that $\Box_j\varphi$ is deducible from $\Box_j X$. By normality of the Θ_i, $\Box_j X \subseteq \bigcup_{i \in I} \Theta_i$. □

Definition 2.9.6. *Let Λ be a normal modal logic and Θ a quasi–normal logic. The locale of normal extensions of Λ is denoted by $\mathcal{E}\,\Lambda$; the locale of quasi–normal extensions of Θ is denoted by $\mathcal{Q}\,\Theta$. We usually speak of the* ***lattice*** *of (normal) extensions, rather than of the* ***locale*** *of (normal) extensions.*

Some authors use *NExt* Λ instead of $\mathcal{E}\,\Lambda$ and *Ext* Λ for $\mathcal{Q}\,\Lambda$. In algebraic terms, the underlying set of $\mathcal{E}\,\Theta$ forms a principal filter in $\mathcal{E}\,\mathbf{K}_\kappa$. As before, the lattice of normal extensions is a complete sublattice of $\mathcal{Q}\,\Theta$, the lattice of quasi–normal extensions. Notice that when Θ is not normal then $\Theta \notin \mathcal{E}\,\Theta$. Nevertheless, $\mathcal{E}\,\Theta$ is a principal filter in $\mathcal{Q}\,\Theta$ induced by the normal closure of Θ, which is unique. We will rarely study lattices of quasi–normal extensions and be concerned only with lattices of normal extensions. The whole lattice $\mathcal{E}\,\mathbf{K}_\kappa$ is extremely complex even for $\kappa = 1$ as we shall see. However, for some strong logics the extension lattices are completely known, such as the logics $\mathbf{K.alt}_1$, **S5** and even **S4.3**. A large part of the study in modal logic has been centered around classifying extensions of certain strong logics in order to gain insight into the structure of the whole lattice $\mathcal{E}\,\mathbf{K}_1$.

Let $\mathfrak{I} = \langle I, \leq \rangle$ be a linearly ordered set. A **chain over** $\mathfrak{I}$ is an indexed family over $\mathfrak{I}$, that is, an order preserving map $j : \mathfrak{I} \to \mathcal{E}\,\mathbf{K}_\kappa$. The chain j is **properly ascending** if for $x, y \in I$ such that $x < y$ we have $j(x) \subsetneq j(y)$. The following is easy to establish.

Proposition 2.9.7. *Let $j : \mathfrak{I} \to \mathcal{E}\,\boldsymbol{K}$ be a properly ascending chain. Then if $\kappa \leq \aleph_0$, $\mathfrak{I}$ is at most countable. And if $\kappa > \aleph_0$, $\mathfrak{I}$ has cardinality $\leq \kappa$.*

Observe namely that a logic can be identified with its set of tautologies. That set is either countably infinite (in case $\kappa \leq \aleph_0$) or of size κ. The same holds for properly descending chains.

Recall from Section 1.1 that an element x is join compact if for every family y_i, $i \in I$, such that $x \leq \bigsqcup_{i \in I} y_i$ there exists a finite $J \subseteq I$ such that $x \leq \bigsqcup_{i \in J} y_i$. A logic is join compact in the lattice of extensions of $\mathcal{E}\,\mathbf{K}_\kappa$ only if it is finitely axiomatizable. For let $\Lambda = \mathbf{K}_\kappa \oplus X$, and $X = \{\varphi_i : i \in I\}$. Then $\Lambda \leq \bigsqcup_{i \in I} \mathbf{K}_\kappa \oplus \varphi_i$. Hence by join compactness there is a finite $J \subseteq I$ such that $\Lambda \leq \bigsqcup_{i \in J} \mathbf{K}_\kappa \oplus \varphi_i$. So $\Lambda = \mathbf{K}_\kappa \oplus \{\varphi_i : i \in J\}$. Hence Λ is finitely axiomatizable. Now let $\Lambda = \mathbf{K}_\kappa \oplus \{\varphi_i : i < n\}$. Assume $\Lambda \leq \bigsqcup_{i \in I} \Theta_i$. So, for each $i < n$, $\varphi_i \in \bigsqcup_{i \in I} \Theta_i$. A proof of φ_i is finite and hence uses only finitely many axioms. Consequently, there is a finite set $J(i) \subseteq I$ such that $\varphi_i \in \bigsqcup_{j \in J(i)} \Theta_i$. Put $J := \bigcup_{i<n} J(i)$. Then $\Lambda \leq \bigsqcup_{j \in J} \Theta_i$. So Λ is join compact.

Theorem 2.9.8. *A logic is join compact in the lattice $\mathcal{E}\,\boldsymbol{K}_\kappa$ iff it is finitely axiomatizable. Every element in $\mathcal{E}\,\boldsymbol{K}_\kappa$ is the join of compact elements; in other words, $\mathcal{E}\,\boldsymbol{K}_\kappa$ is algebraic.*

Let us close with an important concept from lattice theory, that of a *dimension*. In distributive lattices (in fact in modular lattices already) one can show that if x is an element such that there exists a finite chain $y := y_0 < y_1 < y_2 < \ldots < y_n = x$ of length n such that there is no u such that $y_i < u < y_{i+1}$, then any other such chain is finite as well and has length n. n is called the **dimension** of x over y and the **codimension** of y relative to x. If the lattice has a bottom element $\bot$, the **dimension** of x is defined to be the dimension of x over $\bot$. If the lattice has a top element, the **codimension** of x is the codimension of x relative to $\top$. The following theorem is of great theoretical importance and well worth remembering. It has been shown in David Makinson [**146**].

Theorem 2.9.9 (Makinson). *There exist exactly two logics of codimension 1 in the lattice $\mathcal{E}\boldsymbol{K}_1$, namely, the logic of the one–point reflexive frame and the logic of the one–point irreflexive frame.*

Proof. Let us first see that the logics $\Lambda_\bullet = \mathsf{Th}\boxed{\bullet}$ and $\Lambda_\circ = \mathsf{Th}\boxed{\circ}$ are of codimension 1. To see that, we show in turn that $\Lambda_\bullet = \mathbf{K}\oplus\Box\bot$ and $\Lambda_\circ = \mathbf{K}\oplus p \leftrightarrow \Box p$. Note first that these axioms are surely contained in the theory of the corresponding frames, so the axiomatization yields a logic which is possibly weaker, in each of the two cases. For the inclusions '$\subseteq$' observe the following. In $\mathbf{K}\oplus\Box\bot$ we have $\Box\varphi \leftrightarrow \top$, so any formula is equivalent to a nonmodal formula, and in $\mathbf{K} \oplus p \leftrightarrow \Box p$ we have $\Box\varphi \leftrightarrow \varphi$, so again any modal formula is equivalent to a nonmodal formula, by a simple induction. Thus any axiom extending either logic can be written into a form φ, φ nonmodal. But if φ does not hold in either logic, it does not hold in classical logic as well. But there is no strengthening of classical logic which is consistent, by Theorem 1.7.14. So, $\mathbf{K} \oplus \Box\bot$ is maximally consistent, and contained in $\Lambda_\bullet$, so the two must be equal. Likewise, $\mathbf{K} \oplus p \leftrightarrow \Box p = \Lambda_\circ$.

Let $\Theta \not\subseteq \Lambda_\bullet = \mathbf{K} \oplus \Box\bot$, the logic of the one–point irreflexive frame. Then $\Theta \sqcup \mathbf{K} \oplus \Box\bot$ is inconsistent. Thus $\Box\bot$ is inconsistent with Θ and so we have $\Diamond\top \in \Theta$, that is, $\Lambda \supseteq \mathbf{K.D}$. Then any consistent formula without variables is equivalent to either $\bot$ or $\top$. Thus the modal algebra $\mathbf{2}^\circ$ on two elements $0, 1$ such that $\blacksquare 0 = 0$ and $\blacksquare 1 = 1$ is a Θ–algebra. But $\mathbf{2}^\circ \vDash \Box p \leftrightarrow p$, and so $\mathsf{Th}\,\mathbf{2}^\circ = \Lambda_\circ$, so that $\Theta \subseteq \Lambda_\circ$, as required. ∎

Notes on this section. The locales of modal logics are in general very complex. Chapter 7 is devoted to the study of these locales. The results of Wim Blok and

Wolfgang Rautenberg have been groundbreaking in this area. Lately, Kracht [**131**] has investigated the automorphisms of some locales of modal logics. These group measure the homogeneity of the locales. Even though some results have been obtained (concerning, for example, the locale of extensions of **S4.3**), for the standard systems even the size of the groups is still unknown.

Exercise 70. Let $\mathbf{K}.\mathbf{trs}_m$ be the logic defined by the axiom

$$\mathbf{trs}_m := \Box^{\leq m} p \to \Box^{m+1} p \ .$$

Show that

$$\mathbf{K} = \textstyle\bigsqcap_{m\in\omega} \mathbf{K}.trs_m \ .$$

This provides an example of an infinite intersection of logics which cannot be given a canonical axiomatization in terms of the axioms of the individual logics.

Exercise 71. Show that if Λ_i, $i \in I$, all have the finite model property, then so does $\bigsqcap_{i\in I}\Lambda_i$. Similarly for completeness.

Exercise 72. Call a logic Λ *canonical* if $\mathfrak{Can}_\Lambda \vDash \Lambda$. (See Section 3.2.) Show that if Λ_i, $i \in I$, are canonical, so is $\bigsqcup_{i\in I}\Lambda_i$. Moreover, if Λ_1, Λ_2 are canonical, so is $\Lambda_1 \sqcap \Lambda_2$. So, the canonical logics form a sublocale of the locale of normal logics.

Exercise 73. Show that the recursively axiomatizable logics form a sublattice with $\sqcap$ and $\sqcup$. Show that they do not form a sublattice with $\bigsqcup$.

Exercise 74. Show that there is no order preserving injective map from the ordered set of the real numbers into $\mathcal{E}\,\mathbf{K}_1$.

Exercise 75. Show that the finitely axiomatizable logics are closed under finite unions. Moreover, if Θ is weakly transitive, the finitely axiomatizable logics in $\mathcal{E}\Theta$ are also closed under finite intersections.

Exercise 76. Show that a logic is Halldén–complete iff it is not the intersection of two quasi–normal logics properly containing it.

***Exercise 77.** Show that the infinite intersection of decidable logics need not be decidable.

CHAPTER 3

Fundamentals of Modal Logic II

3.1. Local and Global Consequence Relations

With a modal logic Λ typically only the relation $\vdash_\Lambda$ is considered as an associate consequence relation. However, in many applications it is useful to take a stronger one, which we will call the *global consequence relation*. It is denoted by $\Vdash_\Lambda$ and defined as follows.

Definition 3.1.1. *Let Λ be a modal logic. Then $\Delta \Vdash_\Lambda \varphi$ iff φ can be derived from Δ and Λ using the rules (mp.) and (mn.): $\langle\{p\}, \Box_j p\rangle$ ($j < \kappa$). We say that φ* ***follows globally from*** *Δ* ***in*** *Λ if $\Delta \Vdash_\Lambda \varphi$. $\Vdash_\Lambda$ is called the* ***global consequence relation of*** *Λ.*

In the light of the definitions of Section 1.4, $\langle \mathcal{P}_\kappa, \vdash_\Lambda\rangle$ and $\langle \mathcal{P}_\kappa, \Vdash_\Lambda\rangle$ are actually two different logics with identical sets of tautologies. However, since it is customary to identify modal logics with their set of tautologies, we will differentiate $\vdash_\Lambda$ and $\Vdash_\Lambda$ by using the qualifying adjectives **local** and **global**, respectively. In $\vdash_\Lambda$ the rule (mn.) is only admissible, whereas in $\Vdash_\Lambda$ it is derivable.

The geometric intuition behind the notions of *global* versus *local* consequence is as follows. Take a geometrical model $\mathfrak{M} := \langle \mathfrak{F}, \beta, x\rangle$ and a formula φ. We say that φ holds **locally** in $\mathfrak{M}$ if $\langle \mathfrak{F}, \beta, x\rangle \vDash \varphi$ and that φ holds **globally** if $\langle \mathfrak{F}, \beta\rangle \vDash \varphi$. Alternatively, we may distinguish between local and global models. A **local model** is a triple $\langle \mathfrak{F}, \beta, x\rangle$ with $\mathfrak{F}$ a frame, β a valuation into $\mathfrak{F}$ and x a world. A **global model** is a pair $\mathfrak{N} := \langle \mathfrak{F}, \beta\rangle$. A **local extension** of $\mathfrak{N}$ by x is the triple $\mathfrak{N}_x := \langle \mathfrak{F}, \beta, x\rangle$. We say that a local model $\mathfrak{M}$ is a **local Λ–model** for φ if $\mathfrak{M}$ is a model based on a frame for Λ and φ holds locally in it; and we say that a pair $\mathfrak{N}$ is a **global Λ–model** for φ if every local expansion of $\mathfrak{N}$ is a local Λ–model for φ. By the deduction theorem for $\vdash_\Lambda$ and Theorem 2.8.7 we have the following completeness result. $\Delta \vdash_\Lambda \varphi$ iff for every Λ–frame and every local model $\mathfrak{M}$ based on it, φ is true in $\mathfrak{M}$ if Δ is true in $\mathfrak{M}$. It will be shown below that an analogous completeness result holds with respect to $\Vdash_\Lambda$. Fundamental for the completeness is the following fact. (Recall that $\boxtimes^\omega \Delta$ was defined to be the closure of Δ under (mn.).)

Proposition 3.1.2 (Local–Global). *For any given logic Λ, $\Delta \Vdash_\Lambda \psi$ iff $\boxtimes^\omega \Delta \vdash_\Lambda \psi$ iff $\boxplus \Delta_0 \vdash_\Lambda \psi$ for some compound modality $\boxplus$ and a finite set $\Delta_0 \subseteq \Delta$.*

Proof. In Section 2.1 it was proved that any derivation of a formula ψ can be transformed into a derivation of ψ in which all applications of (mn.) are done before (mp.). Hence, if $\Delta \Vdash_\Lambda \varphi$ then φ is derivable from the closure of Δ under (mn.) by means of (mp.) only. This shows the first equivalence. The second holds by compactness and the fact that for any pair $\boxplus, \boxplus'$ of compound modalities there exists a compound modality $\boxplus''$ such that $\boxplus'' p \vdash_\Lambda \boxplus p; \boxplus' p$. ◻

Proposition 3.1.3. *Let Λ be a modal logic. Then $\Delta \Vdash_\Lambda \varphi$ iff for every global model $\mathfrak{N} = \langle \mathfrak{F}, \beta \rangle$ such that $\mathfrak{F} \vDash \Lambda$, we have $\mathfrak{N} \vDash \varphi$ if $\mathfrak{N} \vDash \Delta$.*

Proof. Assume $\Delta \Vdash_\Lambda \varphi$. Then $\boxtimes^\omega \Delta \vdash_\Lambda \varphi$. Now let $\mathfrak{N} = \langle \mathfrak{F}, \beta \rangle$ be a global Λ–model and assume $\mathfrak{N} \vDash \Delta$. Take a local expansion $\mathfrak{N}_x := \langle \mathfrak{F}, \beta, x \rangle$. Then $\mathfrak{N}_x \vDash \boxtimes^\omega \Delta$. Hence $\mathfrak{N}_x \vDash \varphi$. Since this does not depend on the choice of x, $\mathfrak{N} \vDash \varphi$. Now assume that $\Delta \nVdash_\Lambda \varphi$. Then $\boxtimes^\omega \Delta \nvdash_\Lambda \varphi$. Thus the set $\boxtimes^\omega \Delta \cup \{\neg \varphi\}$ is consistent and is therefore contained in a maximally consistent set W. Let $\mathfrak{F}$ be the subframe of $\mathfrak{Can}_\Lambda(var)$ generated by W, and let κ be the natural valuation, defined in Section 2.8. Then $\langle \mathfrak{F}, \kappa \rangle \vDash \Delta$, but $\langle \mathfrak{F}, \kappa \rangle \nvDash \varphi$, as required. ◻

The previous theorem established the correctness of the notion of *global consequence*. From now on the relation $\vdash_\Lambda$ will also be called the **local consequence relation** if that qualification is necessary. Many notions that we have defined so far now split into two counterparts, a local and a global one. However, some care is needed due to the fact that many definitions take advantage of the fact that $\vdash_\Lambda$ admits a deduction theorem whereas $\Vdash_\Lambda$ generally does not (see below). For example, by the definitions of Section 1.6 a logic is called **globally complete** if for every finite set Δ of formulae and each formula φ if $\Delta \nVdash_\Lambda \varphi$ then there exists a Kripke–frame $\mathfrak{f}$ for Λ and a valuation β such that $\langle \mathfrak{f}, \beta \rangle \vDash \Delta$ but $\langle \mathfrak{f}, \beta \rangle \nvDash \varphi$. (Instead of a finite set Δ we may just take a single formula δ, e. g. $\bigwedge \Delta$.) If the frame can always be chosen finite then we say that Λ has the **global finite model property**. Likewise, Λ is **globally decidable** if for finite Δ the problem '$\Delta \Vdash_\Lambda \varphi$' is decidable, that is, we have an algorithm which for any given finite set Δ and formula φ decides (correctly) whether or not $\Delta \Vdash_\Lambda \varphi$. Similarly, for a given complexity class $\mathcal{C}$ we say that Λ is **globally $\mathcal{C}$–computable** (**globally $\mathcal{C}$–hard**, **globally $\mathcal{C}$–complete**) if the problem '$\Delta \Vdash_\Lambda \varphi$?' is in $\mathcal{C}$ (is $\mathcal{C}$–hard, is $\mathcal{C}$–complete). Λ is **locally decidable** (**locally complete** etc.) if it is decidable (complete etc.) simpliciter.

We have seen earlier that the semantics for the local consequence relation leads to pairs $\langle \mathfrak{A}, F \rangle$ where $\mathfrak{A}$ is a modal algebra and F a filter. Since the set of designated elements must be closed under all rules, for $\Vdash_\Lambda$ we must now also require the set of designated elements to be closed under the algebraic counterpart of the rule (mn.).

Definition 3.1.4. *Let $\mathfrak{A}$ be a modal algebra, and $F \subseteq A$ a filter. F is called* **open** *if it satisfies (fi$\blacksquare$.): If $a \in F$ and $j < \kappa$ then also $\blacksquare_j a \in F$.*

Lemma 3.1.5. *Let $\mathfrak{A}$ be an algebra and F be an open filter. Define Θ_F by $a \, \Theta_F \, b$ iff $a \leftrightarrow b \in F$. Then Θ_F is a congruence.*

Proof. In Section 1.7 we have shown that Θ_F is a congruence with respect to the boolean reduct. Hence, we only need to verify that if $a \; \Theta_F \; b$ then also $\blacksquare_j a \; \Theta_F \; \blacksquare_j b$. So, suppose that $a \; \Theta_F \; b$. Then, by definition, $a \leftrightarrow b \in F$. Thus, $a \to b \in F$, from which $\blacksquare_j(a \to b) \in F$, since F is open. Hence, by (mp.) closure, $\blacksquare_j a \to \blacksquare_j b \in F$. Similarly, $\blacksquare_j b \to \blacksquare_j a \in F$ is shown, which together with $\blacksquare_j a \to \blacksquare_j b$ gives $\blacksquare_j a \leftrightarrow \blacksquare_j b \in F$. And that had to be demonstrated. ◻

Usually, we write $\mathfrak{A}/F$ instead of $\mathfrak{A}/\Theta_F$.

Theorem 3.1.6. *Let Λ be a modal logic. The global consequence relation of Λ has a unital semantics.*

Proof. Let $\mathcal{S}$ be the set of all pairs $\langle \mathfrak{A}, F\rangle$ where (i.) $\mathfrak{A}$ is a modal algebra, (ii.) F an open filter in $\mathfrak{A}$, (iii.) $\vdash_{\langle \mathfrak{A}, F\rangle} \supseteq \Vdash_\Lambda$ and (iv.) $\langle \mathfrak{A}, F\rangle$ is reduced. By the results of Section 1.5,

$$\Vdash_\Lambda = \bigcap_{\langle \mathfrak{A}, F\rangle \in \mathcal{S}} \vdash_{\langle \mathfrak{A}, F\rangle}$$

By Lemma 3.1.5 $\langle \mathfrak{A}, F\rangle$ is reduced only when $F = \{1\}$. Thus, $\Vdash_\Lambda$ has a unital semantics. ◻

As a useful consequence we note the following theorem.

Lemma 3.1.7. *Let Λ be a modal logic, φ_1, φ_2 and ψ be modal formulae. Then*

$$\varphi_1 \leftrightarrow \varphi_2 \Vdash_\Lambda \psi(\varphi_1) \leftrightarrow \psi(\varphi_2) \; .$$

The characterization of $\Vdash_\Lambda$ in terms of matrices fits into the geometrical picture as follows. If $\mathfrak{A}$ is a Λ–algebra and $\langle \mathfrak{A}, \{1\}\rangle$ a reduced matrix, then a valuation β into that matrix makes φ true just in case $\overline{\beta}(\varphi) = 1$. If $\mathfrak{A}$ is the algebra of internal sets of a frame $\mathfrak{F}$, then 1 is simply the full underlying set, namely f. So, the corresponding geometrical model is nothing but the global model $\langle \mathfrak{F}, \beta\rangle$.

Now we turn to the interconnection between local and global properties of a logic.

Proposition 3.1.8. *If Λ is globally decidable (has the global finite model property, is globally complete) then Λ is locally decidable (has the local finite model property, is locally complete).*

We will now prove that $\mathbf{K}_\kappa$ has the global finite model property. The proof is an interesting reduction to the local property. Notice that $\mathbf{K}_\kappa$ has the local finite model property, by Theorem 2.7.9.

Lemma 3.1.9. *Let $k := \sharp(sf(\varphi) \cup sf(\psi))$. Then we have*

$$\varphi \Vdash_{\boldsymbol{K}_\kappa} \psi \qquad \textit{iff} \qquad \boxtimes^{\leq 2^k} \varphi \vdash_{\boldsymbol{K}_\kappa} \psi \; .$$

Proof. Surely, if $\boxtimes^{\leq 2^k}\varphi \vdash_{\mathbf{K}_\kappa} \psi$, then also $\varphi \Vdash_{\mathbf{K}_\kappa} \psi$. So, assume $\boxtimes^{\leq 2^k}\varphi \nvdash_{\mathbf{K}_\kappa} \psi$. Then there exists a finite model $\langle \mathfrak{f}, \beta, w_0\rangle \vDash \boxtimes^{\leq 2^k}\varphi; \neg\psi$ rooted at w_0. Moreover, the construction for Theorem 2.7.9 shows that we may assume that $\mathfrak{f}$ is cycle–free, and that between any pair of points there exists at most one path. Let $\Delta := sf(\varphi) \cup sf(\psi)$ and put $S(y) = \{\chi \in \Delta : \langle \mathfrak{f}, \beta, y\rangle \vDash \chi\}$. Let g be the set of all y in f such that along any path from w_0 to y there are no two distinct points v and w such that $S(v) = S(w)$. Then any path from w_0 to $y \in g$ has length $\leq 2^k$, because there are at most 2^k subsets of Δ. Now define $\blacktriangleleft_j$ on g as follows. $y \blacktriangleleft_j z$ iff (1.) $y \lhd_j z$ or (2.) for some $u \notin g$ we have $y \lhd_j u$ and $S(z) = S(u)$. Put $\gamma(p) := \beta(p) \cap g$. We will now show that for every $y \in g$ and $\chi \in \Delta$

$$\langle \mathfrak{g}, \gamma, y\rangle \vDash \chi \quad \Leftrightarrow \quad \langle \mathfrak{f}, \beta, y\rangle \vDash \chi\ .$$

This is true for variables by construction. The steps for negation and conjunction are clear. Now let $\chi = \Diamond_j \delta$. If $\langle \mathfrak{f}, \beta, y\rangle \vDash \Diamond_j \delta$ then for some z such that $y \lhd_j z$ we have $\langle \mathfrak{f}, \beta, z\rangle \vDash \delta$. There are two cases. **Case 1.** $z \in g$. Then by induction hypothesis, $\langle \mathfrak{g}, \gamma, z\rangle \vDash \delta$. From this we conclude $\langle \mathfrak{g}, \gamma, y\rangle \vDash \Diamond_j \delta$, since $y \blacktriangleleft_j z$. **Case 2.** $z \notin g$. Then there is a $u \in g$ such that $S(u) = S(z)$. Therefore, by construction of $\mathfrak{g}$, $y \blacktriangleleft_j u$. Furthermore, $\langle \mathfrak{f}, \beta, u\rangle \vDash \delta$ by definition of $S(-)$. So, $\langle \mathfrak{g}, \gamma, u\rangle \vDash \delta$ by induction hypothesis. From this follows $\langle \mathfrak{g}, \gamma, y\rangle \vDash \Diamond_j \delta$, since $y \blacktriangleleft_j u$. This exhausts the two cases. Now suppose $\langle \mathfrak{g}, \gamma, y\rangle \vDash \Diamond_j \delta$. Then $\langle \mathfrak{g}, \gamma, z\rangle \vDash \delta$ for some z such that $y \blacktriangleleft_j z$. By induction hypothesis, $\langle \mathfrak{f}, \beta, z\rangle \vDash \delta$. If $y \lhd_j z$, then also $\langle \mathfrak{f}, \beta, y\rangle \vDash \Diamond_j \delta$. If, however, $y \not\lhd_j z$, then there is a u such that $y \lhd_j u$ and $S(u) = S(z)$. By definition of $S(-)$, $\langle \mathfrak{f}, \beta, u\rangle \vDash \delta$, from which $\langle \mathfrak{f}, \beta, y\rangle \vDash \Diamond_j \delta$ as well. Now since from w_0 there is always a path of length $\leq 2^k$ to any point $y \in g$, we have $\langle \mathfrak{f}, \beta, y\rangle \vDash \varphi$ for all $y \in g$, and so $\langle \mathfrak{g}, \gamma, y\rangle \vDash \varphi$ for all y. Consequently, $\langle \mathfrak{g}, \gamma, w_0\rangle \vDash \boxtimes^{\omega}\varphi; \neg\psi$, as required. ∎

Theorem 3.1.10. *$\boldsymbol{K}_\kappa$ has the global finite model property.*

Notice that even if $\mathfrak{f}$ was originally cycle–free, we might have inserted cycles into $\mathfrak{g}$. This is in some cases unavoidable. For example, there is no finite cycle–free model against $\Diamond\top \Vdash_{\mathbf{K}} p$, but there are infinite cycle–free models as well as finite models with cycles. The bound for k in the proof can be improved somewhat (see exercises). This theorem has a great number of strong applications as we will see in the next section.

Valentin Goranko and Solomon Passy have shown in [**87**] that the global properties of a logic Λ correspond to the local properties of a logic $\Lambda^{\blacksquare}$ which arises from Λ by adding a so–called *universal modality*. (See Section 2.5.) Recall that $\Lambda^{\blacksquare}$ is defined by

$$\Lambda^{\blacksquare} := \Lambda \otimes \mathbf{S5}(\{\Diamond_j p \to \Diamond_\kappa p : j < \kappa\})$$

Abbreviate $\Box_\kappa$ by $\blacksquare$. It is not hard to show that the canonical frames for $\Lambda^{\blacksquare}$ satisfy two properties. (1.) The relation $\blacktriangleleft$ $(:= \lhd_\kappa)$ is an equivalence relation on the set of points, (2.) For all $j < \kappa$, $\lhd_j \subseteq \blacktriangleleft$. (It also follows from the results of Section 3.2.) By completeness with respect to canonical frames, $\Lambda^{\blacksquare}$ is complete with respect to frames satisfying (1.) and (2.). Moreover, a rooted generated subframe of a frame $\mathfrak{F}$ satisfying (1.) and (2.) actually satisfies (1′.) $\blacktriangleleft = f \times f$. Thus, $\Lambda^{\blacksquare}$ is complete with respect to frames satisfying (1′.) and (2.). It is easy to construct such frames when given a frame $\mathfrak{F}$ for Λ. Namely, put $◀ := f \times f (=: \lhd_\kappa)$, $\mathfrak{f}^{■} := \langle f, \langle \lhd_j : j < \kappa + 1\rangle\rangle$ and $\mathfrak{F}^{■} := \langle \mathfrak{f}^{■}, \mathbb{F}\rangle$. Since $■_\kappa a = \varnothing$ iff $a \neq f$, and $■_\kappa f = f$, $\mathbb{F}$ is closed under $■$. Consequently, $\mathfrak{F}^{■}$ is well–defined. Moreover, if $\mathfrak{F}$ is a frame for Λ, $\mathfrak{F}^{■}$ is a frame for $\Lambda^{■}$ — and conversely.

Recall the definition of $\mathcal{P}_\kappa$ from Section 2.1. Continuing our present notation we write $\mathcal{P}_\kappa(■)$ for the language obtained from $\mathcal{P}_\kappa$ by adding $\Box_\kappa$. Let us agree to call a formula **plain** if it does not contain any universal modality. So, $\varphi \in \mathcal{P}_\kappa(■)$ is plain iff $\varphi \in \mathcal{P}_\kappa$. A **degree 1 combination** of plain formulae is a formula $\varphi \in \mathcal{P}_\kappa(■)$ such that no $■$ occurs in the scope of another occurrence of $■$.

Proposition 3.1.11. *Let Λ be a modal logic, Δ a set of plain formulae, and φ a plain formula. Then $\Delta \Vdash_\Lambda \varphi$ iff $■\Delta \vdash_{\Lambda^{■}} \varphi$. In particular,*

$$\vdash_\Lambda \varphi \qquad \textit{iff} \qquad \vdash_{\Lambda^{■}} \varphi\ .$$

Proof. Suppose that $\Delta \Vdash_\Lambda \varphi$. Then $\boxtimes^\omega\Delta \vdash_\Lambda \varphi$ and so $\boxtimes^\omega\Delta \vdash_{\Lambda^{■}} \varphi$. Now $■\Delta \vdash_{\Lambda^{■}} \boxtimes^\omega\Delta$, as can be shown easily. Hence $■\Delta \vdash_{\Lambda^{■}} \varphi$. Now assume that $\Delta \nVdash_\Lambda \varphi$. Then there exists a global model $\mathfrak{N} := \langle\mathfrak{F},\beta\rangle$ such that $\mathfrak{N} \vDash \Delta$ and $\mathfrak{N} \nvDash \varphi$. Thus for some local extension $\mathfrak{N}_x$, $\mathfrak{N}_x \nvDash \varphi$. Then $\mathfrak{M} := \langle\mathfrak{F}^{■},\beta, x\rangle$ is a local $\Lambda^{■}$–model such that $\mathfrak{M} \vDash ■\Delta; \neg\varphi$, as required. ◻

A sharper theorem can be established. Before we can prove it, however, we need the following auxiliary theorem concerning simplifications of formulae.

Lemma 3.1.12. *Let Λ be a polymodal logic and $\Lambda^{■}$ be the extension by a universal modality. Then any formula in $\mathcal{P}_\kappa(■)$ is deductively equivalent to a degree 1 combination of plain formulae.*

Proof. Let φ be given. We can assume φ to be in normal form. The following equivalences are theorems of $\mathbf{K}_\kappa^{■}$ for $\boxplus = \Box_j$, $j < \kappa$, or $\boxplus = ■$.

$$\begin{array}{lcl}
\boxplus■p & \leftrightarrow & \boxplus\bot \vee ■p \\
\boxplus♦p & \leftrightarrow & \boxplus\bot \vee ♦p \\
\boxplus(q \vee ■p) & \leftrightarrow & \boxplus q \vee ■p \\
\boxplus(q \vee ♦p) & \leftrightarrow & \boxplus q \vee ♦p \\
\boxplus(q \wedge ■p) & \leftrightarrow & \boxplus q \wedge \boxplus■p \\
\boxplus(q \wedge ♦p) & \leftrightarrow & \boxplus\bot \vee (\boxplus q \wedge ♦p)
\end{array}$$

Analogous equivalences can be derived for the dual operator from these upper six equivalences. The lemma now follows by induction on the degree of φ. ◻

Theorem 3.1.13 (Goranko & Passy). *For the following properties $\mathfrak{P}$, Λ has $\mathfrak{P}$ globally iff $\Lambda^{■}$ has $\mathfrak{P}$ locally: **decidability, finite model property, completeness.***

Proof. One direction follows from Proposition 3.1.11; namely, if $\Lambda^{■}$ has $\mathfrak{P}$ locally, Λ has $\mathfrak{P}$ globally. So we have to prove the converse direction. The idea to the proof is to reduce a statement of the form '$\vdash_{\Lambda^{■}} \psi$' to a boolean combinations of problems of the form '$\chi \Vdash_\Lambda \varphi$'. (Moreover, this reduction will be effective, so it is enough for the proof of decidability of '$\vdash_{\Lambda^\blacksquare} \psi$' if we show the problems '$\chi \Vdash_\Lambda \varphi$' to be decidable.) Now start with '$\vdash_{\Lambda^\blacksquare} \psi$'. Transform ψ into conjunctive normal form. This does not affect theoremhood of ψ. So, without loss of generality ψ can be asumed to be already in conjunctive normal form. Moreover, we have seen in Lemma 3.1.12 that ψ is deductively equivalent to a degree 1 combination of plain formulae. So, we may as well assume that it is already of this form. If ψ is of the form $\psi_1 \wedge \psi_2$ the problem '$\vdash_{\Lambda^\blacksquare} \psi$' is equivalent to the conjunction of '$\vdash_{\Lambda^\blacksquare} \psi_1$' and '$\vdash_{\Lambda^\blacksquare} \psi_2$'. Hence, assume now that ψ is not of that form. Then ψ is of the form $\bigvee_{i<n} \blacksquare\rho_i \vee \blacklozenge\sigma \vee \tau$, where all ρ_i, σ and τ are plain. We claim that $\vdash_{\Lambda^\blacksquare} \bigvee_{i<n} \blacksquare\rho_i \vee \blacklozenge\sigma \vee \tau$ iff either for some $i < n$ $\neg\sigma \Vdash_\Lambda \rho_i$, or $\neg\sigma \Vdash_\Lambda \tau$. To see this, assume that the left hand side fails. Then there exists a local $\Lambda^\blacksquare$–model $\langle \mathfrak{F}, \beta, x\rangle \vDash \bigwedge_{i<n} \neg\blacksquare\rho_i; \blacksquare\neg\sigma; \neg\tau$. We may assume that $\mathfrak{F} = \mathfrak{G}^\blacksquare$ for some Λ–frame $\mathfrak{G}$. Then $\mathfrak{M} := \langle \mathfrak{G}, \beta\rangle$ is a global Λ–model and $\mathfrak{M} \vDash \neg\sigma$ as well as $\mathfrak{M} \nvDash \rho_i$ for all $i < n$ and $\mathfrak{M} \nvDash \tau$, as required. Now assume that the right hand side fails. Then there exist global Λ–models $\mathfrak{M}_i = \langle \mathfrak{F}_i, \beta_i\rangle$ such that $\mathfrak{M}_i \vDash \neg\sigma$, $\mathfrak{M}_i \nvDash \rho_i$ for all $i < n$ and a global Λ–model $\mathfrak{N} = \langle \mathfrak{G}, \gamma\rangle$ such that $\mathfrak{N} \vDash \sigma$ and $\mathfrak{N} \nvDash \tau$. In particular, $\mathfrak{N}_x \vDash \neg\tau$ for some local extension $\mathfrak{N}_x$ of $\mathfrak{N}$. Put $\mathfrak{H} := \bigoplus_{i<n} \mathfrak{F}_i \oplus \mathfrak{G}$. Let $\delta := \bigoplus_{i<n} \beta_i \oplus \gamma$. Then $\langle \mathfrak{H}^\blacksquare, \delta, x\rangle \vDash \bigwedge_{i<n} \neg\blacksquare\rho_i; \blacksquare\neg\sigma; \neg\tau$. This concludes the proof in the case of decidability. For completeness and finite model property, notice that in the previous construction if $\mathfrak{F}_i$ and $\mathfrak{G}$ are (finite) Kripke–frames, so is $\mathfrak{H}^\blacksquare$. ◻

Notes on this section. The universal modality has enjoyed great popularity in modal logic. It was observed in [**86**] that the universal modality has in conjunction with *nominals* the same expressive power as the difference operator, explored by Maarten de Rijke in [**177**]. It was proved subsequently that $\Lambda^\blacksquare$ shares few properties with Λ (apart from those which they must share by virtue of the results of this section). In passing from Λ to $\Lambda^\blacksquare$ finite model property can get lost (Frank Wolter [**235**]), decidability (Edith Spaan [**202**]) and even completeness, see Section 9.6. The number k in Lemma 3.1.9 cannot be significantly reduced. In Marcus Kracht [**130**] it is shown that asymptotically k must be at least as large as $2^{c\sqrt{n}}$, where n is the size of φ. Moreover, Edith Spaan [**202**] has shown that $\mathbf{K.alt}_1$ is globally PSPACE–complete and that $\mathbf{K}_\kappa$ is globally EXPTIME–complete. So, adding the universal modality can raise the complexity to any higher degree.

Exercise 78. Prove Proposition 3.1.8.

Exercise 79. Show the following improved bounds for global to local reduction. Define $\Delta := sf(\varphi) \cup var(\psi)$. Further, let $\mu := max\{dp(\psi), 2^{\sharp\Delta}\}$. Show that

$$\varphi \Vdash_{\mathbf{K}_\kappa} \psi \qquad \Leftrightarrow \qquad \boxtimes^{\leq\mu}\varphi \vdash_{\mathbf{K}_\kappa} \psi$$

Exercise 80. Let Λ be a logic. Let $r(\varphi, \psi)$ be such that $\varphi \Vdash_\Lambda \psi$ iff $\boxtimes^{\leq r(\varphi,\psi)}\varphi \vdash_\Lambda \psi$. Now let Λ be a logic which is locally decidable, but not globally decidable. (Such logics exist, see Section 9.4.) Show that $r(\varphi, \psi)$ is not computable.

Exercise 81. (R. E. Ladner [**137**].) Let $\Lambda \subseteq$ **S5** be consistent. Show that satisfiability of a formula is NP–complete. *Hint.* Clearly, the problem is NP–hard. To show that it is in NP, show that any formula can be reduced to a formula of depth 1.

Exercise 82. Show that a modal logic is locally tabular iff it is globally tabular.

3.2. Completeness, Correspondence and Persistence

Clearly, Kripke–models are easier to handle than canonical frames. Mostly, it is easier to reason in a Kripke–structure than to reason syntactically by shuffling formulae. Moreover, canonical models are very difficult structures. All the knowledge of a logic is coded in the canonical frame, so there is little hope that we can use the canonical frame in any effective way. However, the abstract existence of such a frame alone can provide us with many important results. Consider the basic logic $\mathbf{K}_\kappa$. We know that Kripke frames satisfy all postulates of $\mathbf{K}_\kappa$. Now, if φ is not a theorem of $\mathbf{K}_\kappa$, then we can base a countermodel for φ on the canonical frame, that is we have a world X such that

$$\langle \mathfrak{Can}_{\mathbf{K}_\kappa}(var), v, X\rangle \vDash \neg\varphi$$

where $v(p) := \{X : p \in X\}$. However, as the Kripke structure underlying that frame satisfies $\mathbf{K}_\kappa$, we can actually forget the internal sets. Then, using the same valuation we have

$$\langle \mathfrak{can}_{\mathbf{K}_\kappa}(var), v, X\rangle \vDash \neg\varphi$$

We have established now that $\mathbf{K}_\kappa$ is complete by using the canonical frame and 'forgetting' the internal sets. If this is possible, a logic is said to be *canonical* or *c–persistent*.

Definition 3.2.1. *A logic Λ is called α–**canonical** if for every $\beta < \alpha$, $\Lambda \subseteq \mathsf{Th}\,\mathfrak{can}_\Lambda(\beta)$. A logic Λ is **canonical** or **c–persistent** if it is α–canonical for every α, and it is called **weakly canonical** if it is $\aleph_0$–canonical.*

The following is an easy consequence of the definition.

Proposition 3.2.2. *Let α, β be cardinal numbers and $\alpha \leq \beta$. Let Λ be β–canonical. Then Λ is also α–canonical.*

Proof. By assumption, there exists an embedding $f : \alpha \rightarrowtail \beta$. By Theorem 2.8.11 this induces a contraction $X_f : \mathfrak{Can}_\Lambda(\beta) \twoheadrightarrow \mathfrak{Can}_\Lambda(\alpha)$. By assumption, $\mathfrak{can}_\Lambda(\beta) \vDash \Lambda$. Then also $\mathfrak{can}_\Lambda(\alpha) \vDash \Lambda$. ∎

Definition 3.2.3. *A logic Λ is called **complete** if it is the logic of its Kripke–structures. Λ is called α–**compact** if every consistent set based on $< \alpha$ many variables has a model based on a Kripke–frame for Λ. Λ is called **strongly compact** or simply **compact** if Λ is α–compact for every α and Λ is called **weakly compact** if it is $\aleph_0$–compact.*

Obviously, if a logic is compact, it is also weakly compact, and if it is weakly compact it is complete. Neither of the converses hold; there are complete logics which are not weakly compact and there are logics which are weakly compact but not strongly compact. (This has been shown first in [**66**], who also defined the notions of weak and strong compactness.) The reader may verify that logics axiomatized by constant axioms are 1–compact. Compactness is rather closely connected with a different property of logics, called *complexity* (see [**82**]).

Definition 3.2.4. *A logic Λ is α–**complex** if every β–generable algebra, where $\beta < \alpha$, is isomorphic to a subalgebra of the algebra of all subsets of a Kripke–frame for Λ.*

α–complexity is not directly equivalent with α–compactness if α is finite. Rather, the right notion to choose here is *global α–compactness*. A logic Λ is **globally α–compact** if for every pair consisting of a set Φ and a formula φ, based together on β–many variables, $\beta < \alpha$, if $\Phi \nVdash_\Lambda \varphi$ then there exists a Kripke–frame $\mathfrak{f} \vDash \Lambda$, and a valuation β such that $\langle \mathfrak{f}, \beta \rangle \vDash \Phi$ but $\langle \mathfrak{f}, \beta \rangle \nvDash \varphi$. If a logic is locally $\alpha + 1$–compact then it is also globally α–compact.

Theorem 3.2.5 (Wolter). *Let Λ be a κ–modal logic. Λ is globally α–compact iff it is α–complex. Moreover, if Λ is globally $\alpha + 1$–compact, it is locally α–compact.*

Proof. Suppose Λ is α–complex and take a set Φ and a formula φ such that $\Phi \nVdash_\Lambda \varphi$. Assume that Φ and φ are based on β many variables, $\beta < \alpha$. Then there exists a model $\langle \mathfrak{A}, \gamma \rangle \vDash \Phi$ such that $\langle \mathfrak{A}, \gamma \rangle \nvDash \varphi$. By assumption, there exists a Kripke–frame $\mathfrak{f}$ such that $\mathfrak{A}$ is a subalgebra of the algebra $\mathfrak{B}$ of subsets of $\mathfrak{f}$. Let $\iota : \mathfrak{A} \rightarrowtail \mathfrak{B}$ be an embedding. Then $\delta := \iota \circ \gamma$ is a valuation on $\mathfrak{f}$. Now, for every $\psi \in \Phi$ we have $\overline{\gamma}(\psi) = 1$, and so $\overline{\delta}(\psi) = 1$ as well. Thus $\langle \mathfrak{f}, \delta \rangle \vDash \Phi$. However, since $\overline{\gamma}(\varphi) \neq 1$, also $\langle \mathfrak{f}, \delta \rangle \nvDash \varphi$. For the converse assume that Λ is globally α–compact. Let $\mathfrak{A}$ be a β–generable algebra, where $\beta < \alpha$. Then choose a generating set X such that $\sharp X = \beta$. For each $a \in X$ take a variable p_a and let γ be the valuation defined by $\gamma(p_a) := a$. $\overline{\gamma}$ is surjective by choice of X. Let $U(\mathfrak{A})$ be the collection of ultrafilters of $\mathfrak{A}$. For every $U \in U(\mathfrak{A})$, $\overline{\gamma}^{-1}[U] := \{\varphi : \overline{\gamma}(\varphi) \in U\}$ is a maximally Λ–consistent set in the variables p_i, $i < \beta$. Now Λ is α–compact and therefore for any $\overline{\gamma}^{-1}[U]$ there is a Λ–Kripke–model $\langle \mathfrak{g}_U, \delta_U, x_U \rangle \vDash \overline{\gamma}^{-1}[U]$. We assume that the $\mathfrak{g}_U$ are disjoint and take $\mathfrak{g} := \bigoplus \mathfrak{g}_U$ and $\delta(p) := \bigoplus_U \delta_U(p)$. It then follows that $\mathfrak{g} \vDash \Lambda$ and that the algebra $\mathfrak{B}$ generated by the $\gamma(p_a)$, $a \in A$, is a subalgebra of $\mathfrak{A}$. We show that in fact $\mathfrak{B} \cong \mathfrak{A}$. Namely, this holds if

$$\{\varphi : \overline{\gamma}(\varphi) = g\} = \{\varphi : \overline{\delta}(\varphi) = 1\} .$$

In fact,

$$\begin{array}{rcl}\overline{\gamma}(\varphi) = 1 & \Leftrightarrow & (\forall\boxplus)\overline{\gamma}(\boxplus\varphi) = 1 \\ & \Leftrightarrow & (\forall\boxplus)(\forall U \in U(\mathfrak{A}))\overline{\gamma}(\boxplus\varphi) \in U \\ & \Leftrightarrow & (\forall\boxplus)(\forall U \in U(\mathfrak{A}))\boxplus\varphi \in \overline{\gamma}^{-1}[U] \\ & \Leftrightarrow & (\forall\boxplus)(\forall U \in U(\mathfrak{A}))\langle\mathfrak{g}_U, \delta_U, x_U\rangle \vDash \boxplus\varphi \\ & \Leftrightarrow & \langle\mathfrak{g}, \delta\rangle \vDash \varphi \\ & \Leftrightarrow & \overline{\delta}(\varphi) = g\end{array}$$

Now for the second claim. Assume that Φ is a locally consistent set, based on β many variables, $\beta < \alpha$. Then let $\Psi := \{p_\alpha \to \varphi : \varphi \in \Phi\}$. $\boxtimes^\omega\Psi; p_\alpha$ is based on $< \alpha + 1$ variables and is also consistent. For take a finite subset S; S is without loss of generality of the form $\boxplus(p_\alpha \to \Phi_0); p_\alpha$, $\boxplus$ a compound modality and Φ_0 a finite subset of Φ. By assumption, Φ_0 has a Kripke–model $\langle\mathfrak{f}, \beta, x\rangle$, since Φ_0 is locally consistent. (If a logic is globally $\alpha + 1$–compact, it is also complete for formulas with $\leq \alpha$ many variables.) Now put $\beta^+(p_\alpha) := \{x\}$. Then

$$\langle\mathfrak{f}, \beta^+, x\rangle \vDash \boxplus(p_\alpha \to \Phi_0); p_\alpha\ .$$

Hence S is consistent. Since S was arbitrarily chosen, $\boxplus^\omega\Psi; p_\alpha$ is consistent. Consequently, $\boxtimes^\omega\Psi \nvdash_\Lambda \neg p_\alpha$, from which $\Psi \nVdash_\Lambda \neg p_\alpha$. By $\alpha + 1$–compactness, there exists a Kripke–model $\langle\mathfrak{g}, \gamma\rangle \vDash \Psi$ such that $\langle\mathfrak{g}, \gamma, y\rangle \vDash p_\alpha$ for some y. Then $\langle\mathfrak{g}, \gamma, y\rangle \vDash \Phi$, by construction of Ψ. ∎

Corollary 3.2.6 (Wolter). *Let α be infinite, Λ a normal modal logic. Then Λ is α–complex iff it is globally α–compact iff it is locally α–compact.*

If $\alpha \leq \beta$ and Λ is α–canonical then Λ is not necessarily β–canonical. It is an open problem, for example, whether $\aleph_1$–canonicity implies $\aleph_2$–canonicity. It is possible to show that we cannot give any finite bound n_0 such that a logic is canonical if it is n_0–compact. A simple but instructive example is the case of logics extending **K.D**. The 1–canonical frame consists of a single reflexive point (since there are only trivial constant propositions) and so every extension is 1–canonical. But they are not all $\aleph_1$–canonical. For example, take the logic **Grz.3**. By a theorem of Kit Fine in [**63**] **Grz.3** is weakly compact. (An exercise with hints how to prove this theorem is provided in Section 6.5.) We show that it is not $\aleph_1$–compact, from which follows (with the next theorem) that it is not $\aleph_1$–canonical.

$$Y := \{p_0\} \cup \{\Box(p_i \to \Diamond p_{i+1}) : i \in \omega\} \cup \{\Box(p_j \to \Box\neg p_i) : i < j < \omega\}$$

Each finite subset is satisfiable, taking a suitably large chain $\langle n, \leq\rangle$. For take a finite subset X. Without loss of generality we assume that X is the subset containing the formulas in which all and only the variables up to number n_0 occur. If $i < n_0$, then if p_i is true at a point x, it must have a successor y at which p_{i+1} holds. Moreover, $y \ntriangleleft x$. So let us take $\mathfrak{f}(n_0) := \langle\{0, 1, \ldots, n_0\}, \leq\rangle$. This is a **Grz.3**–frame. (For let $\langle\mathfrak{f}(n_0), \beta, k\rangle \vDash \Diamond p$. Let ℓ be the largest number such that $\ell \in \beta(p)$. Then $\langle\mathfrak{f}(n_0), \beta, \ell\rangle \vDash p; \Box(\neg p \to \Box\neg p)$, since for any successor $j \geq \ell$ if $j \vDash \neg p$ then $j > \ell$ and $j \vDash \Box\neg p$, by choice of ℓ. Hence $\langle\mathfrak{f}(n_0), \beta, k\rangle \vDash \Diamond(p \wedge \Box(\neg p \to \Box\neg p))$, and that had to be shown.) If we put $\beta(p_i) := \{i\}$, then $\langle \mathfrak{f}(n_0), \beta, 0\rangle \vDash X$. Now let $\mathfrak{f}$ be a **Grz.3**–frame and $\langle \mathfrak{f}, \beta, x\rangle \vDash Y$. Then pick $x_i \in \beta(p_i)$. Then for all i, $x_i \lhd x_{i+1}$ (by linearity of the frame) and $x_j \ntriangleleft x_i$ for all $j > i$. Thus we have a strictly ascending chain of points. Put $\gamma(p) := \{x_{2k} : k \in \omega\}$. We have

$$\langle \mathfrak{f}, \gamma, x_0\rangle \vDash \Diamond p; \neg\Diamond(p \wedge \Box(\neg p \to \Diamond\neg p)) .$$

Thus, $\mathfrak{f}$ is not a frame for **Grz**. Contradiction.

In the exercise we have put a proof that **G** is not 2–compact. This has been shown by Warren Goldfarb (see [**32**]) and also in [**66**]. However, **G** is 1–compact.

Proposition 3.2.7. *Let Λ be a logic. If Λ is α–canonical, it is α–compact. In particular, if Λ is (weakly) canonical it is (weakly) compact.*

The proof is easy, once we know that logics are generally complete with respect to their canonical frame. The converse of this statement is false. Frank Wolter [**240**] shows that the tense logic of the reals is compact but not canonical. So, which logics are canonical? This question has been answered for all standard systems. Generally, failure of canonicity is hard to demonstrate, whereas the contrary is normally straightforward. **K**.**D**, **K4**, **K5**, **K**.**alt**$_\mathbf{1}$, **S4**.**3**, **S5** are all canonical, **G** and **Grz** are not.

Example. We show that **S4** is canonical. The proof for canonicity is in two stages. First we show so called *correspondence* of the axioms with properties of Kripke frames. Let $\mathfrak{f} = \langle f, \lhd\rangle$ be a Kripke frame. Then the following holds.

(cot.)	$\mathfrak{f} \vDash p \to \Diamond p$	iff	$\mathfrak{f}$ is reflexive
(co4.)	$\mathfrak{f} \vDash \Diamond\Diamond p \to \Diamond p$	iff	$\mathfrak{f}$ is transitive

First (cot.). Suppose that $\mathfrak{f}$ is reflexive, β is a valuation and x a world. Then $\langle \mathfrak{f}, \beta, x\rangle \vDash p$ implies $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond p$. Suppose now that $\mathfrak{f}$ is not reflexive, say $x \ntriangleleft x$. Then let $\beta(p) := \{x\}$. Then we have $\langle \mathfrak{f}, \beta, x\rangle \vDash p; \neg\Diamond p$. Now (co4.). Suppose $\mathfrak{f}$ is transitive, then it satisfies **4**. For let $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond\Diamond p$. Then there are y and z such that $x \lhd y \lhd z$ $\langle \mathfrak{f}, \beta, z\rangle \vDash p$. By transitivity, $x \lhd z$ and so $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond p$. Now assume $\mathfrak{f}$ is not transitive. Then there are points $x \lhd y \lhd z$ such that $x \ntriangleleft z$. Choose $\beta(p) := \{z\}$. Then $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond\Diamond p; \neg\Diamond p$.

So, if we can show that the Kripke structure underlying the canonical frame is reflexive and transitive, we have proved that **S4** is canonical. Assume then that the canonical stucture is not reflexive. Then for some set X, $X \ntriangleleft X$; by definition, there must be a formula φ such that $\varphi \in X$ but $\Diamond\varphi \notin X$. Hence $\varphi \wedge \neg\Diamond\varphi \in X$, and so the canonical frame does not satisfy the axiom $p \to \Diamond p$. Contradiction. Hence the structure is reflexive. Now assume it is not transitive, that is, there are sets $X \lhd Y \lhd Z$ such that $X \ntriangleleft Z$. Then there is a formula φ such that $\varphi \in Z$ but $\Diamond\varphi \notin X$. However, again by definition, $\Diamond\varphi \in Y$ and so $\Diamond\Diamond\varphi \in X$, showing that $\neg\Diamond\varphi \wedge \Diamond\Diamond\varphi \in X$, in contradiction to the assumption that our frame satisfies the axiom (**4**.).

Lemma 3.2.8. *Let s be a finite set of finite sequences over κ. Let X and Y be worlds such that for all φ we have $\Box^s\varphi \in X$ only if $\varphi \in Y$. Then $X \lhd^s Y$ in the canonical frame for $\mathbf{K}_\kappa$.*

Proof. First we show the claim for sequences. This in turn is shown by induction on the length of the sequence. If σ has length 0 the claim is immediate. So let σ be a sequence and $\sigma = j^\frown\tau$ for some sequence τ and $j < \kappa$. Let $A_0 := \{\psi : \Box_j\psi \in X\}$, $A_1 := \{\Diamond^\tau\psi : \psi \in Y\}$ and $A := A_0 \cup A_1$. Then $\Box_j A_0 \subseteq X$ and $\Diamond_j A_1 \subseteq X$. We claim that A is consistent. To that end, let $\Delta_0 \subseteq A_0$ and $\Delta_1 \subseteq A_1$ be finite sets. There exists a $\delta \in A_1$ such that $\delta \vdash_\Lambda \delta'$ for all $\delta' \in \Delta_1$, as is easily shown. Now assume that $\Delta_0; \Delta_1$ is Λ–inconsistent. Then $\Delta_0; \delta \vdash_\Lambda \bot$. From this it follows that $\Box_j\Delta_0 \vdash_\Lambda \Box_j\neg\delta$ and so $\Box_j\Delta_0; \Diamond_j\delta$ is inconsistent in Λ. However, $\Box_j\Delta; \Diamond_j\delta \subseteq X$ and X is Λ–consistent. Contradiction. Therefore, A is consistent and there exists a world Z containing A. Then $X \lhd_j Z$ by definition of the canonical frame, and $Z \lhd^\tau Y$ by induction hypothesis. Hence $X \lhd^\sigma Y$. Now let $s := \{\sigma_i : i < n\}$ be a set of finite sequences over κ and assume that $X \not\lhd^s Y$. Then for every $i < n$ there exists a φ_i such that $\Box^{\sigma_i}\varphi \in X$ but $\varphi \notin Y$. Put $\varphi := \bigvee_{i<n}\varphi_i$. Then $\Box^s\varphi \in X$ but $\varphi \notin Y$. ∎

Theorem 3.2.9. *Let s and t be finite sets of finite sequences over κ. Then $\mathbf{K}_\kappa \oplus \Box^s p \to \Box^t p$ is canonical. Moreover, the canonical frame satisfies $\lhd^t \subseteq \lhd^s$.*

Proof. We prove the second claim first. Let X and Y be worlds such that $X \lhd^t Y$. Assume that $\Box^s\varphi \in X$. Then also $\Box^t\varphi \in X$. Since $X \lhd^t Y$ we have $\varphi \in Y$. Hence by the previous lemma $X \lhd^s Y$. So, $\lhd^t \subseteq \lhd^s$. To show that $\mathbf{K}_\kappa \oplus \Box^s p \to \Box^t p$ is canonical it is enough to show that if $\mathfrak{f}$ is a Kripke–frame such that $\lhd^t \subseteq \lhd^s$ then $\mathfrak{f} \vDash \Box^s p \to \Box^t p$. But this is straightforward. ∎

As an example, let $\mathfrak{f} = \langle f, \lhd, \blacktriangleleft\rangle$ be a bimodal Kripke–structure. $\mathfrak{f}$ satisfies $\Diamond p \to \blacklozenge p$ iff $\lhd \subseteq \blacktriangleleft$. $\mathfrak{f}$ satisfies $\Diamond\blacklozenge p \to \blacklozenge\Diamond p$ iff $\lhd\circ\blacktriangleleft \subseteq \blacktriangleleft\circ\lhd$. Also, $\mathfrak{f}$ satisfies $\Diamond\blacklozenge p \leftrightarrow \blacklozenge\Diamond p$ iff $\lhd$ and $\blacktriangleleft$ commute.

Definition 3.2.10. *Let φ a modal formula and α be a first–order formula in the language of predicate logic with $\lhd_i$ and equality. φ **corresponds** to α iff a Kripke structure $\mathfrak{f}$ satisfies φ exactly if it satisfies α.*

In light of this definition the following correspondences are valid.

Axiom	First–Order Property	Description
$\Diamond\top$	$(\forall x)(\exists y)(x \lhd y)$	definality
$p \to \Diamond p$	$(\forall x)(x \lhd x)$	reflexivity
$p \to \Box\Diamond p$	$(\forall xy)(x \lhd y \to y \lhd x)$	symmetry
$\Diamond\Diamond p \to \Diamond p$	$(\forall xyz)(x \lhd y \wedge y \lhd z. \to .x \lhd z)$	transitivity
$\Diamond\Box p \to \Box\Diamond p$	$(\forall xyz)(\exists w)(x \lhd y \wedge x \lhd z. \to .$ $y \lhd w \wedge z \lhd w)$	convergence
$\Diamond p \to \Box p$	$(\forall xyz)(x \lhd y \wedge x \lhd z. \to .y = z)$	partial functionality
$\Diamond p \to \Box\Diamond p$	$(\forall xyz)(x \lhd y \wedge x \lhd z. \to .y \lhd z)$	euclideanness
.3	$(\forall xyz)(x \lhd y \wedge x \lhd z. \to .$ $y \lhd z \vee y = z \vee z \lhd y)$	local connectedness

For polymodal logics similar correspondences can be established. Important are the tense postulates $p \to \Box_0\Diamond_1 p$ and $p \to \Box_1\Diamond_0 p$. A frame satisfies the first iff the relation $\lhd_0$ is contained in the converse of $\lhd_1$, that is, if $x \lhd_0 y$ implies $y \lhd_1 x$. For a proof assume that $x \lhd_0 y \not\lhd_1 x$. Then put $\beta(p) := \{x\}$. Then $\langle \mathfrak{f}, \beta, y\rangle \vDash \Box_1\neg p$ and so $\langle \mathfrak{f}, \beta, x\rangle \vDash p; \Diamond_0\Box_1\neg p$. So the axiom is violated. Assume then that the frame satisfies the first order property. Pick a valuation β and a point x. Assume $\langle \mathfrak{f}, \beta, x\rangle \vDash p$. Take a y such that $x \lhd_0 y$. Then $y \lhd_1 x$ and so $\langle \mathfrak{f}, \beta, y\rangle \vDash \Diamond_1 p$. Thus, as y was arbitrary, $\langle \mathfrak{f}, \beta, x\rangle \vDash \Box_0\Diamond_1 p$, which had to be shown.

Theorem 3.2.11. *A bimodal Kripke structure satisfies $\boldsymbol{K.t}$ iff $\lhd_0$ is the converse of $\lhd_1$.*

We conclude with a general theorem on logics of bounded alternative.

Theorem 3.2.12 (Bellissima). *Every logic of bounded alternative is canonical and hence compact and complete.*

Proof. Let Λ be of alternative α and φ an axiom of Λ. We want to show that $\mathfrak{can}_\Lambda(\aleph_0) \vDash \varphi$. To see that, assume that there is a point x and a valuation β such that $\langle \mathfrak{can}_\Lambda(\aleph_0), \beta, x\rangle \vDash \neg\varphi$. Let d be the modal depth of φ. Then there are at most $1 + \alpha + \alpha^2 + \ldots + \alpha^d = \frac{\alpha^{d+1}-1}{\alpha-1}$ points reachable in at most d steps from x. Let the set of these points be X. We claim now that for any set $T \subseteq X$ there is an internal set T° in $\mathfrak{Can}_\Lambda(\aleph_0)$ such that T° coincides with T on the d–transit of x. Thus if we put $\gamma(p) := \beta(p)^\circ$ then we have $\langle \mathfrak{can}_\Lambda(\aleph_0), \gamma, x\rangle \vDash \neg\varphi$, and since the $\gamma(p)$ are internal, we now have $\langle \mathfrak{Can}_\Lambda(\aleph_0), \gamma, x\rangle \nvDash \varphi$, which had to be shown. Now for the existence of these sets. Let $X = \{x_0, \ldots, x_{n-1}\}$. For each pair i, j we have a set $S(i, j)$ containing x_i but not x_j. (For if x_i and x_j are different, they represent different ultrafilters, and so there is a formula φ such that $\varphi \in x_i$ but $\varphi \notin x_j$. Now let $S(i, j) := \widehat{\varphi} = \{X : \varphi \in X\}$ in the canonical frame.) Then let $S(i) = \bigcap_{j \neq i} S(i, j)$. $S(i)$ contains x_i but none of the j. Thus, for any subset $Y \subseteq X$ put $Y^0 = \bigcup\{S(i) : i \in Y\}$. Y^0 is internal. This concludes the proof. ∎

Exercise 83. Prove the remaining correspondence properties.

Exercise 84. Show that a frame $\mathfrak{G}$ satisfies **D** iff the underlying Kripke structure satisfies $(\forall x)(\exists y)(x \lhd y)$. Show that **K.D** is canonical.

Exercise 85. Show that **S5** is canonical. *Hint.* Show that the axiom **B** is valid in a Kripke structure iff that structure is symmetric. Proceed as with **S4**.

Exercise 86. Show that the tense logic **K**.t is canonical and complete.

Exercise 87. Show that a finite Kripke–frame satisfies **G** iff it is transitive and irreflexive.

Exercise 88. Show that a finite Kripke–frame satisfies **Grz** iff it is reflexive, transitive and antisymmetric, that is, if $x \lhd y$ and $y \lhd x$ then $x = y$.

Exercise 89. Show that if $\mathbf{K} \oplus \varphi$ is canonical for all $\varphi \in X$, then $\mathbf{K} \oplus X$ is canonical as well.

Exercise 90. Let Λ be complete. Define the *Kripke–consequence* $\vdash^k_\Lambda$ for Λ as follows. $\Phi \vdash^k_\Lambda \varphi$ iff for every Λ–Kripke–frame $\mathfrak{f}$ if $\langle \mathfrak{f}, \beta, x\rangle \vDash \Phi$ then also $\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi$. Show that $\vdash_\Lambda \subseteq \vdash^k_\Lambda$ and that equality holds iff Λ is compact iff $\vdash^k_\Lambda$ is finitary (or compact).

***Exercise 91.** Show that if Λ is not complete, then it may well be that $\vdash^k_\Lambda$ is finitary. In that case, however, $\vdash^k_\Lambda$ must be strictly stronger than $\vdash_\Lambda$.

***Exercise 92.** Show that **G** is 1–compact but not 2–compact. *Hint.* To show the first claim show that one can only define boolean combinations of statements of the form $\Box^n\bot \wedge \neg\Box^{n-1}\bot$, stating that there exists no upgoing chain of points of length $n + 1$. Show then that any infinite set of such formulae if jointly consistent has a model based on a Kripke–frame. For the second claim, consider the following formulae. Put

$$\varphi_i := \neg\Diamond(p \wedge \Box^i\bot) \wedge \Diamond(p \wedge \Box^{i+1}\bot) \;.$$

Show that the following set is consistent, but has no model based on a Kripke–frame

$$\{\Diamond\varphi_0\} \cup \{\Box(\varphi_i \to \Diamond\varphi_{i+1}) : i \in \omega\} \;.$$

(This is the solution of Warren Goldfarb as given in [**32**].)

***Exercise 93.** Let us consider a fixed set $\{p_i : i < n\}$ of sentence letters and define for a set $S \subseteq n$ the formulae $\chi_S := \bigwedge_{i\in S} p_i \wedge \bigwedge_{i\notin S} \neg p_i$. Consider the following formulae

$$\varphi_n := \bigwedge_{S\subseteq n} \Diamond(\chi_S \wedge \Box\bot). \to .\Box(\Box p_n \to p_n) \to \Box p_n \;.$$

Show that the logic $\mathbf{K}_1 \oplus \varphi_n$ is $n - 1$–canonical but not n–compact.

3.3. Frame Constructions II

This chapter is devoted to the question of making models as small as possible or to produce models that satisfy certain properties. Before we enter this discussion, we introduce some very important terminology. First, let $\mathfrak{f} = \langle f, \langle \lhd_i : i < \kappa \rangle \rangle$ be a Kripke–frame and $g \subseteq f$. Let $\mathfrak{g} = \langle g, \langle \lhd_i^{\mathfrak{g}} : i < \kappa \rangle \rangle$ with $\lhd_i^{\mathfrak{g}} = \lhd_i \cap g \times g$. Then we call $\mathfrak{g}$ a **subframe of** $\mathfrak{f}$, in symbols $\mathfrak{g} \sqsubseteq \mathfrak{f}$. For a general frame $\mathfrak{F}$, $\mathfrak{G}$ is a subframe if $g \in \mathbb{F}$, $\mathfrak{g} \sqsubseteq \mathfrak{f}$, and $\mathbb{G} = \{a \cap g : a \in \mathbb{F}\}$. In general, for any subset $p \subseteq f$, $\{p \cap a : a \in \mathbb{F}\}$ is called the **trace algebra** of $\mathfrak{F}$ in $\mathfrak{G}$. The trace algebra is actually closed under complement and union, as is verified. Moreover,

$$\begin{aligned} \blacksquare_j^g b &= \{x \in g : (\forall y \rhd_j^g x)(y \in b)\} \\ &= \{x \in g : (\forall y \rhd_j^f x)(y \in g. \Rightarrow .y \in b)\} \\ &= g \cap \blacksquare_j^f (g \to b) \end{aligned}$$

Hence, the subframe based on an internal set is always well–defined. A valuation β on $\mathfrak{f}$ defines a valuation γ on $\mathfrak{g}$ in the natural way, by $\gamma(p) := \beta(p) \cap g$; this valuation is often also denoted by β. Let us be given a Kripke–frame $\mathfrak{f}$ and a subset S. We put $suc_j(S) = \{y : (\exists x \in S)(x \lhd_j y)\}$. The m–**wave** $Wave_{\mathfrak{f}}^m(S)$ and the m–**transit** $Tr_{\mathfrak{f}}^m(S)$ of S in $\mathfrak{f}$ are defined as follows.

$$\begin{aligned} Wave_{\mathfrak{f}}^0(S) &:= S \\ Wave_{\mathfrak{f}}^1(S) &:= \textstyle\bigcup_{j<\kappa} suc_j(S) \\ Wave_{\mathfrak{f}}^{m+1}(S) &:= Wave_{\mathfrak{f}}^1(Wave_{\mathfrak{f}}^m(S)) \\ Tr_{\mathfrak{f}}^m(S) &:= \textstyle\bigcup_{i<m} Wave_{\mathfrak{f}}^i(S) \\ Tr_{\mathfrak{f}}(S) &:= \textstyle\bigcup_{i\in\omega} Tr_{\mathfrak{f}}^i(S) \end{aligned}$$

$Tr_{\mathfrak{f}}(S)$ is called the **transit** of S in $\mathfrak{f}$. All these definitions just define subsets of frames; but if $\mathfrak{f} = \langle f, \langle \lhd_j : j < \kappa \rangle \rangle$ is a polyframe and $h \subseteq f$ then we can regard h naturally as a subframe $\mathfrak{h} = \langle h, \langle \lhd_j^{\mathfrak{h}} : j < \kappa \rangle \rangle$ where $\lhd_j^{\mathfrak{h}} := \lhd_j \cap h \times h$. Similarly, we write $\mathfrak{Tr}_{\mathfrak{f}}^m(S)$ for the subframe based on the m–transit of S. If there exists an element w such that $\mathfrak{f} = \mathfrak{Tr}_{\mathfrak{f}}(\{w\})$ (which we will also denote by $\mathfrak{Tr}_{\mathfrak{f}}(w)$) then $\mathfrak{f}$ is called **rooted** and w the **root** of $\mathfrak{f}$. Rooted frames are sometimes also called **one–generated**. However, we will avoid this terminology. All definitions apply equally to generalized frames, where the waves and transits are computed with reference to the underlying Kripke–frames. Put $x \preccurlyeq y$ if $y \in Tr_{\mathfrak{f}}(x)$ and $x \prec y$ if $Tr_{\mathfrak{f}}(y) \subsetneq Tr_{\mathfrak{f}}(x)$. Put $Cy(x) = \{y : x \preccurlyeq y \preccurlyeq x\}$ and call it the **cycle** of x. A set M is a **cycle** if it is of the form $Cy(x)$ for some x. A frame is **cyclic** if its underlying set is a cycle. The **depth** of a point in a frame $\mathfrak{F}$ is defined as follows.

$$dp(x) := \{dp(y) : x \prec y\}$$

This is generally not a good definition, e. g. in the frame $\langle \omega, \leq \rangle$. Hence, we require that the definition is applied only to points for which there is no infinite chain $\langle x_i : i \in \omega \rangle$ such that $x_0 = x$ and $x_i \prec x_{i+1}$. In this case the depth is well–defined and yields an ordinal number. (This number is not necessarily finite.) In all other cases, x is said to have no depth.

Definition 3.3.1. *Let $\mathfrak{f}$ be a Kripke–frame. A **precluster** is a maximal set P of points such that for all $x, y \in P$ and all $j < \kappa$ we have $suc_j(x) = suc_j(y)$. A **cluster** is a maximal connected set C of points such that for all $x, y \in C$ and all $j < \kappa$ we have $suc_j(x) = suc_j(y)$.*

A cluster is a cycle in a precluster, but not conversely. Given a frame $\mathfrak{f}$ and a precluster P the map collapsing P into a single point can be turned into a p–morphism. This can be derived from a more general theorem which we will now present. Let $\mathfrak{F}$ be a frame and $\mathfrak{G} \sqsubseteq \mathfrak{F}$ be a subframe. Assume that for all $x, y \in g$ and all $j < \kappa$ we have $suc_j(x) \cap (f - g) = suc_j(y) \cap (f - g)$. Then we call $\mathfrak{G}$ a **local subframe** of $\mathfrak{F}$.

Proposition 3.3.2 (Net Extension III). *Let $\mathfrak{F}$ be a frame and $\mathfrak{G} \sqsubseteq \mathfrak{F}$ be a local subframe. Let $\sim$ be a net on $\mathfrak{G}$. Put $x \approx y$ if (i.) $x, y \in g$ and $x \sim y$ or (ii.) $x, y \in f - g$ and $x = y$. Then $\approx$ is a net on $\mathfrak{F}$.*

Proof. Let $x \lhd_j y$ and $x \approx x'$. If $x \notin g$ then $x = x'$ and then trivially $x' \lhd_j y$. So, suppose that $x \in g$. Then $x' \in g$ and $x \sim x'$. Assume $y \notin g$. By assumption on g, $suc_j(x) \cap (f - g) = suc_j(x') \cap (f - g)$, so that $x' \lhd_j y$. Assume now $y \in g$. Since $\sim$ is a net there exists a y' such that $y \sim y'$ and $x' \lhd_j y'$. This shows that $\approx$ is a net on $\mathfrak{g}$. Next, let $a \in \mathbb{F}$. Then $a = b \cup c$ where $b := a \cap g$ and $c := a \cap (f - g)$. Now $[a]_\approx = [b]_\approx \cup [c]_\approx = [b]_\sim \cup c \in \mathbb{F}$, since $[b]_\sim \in \mathbb{F}$ by the fact that $\sim$ is a net on $\mathfrak{F}$. ◻

This theorem allows several generalizations, but in this form it is most useful (and simple). There is another very useful operation that allows to replace a subframe by a larger subframe. In this construction we say that $\mathfrak{K}$ is obtained by **blowing up** $\mathfrak{F}$ by a p–morphism.

Theorem 3.3.3 (Blowing Up). *Let $\mathfrak{G} \sqsubseteq \mathfrak{F}$ and $c : \mathfrak{H} \twoheadrightarrow \mathfrak{G}$. Assume that h is disjoint with f (and so also with g). Define a new frame $\mathfrak{K}$ as follows.*

$$\begin{array}{lll}
k & := & (f - g) \cup h \\
\lhd_j^k & := & \lhd_j^f \cap (f - g)^2 \cup \lhd_j^h \\
 & & \cup\{\langle x, y\rangle : x \in f, y \in h, x \lhd_j^f c(y)\} \\
 & & \cup\{\langle x, y\rangle : x \in h, y \in f, c(x) \lhd_j^f y\} \\
\mathbb{K} & := & \{a \cup b : a \in \mathbb{F}, b \in \mathbb{H}, a \cap g = \varnothing\}
\end{array}$$

If $\mathbb{G} = \{c[a] : a \in \mathbb{H}\}$ then $\mathfrak{K}$ is a frame. The map d defined by $d(x) := x$ if $x \notin h$ and $d(x) := c(x)$ otherwise is a p–morphism of $\mathfrak{K}$ onto $\mathfrak{F}$.

Proof. First we verify that d is a contraction of the Kripke–frames. To see that, assume $x \lhd_j^k y$. If $\{x, y\} \subseteq k - h$ then $x \lhd_j^f y$ and $d(x) = x$ as well as $y = d(y)$. If $\{x, y\} \subseteq h$ then $x \lhd_j^h y$. Then $d(x) = c(x)$ and $d(y) = c(y)$ and so by assumption on

c, $c(x) \lhd^g_j c(x)$. Hence $d(x) \lhd^f_j d(y)$. Next, let $x \in k - h$ and $y \in h$. Then $x \lhd^k_j y$ iff $x \lhd^f_j c(y)$. Since $x = d(x)$ and $c(y) = d(y)$, the claim follows. Likewise for the last case, that $x \in h$ and $y \in k - h$. This shows the first p–morphism condition. Next, let $x \in k$, $u \in f$ and assume $d(x) \lhd^f_j u$. Let $\{x, u\} \subseteq f - g$. Then $x = d(x)$ and $u = d(u)$ and $x \lhd^k_j u$, as required. Next assume $\{x, u\} \subseteq g$. Then, since c is a p–morphism, there exists a $y \in h$ such that $x \lhd^h_j y$ and $c(y) = u$. Then also $d(y) = u$ and $x \lhd^k_j y$. Third, assume $x \in g$ and $u \in f - g$. Then $d(u) = u$ and $x \lhd^k_j u$ by construction. The remaining case is also straightforward. To see that $\mathfrak{K}$ is a frame, we must show that $\mathbb{K}$ is closed under boolean operations and $\blacklozenge^k_j$. The boolean operations are straightforward. To show closure under $\blacklozenge^k_j$, we take a set $a \in \mathbb{K}$. It is of the form $a_1 \cup a_2$ where $a_1 \in \mathbb{F}$ and $a_1 \subseteq f - g$ and $a_2 \in \mathbb{H}$. By assumption, $b_2 := c[a_2] \in \mathbb{G}$ and so also $b_2 \in \mathbb{F}$. (1.) $\blacklozenge^k_j a_1 = ((k - h) \cap \blacklozenge^k_j a_1) \cup (h \cap \blacklozenge^k_j a_1)$. $(k - h) \cap \blacklozenge^k_j a_1 = (f - g) \cap \blacklozenge^f_j a_1 \in \mathbb{F}$. $h \cap \blacklozenge^k_j a_1 = c^{-1}[g \cap \blacklozenge^f_j a_1] \in \mathbb{H}$. So this case is settled. (2.) $\blacklozenge^k_j a_2 = ((k-h) \cap \blacklozenge^k_j a_2) \cup (h \cap \blacklozenge^k_j a_2)$. $(k-h) \cap \blacklozenge^k_j a_2 = c^{-1}[(f-g) \cap \blacklozenge^f_j b_2] \in \mathbb{F}$. $h \cap \blacklozenge^k_j a_2 = c^{-1}[g \cap \blacklozenge^f_j b_2]$. This concludes the proof that $\mathfrak{K}$ is a frame. Finally, $c^{-1} : \mathbb{F} \to \mathbb{K}$ is clearly injective. This concludes the proof of the theorem. ◻

Corollary 3.3.4 (Multiplication). *Let $\mathfrak{F}$ be a frame, and $\mathfrak{G} \sqsubseteq \mathfrak{F}$. There exists a natural p–morphism $c : \bigoplus_{i \in I} \mathfrak{G} \twoheadrightarrow \mathfrak{G} : \langle x, i \rangle \mapsto x$. The result $\mathfrak{K}$ of blowing up $\mathfrak{F}$ by c is a frame. We say that $\mathfrak{F}$ has been obtained from $\mathfrak{F}$ by **multiplying** $\mathfrak{G}$ (**$\sharp I$ times**).*

The *depth* of a modal formula corresponds quite directly to the bounded transits.

Proposition 3.3.5. *Let $\langle \mathfrak{f}, \beta, x \rangle \vDash \varphi$ and $k = dp(\varphi)$. Then*

$$\langle \mathfrak{Tr}^k_{\mathfrak{f}}(x), \beta, x \rangle \vDash \varphi \ .$$

Proof. By induction on $k = dp(\varphi)$. If $k = 0$, then it is easy to check that $\langle \{x\}, \beta, x \rangle \vDash \varphi$ as well, by the fact that the evaluation clauses are only $(md\neg.)$, $(md\wedge.)$ and so do not change the world at which we evaluate. Now suppose that $\varphi = \Box_j \psi$, $dp(\psi) = k$. Then

$$\begin{array}{rcll} \langle \mathfrak{f}, \beta, x \rangle \vDash \Box_j \psi & \Leftrightarrow & \text{for all } y \rhd_j x & \langle \mathfrak{f}, \beta, y \rangle \vDash \psi \\ & \Leftrightarrow & \text{for all } y \rhd_j x & \langle \mathfrak{Tr}^k_{\mathfrak{f}}(y), \beta, y \rangle \vDash \psi \\ & \Leftrightarrow & \text{for all } y \rhd_j x & \langle \mathfrak{Tr}^{k+1}_{\mathfrak{f}}(x), \beta, y \rangle \vDash \psi \\ & \Leftrightarrow & \langle \mathfrak{Tr}^{k+1}_{\mathfrak{f}}(x), \beta, x \rangle \vDash \Box_j \psi & . \end{array}$$

◻

Having established that a consistent formula has a finite model we can also give some bounds for the size of such a model. We can actually use the bounds obtainable from the proof directly, but we seize the opportunity to introduce the method of *filtration*. By itself it is a rather crude method for proving finite model property, and many people have found ingenious ways of refining it so that it allows to give proofs for many standard systems. Here we only present the most basic variant. Suppose that we have a model $\langle \mathfrak{F}, \beta, x\rangle \vDash \varphi$. Assume β to be defined only for the variables of φ. Now let $X := sf(\varphi)$. Let $v \sim_X w$ iff they satisfy the same formulae of X, that is, for all $\psi \in X$ $\langle \mathfrak{F}, \beta, v\rangle \vDash \psi \Leftrightarrow \langle \mathfrak{F}, \beta, w\rangle \vDash \psi$. This is an equivalence relation. Let $[v] := \{w : v \sim_X w\}$ and $[f] = \{[v] : v \in f\}$. There are at most $2^{\sharp X}$ distinct classes, so $\sharp[f] \leq 2^{\sharp X}$. Now put $[v]\overline{\lhd}_j[w]$ iff there exists $\widehat{v} \in [v]$ and $\widehat{w} \in [w]$ such that $\widehat{v} \lhd_j \widehat{w}$. Next define γ by $\gamma(p) := \{[v] : \langle \mathfrak{F}, \beta, v\rangle \vDash p\}$. This is well–defined by the fact that if $v \vDash p$ then $p \in var(\varphi)$ and if $v \sim_X w$ then also $w \vDash p$. This model is said to be obtained by **filtrating** the original model. Then

$$\langle\langle [f], \langle \overline{\lhd}_j : j < \kappa\rangle\rangle, \gamma, [x]\rangle \vDash \varphi$$

Namely, by induction on $\psi \in X$ we show that $[v] \vDash \psi$ iff $v \vDash \psi$. This is straightforward for variables, and the only critical step is $\psi = \Diamond_j\chi$. Here, assume $[v] \vDash \Diamond_j\chi$. Then $[w] \vDash \chi$ for some w such that $[v]\overline{\lhd}_j[w]$. Then there exist $\widehat{v}$ and $\widehat{w}$ such that $\widehat{v} \sim_X v$, $\widehat{w} \sim_X w$ and $\widehat{v} \lhd_j \widehat{w}$. Thus $\widehat{w} \vDash \chi$ by construction and induction hypothesis and so $\widehat{v} \vDash \Diamond_j\chi$, showing $v \vDash \Diamond_j\chi$ by $v \sim_X \widehat{v}$. Conversely, if $v \vDash \Diamond_j\chi$ then for some j–successor w we have $w \vDash \chi$ and so $[w] \vDash \chi$ by induction hypothesis. By construction, $[v]\overline{\lhd}_j[w]$ and so $[v] \vDash \Diamond_j\chi$, as required.

Theorem 3.3.6. *Let φ be consistent in $\boldsymbol{K}_\kappa$. Then there exists a model with at most 2^k points, where $k = \sharp sf(\varphi)$.*

Normal forms are closely connected with a technique called *unravelling*. There are more or less cautious variants of unravelling. The most basic method is the following. Suppose we have a pointed Kripke–frame $\langle \mathfrak{f}, w_0\rangle$. A **path** of length r in $\langle \mathfrak{f}, w_0\rangle$ is a function $\pi : r + 1 \to f$ such that $\pi(0) = w_0$ and for all $i < r + 1$ we have $\pi(i) \lhd_j \pi(i + 1)$ for some $j < \kappa$. We say that $\pi(0)$ is the **begin point** and $\pi(r)$ the **end point** of π, denoted by $ep(\pi)$. The length of π is denoted by $\ell(\pi)$. We say that $\pi^+ : m \to f$ **extends** $\pi : n \to f$ if $n \leq m$ and for all $i \leq n$, $\pi(i) = \pi^+(i)$. The **unravelling** of degree r of $\langle \mathfrak{f}, w_0\rangle$ is denoted by $\mathfrak{u}_r(\mathfrak{f}, w_0)$, and defined as follows. It is a frame based on the set of paths of length $\leq r$. The relation $\lhd_j$ is defined by $\pi \lhd_j \pi^+$ iff (i.) $\ell(\pi^+) = \ell(\pi) + 1$, (ii.) π^+ extends π and (iii.) $ep(\pi) \lhd_j ep(\pi^+)$. So, $\pi \lhd_j \pi^+$ if π^+ is the path that goes just one step further than π and to a point which is j–accessible from the end point of π. $\mathfrak{u}_r$ thus defined has a unique generating point, namely the path $\pi_0 : 1 \to f$, sending 0 to w_0; we denote it by $\langle w_0\rangle$. The map $ep(-) : \pi \to ep(\pi)$ sending each path to its end point is not quite a p–morphism. However, with respect to points of restricted depth, it is as good as a p–morphism. We formalize this as follows. Say that $h : f \to g$ is n–**localic** with respect to $S \subseteq f$ if the following holds.

(gmf.) If $x, y \in Tr^n(S)$ and $x \lhd_j y$ then $h(x) \lhd_j h(y)$

(gmb.) If $x \in Tr^{n-1}(S)$ and $h(x) \lhd_j u$ then
$x \lhd_j y$ for some y such that $h(y) = u$

It is easy to verify that the map $ep : \mathfrak{u}_r(\mathfrak{f}, w_0) \to \mathfrak{f}$ is r–localic with respect to $\langle w_0\rangle$. The next theorem then says that whenever we can satisfy a formula φ of depth $\leq r$ at w_0 then it can be satisfied in $\mathfrak{u}_r(\mathfrak{f}, w_0)$ at the path π_0.

Proposition 3.3.7. *Let $\mathfrak{f}$ and $\mathfrak{g}$ be Kripke–frames. Suppose that $h : f \to g$ is n–localic with respect to S and that $w \in S$. Let β, γ be valuations such that $\gamma(p) = h[\beta(p)]$. Then for all formulae φ of degree $\leq n$*

$$\langle \mathfrak{f}, \beta, w\rangle \vDash \varphi \qquad \Leftrightarrow \qquad \langle \mathfrak{g}, \gamma, h(w)\rangle \vDash \varphi$$

The proof is an easy induction on φ and is left to the reader. The frame $\mathfrak{u}_m$ is a subframe of $\mathfrak{u}_n$ if $m \leq n$. It consists of the m–transit of π_0 in $\mathfrak{u}_n$. The total unravelling $\mathfrak{u}_\omega$ is the union of all $\mathfrak{u}_n$, $n \in \omega$. This is well–defined.

Theorem 3.3.8. *Let $\langle \mathfrak{f}, w_0\rangle$ be a pointed Kripke–frame. The map ep, sending each path beginning at w_0 to its end point, is a p–morphism from $\langle \mathfrak{u}_\omega, \langle w_0\rangle\rangle$ onto $\langle \mathfrak{f}, w_0\rangle$.*

Proof. The map is onto by definition of the transit. Now take π, π^+ such that $\pi \lhd_j \pi^+$. Then $ep(\pi) \lhd_j ep(\pi^+)$, by construction. Next assume $ep(\pi) \lhd_j u$. Then the path π^+ defined by extending π by just one more point, namely u, is a well–defined path and we have $\pi \lhd_j \pi^+$ by construction. ◻

The method of unravelling can be used to show that $\mathbf{K}_\kappa$ is complete with respect to completely intransitive trees, by first showing that it is complete and then using unravelling to get a totally intransitive tree from a model. This is somewhat better than the proof via normal forms, which established completeness with respect to acyclic frames only. Finally, let us note that if $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ is a frame and $q : \mathfrak{u}_\omega(\mathfrak{f}, w_0) \twoheadrightarrow \mathfrak{f}$ a total unravelling, then we can define a system $\mathbb{U}$ of sets by $\mathbb{U} := \{q^{-1}[a] : a \in \mathbb{F}\}$. By the fact that we have a p-morphism, $\mathfrak{U}_\omega(\mathfrak{F}, w_0) := \langle \mathfrak{u}_\omega(\mathfrak{f}, w_0), \mathbb{U}\rangle$ has the same modal theory as $\mathfrak{F}$. This fact is of some importance.

We remark here that there is a more extreme variant of unravelling, which is as follows. Define a κ–path to be a sequence $\pi = \langle w_0, \lambda_0, w_1, \lambda_1, w_2, \dots, w_n\rangle$ such that $w_i \lhd_{\lambda_i} w_{i+1}$ for every $i < n$. We say that π *starts* in w_0 and ends in w_n. Put $\pi \lhd_j \pi'$ if $\pi' = \langle w_0, \lambda_0, w_1, \lambda_1, \dots, w_n, \lambda_n, w_{n+1}$ for some $\lambda_n < \kappa$ and some $w_{n+1} \in f$. Then we can form the frame $\mathfrak{x}_\omega(\mathfrak{f}, w_0)$ of all κ–paths in $\mathfrak{f}$ starting at w_0. (The definition of $\mathfrak{x}_n(\mathfrak{f}, w_0)$ is analogous. It is the subframe of κ–paths of length $\leq n$.) The map ζ sending π to the sequence $\langle w_i : i < n + 1\rangle$ is a p–morphism onto $\mathfrak{u}_\omega(\mathfrak{f}, w_0)$. It follows that there is a p–morphism from $\mathfrak{x}_\omega(\mathfrak{f}, w_0)$ onto the transit of w_0 in $\mathfrak{f}$, namely the map sending each κ–path to ints end point. In $\mathfrak{x}_\omega(\mathfrak{f}, w_0)$ for any pair x and y of points there is at most one relation $\lhd_j$ such that $x \lhd_j y$. This is not so in $\mathfrak{u}_\omega(\mathfrak{f}, w_0)$. For practical purposes the difference between these constructions is only marginal.

Exercise 94. Show by means of filtration that $\mathbf{K}_\kappa$ has the global finite model property.

Exercise 95. Generalize the method of unravelling to unravellings generated by *sets* rather than *points*. That is, form the frame consisting of paths starting in a given set S.

Exercise 96. A *forest* is a disjoint union of trees. Show that every Kripke–frame is the p–morphic image of a forest of intransitive, irreflexive trees.

3.4. Weakly Transitive Logics I

The notion of a weakly transitive logic plays a pivotal role in modal logic. Many theorems hold only for weakly transitive logics. In this section we will collect some elementary facts about them. In Section 4.3 we will return to that subject matter again. First recall that a logic is weakly transitive iff there exists a maximal modality with respect to $\leq_\Lambda$, where $\leq_\Lambda$ is defined by $\boxplus \leq_\Lambda \boxplus'$ iff $\boxplus' p \to \boxplus p \in \Lambda$ (see Section 2.1). Moreover, we have seen in Theorem 3.2.9 that logics axiomatized by axioms of the form $\boxplus p \to \boxplus' p$ are canonical and that the frame property determined by them is elementary.

Proposition 3.4.1. *Let Λ be a weakly transitive logic with master modality $\boxplus$. Then $\boxplus = \Box^s$ for some finite set of paths s and Λ is complete with respect to frames satisfying*

$$y \in Tr(x) \quad \Leftrightarrow \quad x \lhd^s y.$$

Proof. The first claim has been shown in Section 2.1. For the second claim we show that the canonical frame for Λ satisfies the property. Suppose that $Y \in Tr(X)$ in $\mathfrak{Can}_\Lambda(var)$. Then there exists a sequence σ of numbers $< \kappa$ such that $X \lhd^\sigma Y$. Now $\boxplus^s p \to \boxplus^\sigma p \in \Lambda$. Thus, by Theorem 3.2.9, $\lhd^\sigma \subseteq \lhd^s$, which means that $X \lhd^s Y$, as desired. ◻

Proposition 3.4.2. *The weakly transitive κ–modal logics form a filter in the lattice $\mathcal{E}\,\mathbf{K}_\kappa$. This filter is not principal. For a given compound modality $\boxplus$ there exists a least logic such that $\boxplus$ is maximal with respect to $\leq_\Lambda$.*

Proof. The first claim is straightforward. To see that there is no least weakly transitive logic observe that $\mathbf{K}_\kappa$ has the finite model property. So, if $\varphi \notin \mathbf{K}_\kappa$ there exists a finite $\mathfrak{f}$ such that $\mathfrak{f} \nvDash \varphi$. For some n, $\mathfrak{f}$ is n–transitive. Hence, $\varphi \notin \mathbf{K}_\kappa.trs_n$. This shows that

$$\mathbf{K}_\kappa = \bigcap_{n \in \omega} \mathbf{K}_\kappa.\mathbf{trs}_n$$

However, $\mathbf{K}_\kappa$ is not weakly transitive since we standardly assume $\kappa > 0$. Finally, let $\boxplus$ be given. Then put

$$\Lambda := \mathbf{K}_\kappa \oplus \{\boxplus p \to \boxplus' p : \boxplus' \text{ compound}\}$$

This is the smallest logic in which $\boxplus$ is maximal with respect to $\leq_\Lambda$. ◻

Proposition 3.4.3. *Let Λ be a logic. Λ is weakly transitive iff for every Λ–algebra and every principal open filter F of $\mathfrak{A}$, F is principal as a boolean filter.*

Proof. Suppose that Λ is weakly transitive with master modality $\boxplus$. Let F be a principal open filter, generated by the element a. We claim that $\boxplus a$ is the smallest element of F. To see this, observe that the least open filter containing a set E is the least boolean filter containing the set $E^{\Box} := \{\boxplus' e : e \in E, \boxplus' \text{ compound}\}$. This is the algebraic analogue of Proposition 3.1.2. Since $\boxplus$ is the master modality, we have $\boxplus e \leq \boxplus' e$ for all $\boxplus'$. Hence, the open filter generated by E is the boolean filter generated by $\boxplus E := \{\boxplus e : e \in E\}$. In particular, if $E = \{a\}$ this shows that the open filter is principal as a boolean filter. Now assume that Λ is not weakly transitive. Then let $\mathfrak{A} := \mathfrak{F}_\Lambda(\{p\})$. The open filter generated by (the equivalence class of) p in $\mathfrak{A}$ is not principal as a boolean filter; otherwise it has a smallest element. This element is of the form $\boxplus p$ for some compound modality $\boxplus$. Then, for any compound modality $\boxplus'$, $\boxplus p \leq \boxplus' p$, which is the same as $\boxplus p \to \boxplus' p \in \Lambda$. Hence Λ is weakly transitive, contrary to our assumption. ◻

The following is proved in [**30**] using algebraic methods.

Theorem 3.4.4 (Blok & Pigozzi). *Let Λ be a modal logic. $\Vdash_\Lambda$ admits a deduction theorem iff Λ is weakly transitive.*

Proof. Suppose $\Vdash_\Lambda$ admits a deduction theorem. Then there exists a term $p \twoheadrightarrow q$ such that for all sets of formulae Δ and formulae φ and ψ

$$(\ddagger) \qquad \Delta; \varphi \Vdash_\Lambda \psi \quad \Leftrightarrow \quad \Delta \Vdash_\Lambda \varphi \twoheadrightarrow \psi$$

Now since $p \twoheadrightarrow q \Vdash_\Lambda p \twoheadrightarrow q$ we deduce that $p \twoheadrightarrow q; p \Vdash_\Lambda q$. By Theorem 3.1.2 there exists a compound modality $\boxplus$ such that $\boxplus(p \twoheadrightarrow q); \boxplus p \vdash_\Lambda q$. By the Deduction Theorem for the local consequence of Λ, $\boxplus(p \twoheadrightarrow q) \vdash_\Lambda \boxplus p \to q$. Now let $\boxplus'$ be an arbitrary compound modality. Replacing q by $\boxplus' p$ we get $\boxplus(p \twoheadrightarrow \boxplus' p) \vdash_\Lambda \boxplus p \to \boxplus' p$. Notice now that $\boxplus(p \twoheadrightarrow \boxplus' p)$ is a theorem of Λ; for since $p \Vdash_\Lambda \boxplus' p$ we immediately get $\Vdash_\Lambda p \twoheadrightarrow \boxplus' p$, using $(\ddagger)$. And so $\boxplus(p \twoheadrightarrow \boxplus' p)$ is a theorem as well. Hence we have $\vdash_\Lambda \boxplus p \to \boxplus' p$. This shows that Λ is weakly transitive. For the converse, assume Λ is weakly transitive. Then there exists a compound modality $\boxplus$ such that $\boxplus p \to \boxplus' p \in \Lambda$ for all compound modalities $\boxplus'$. Put $p \twoheadrightarrow q := \boxplus p \to q$. We claim that $(\ddagger)$ holds with respect to $\twoheadrightarrow$. For assume $\Delta; \varphi \Vdash_\Lambda \psi$. By Theorem 3.1.2 there exists a $\boxplus'$ such that $\boxplus' \Delta; \boxplus' \varphi \vdash_\Lambda \psi$. Since $\boxplus \varphi \vdash_\Lambda \boxplus' \varphi$, we also have $\boxplus' \Delta; \boxplus \varphi \vdash_\Lambda \psi$ and so $\boxplus' \Delta \vdash_\Lambda \boxplus \varphi \to \psi$. Hence $\Delta \Vdash_\Lambda \boxplus \varphi \to \psi$, as desired. The other direction of $(\ddagger)$ is straightforward. ◻

We close with the following useful observation. Given that Λ is m–transitive and that we have finitely many operators, then $\boxplus p := \boxtimes^{\leq m} p$ for some m is a master modality (though clearly not the only one). Hence if $\kappa < \aleph_0$ a weakly transitive logic is m–transitive for some m. There is an analogue of the next theorem without the assumption $\kappa < \aleph_0$, but we leave the generalization to the reader.

Theorem 3.4.5. ($\kappa < \aleph_0$.) *If a modal logic is m–transitive then every extension of Λ can be axiomatized by formulae of modal depth $\leq m + 1$.*

Proof. Suppose that Λ is m–transitive and let φ be a formula. Take a fresh variable q_ψ for each subformula and define ω as follows.

$$\begin{array}{rll} \omega := & & \langle q_p \leftrightarrow p : p \in var(\varphi)\rangle \\ & \wedge & \bigwedge\langle q_{\neg\psi} \leftrightarrow \neg q_\psi : \neg\psi \in sf(\varphi)\rangle \\ & \wedge & \bigwedge\langle q_{\psi\wedge\chi} \leftrightarrow q_\psi \wedge q_\chi : \psi \wedge \chi \in sf(\varphi)\rangle \\ & \wedge & \bigwedge\langle q_{\Box_j\psi} \leftrightarrow \Box_j q_\psi : \Box_j\psi \in sf(\varphi)\rangle \end{array}$$

(Obviously, the variables q_ψ must be pairwise distinct and distinct from the variables of φ.) Consider a model $\langle \mathfrak{F}, \beta, w_0\rangle \vDash \boxtimes^{\leq m}\omega \wedge \neg q_\varphi$. Let $\mathfrak{F}$ be rooted at w_0. Then it follows that $\langle \mathfrak{F}, \beta\rangle \vDash \omega$, since Λ is m–transitive. By induction on ψ it is shown using Lemma 3.1.7 that $\langle \mathfrak{F}, \beta\rangle \vDash q_\psi \leftrightarrow \psi$. Hence $\langle \mathfrak{F}, \beta, w_0\rangle \vDash \neg\varphi$. Conversely, assume $\langle \mathfrak{F}, \gamma, x\rangle \vDash \neg\varphi$. Put $\gamma(q_\psi) := \overline{\beta}(\psi)$. Then $\langle \mathfrak{F}, \gamma, x\rangle \vDash \boxtimes^{\leq m}\omega \wedge q_\psi$. Thus $\Lambda \oplus \varphi = \Lambda \oplus \boxtimes^{\leq m}\omega \to q_\varphi$, and we have $dp(\boxtimes^{\leq m}\omega \to q_\varphi) = m + 1$. ◻

Exercise 97. Show that a weakly transitive logic is globally decidable iff it is locally decidable. Likewise for *globally complete* and *global finite model property*.

Exercise 98. Formulate and prove a version of Theorem 3.4.5 that does not restrict κ to be finite.

3.5. Subframe Logics

In [**66**], Kit Fine introduced the notion of a *subframe logic* for logics extending **K4** and proved that all subframe logics have the finite model property. This will be shown again in Chapter 8.3. In Wolter [**244**] this concept was extended to general logics and it was shown that there exist subframe logics without the finite model property. Nevertheless, subframe logics have been established as an important tool in modal logic. The notion of a subframe logic is based on the concept of a subframe as defined previously. The algebraic concept corresponding to it is the notion of a *relativization*.

Definition 3.5.1. *Let* $\mathfrak{A} = \langle A, 1, -, \cap, \langle \blacksquare_j : j < \kappa\rangle\rangle$ *and* $b \in A$. *Put* $A_b := \{c : c \leq b\}$ *and* $\mathfrak{A}_b := \langle A_b, b, -, \cap, \langle \blacksquare_j^b : j < \kappa\rangle\rangle$ *where* $-$ *is the relative complement and* $\blacksquare_j^b c := b \cap \blacksquare_j(b \to c)$. *An algebra* $\mathfrak{B}$ *is called a* ***relativization*** *of* $\mathfrak{A}$ *if* $\mathfrak{B} = \mathfrak{A}_b$ *for some* $b \in A$.

Definition 3.5.2. *A logic is called a* ***subframe logic*** *if its class of frames is closed under taking subframes. Alternatively, a logic is a subframe logic if its class of algebras is closed under relativizations.*

Theorem 3.5.3 (Wolter). ($\kappa < \aleph_0$.) *Every subframe logic of bounded operator alternative has the finite model property.*

Proof. Every logic of bounded alternative is complete by Theorem 3.2.12. Hence if $\varphi \notin \Lambda$ then there is a Kripke–frame $\mathfrak{f}$ such that $\langle \mathfrak{f}, x\rangle \nvDash \varphi$ for some x and $\mathfrak{f} \vDash \Lambda$. Let

d be the modal depth of φ. Then $\mathfrak{Tr}^d_{\mathfrak{f}}(x) \vDash \varphi$. $\mathfrak{Tr}^d_{\mathfrak{f}}(x)$ is finite, and by the fact that Λ is a subframe logic it is also a frame for Λ. □

In Chapter 8.3 we will show the subframe theorem of [**66**]. The proof is somewhat long and tedious. However, there are some restricted variants which can be proved with less effort. We present one here. It is illustrative in the sense that it demonstrates that the subframe property is not straightforward in case we fail to know about completeness.

Theorem 3.5.4 (Fine). *Every subframe logic extending* $\boldsymbol{G}$ *has the finite model property.*

Proof. Observe that the **G**–axiom states that for every set a and $x \in a$ there is a maximal successor, that is, a point y such that $x \lhd y \in a$ but no successor of y is in a. Furthermore, by transitivity, if u has a successor in a then it has a maximal successor in a as well. The maximal points of a can be defined by $a \cap \blacksquare - a$. Let $\varphi \notin \Lambda$. Then for $n := \sharp var(\varphi)$ we have $\langle \mathfrak{Can}_\Lambda(n), \beta, w_0\rangle \vDash \neg\varphi$ for some β and x. Now let M be the set of points x such that $x \vDash \psi$ for some $\Diamond\psi \in X \cup \{\Diamond\varphi\}$ but $x \vDash \Box\neg\psi$. M is an internal set and so defines a subframe $\mathfrak{M}$. There exists a $w^* \in M$ such that $w^* \vDash \varphi$. For if $w_0 \notin M$ then $w_0 \vDash \Diamond\varphi$ and so $w_0 \vDash \Diamond(\varphi \wedge \Box\neg\varphi)$. Hence there exists a w^* such that $w_0 \lhd w^*$ and $w^* \vDash \varphi; \Box\neg\varphi$. From this follows $w^* \in M$. Let γ be the restriction of β to M. Then $\langle \mathfrak{M}, \gamma, w^*\rangle \vDash \varphi$. This is proved by showing that for all $\chi \in X$ and $x \in M$

$$\langle \mathfrak{M}, \gamma, x\rangle \vDash \chi \qquad \Leftrightarrow \qquad \langle \mathfrak{Can}_\Lambda(n), \beta, x\rangle \vDash \chi \ .$$

This holds for $\chi = p$ by definition of γ. The steps for $\wedge$ and $\neg$ are straightforward. Now let $\chi = \Diamond\psi$. From left to right is immediate. Now assume that $\langle \mathfrak{Can}_\Lambda(n), \beta, x\rangle \vDash \Diamond\psi$. Then there exists a y such that $x \lhd y$ and $\langle \mathfrak{Can}_\Lambda(n), \beta, y\rangle \vDash \psi; \Box\neg\psi$. It follows that $y \in M$; by induction hypothesis, $\langle \mathfrak{M}, \gamma, y\rangle \vDash \psi$. By assumption on Λ, $\mathfrak{M} \vDash \Lambda$. $\mathfrak{M}$ is not necessarily finite. However, $\mathfrak{M} \vDash \Box^m\bot$ for some m. For let $x \in M$; put $\mathbb{P}(x) := \{\Diamond\chi \in X : \langle \mathfrak{M}, \gamma, x\rangle \vDash \Diamond\chi\}$. If $x \lhd y$ then $\mathbb{P}(y) \subsetneq \mathbb{P}(x)$. Let $m := \sharp\{\Diamond\chi : \Diamond\chi \in X\}$. Then $\mathfrak{M} \vDash \Box^m\bot$. By Theorem 2.7.14, the set of internal sets is finite. By Theorem 2.4.11 the refinement map is a p–morphism. It has a finite image. Hence Λ has the finite model property. □

Now we show that the satisfiability of a formula φ on a subframe of $\mathfrak{F}$ can be translated into the satisfiability of another formula, which can be derived syntactically from φ.

$$\begin{array}{lcl} q \downarrow \chi & := & q \wedge \chi \\ (\neg\varphi) \downarrow \chi & := & \chi \wedge \neg(\varphi \downarrow \chi) \\ (\varphi \wedge \psi) \downarrow \chi & := & (\varphi \downarrow \chi) \wedge (\psi \downarrow \chi) \\ (\Box_j\varphi) \downarrow \chi & := & \chi \wedge \Box_j(\chi \to (\varphi \downarrow \chi)) \end{array}$$

We call $\varphi \downarrow \chi$ the **localization of** φ to χ.

Lemma 3.5.5. *Let* $\mathfrak{F}$ *be a frame,* $\mathfrak{G} = \mathfrak{F}_g$. *Let* $\beta : var \to f$ *and* $\gamma : var \to g$ *such that* $\beta(p) = g$ *and* $\gamma(q) = \overline{\beta}(q \wedge p)$ *for all* $q \neq p$. *Then* $\overline{\beta}(\varphi \downarrow p) = \overline{\gamma}(\varphi)$.

Proof. For a variable q, $x \in \overline{\beta}(q \downarrow p)$ iff $x \in \overline{\beta}(q \wedge p)$ iff $x \in \beta(q) \cap \beta(p)$ iff $x \in g$ and $x \in \gamma(q)$. Further, $x \in \overline{\beta}((\neg\varphi) \downarrow p)$ iff $x \in \overline{\beta}(p \wedge \neg(\varphi \downarrow p))$ iff $x \in g$ and $x \notin \overline{\gamma}(\varphi)$ iff $x \in g$ and $x \in \overline{\gamma}(\neg\varphi)$. The step for conjunction is straightforward. Now we turn to $\Box_j$.

$$\begin{array}{ll} & x \in \overline{\beta}((\Box_j\varphi) \downarrow p) \\ \text{iff} & x \in \overline{\beta}(p \wedge \Box_j(p \to (\varphi \downarrow p))) \\ \text{iff} & x \in g \text{ and } x \in \blacksquare_j\overline{\beta}(p \to (\varphi \downarrow p)) \\ \text{iff} & x \in g \text{ and for every } y \in g \text{ such that } y \rhd_j x : y \in \overline{\beta}(\varphi \downarrow p) \\ \text{iff} & x \in g \text{ and for every } y \in g \text{ such that } y \rhd_j x : y \in \overline{\gamma}(\varphi) \\ \text{iff} & x \in \overline{\gamma}(\Box_j\varphi) \end{array}$$

This ends the proof. ∎

Consequently, $p \to (\varphi \downarrow p)$ holds in $\mathfrak{F}$ iff φ holds in all subframes of $\mathfrak{F}$. Now define $\varphi^{sf} := p \to (\varphi \downarrow p)$, where p is a variable not occurring in φ. (For our purposes it will not matter which variable gets chosen.)

Theorem 3.5.6. *Let $\Lambda = \mathbf{K}_\kappa \oplus \Delta$ be a logic. The smallest subframe logic containing Λ, Λ^{sf} is axiomatizable by $\mathbf{K}_\kappa \oplus \Delta^{sf}$. The largest subframe logic contained in Λ is equal to $\Lambda_{sf} = \mathbf{K}_\kappa \oplus \{\varphi^{sf} : \varphi^{sf} \in \Lambda\}$.*

If follows immediately that a logic Λ is a subframe logic iff for every $\varphi \in \Lambda$ also $\varphi^{sf} \in \Lambda$.

Corollary 3.5.7 (Wolter). *Let $S\,F\,\boldsymbol{K}_\kappa$ denote the set of κ–modal subframe logics. This set forms a complete lattice with the operations of $\mathcal{E}\,\boldsymbol{K}_\kappa$. The natural embedding of $\mathcal{SF}\,\boldsymbol{K}_\kappa$ into $\mathcal{E}\,\boldsymbol{K}_\kappa$ commutes with infinite meets and joins.*

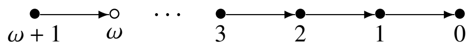
Proof. Let Λ_i, $i \in I$, be a set of subframe logics. Let $\mathfrak{F}$ be a frame and $\mathfrak{G}$ a subframe of $\mathfrak{F}$. Then if $\mathfrak{F} \vDash \bigsqcup_I \Lambda_i$, then $\mathfrak{F} \vDash \Lambda_i$ for all $i \in I$. By assumption, $\mathfrak{G} \vDash \Lambda_i$ for all $i \in I$, and so $\mathfrak{G} \vDash \bigsqcup_I \Lambda_i$. This shows that the infinite join is a subframe logic. Next assume that $\varphi \in \bigsqcap_I \Lambda_i$. Then for all $i \in I$, $\varphi \in \Lambda_i$. By assumption on the Λ_i, $\varphi^{sf} \in \Lambda_i$ for all $i \in I$. Hence $\varphi^{sf} \in \bigsqcap_I \Lambda_i$. Hence, the infinite meet is a subframe logic. ∎

In Wolter [234] an infinite series of incomplete subframe logic has been constructed. The simplest is the following logic, which is actually one of the earliest examples of an incomplete logic, taken from van Benthem [9]. Moreover, in Cresswell [48] it is shown that this logic is decidable, thus showing that there exist decidable, but incomplete logics. Cresswell uses Rabin's Theorem, but the result follows easily from Corollary 2.6.7. Take the frame $\Omega_{\omega+1}$ to be $\langle \omega + 2, \lhd, \mathbb{O} \rangle$ with

$$\alpha \lhd \beta \quad \text{iff} \quad \begin{cases} \beta < \alpha & \text{and} \quad \beta, \alpha < \omega + 2 \\ \beta \leq \alpha & \text{and} \quad \beta < \alpha = \omega \end{cases}$$

For the algebra $\mathbb{O}$ of sets we take the smallest algebra of sets on that frame. This is an infinite algebra. We will approach its structure in stages. First of all, take the

FIGURE 3.1.



generated subframe of finite numbers and let $\mathbb{F}$ be the trace algebra on that set. We claim that $\mathbb{F}$ is nothing but the algebra of finite and cofinite sets. To show this, two things are required. We have to show that it is closed under all operations, and that each set is definable by a constant formula. The first is not so hard. The finite and cofinite sets are closed under intersection and complement. Furthermore, if $a \subseteq \omega$ is an arbitrary subset, let n be the largest number such that $[0, n] \subseteq a$; n exists iff $a \neq \omega$. Then $\blacksquare a = [0, n+1]$, as is readily checked. Thus, $\blacksquare a$ is finite whenever $a \neq \omega$. Moreover, $\blacksquare\omega = \omega$, which is cofinite. This shows the closure under the operations. Now, let us define the following constant formulae.

$$\begin{array}{lcl} \mathsf{f}(0) & := & \Box\bot \\ \mathsf{f}(n+1) & := & \Diamond\mathsf{f}(n). \wedge .\Box\neg\Diamond\mathsf{f}(n) \end{array}$$

It is straightforward to verify that $\mathsf{f}(n)$ can only be true at n, so that all singleton sets are 0–definable, that is, definable by means of a constant formula. The smallest boolean algebra containing them is the algebra of finite and confinite sets. Lets go one step further and take the subframe generated by ω. Here it turns out that the trace of $\mathbb{O}$ contains all finite sets which do not contain ω, and all infinite sets which do contain ω. For the extension of the formulae $\mathsf{f}(n)$ is still $\{n\}$, $n < \omega$. The set of these sets is closed under the operations, as is easily verified. Finally, we let us consider $\mathbb{O}$. Observe that the extension of the formula $\mathsf{f}(\omega+1)$ is exactly $\{\omega+1\}$ where

$$\mathsf{f}(\omega+1) := \neg\Diamond\Box\bot\ .$$

This means that the full algebra consists of all sets whose trace relative to the frame generated by ω is either finite and does not contain ω, or is infinite and contains ω.

We claim that $\mathsf{Th}\,\Omega_{\omega+1}$ is a subframe logic. To that end take an internal set g in that frame. If it is finite, it does not contain the point ω and the trace algebra is the powerset algebra. Therefore, the frame is isomorphic to a generated subframe. If g is infinite, however, it contains ω, and the subframe is isomorphic to either $\Omega_{\omega+1}$ or the subframe generated by ω. We conclude that $\mathsf{Th}\,\Omega_{\omega+1}$ is a subframe logic.

THEOREM 3.5.8 (Wolter). $\mathsf{Th}\,\Omega_{\omega+1}$ *is an incomplete subframe logic.*

PROOF. To begin, **G.3** is a subframe logic; it has the finite model property and is therefore complete. Now, take the logic Λ with the following axioms.

$$\Box(\Diamond p \to \Diamond(p \wedge \neg\Diamond p)), \qquad \Diamond p \wedge \Diamond q. \to .\Diamond(p \wedge \Diamond q) \vee \Diamond(q \wedge \Diamond p) \vee \Diamond(p \wedge q)\ .$$

Take a frame $\mathfrak{F}$ for Λ, and a point x. Then — by choice of the axioms — for every successor y of x the subframe generated by y satisfies **G.3**. $\Omega_{\omega+1}$ satisfies the axioms of Λ, since the subframe generated by ω satisfies **G.3**. Now consider the logic $\Lambda \oplus \varphi$ with

$$\varphi := \Diamond q \wedge \neg\Diamond(q \wedge \Box\neg q) \wedge \Diamond p. \to .\Diamond\Diamond p \ .$$

This formula implies in conjunction with the other axioms that when we have a violation of the **G**–axiom at a point x then x must have a reflexive successor. (This successor can of course be x.) $\varphi \in \mathsf{Th}\,\Omega_{\omega+1}$. Now, any Kripke–frame for $\Lambda \oplus \varphi$ must be a frame for **G.3**. For if we have a Kripke–frame $\mathfrak{f}$ for Λ and x a point then either the transit of x is a **G.3**–frame or only the transit without x is. In the latter case x has a reflexive successor; let it be y. Then since $\langle \mathfrak{f}, x\rangle \vDash \Box(\Diamond p \to \Diamond(p \wedge \neg\Diamond p))$ and $x \lhd y$ we have $\langle \mathfrak{f}, y\rangle \vDash \Diamond p \to \Diamond(p \wedge \neg\Diamond p)$. This enforces that the transit of y in $\mathfrak{f}$ is a frame for **G**. But this cannot be, since then we must have $y \ntriangleleft y$. Contradiction. The proof is now almost complete. First, we have

$$\Lambda \oplus \varphi \subseteq \mathsf{Th}\,\Omega_{\omega+1} \subseteq \mathbf{G.3}.$$

Equality of the last two cannot hold, because the formula $\Diamond p \wedge \neg\Diamond(p \wedge \neg\Diamond p))$ is satisfiable in $\Omega_{\omega+1}$. All logics in between $\Lambda \oplus \varphi$ and **G.3** have the same Kripke–frames. Hence, any such logic if not equal to **G.3** is incomplete. $\mathsf{Th}\,\Omega_{\omega+1}$ is such a logic. ◻

Now, even if subframe logics may be incomplete, in case of their completeness we can show that they are complete with respect to frames of size $\aleph_0$ if $\kappa < \aleph_1$ and κ if $\aleph_1 \leq \kappa$. This may not seem such an improvement. However, it is a priori not clear that complete logics are complete with respect to countable models even in the case $\kappa = 1$. Secondly, the proof method itself is well–worth remembering. It will be used in many variations throughout this book. For extensions of **K4** this theorem has first been proved in Kit Fine [**66**].

Theorem 3.5.9. ($\kappa < \aleph_0$.) *Let Λ be a subframe logic and suppose that Λ is complete with respect to Kripke–frames. Then Λ is complete with respect to Kripke–frames of cardinality $\leq \aleph_0$.*

Proof. Let $\neg\varphi \notin \Lambda$. Then there exists a Λ–model $\langle \mathfrak{f}, \beta, w_0\rangle \vDash \varphi$, $\mathfrak{f}$ a Kripke–frame. Put $S_0 := \{w_0\}$; let $\mathfrak{s}_0$ be the subframe of $\mathfrak{f}$ based on S_0, and let $\gamma_0(p) := \beta(p) \cap S_0$. Inductively we define sets S_n; given S_n, $\mathfrak{s}_n$ is the subframe based on S_n and $\gamma_n(p) := \beta(p) \cap S_n$. The construction of the S_n is as follows. Suppose that there exists a $\Box_j\psi \in sf(\varphi)$ and a $x \in S_n$ such that $\langle \mathfrak{f}, \beta, x\rangle \vDash \neg\Box_j\psi$ but $\langle \mathfrak{s}_n, \gamma_n, x\rangle \vDash \Box_j\psi$. Then let $y := y(x, \Box_j\psi) \in f$ be a point such that $x \lhd_j y$ and $\langle \mathfrak{f}, \beta, y\rangle \vDash \neg\psi$. Then let

$$S_{n+1} := S_n \cup \{y(x, \Box_j\psi) : x \in S_n, \Box_j\psi \in sf(\varphi), \langle \mathfrak{s}_n, \gamma_n, x\rangle \vDash \Box_j\psi\} \ .$$

Finally, we put $g := \bigcup_{i\in\omega} S_i$, $\delta(p) := \beta(p) \cap g$. g is finite if $S_{n+1} = S_n$ for some n, else g is countably infinite. We claim that for every $x \in g$ and every $\psi \in sf(\varphi)$

$$\langle \mathfrak{g}, \delta, x\rangle \vDash \psi \qquad \Leftrightarrow \qquad \langle \mathfrak{f}, \beta, x\rangle \vDash \psi$$

The proof is by induction on ψ. For variables this is immediate; for $\psi = \psi_1 \wedge \psi_2$ and $\psi = \neg\psi_1$ this is also immediate. Now assume finally that $\psi = \Box_j\tau$. Suppose that $\langle \mathfrak{g}, \delta, x\rangle \nvDash \Box_j \tau$. Let n be the smallest number such that $x \in S_n$. Then, by construction, there exists a point $y \in S_{n+1}$ such that $x \lhd_j y$ and $\langle \mathfrak{f}, \beta, y\rangle \vDash \neg\tau$. By induction hypothesis, $\langle \mathfrak{g}, \delta, y\rangle \vDash \neg\tau$. Since $y \in g$ we have $\langle \mathfrak{g}, \delta, x\rangle \vDash \neg\Box_j \tau$. The converse direction is straightforward. □

We give an example to show that completeness is necessary for the proof method to work properly. Take the logic $\mathsf{Th}\,\Omega_{\omega+1}$ defined above. Let $\beta(p)$ be a set containing the point ω. Then $\langle \Omega_{\omega+1}, \beta, \omega\rangle \vDash p$. It is clear that $S_1 = S_0$ in the construction. However, the frame based on a single reflexive point is not a frame for the logic. This shows that in the incomplete case we cannot get rid of the restriction that the subframe is based on an internal set. Since internal sets need not be countable, the proof methods fails in this case as it stands.

Let us define the **Kuznetsov–index**, $Kz(\Lambda)$, of a complete logic Λ to be the supremum of the cardinalities of minimal frames refuting nontheorems of Λ.

$$Kz(\Lambda) := sup_{\varphi\notin\Lambda} inf\{\sharp f : \mathfrak{f} \nvDash \varphi\}$$

The Kuznetsov–index is finite if Λ is tabular, and $\geq \aleph_0$ otherwise. We have shown that if Λ is a complete subframe logic, the Kuznetsov–index does not exceed $\aleph_0$. In general, following Chagrov and Zakharyaschev the complexity of a logic is a function f_Λ such that $f_\Lambda(n)$ is the supremum of the cardinalities of minimal models refuting nontheorems of length n. If Λ has the finite model property and κ is finite then $f_\Lambda(n)$ is finite for every n. Clearly, the Kuznetsov–index is the supremum of all $f_\Lambda(n)$. It is shown in Chagrov and Zakharyaschev [**43**] that there exists a logic with Kuznetsov–Index $\beth_\omega$, where $\beth$ is the so–called *beth–function*; roughly, $\beth_\lambda$ is the λ–fold iteration of the exponentiation function. In Kracht [**129**], for each countable ordinal λ a logic with Kuznetsov–Index $\beth_\lambda$ is constructed. Furthermore, it is shown that there exists a logic whose Kuznetsov–Index is the least strongly inaccessible cardinal.

Exercise 99. Show that **S5** is a subframe logic.

Exercise 100. Show that **G.3** is a subframe logic.

Exercise 101. Show that $\Lambda \oplus \varphi = \mathsf{Th}\,\Omega_{\omega+1}$ and that they are immediately below **G.3**, i. e. there is no logic Θ with $\Lambda \oplus \varphi \subsetneq \Theta \subsetneq$ **G.3**.

***Exercise 102.** Show that there is a descending chain of $\aleph_0$ many incomplete subframe logics. *Hint.* Extend the construction of $\Omega_{\omega+1}$ above to frames Ω_α based on ordinal numbers $\alpha \leq \omega \times \omega$. You have to put $\beta \lhd \gamma$ iff $\beta = \omega \times k + m$ and (a.) $\gamma = \omega \times k + n$ and $m < n$ or (b.) $\gamma = \omega \times (k+1)$. Let the algebra of sets be the minimal algebra. Now show that all the logics of Ω_α are different.

Exercise 103. Let $\kappa \geq \aleph_0$. Show that if a subframe logic is complete with respect to

Kripke–frames, it is complete with respect to Kripke–frames of cardinality $\leq \kappa$.

Exercise 104. Assume that $\kappa < \aleph_1$. Show that a subframe logic is globally complete with respect to Kripke–frames if it is globally complete with respect to Kripke–frames which are finite or countably infinite. Show the same for $\aleph_1$–compactness.

Exercise 105. Let $\mathfrak{F}$ be a frame, and $S = Tr_{\mathfrak{F}}(S)$. Show that $a \mapsto a \cap S$ is a homomorphism of $\langle \mathbb{F}, 1, -, \cap, \langle \blacksquare_j : j < \kappa \rangle \rangle$ onto the trace algebra over S. *Remark.* This shows that in contrast to arbitrary subframes, the generated subframes need not be based on internal sets.

Exercise 106. Let κ be countable and Λ a canonical κ–modal logic. Show that the Kuznetsov–index of Λ is $\leq 2^{\aleph_0}$.

3.6. Constructive Reduction

In Section 3.1 we have proved the global finite model property for the basic logic $\mathbf{K}_\kappa$. We will now use this proof to obtain a number of other results on (global) finite model property using a technique which we call *constructive reduction*. This technique is syntactic. The standard situation is that certain properties have been established for a logic Λ, for example $\mathbf{K}$, and that we consider an extension $\Lambda \oplus A$ for some set of axioms A. It would be rather unfortunate not to be able to use knowledge about Λ for $\Lambda \oplus A$. However, in the overwhelming number of cases nothing can be inferred for $\Lambda \oplus A$ from Λ. On the other hand, many standard systems are an exception to this. Before we investigate the formal background of this method, let us see some nontrivial applications.

Theorem 3.6.1. *Let Λ have the global finite model property and let χ be a constant formula. Then $\Lambda \oplus \chi$ has the global finite model property as well.*

Proof. We show that

$$\varphi \Vdash_{\Lambda \oplus \chi} \psi \quad \Leftrightarrow \quad \varphi; \chi \Vdash_\Lambda \psi \ .$$

From right to left is trivial. From left to right, take a proof of ψ from φ in $\Lambda \oplus \chi$. We know that we can move substitutions at the beginning of the proof. Now χ is constant, so we cannot derive anything but χ from χ using substitutions. Hence the proof is a proof in Λ of ψ from φ together with χ using (mn.) and (mp.), as required. ◻

Theorem 3.6.2. ***K4** has the global finite model property.*

Proof. Let Δ be a set of formulae. Put

$$X_4(\Delta) := \{\Box\chi \to \Box\Box\chi : \Box\chi \in sf[\Delta]\}.$$

We show that

$$\varphi \Vdash_{\mathbf{K4}} \psi \quad \Leftrightarrow \quad \varphi; X_4(\{\varphi, \psi\}) \Vdash_{\mathbf{K}} \psi$$

From right to left is straightforward. For the direction from left to right, assume $\varphi; X_4(\{\varphi, \psi\}) \Vdash_{\mathbf{K}} \psi$ is not the case. Then there exists a finite model $\langle \mathfrak{f}, \beta, x\rangle$ such that $\langle \mathfrak{f}, \beta\rangle \vDash \varphi; X_4(\{\varphi, \psi\})$ but $\langle \mathfrak{f}, \beta, x\rangle \nvDash \psi$. $\mathfrak{f} = \langle f, \lhd\rangle$. Let $\blacktriangleleft$ be the transitive closure of $\lhd$. We show that for all subformulae χ of φ or ψ and all worlds y

$$(\dagger) \qquad \langle f, \blacktriangleleft, \beta, y\rangle \vDash \chi \quad \Leftrightarrow \quad \langle f, \lhd, \beta, y\rangle \vDash \chi$$

This then establishes $\langle f, \blacktriangleleft, \beta\rangle \vDash \varphi$ and $\langle f, \blacktriangleleft, \beta, x\rangle \vDash \neg\psi$. $\langle f, \blacktriangleleft\rangle$ is transitive; therefore $\langle f, \blacktriangleleft\rangle \vDash \mathbf{K4}$. We show $(\dagger)$ by induction on χ. For variables there is nothing to show. The steps for $\neg$ and $\wedge$ are straightforward. Now let $\chi = \Box\chi'$. Assume $\langle f, \blacktriangleleft, \beta, y\rangle \nvDash \Box\chi'$. Then there is a z such that $y \blacktriangleleft z$ and $\langle f, \blacktriangleleft, \beta, z\rangle \vDash \neg\chi'$. By induction hypothesis, $\langle f, \lhd, \beta, z\rangle \vDash \neg\chi'$. By definition of $\blacktriangleleft$ there is a chain $y = y_0 \lhd y_1 \lhd \ldots \lhd y_n = z$. Now $\langle f, \lhd, \beta, y_{n-1}\rangle \vDash \neg\Box\chi'$. If $n - 1 > 0$ then $\langle f, \lhd, \beta, y_{n-2}\rangle \vDash \neg\Box\Box\chi'$. Since $\Box\chi' \to \Box\Box\chi' \in X_4(\{\varphi, \psi\})$ and $\langle f, \lhd, \beta, y_{n-2}\rangle \vDash X_4(\{\varphi, \psi\})$ we must have $\langle f, \lhd, \beta, y_{n-2}\rangle \vDash \neg\Box\chi'$. Iterating this argument we get $\langle f, \lhd, \beta, y\rangle \vDash \neg\Box\chi'$. So, $\langle f, \lhd, \beta, y\rangle \nvDash \Box\chi'$. Clearly, if $\langle f, \lhd, \beta, y\rangle \nvDash \Box\chi'$ then $\langle f, \blacktriangleleft, \beta, y\rangle \nvDash \Box\chi'$, since $\lhd \subseteq \blacktriangleleft$. ∎

A note on the reduction sets. Since $\Diamond$ is not a primitive symbol, some care is needed in the formulation of the reduction sets. The following definition of a reduction set for **K4** will not do.

$$Y_4(\Delta) := \{\Diamond\Diamond\chi \to \Diamond\chi : \Diamond\chi \in sf[\Delta]\}$$

(Here, $\Diamond$ abbreviates $\neg\Box\neg$.) Take for example the set $\Delta = \{\neg\Box\Box p, \Box p\}$. It is **K4**–consistent. Yet, there is no subformula of a formula of Δ that matches $\Diamond\chi$ for some χ. Hence, $Y_4(\Delta) = \varnothing$. But Δ is clearly **K**–consistent. So, this definition of the reduction sets does not work. The reader may pause to reflect on why the chosen reduction sets actually avoid this problem.

Theorem 3.6.3. *The basic tense logic $\boldsymbol{K.t}$ has the global finite model property.*

Proof. Let

$$\begin{aligned} X_t(\Delta) := & \quad \{\neg\chi \to \Box_0\neg\Box_1\chi : \Box_1\chi \in sf[\Delta]\} \\ & \cup \quad \{\neg\chi \to \Box_1\neg\Box_0\chi : \Box_0\chi \in sf[\Delta]\} . \end{aligned}$$

We show that

$$(\ddagger) \qquad \varphi \Vdash_{\mathbf{K.t}} \psi \quad \Leftrightarrow \quad \varphi; X_t(\{\varphi, \psi\}) \Vdash_{\mathbf{K}} \psi$$

Proceed as in the previous proof. Let $\mathfrak{M} = \langle \mathfrak{f}, \beta, w_0\rangle$ be a local model where $\mathfrak{f} = \langle f, \lhd_0, \lhd_1\rangle$ is a finite $\mathbf{K}_2$–Kripke–frame such that $\langle \mathfrak{f}, \beta\rangle \vDash \varphi; X_t(\{\varphi, \psi\})$ and $\langle \mathfrak{f}, \beta, w_0\rangle \vDash \neg\psi$. Let $\blacktriangleleft_0 := \lhd_0 \cup \lhd_1^{\smile}$ and $\blacktriangleleft_1 := \lhd_1 \cup \lhd_0^{\smile}$. Then the frame $\langle f, \blacktriangleleft_0, \blacktriangleleft_1\rangle$ is a tense frame, for $\blacktriangleleft_0^{\smile} = (\lhd_0 \cup \lhd_1^{\smile})^{\smile} = \lhd_0^{\smile} \cup \lhd_1 = \blacktriangleleft_1$. For all $\chi \in sf(\varphi) \cup sf(\psi)$ we have

$$(\dagger) \qquad \langle f, \blacktriangleleft_0, \blacktriangleleft_1, \beta, y\rangle \vDash \chi \quad \Leftrightarrow \quad \langle f, \lhd_0, \lhd_1, \beta, y\rangle \vDash \chi .$$

This is clear for variables; the steps for $\neg$ and $\wedge$ are straightforward. Now let $\chi = \Box_0\tau$. From left to right is clear. Now the direction from right to left. Assume that $\langle f, \blacktriangleleft_0, \blacktriangleleft_1, \beta, y\rangle \nvDash \Box_0\tau$. Then there is a w such that $y \blacktriangleleft_0 w$ and $\langle f, \blacktriangleleft_0, \blacktriangleleft_1, \beta, w\rangle \vDash \neg\tau$. By induction hypothesis, $\langle f, \lhd_0, \lhd_1, \beta, w\rangle \vDash \neg\tau$. If $y \lhd_0 w$, we are done; for then

$\langle f, \lhd_0, \lhd_1, \beta, y\rangle \nvDash \Box_0\tau\,(= \chi)$. Otherwise $w \lhd_1 y$. Now, $\langle f, \lhd_0, \lhd_1, \beta, w\rangle \vDash \Box_1\neg\Box_0\tau$, since $\langle f, \lhd_0, \lhd_1, \beta, w\rangle \vDash X_t(\{\varphi, \psi\})$. Thus $\langle f, \lhd_0, \lhd_1, y\rangle \vDash \neg\Box_0\tau$. So, $\langle f, \lhd_0, \lhd_1, \beta, y\rangle \nvDash \Box_0\tau$. The step $\chi = \Box_1\tau$ is analogous. □

Informally, we say that a property $\mathfrak{P}$ *pushes up* or *can be pushed up* from Λ to $\Lambda \oplus A$ if we can prove that $\Lambda \oplus A$ has $\mathfrak{P}$ on the condition that Λ has $\mathfrak{P}$. Properties that can be dealt with in this way are among other decidability, finite model property, completeness and interpolation. The fundamental property in this connection is decidability. Notice that given Λ, A, Δ and φ, there is a set $Y \subseteq \boxtimes^\omega A^s$ such that

$$\Delta \vdash_{\Lambda\oplus A} \varphi \qquad \Leftrightarrow \qquad Y; \Delta \vdash_\Lambda \varphi$$

(Recall that A^s denotes the closure of A under substitution.) We call Y a **local reduction set** for Δ and φ. Moreover, if Δ is finite then Y can be chosen finite. A **local reduction function for $\Lambda \oplus A$ with respect to Λ** is a function $X : \wp(\mathcal{P}_\kappa) \to \wp(\mathcal{P}_\kappa)$ such that (i) $X(\Delta \cup \{\varphi\})$ is a local reduction set for Δ and φ and (ii) $X(\Delta)$ is finite whenever Δ is finite. We will write $X(\varphi)$ rather than $X(\{\varphi\})$. Note that there are two cases. (a.) $\varphi \notin \Lambda \oplus A$. Then $Y = \varnothing$ is a reduction set, (b.) $\varphi \in \Lambda \oplus A$. Then there is a proof of φ from Λ in $\boxtimes^\omega A^s$ using only modus ponens. Obviously, this proof uses only a finite subset of $\boxtimes^\omega A^s$, and we take Y to be this subset. Similarly, there is a finite set $Y \subseteq A^s$ such that

$$\Delta \Vdash_{\Lambda\oplus A} \varphi \qquad \Leftrightarrow \qquad \Delta; Y \Vdash_\Lambda \varphi$$

Such a Y is called a **global reduction set** for Δ and φ. A **global reduction function for $\Lambda \oplus A$ with respect to Λ** is a function $X : \wp(\mathcal{P}_\kappa) \to \wp(\mathcal{P}_\kappa)$ such that (i) $X(\Delta \cup \{\varphi\})$ is a global reduction set for Δ and φ and (ii) $X(\Delta)$ is finite whenever Δ is finite. We note the following properties of reduction sets. The proof is left as an exercise.

Proposition 3.6.4. *(1.) There exists a reduction function X for $\Lambda \oplus A$ with respect to Λ such that (a.) $var[X(\Delta)] \subseteq var[\Delta]$, (b.) $X(\Delta) = \bigcup\langle X(\Delta_0) : \Delta_0 \subseteq \Delta, \Delta_0$ finite$\rangle$ (2.) Let $X, Y : \wp(\mathcal{P}_\kappa) \to \wp(\mathcal{P}_\kappa)$ be functions mapping finite sets to finite sets such that $X(\Delta) \subseteq Y(\Delta)$ for all Δ. Then if X is a (global/local) reduction function for $\Lambda \oplus A$ with respect to Λ then so is Y.*

As (1b.) shows, we may always assume that the reduction function is determined by its values on finite sets. This means that we may actually restrict our attention to functions from finite subsets of $\mathcal{P}_\kappa$ to finite subsets of $\mathcal{P}_\kappa$. If $\Lambda \oplus A$ is decidable and A enumerable, a local reduction set can always be constructed. For suppose that ψ is given. Then start enumerating the proofs of $\Lambda \oplus A$ in which (sub.) is applied before (mn.) and (mn.) before (mp.); in parallel, enumerate the nontheorems of $\Lambda \oplus A$. If φ is a theorem, it will occur at end of a proof Π. The reduction set will then consist in all formulae to which only (mp.) is applied in Π. (This is more than necessary, but certainly a sufficient set.) If φ is a nontheorem, then the empty set is a reduction set for φ. Similarly, if $\Lambda \oplus A$ is globally decidable, it allows the construction of global reduction sets. Conversely, suppose we are able to produce for each φ a local reduction set. Then decidability can be pushed up. To see the first, assume

Λ is decidable and let φ be given. Construct $X(\{\varphi\})$. Since Λ is decidable, we can decide $X(\{\varphi\}) \vdash_{\Lambda} \varphi$, which by definition of the reduction sets is nothing but $\vdash_{\Lambda\oplus A} \varphi$. Similarly for global decidability. For the purpose of the next theorem, a *computable function* from $\wp(\mathcal{P}_\kappa)$ to $\wp(\mathcal{P}_\kappa)$ is a function f such that (i.) $f(\Delta) = \bigcup\langle f(\Delta_0) : \Delta_0 \subseteq \Delta, \Delta_0 \text{ finite}\rangle$ and (ii.) there exists an algorithm computing $f(\Delta)$ for any given finite Δ.

Definition 3.6.5. *A logic $\Lambda \oplus A$ is said to be **locally constructively reducible** to Λ if there is a computable local reduction function for $\Lambda \oplus A$ and Λ. $\Lambda \oplus A$ is said to be **globally constructively reducible** to Λ if there is a computable global reduction function.*

Theorem 3.6.6. *Suppose $\Lambda \oplus A$ is globally (locally) constructively reducible to Λ. Then $\Lambda \oplus A$ is globally (locally) decidable if Λ is.*

In many cases, it is possible to show that other properties can be pushed up as well. For example, for transitivity we have established that $X_4 : \Delta \mapsto \{\Box\chi \to \Box\Box\chi : \Box\chi \in sf[\Delta]\}$ is a global reduction function, and that furthermore any frame satisfying both φ locally and $X_4(\{\varphi\})$ globally satisfies φ also when the relation $\lhd$ is replaced by the transitive closure. It then is a **K4**–frame and a finite transitive model for φ. So if Λ is a monomodal logic whose frames are closed under passing from $\lhd$ to its transitive closure then we can push up the global finite model property from Λ to Λ.**4**. We will show here that many of the standard systems mentioned in Section 2.5 have global finite model property. The following are global reduction functions.

$$\begin{array}{lcl}
X_4(\Delta) & := & \{\Box\chi \to \Box\Box\chi : \Box\chi \in sf[\Delta]\} \\
X_T(\Delta) & := & \{\Box\chi \to \chi : \Box\chi \in sf[\Delta]\} \\
X_B(\Delta) & := & \{\neg\chi \to \Box\neg\Box\chi : \Box\chi \in sf[\Delta]\} \\
X_G(\Delta) & := & \{\neg\Box\chi \to \neg\Box(\chi \vee \neg\Box\chi) : \Box\chi \in sf[\Delta]\} \\
X_{Grz}(\Delta) & := & \{\neg\Box\chi \to \neg\Box(\chi \vee \neg\Box(\chi \to \Box\chi)) : \Box\chi \in sf[\Delta]\} \\
X_{alt1}(\Delta) & := & \{\neg\Box\chi \to \Box\neg\chi : \Box\chi \in sf[\Delta]\}
\end{array}$$

The reader may check that the formulae are indeed axioms. The reduction of Λ.**4** to Λ has been proved for those logics whose class of frames is closed under passing from $\lhd$ to the transitive closure. Now, for reflexivity we claim that if the class of frames for Λ is closed under passing from $\lhd$ to its reflexive closure, denoted by $\lhd^\bullet$, then the above function achieves global reduction. Namely, suppose that we have a Λ–frame $\mathfrak{f}$ and $\langle\mathfrak{f},\beta\rangle \vDash X_T(\{\psi;\varphi\})$. Let $\mathfrak{f}^\bullet$ be obtained by changing $\lhd$ to its reflexive closure. By definition, $\mathfrak{f}^\bullet \vDash \Lambda$ and so $\mathfrak{f}^\bullet \vDash \Lambda.\mathbf{T}$. By induction on the set $sf(\varphi)$ we show that for all w in the transit of x

$$\langle\mathfrak{f}^\bullet,\beta,w\rangle \vDash \chi \qquad \Leftrightarrow \qquad \langle\mathfrak{f},\beta,w\rangle \vDash \chi$$

The only critical step is $\chi = \Box\tau$. From left to right this follows from the fact that if $x \lhd y$ then also $x \lhd^\bullet y$. For the other direction, assume we have $\langle\mathfrak{f}^\bullet,\beta,w\rangle \nvDash \Box\tau$. Then there is a v such that $w \lhd^\bullet v$ and $\langle\mathfrak{f}^\bullet,\beta,v\rangle \vDash \neg\tau$. If $v \neq w$, we are done for then also $w \lhd v$. So assume the only choice for v is $v = w$ and that $w \ntriangleleft w$. Then we have

$\langle \mathfrak{f}, \beta, w \rangle \vDash \Box\tau$. But $\langle \mathfrak{f}, \beta, w \rangle \vDash \Box\tau \to \tau$, by choice of the reduction function. Hence $\langle \mathfrak{f}, \beta, w \rangle \vDash \tau$, and so $\langle \mathfrak{f}^{\bullet}, \beta, w \rangle \vDash \tau$, a contradiction. So there always is a successor $v \neq w$, and it is safe to close $\lhd$ reflexively.

The proof for **B** is the same as for tense logic, so we will omit it here. We can use this technique iteratively to show that a logic defined by a mixture of reflexivity, transitivity or symmetry axioms has the finite model property. However, since each particular pushing up has its preconditions, some care is called for. The idea is always the following. Assume Λ has the global finite model property; then constructively reduce $\Lambda \oplus \varphi$ to Λ. This works if we can be sure that the procedure that turns a frame $\mathfrak{f}$ for $\mathbf{K}_\kappa$ into a frame $\mathfrak{f}^\varphi$ for $\mathbf{K}_\kappa \oplus \varphi$ also turns a frame for Λ into a frame for $\Lambda \oplus \varphi$. The proof of the following theorem illustrates this.

Theorem 3.6.7. *Let Λ be a finitely axiomatizable polymodal logic based on postulates of reflexivity, transitivity and symmetry for its operators, in any combination. Then Λ has the global finite model property.*

Proof. Let $\Lambda = \mathbf{K}_\kappa \oplus R \oplus S \oplus T$, where R is a set of reflexivity postulates, S a set of symmetry postulates and T a set of transitivity postulates. First of all $\mathbf{K}_\kappa$ has the global finite model property, so we need to consider finite Kripke–frames only. We will start with $\mathbf{K}_\kappa$ and add the postulates one by one. First, we add all $\varphi \in R$. The map $\mathfrak{f} \mapsto \mathfrak{f}^\varphi$ is defined by taking the reflexive closure of the relation $\lhd_j$ for some j. Since each of the reflexivity postulates concerns a different operator, it does not matter in which order we add the axioms. In the end we obtain a frame $\mathfrak{G}$ satisfying all $\varphi \in R$. Next we turn to S. The map $\mathfrak{f} \mapsto \mathfrak{f}^\varphi$, $\varphi \in S$, is now the map which turns $\lhd_i$ into its symmetric closure. Since the symmetric closure of a reflexive relation is again reflexive, $\mathfrak{f}^\varphi$ satisfies all postulates of R. Moreover, the symmetric closure of one relation does not interfere with any other relation, so $\mathfrak{f}^\varphi$ satisfies all symmetry postulates that $\mathfrak{f}$ satisfies. Thus we can construct a frame for $R \cup S$. Now we turn to T; the transitive closure of a reflexive relation is reflexive, and the transitive closure of a symmetric relation is again symmetric. ◻

There remain the sets for **G**, **Grz** and $\mathbf{alt}_1$. Now, both **G** and **Grz** are transitive logics. (This is the content of some exercises in Section 2.5.) We will now show that the functions above establish a reduction from **G** to **K4** and a reduction from **Grz** to **S4**. The first of these has been shown by Philippe Balbiani and Andreas Herzig in [**3**].

Theorem 3.6.8. ***G** has the global finite model property.*

Proof. We establish that the reduction function is a reduction function from the logic to **K4**, which has global finite model property by Theorem 3.6.2. Moreover, **K4** is transitive, so we only need to consider reductions where the antecedent is identical to $\top$. Thus let $\mathfrak{f}$ be a finite transitive frame and

$$\langle \mathfrak{f}, \beta, w_0 \rangle \vDash \varphi; \Box^{\leq 1}\{\neg\Box\chi \to \neg\Box(\chi \vee \neg\Box\chi) : \Box\chi \in sf(\varphi)\} .$$

Now, pick points from the frame as follows. Put $S_0 := \{w_0\}$. The sets S_n are now defined inductively. Let $x \in S_n$ and $\Box\chi \in sf(\varphi)$ such that $\langle \mathfrak{f}, \beta, x\rangle \nvDash \Box\chi$, and no successor of x in $S_n - \{w_0\}$ exists such that $\langle \mathfrak{f}, \beta, y\rangle \vDash \neg\chi$. Then, by assumption on the reduction function, $\langle \mathfrak{f}, \beta, x\rangle \vDash \Diamond(\neg\chi \wedge \Box\chi)$. Hence there exists a $\widehat{x}$ such that $\widehat{x} \vDash \neg\chi; \Box\chi$. (Moreover, if $x = w_0$, then $\widehat{x} \neq w_0$. For $\widehat{x}$ is irreflexive, and so $w_0 \lhd \widehat{x}$ implies $w_0 \neq \widehat{x}$.) It follows that $\widehat{x}$ is irreflexive. Put $S_{n+1} := S_n \cup \{\widehat{x}\}$. The selection ends after some steps, since f is finite. Call the resulting set g. Let $\lhd^g := \lhd \cap (g \times g) - \{\langle w_0, w_0\rangle\}$. Then put $\mathfrak{g} := \langle g, \lhd^g\rangle$. (Alternatively, we might simply take $\mathfrak{g}$ to be the subframe consisting of w_0 and all irreflexive points from $\mathfrak{f}$, with the transition $w_0 \to w_0$ being removed.) $\mathfrak{g}$ is transitive and irreflexive, hence it is a frame for **G**. Put $\gamma(p) := \beta(p) \cap g$. We now show that for every subformula ψ of φ and every point $y \in g$, $\langle \mathfrak{g}, \gamma, y\rangle \vDash \psi$ iff $\langle \mathfrak{f}, \beta, y\rangle \vDash \psi$. This holds for variables by construction, and the steps $\neg$, $\wedge$ are straightforward. Now let $\langle \mathfrak{g}, \gamma, y\rangle \nvDash \Box\chi$. Then also $\langle \mathfrak{f}, \beta, y\rangle \nvDash \Box\chi$. Conversely, suppose that $\langle \mathfrak{f}, \beta, y\rangle \nvDash \Box\chi$, for some $\Box\chi \in sf(\varphi)$. Then also $\langle \mathfrak{g}, \beta, y\rangle \nvDash \Box\chi$, since a successor z for y has been chosen such that $\langle \mathfrak{f}, \beta, z\rangle \vDash \chi; \Box\neg\chi$. By induction hypothesis, $\langle \mathfrak{g}, \gamma, z\rangle \vDash \chi$. Moreover, $y \lhd^g z$. For if $y \neq w_0$ this holds by definition of $\lhd^g$. For $y = w_0$ observe that either $w_0 \lhd^f w_0$, and then $z \neq w_0$, since $z \ntriangleleft z$. From this follows $w_0 \lhd^g z$. Or else, $w_0 \ntriangleleft^f w_0$, in which case $w_0 \lhd^g z$ anyway. And so $\langle \mathfrak{g}, \gamma, y\rangle \nvDash \Box\chi$, as required. ◻

Theorem 3.6.9. *Grz has the global finite model property.*

Proof. As in the previous proof, this time reducing to **S4**. By Theorem 3.6.7, **S4** has the (global) finite model property. Let $\langle \mathfrak{f}, \beta, w_0\rangle$ a finite **S4**–model such that

$$\langle \mathfrak{f}, \beta, w_0\rangle \vDash \varphi; \Box^{\leq 1}\{\neg\Box\chi \to \neg\Box(\chi \vee \neg\Box(\chi \to \Box\chi)) : \Box\chi \in sf(\varphi)\} .$$

We select a subset g of f in the following way. We start with $S_0 := \{w_0\}$. S_{n+1} is defined inductively as follows. If $x \in S_n$ and $\langle \mathfrak{f}, \beta, x\rangle \vDash \neg\Box\chi$, but no y exists in S_n such that $x \lhd y$ and $\langle \mathfrak{f}, \beta, y\rangle \vDash \neg\chi$, then we choose a successor $\widehat{y}$ of x as follows. By choice of the reduction function there is a successor y of x such that $y \vDash \neg\chi; \Box(\chi \to \Box\chi)$. Therefore (i) the entire cluster $C(y)$ satisfies $\neg\chi$, (ii) no point in a cluster succeeding $C(y)$ and different from $C(y)$ satisfies χ. Then $S_{n+1} := S_n \cup \{y\}$. This procedure comes to a halt after finitely many steps. The resulting set is called g, and the subframe based on it $\mathfrak{g}$. It is directly verified that g contains at most one point from each cluster. (Moreover, the selection procedure produces a model whose depth is bounded by the number of formulae in $sf(\varphi)$ of the form $\Box\chi$ as can easily be seen.) So all clusters have size 1. $\mathfrak{g}$ is reflexive and transitive, being a subframe of $\mathfrak{f}$. So, $\mathfrak{g}$ is a **Grz**–frame. Let $\gamma(p) := \beta(p) \cap g$. It is shown as in the previous proof that for every subformula χ of φ and every $x \in g$, $\langle \mathfrak{g}, \gamma, x\rangle \vDash \chi$ exactly when $\langle \mathfrak{f}, \beta, x\rangle \vDash \chi$. In particular, $\langle \mathfrak{g}, \gamma, w_0\rangle \vDash \varphi$. This concludes the proof. ◻

Obviously, to have local reduction functions is much stronger than to have global reduction functions. Yet, for practical purposes it is enough to compute global reduction functions, since most standard systems are globally decidable. Thus, the additional gain in establishing a local reduction, which is possible in many cases, is rather marginal. Moreover, as we will see, many properties can be pushed up even when we have global reduction functions and no local functions. We shall end the section by a few remarks of complexity. We shall state here without proof that the size of the reduction sets of this section is quadratic in the size of the initial set (if the size is simply the sum of the lengths of the formulae contained in it). This is actually easy to verify. However, if we change to the packed representation (see the exercises of Section 1.8) then the increase is only linear. To verify this is left as an exercise. It follows that the logics discussed in this section are globally EXPTIME, since **K** is. One has to take note here that the typical complexity measures are established with respect to the length of the set, not with respect to the length of the packed representation, which can in extreme cases be exponentially smaller. Yet, they can typically be redone with respect to the length of the packed representation. To verify that **K** is globally EXPTIME even with respect to the packed representation, tableaux can be used. Moreover, using tableaux one can show that **K4** is globally PSPACE from which the same follows for the systems extending **K4**. Unfortunately, reduction sets do not allow to show that **K4** is in PSPACE.

Exercise 107. Show Proposition 3.6.4.

Exercise 108. Show that the symmetric closure of a transitive relation does not need not be transitive again.

Exercise 109. In the next three exercises we will show a rather general theorem on reduction of Λ.**G** and Λ.**Grz** to Λ for logics containing **K4**. Let $\Lambda \supseteq$ **K4** have finite model property. Say that a Λ is a *cofinal subframe logic* if it is closed under taking away from a finite frame any set of points which is not final. Here a point x is *final* if for all y $x \lhd y$ implies $y \lhd x$. Now let $\langle \mathfrak{f}, \beta \rangle$ and φ be given, and $\mathfrak{f}$ be finite. Say that x is φ–**maximal** if for some subformula χ, χ is satisfied at x and whenever y satisfies χ and $x \lhd y$, then also $y \lhd x$. Show now that if $\langle \mathfrak{f}, \beta, x \rangle \vDash \varphi$, and if we take the subframe $\mathfrak{g}$ of all φ–maximal points, then $\langle \mathfrak{g}, \beta, y \rangle \vDash \varphi$ for some y. Show that every final point of $\mathfrak{f}$ is in $\mathfrak{g}$. Thus, if Λ is a cofinal subframe logic, and $\mathfrak{f}$ is a finite frame for Λ, so is $\mathfrak{g}$. Give the global reduction sets!

Exercise 110. (Continuing the previous exercise.) Now show that the reduction function given above for **G** establishes that if Λ is a cofinal subframe logic, then there exists a global reduction function for Λ.**G** to Λ.

Exercise 111. (Continuing the previous exercise.) Show that the reduction function for **Grz** establishes that there is a reduction function any Λ.**Grz** to Λ, provided that Λ is a subframe logic.

Exercise 112. Using the Lemma 3.1.9, produce local reduction sets for the logics for which global reduction sets have been given.

Exercise 113. Let X be any of the reduction functions of this section. Show that there is a constant c_X such that for every set Δ: $|X(\Delta)^\spadesuit| \leq c|\Delta^\spadesuit|$, where $\Delta^\spadesuit$ is the packed representation of Δ.

3.7. Interpolation and Beth Theorems

Recall from Section 1.6 the definition of *interpolation*. Interpolation is defined with respect to the consequence relation. Since a modal logic admits several consequence relations, we have several notions of interpolation, in particular *global* and *local* interpolation.

Definition 3.7.1. *A modal logic Λ has* **local interpolation** *if for every pair φ and ψ of formulae with $\varphi \vdash_\Lambda \psi$ there is a χ such that $var(\chi) \subseteq var(\varphi) \cap var(\psi)$ and $\varphi \vdash_\Lambda \chi$ as well as $\chi \vdash_\Lambda \psi$. Λ has* **global interpolation** *if for every pair φ, ψ of formulae with $\varphi \Vdash_\Lambda \psi$ there is a χ such that $var(\chi) \subseteq var(\varphi) \cap var(\psi)$ and $\varphi \Vdash_\Lambda \chi$ as well as $\chi \Vdash_\Lambda \psi$.*

These definitions are taken from [**151**], though the terminology used here is more systematic. Since we have a deduction theorem for local deducibility, we can reformulate local interpolation in such a way that it depends only on the set of theorems. Λ has the **Craig Interpolation Property** if whenever $\varphi \to \psi \in \Lambda$ there exists a χ which is based on the common variables of φ and ψ such that $\varphi \to \chi; \chi \to \psi \in \Lambda$. A logic has the Craig Interpolation Property iff it has local interpolation.

Proposition 3.7.2. *If Λ has local interpolation it also has global interpolation.*

Proof. Suppose that Λ has local interpolation. Let $\varphi \Vdash_\Lambda \psi$. Then for some compound modality $\boxplus$ we have $\boxplus\varphi \vdash_\Lambda \psi$. Whence by local interpolation there is a χ with $var(\chi) \subseteq var(\varphi) \cap var(\psi)$ such that $\boxplus\varphi \vdash_\Lambda \chi$ and $\chi \vdash_\Lambda \psi$. Hence $\varphi \Vdash_\Lambda \chi$ as well as $\chi \Vdash_\Lambda \psi$. □

The converse implication does not hold, as has been shown in [**151**]. Interpolation is closely connected with the so–called *Beth property*. It says, in intuitive terms, that if we have defined p implicitly, then there also is an explicit definition of p. An *explicit definition* is a statement of the form $\chi \leftrightarrow p$ where $p \notin var(\chi)$. An *implicit definition* is a formula $\psi(p, \vec{q})$, such that the value of p in a model is uniquely defined by the values of the variables $\vec{q}$. The latter can be reformulated syntactically. In a logic $\vdash$, $\varphi(p, \vec{q})$ **implicitly defines** p if $\varphi(p, \vec{q}); \varphi(r, \vec{q}) \vdash p \leftrightarrow r$. Given Λ, we may choose $\vdash$ to be either $\vdash_\Lambda$ or $\Vdash_\Lambda$. This gives rise to the notions of **local** and **global** implicit definitions.

Definition 3.7.3. *Λ is said to have the* **local Beth Property** *if the following holds. Suppose $\varphi(p, \vec{q})$ is a formula and*

$$\varphi(p, \vec{q}); \varphi(r, \vec{q}) \vdash_\Lambda p \leftrightarrow r.$$

Then there exists a formula $\chi(\vec{q})$ not containing p as a variable such that

$$\varphi(p, \vec{q}) \vdash_\Lambda p \leftrightarrow \chi(\vec{q}) .$$

Analogously, the $\boldsymbol{global\ Beth\ property}$ *is defined by replacing* $\vdash_\Lambda$ *by* $\Vdash_\Lambda$.

The notion of definability was introduced by Beth in [**16**] under the name *Padoa's Method*. The lack of the deduction theorem for the global consequence makes the global Beth property somewhat more difficult to handle than the local equivalent. For the local Beth property we can actually prove that it is equivalent to the Craig Interpolation Property.

Theorem 3.7.4 (Maksimova). *Let* Λ *be a classical modal logic. Then* Λ *has local interpolation iff it has the local Beth property.*

Proof. Suppose first that Λ has local interpolation. Assume that φ defines q implicitly, that is,

$$(\dagger) \qquad \varphi(p,\vec{q});\varphi(r,\vec{q}) \vdash_\Lambda p \leftrightarrow r.$$

Then we also have $\varphi(p,\vec{q}); p \vdash_\Lambda \varphi(r,\vec{q}) \to r$ and thus by interpolation there is a $\chi(\vec{q})$ such that

$$(\ddagger) \qquad \varphi(p,\vec{q}); p \vdash_\Lambda \chi(\vec{q}) \vdash_\Lambda \varphi(r,\vec{q}) \to r$$

We claim that $\chi(\vec{q})$ is the desired explicit definition, that is, that the following holds.

$$\varphi(p,\vec{q}) \vdash_\Lambda p \leftrightarrow \chi(\vec{q}).$$

One implication holds by definition of the interpolant; for $\varphi(p,\vec{q}); p \vdash_\Lambda \chi(\vec{q})$. For the other direction, observe that we have $\chi(\vec{q}) \vdash_\Lambda \varphi(r,\vec{q}) \to r$. Using the deduction theorem we can derive $\varphi(p,\vec{q}) \vdash_\Lambda \chi(\vec{q}) \to r$. Now replace r by p, and the desired conclusion follows. Now for the converse, assume that Λ has the local Beth Property. We will show that if $\varphi(p,\vec{q}) \vdash_\Lambda \psi(r,\vec{q})$ then there is a $\chi(\vec{q})$ such that $\varphi(p,\vec{q}) \vdash_\Lambda \chi(\vec{q}) \vdash_\Lambda \psi(r,\vec{q})$. Let us call this the 1–interpolation property, since we can get rid of a single variable in φ and a single variable in ψ. The n–interpolation property is formulated similarly, but with the difference that we can eliminate up to n variables in the premiss and up to n in the conclusion. One can easily show that 1–interpolation property implies n–interpolation for every n, and hence the local interpolation property. We leave this to the reader. The hard part is to show 1–interpolation. Thus, assume $\varphi(p,\vec{q}) \vdash_\Lambda \psi(r,\vec{q})$. Define

$$\delta_1(p,\vec{q}) := (p \to \varphi(p,\vec{q})) \wedge (\psi(p,\vec{q}) \to p).$$

Since we have $\delta_1(p,\vec{q}); \delta_1(r,\vec{q}) \vdash_\Lambda p \leftrightarrow r$, we get a formula $\chi_1(\vec{q})$ such that

$$\delta_1(p,\vec{q}) \vdash_\Lambda p \leftrightarrow \chi_1(\vec{q}).$$

We then have

$$\varphi(p,\vec{q}); p \vdash_\Lambda \chi_1(\vec{q}) \qquad \chi_1(\vec{q}) \vdash_\Lambda p \vee \psi(p,\vec{q}).$$

Now define

$$\delta_2(p,\vec{q}) := (\neg p \to \varphi(p,\vec{q})) \wedge (\psi(p,\vec{q}) \to \neg p).$$

Again, it is checked that δ_2 is an implicit definition and so we can use Beth's property again to get a $\chi_2(\vec{q})$ with $\varphi(p,\vec{q}) \vdash_\Lambda p \leftrightarrow \chi_2(\vec{q})$. After some boolean rewriting

$$\varphi(p,\vec{q}); \neg p \vdash_\Lambda \neg\chi_2(\vec{q}) \ , \qquad \neg\chi_2(\vec{q}) \vdash_\Lambda \neg p \vee \psi(p,\vec{q}) \ .$$

Next we define the following formula

$$\delta_3(p, \vec{q}) := (p \to \varphi(\neg p, \vec{q})) \wedge (\psi(p, \vec{q}) \to p).$$

We have $\varphi(\neg p, \vec{q}) \vdash_\Lambda \psi(r, \vec{q})$ and so it is checked that $\delta_3(p, \vec{q})$ also provides an implicit definition of p. And so we get a third formula $\chi_3(\vec{q})$, such that $\delta_3(p, \vec{q}) \vdash_\Lambda p \leftrightarrow \chi_3(\vec{q})$. From this we get

$$\varphi(\neg p, \vec{q}); p \vdash_\Lambda \chi_3(\vec{q}) \ , \qquad \chi_3(\vec{q}) \vdash_\Lambda p \vee \psi(p, \vec{q}).$$

Substituting $\neg p$ for p in the first statement we get

$$\varphi(p, \vec{q}); \neg p \vdash_\Lambda \chi_3(\vec{q}).$$

Finally, define

$$\delta_4(p, \vec{q}) := (\neg p \to \varphi(\neg p, \vec{q})) \wedge (\psi(p, \vec{q}) \to \neg p).$$

This defines p implicitly, and so we have a $\chi_4(\vec{q})$ with $\delta_4(p, \vec{q}) \vdash_\Lambda p \leftrightarrow \chi_4(\vec{q})$. We get after rewriting

$$\varphi(\neg p, \vec{q}); \neg p \vdash_\Lambda \neg\chi_4(\vec{q}) \qquad \neg\chi_4(\vec{q}) \vdash_\Lambda \neg p \vee \psi(p, \vec{q}).$$

Substituting $\neg p$ for p in the first statement we get

$$\varphi(p, \vec{q}); p \vdash_\Lambda \neg\chi_4(\vec{q}).$$

The desired interpolant is

$$\chi(\vec{q}) := (\chi_1(\vec{q}) \wedge \neg\chi_4(\vec{q})) \vee (\neg\chi_2(\vec{q}) \wedge \chi_3(\vec{q})).$$

Namely, $\varphi(p, \vec{q}); p \vdash_\Lambda \chi_1(\vec{q}) \wedge \neg\chi_4(\vec{q})$ and so $\varphi(p, \vec{q}); \neg p \vdash_\Lambda \neg\chi_2(\vec{q}) \wedge \chi_4(\vec{q})$, so that $\varphi(p, \vec{q}) \vdash_\Lambda \chi(\vec{q})$. Furthermore, $\chi_1(\vec{q}) \wedge \neg\chi_4(\vec{q}) \vdash_\Lambda \psi(p, \vec{q})$ and in addition $\neg\chi_2(\vec{q}) \wedge \chi_3(\vec{q}) \vdash_\Lambda \psi(p, \vec{q})$, from which $\chi(\vec{q}) \vdash_\Lambda \psi(p, \vec{q})$, as required. □

Theorem 3.7.5. *A classical modal logic with local interpolation also has the global Beth–property.*

Proof. Assume that $\varphi(p, \vec{q}); \varphi(r, \vec{q}) \Vdash_\Lambda p \leftrightarrow r$. Then for some compound modality $\boxplus$ we have

$$\boxplus\varphi(p, \vec{q}); \boxplus\varphi(r, \vec{q}) \vdash_\Lambda p \leftrightarrow r.$$

This can now be rearranged to

$$\boxplus\varphi(p, \vec{q}); p \vdash_\Lambda \boxplus\varphi(r, \vec{q}) \to r.$$

We get an interpolant $\chi(\vec{q})$ and so we have $\boxplus\varphi(p, \vec{q}); p \vdash_\Lambda \chi(\vec{q})$, from which $\boxplus\varphi(p, \vec{q}) \vdash_\Lambda p \to \chi(\vec{q})$. So $\varphi(p, \vec{q}) \vdash_\Lambda p \to \chi(\vec{q})$. And we moreover have $\chi(\vec{q}) \vdash_\Lambda \boxplus\varphi(r, \vec{q}) \to r$, from which we get $\boxplus\varphi(r, \vec{q}) \vdash_\Lambda \chi(\vec{q}) \to r$, and so $\varphi(r, \vec{q}) \vdash_\Lambda \chi(\vec{q}) \to r$. Replacing r by p we get the desired result. □

The picture obtained thus far is the following.

$$\begin{tikzcd} \text{local interpolation} \arrow[r] \arrow[d, shift left] & \text{global interpolation} \\ \text{local Beth Property} \arrow[u, shift left] \arrow[r] & \text{global Beth Property} \end{tikzcd}$$

It can be shown that there exist logics without global interpolation while having the global Beth Property and that there exist logics with global interpolation without the global Beth Property. An example of the first kind is the logic **G.3**. See [**150**].

Recall from Section 1.6 the notion of Halldén–completeness of a logic. As with interpolation the notion of Halldén–completeness of Λ splits into (at least) two different concepts.

Definition 3.7.6. *Let Λ be a modal logic. Λ is **locally Halldén–complete** if whenever $\varphi \vdash_\Lambda \psi$ and $var(\varphi) \cap var(\psi) = \varnothing$ we have $\varphi \vdash_\Lambda \bot$ or $\vdash_\Lambda \psi$. Λ is **globally Halldén–complete** if whenever $\varphi \Vdash_\Lambda \psi$ and $var(\varphi) \cap var(\psi) = \varnothing$ we have $\varphi \Vdash_\Lambda \bot$ or $\Vdash_\Lambda \psi$.*

Global Halldén–completeness is called the **Pseudo Relevance Property** in [**153**]. In the literature, a logic Λ is called *Halldén–complete* if for φ and ψ disjoint in variables, if $\varphi \vee \psi \in \Lambda$ then $\varphi \in \Lambda$ or $\psi \in \Lambda$. Clearly, this latter notion of Halldén–completeness coincides with local Halldén–completeness. This follows from the deduction theorem, since $\varphi \vdash_\Lambda \psi$ is equivalent to $\vdash_\Lambda \neg\varphi \vee \psi$. Local (global) Halldén–completeness nearly follows from the corresponding interpolation property. Namely, if we have $\vdash_\Lambda \varphi \vee \psi$ then $\neg\varphi \vdash_\Lambda \psi$. We then get a constant formula χ such that $\neg\varphi \vdash_\Lambda \chi$ and $\chi \vdash_\Lambda \psi$. On the condition that we can choose χ to be either $\top$ or $\bot$ we get our desired conclusion. For $\chi = \top$ yields $\vdash_\Lambda \psi$ and $\chi = \bot$ yields $\vdash_\Lambda \varphi$. Thus, if Λ has trivial constants (see Section 2.6) then interpolation implies Halldén–completeness. But this is exactly the condition we need anyway to have Halldén–completeness. For notice that always $\vdash_\Lambda \neg\Box_j\bot \vee \Box_j\bot$. So if Λ is Halldén–complete then we have either $\vdash_\Lambda \neg\Box_j\bot$ or $\vdash_\Lambda \Box_j\bot$. These are exactly the conditions under which Λ has trivial constants.

Proposition 3.7.7. *A logic Λ is (locally/globally) Halldén–complete only if it has trivial constants. If Λ has trivial constants and has (local/global) interpolation then it is (locally/globally) Halldén–complete.*

Finally, we will establish some criteria for interpolation. Assume that we have a logic $\Lambda \oplus A$ and global reduction sets for $\Lambda \oplus A$ with respect to Λ. Let us say that the reduction sets **split** if there exists a reduction function X such that (i.) for all sets Δ, $var[X(\Delta)] \subseteq var[\Delta]$ and (ii.) $X(\varphi \to \psi) = X(\varphi) \cup X(\psi)$.

Theorem 3.7.8. *Suppose that $\Lambda \oplus A$ can be globally reduced to Λ with splitting reduction sets. Then $\Lambda \oplus A$ has local (global) interpolation if Λ has local (global) interpolation. Moreover, $\Lambda \oplus A$ is locally (globally) Halldén–complete if Λ is.*

Proof. Assume $\varphi \vdash_{\Lambda \oplus A} \psi$. Then $\vdash_{\Lambda \oplus A} \varphi \to \psi$ and a fortiori $\Vdash_{\Lambda \oplus A} \varphi \to \psi$. By global reduction we get $X(\varphi \to \psi) \Vdash_{\Lambda} \varphi \to \psi$ and so for some compound modality $\boxplus$

$$\boxplus X(\varphi \to \psi) \vdash_{\Lambda} \varphi \to \psi.$$

This is the same as

$$\boxplus X(\varphi); \boxplus X(\psi) \vdash_{\Lambda} \varphi \to \psi,$$

by the fact that the reduction sets split. We can rearrange this into

$$\boxplus X(\varphi); \varphi \vdash_{\Lambda} \boxplus X(\psi) \to \psi.$$

By assumption on X, $var[X(\varphi)] \subseteq var(\varphi)$ and $var[X(\psi)] \subseteq var(\psi)$. By local interpolation for Λ we obtain a χ in the common variables of φ and ψ such that

$$\varphi; \boxplus X(\varphi) \vdash_{\Lambda} \chi \vdash_{\Lambda} \boxplus X(\psi) \to \psi.$$

From this follows that $\varphi \vdash_{\Lambda \oplus A} \chi \vdash_{\Lambda \oplus A} \psi$, by the fact that the reduction sets only contain instances of theorems. Pushing up global interpolation works essentially in the same way. Now for Halldén–completeness, assume that $\varphi \vdash_{\Lambda \oplus A} \psi$ for φ and ψ disjoint in variables. Then

$$\varphi; \boxplus X(\varphi) \vdash_{\Lambda} \boxplus X(\psi) \to \psi.$$

The left hand side is disjoint in variables from the right hand side, and so either the left hand side is inconsistent or the right hand side a theorem. In the first case, $\varphi \vdash_{\Lambda \oplus A} \bot$. In the second case $\vdash_{\Lambda \oplus A} \psi$, as required. The proof for global Halldén–completeness is analogous. ◻

In the next section we will prove that **K** has local interpolation. We conclude from that the following theorem.

Corollary 3.7.9. *The logics $\mathbf{K.alt}_1$, $\mathbf{K4}$, $\mathbf{K.B}$, $\mathbf{K.T}$, $\mathbf{K.BT}$, $\mathbf{S4}$, $\mathbf{S5}$, $\mathbf{G}$ and $\mathbf{Grz}$ have local interpolation.*

This can be generalized to polymodal logics as well, namely to those which have no interaction postulates for the operators, and whose logical fragments for the individual operators is one of the above logics. This, however, will be a consequence of a far more general result on so–called *independent fusions* to be developed in Chapter 6. As an application we will prove a rather famous theorem, the so–called *Fixed Point–Theorem* of Sambin and de Jongh (see [**184**]; for the history see also [**32**]). It is a theorem about the provability logic **G**. We say that a formula $\psi(\vec{q})$ is a **fixed point** of $\varphi(p, \vec{q})$ with respect to p in a logic Λ if

$$\vdash_{\Lambda} \psi(\vec{q}) \leftrightarrow \varphi(\psi(\vec{q}), \vec{q}).$$

.

Theorem 3.7.10 (Sambin, de Jongh). *Let $\varphi(p, \vec{q})$ be a formula such that every occurrence of p is modalized. Then $\varphi(p, \vec{q})$ has a fixed point for p in $\mathbf{G}$.*

PROOF. Consider a formula $\varphi(p,\vec{q})$ in which the sentence letter p occurs only modalized, that is, in the scope of an $\Box$. We show that for every finite Kripke–frame and valuation β on $\vec{q}$ there exists one and only one extension β^+ such that $\langle \mathfrak{f}, \beta^+\rangle \vDash p \leftrightarrow \varphi(p,\vec{q})$. To see this take a finite Kripke–frame $\mathfrak{f}$. The accessibility relation is transitive and cycle–free, that is, irreflexive. β^+ will be defined by induction on the depth of a point, that is, the length of a maximal ascending chain starting at that point. To start, consider a point x without successors. Then, since p occurs only modalized, $\varphi(p,\vec{q})$ holds at x iff $\varphi(\bot,\vec{q})$ holds at x; so the value of p at x is well–defined and does not depend on p. Put $x \in \beta^+(p)$ iff $x \in \overline{\beta^+}(\varphi(\bot,\vec{q}))$. Now assume that the claim has been established for points of depth d. Let x be of depth $d+1$. We have to show that $x \in \beta^+(p)$ iff $x \in \overline{\beta^+}(\varphi(p,\vec{q}))$. We can regard φ as a boolean combination of formulae not containing p and formulae of the form $\Box\chi$. Since the value of p is already fixed on points of lesser depth, we know whether or not $x \in \overline{\beta^+}(\Box\chi)$; also, the value of formulae not containing p is fixed at x. Hence there is a unique way to assign p at x. This completes the proof of the existence and uniqueness of β^+.

The first consequence is that because the extension is unique on finite frames, we have

$$p \leftrightarrow \varphi(p,\vec{q}); r \leftrightarrow \varphi(r,\vec{q}) \Vdash_{\mathbf{G}} p \leftrightarrow r \ .$$

Taking $\varphi_1(p,\vec{q}) := p \leftrightarrow \varphi(p,\vec{q})$ we now have a global implicit definition of p. Since **G** has local interpolation it has the global Beth–property and so there exists an explicit definition, that is, a $\psi(\vec{q})$ such that

$$p \leftrightarrow \varphi(p,\vec{q}) \Vdash_{\mathbf{G}} p \leftrightarrow \psi(\vec{q}) \ .$$

From this we deduce that

$$p \leftrightarrow \varphi(p,\vec{q}) \Vdash_{\mathbf{G}} \psi(\vec{q}) \leftrightarrow \varphi(\psi(\vec{q}),\vec{q}) \ .$$

(Simply replace p by $\psi(\vec{q})$.) We claim now that in fact

$$\vdash_{\mathbf{G}} \psi(\vec{q}) \leftrightarrow \varphi(\psi(\vec{q}),\vec{q}) \ .$$

To see this, take a finite frame $\mathfrak{f}$ for **G** and a valuation of $\vec{q}$. Then there exists an extension β^+ of β giving a value to p such that $\langle \mathfrak{f}, \beta^+\rangle \vDash p \leftrightarrow \varphi(p,\vec{q})$. In this model we have $\langle \mathfrak{f}, \beta^+\rangle \vDash \psi(\vec{q}) \leftrightarrow \varphi(\psi(\vec{q}),\vec{q})$. But then also $\langle \mathfrak{f}, \beta\rangle \vDash \psi(\vec{q}) \leftrightarrow \varphi(\psi(\vec{q}),\vec{q})$, as required. ∎

The proof via the Beth–property has first been given by SMORYŃSKI [**200**]. It is worthwile to note that many direct proofs of this theorem have been given in the past, e. g. SAMBIN [**184**], [**188**], REIDHAAR–OLSON [**175**] and GLEIT and GOLDFARB [**76**]. The difficulty of these proofs varies. This proof via the Beth–property reduces the problem to that of the interpolation of **G** modulo the Theorem 3.7.5. The difficulty that most proofs face is that the construction of the interpolant is not disentangled from the actual fixed point property. For notice that the special property of **G** is that for formulae $\varphi(p,\vec{q})$ in which p occurs only modalized, the fixed point equation $p \leftrightarrow \varphi(p,\vec{q})$ can be globally satisfied in one and only one way for given $\vec{q}$. This is proved semantically. The uniqueness of the solution then gives rise to an explicit definition because of the Beth–property. That means that the solution is *effable*. The existence of a solution then allows to deduce that this explicit definition is a fixed point of φ. Note that the previous proof did not depend on **G**, only on some critical features of **G**. It trivially also holds for all extensions of **G**. It has been observed by Maksimova that this can be used to show that all extensions of **G** have the global Beth–property (see [**149**]). To show this, two auxiliary facts must be established, which we shall give as exercises. Call a formula φ $\vec{q}$–**boxed** if every occurrence of a variable from $\vec{q}$ is in the scope of some modal operator.

Lemma 3.7.11. *Let Λ be a logic containing $\boldsymbol{G}$. Let q_i, $i < n$, be distinct variables and p a variable not contained in $\vec{q}$. For a set $S \subseteq n$ define χ_S by $\chi_S := \bigwedge_{i \in S} q_i \wedge \bigwedge_{i \in n-S} \neg q_i$. If $\varphi(p, \vec{q})$ is $\vec{q}$–boxed and*

$$\Lambda \vdash \chi_S \to \varphi(p, \vec{q})$$

then already $\Lambda \vdash \varphi(p, \vec{q})$.

Lemma 3.7.12. *Let $\varphi(p, \vec{q})$ be a formula. Then there exist $\vec{q}$–boxed formulae $\psi_1(p, \vec{q})$, $\psi_2(p, \vec{q})$, $\chi_1(p, \vec{q})$ and $\chi_2(p, \vec{q})$ such that*

$$\begin{aligned} \boldsymbol{G} &\vdash \varphi(p, \vec{q}) \leftrightarrow ((p \vee \psi_1(p, \vec{q})) \wedge (\neg p \vee \psi_2(p, \vec{q}))) \\ \boldsymbol{G} &\vdash \varphi(p, \vec{q}) \leftrightarrow ((p \wedge \chi_1(p, \vec{q})) \vee (\neg p \wedge \chi_2(p, \vec{q}))) \end{aligned}$$

Theorem 3.7.13 (Maksimova). *Let Λ be a logic containing $\boldsymbol{G}$. Then Λ has the global Beth property.*

Proof. Suppose that $\varphi(p, \vec{q})$ is a global implicit definition of p in Λ. Then $\varphi(p, \vec{q}); \varphi(r, \vec{q}) \Vdash_\Lambda p \leftrightarrow r$. Using Lemma 3.7.12 we get $\vec{q}$–boxed formulae $\chi_1(p, \vec{q})$ and $\chi_2(p, \vec{q})$ such that

$$(\dagger) \qquad \Lambda \vdash \varphi(p, \vec{q}) \leftrightarrow ((p \wedge \chi_1(p, \vec{q})) \vee (\neg p \wedge \chi_2(p, \vec{q})))$$

Since we also have

$$\Lambda \vdash \Box^{\leq 1}\varphi(p, \vec{q}) \wedge \Box^{\leq 1}\varphi(r, \vec{q}) \to (p \leftrightarrow r)$$

we now get

$$\Lambda \vdash (\Box\varphi(p, \vec{q}) \wedge \Box\varphi(r, \vec{q}) \wedge p \wedge \chi_1(p, \vec{q}) \wedge \neg r \wedge \chi_2(r, \vec{q})) \to (p \to r)$$

This formula has the form $(\mu \wedge p \wedge \neg r) \to (p \to r)$, where μ is $\vec{q}$–boxed. This is equivalent to $\neg\mu \vee \neg p \vee r$, or $(\neg r \wedge p) \to \neg\mu$. By use of Lemma 3.7.11 we deduce that $\Lambda \vdash \mu$, that is,

$$\Lambda \vdash \Box\varphi(p, \vec{q}) \wedge \Box\varphi(r, \vec{q}) \to (\chi_1(p, \vec{q}) \to \neg\chi_2(r, \vec{q})) \; .$$

We substitute p for r and obtain

$$\Lambda \vdash \Box\varphi(p, \vec{q}) \to (\chi_1(p, \vec{q}) \to \neg\chi_2(p, \vec{q})) \; .$$

Now from this and (†) it follows after some boolean manipulations

$$\Lambda \vdash \Box^{\leq 1}\varphi(p, \vec{q}) \to \Box^{\leq 1}(p \leftrightarrow \chi_1(p, \vec{q}))$$

By the fixed point theorem for **G** there is a $\psi(\vec{q})$ such that

$$\mathbf{G} \vdash \Box^{\leq 1}(p \leftrightarrow \chi_1(p, \vec{q})) \to (p \leftrightarrow \psi(\vec{q})) .$$

So we obtain

$$\Lambda \vdash \Box^{\leq 1}\varphi(p, \vec{q}) \to (p \leftrightarrow \psi(\vec{q})) ,$$

which is nothing but

$$\varphi(p, \vec{q}) \Vdash_\Lambda p \leftrightarrow \psi(\vec{q}) .$$

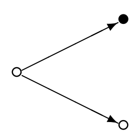
∎

Exercise 114. Show that if a logic has 1–interpolation then it has n–interpolation for every $n \in \omega$.

Exercise 115. (Rautenberg [**171**].) Let Λ have local interpolation and let A be a set of constant formulae. Then $\Lambda \oplus A$ has local interpolation.

Exercise 116. As in the previous exercise, but with global interpolation.

Exercise 117. Show that the logic of the following frame is not Halldén–complete but has interpolation.



Exercise 118. Show that any quasi–normal logic determined by a single rooted frame $\langle \mathfrak{F}, x \rangle$ is Halldén–complete.

Exercise 119. Show that if a logic has the local Beth–property it also has the global Beth–property.

Exercise 120. (Maksimova [**149**].) Show Lemma 3.7.11.

Exercise 121. (Maksimova [**149**].) Show Lemma 3.7.12. *Hint.* You may assume that $\varphi(p, \vec{p})$ is is normal form. Now reduce the case where it is a disjunction to the case where it is not.

3.8. Tableau Calculi and Interpolation

In this chapter we will introduce the notion of a *tableau* (plural *tableaux*), a popular method for showing the decidability of logics. Since we have already established the decidability via the finite model property, we do not need tableaux for this purpose. However, there are additional reasons for studying tableau methods. One is the connection between tableau calculi and interpolation. This connection has been explored by Wolfgang Rautenberg in [**171**]. However, we will show at the end that a proof via normal forms is also possible. Tableaux are also an efficient technique for checking the satisfiability of a formula in a logic. They will allow to give bounds on the complexity of the satisfiability problem. Tableaux can be described as a method of deriving a contradiction as fast and efficiently as possible. A tableau can be seen as computing along branches in a model to see at which point we reach an inconsistency. To help in understanding these remarks, let us describe a tableau calculus for $\mathbf{K}_\kappa$, which we denote here by $\mathcal{C}_\mathbf{K}$. To keep the calculus short, we assume to have only the symbols $\neg$, $\wedge$ and $\Box_j$. The other symbols are treated as abbreviations. The calculus operates on *sets* of formulae. As usual, $X;\varphi$ denotes $X \cup \{\varphi\}$ and $X;Y$ denotes $X \cup Y$. Thus $X;\varphi;\varphi$ is the same as $X;\varphi$. The rules are as follows.

$$(\neg E)\quad \frac{X;\neg\neg\varphi}{X;\varphi} \qquad\qquad (\wedge E)\quad \frac{X;\varphi\wedge\psi}{X;\varphi;\psi}$$

$$(\vee E)\quad \frac{X;\neg(\varphi\wedge\psi)}{X;\neg\varphi|X;\neg\psi} \qquad\qquad (\Box_j E)\quad \frac{\Box_j X;\neg\Box_j\varphi}{X;\neg\varphi}$$

$$(w)\quad \frac{X;Y}{X}$$

The last rules is known as **weakening**. We abbreviate by $(\Box E)$ the set of rules $\{(\Box_j E) : j < \kappa\}$. A $\mathcal{C}_\mathbf{K}$–**tableau** for X is a rooted labelled tree, the labels being sets of formulae, such that (i) the root has label X, (ii) if a node x has label Y and a single daughter y then the label of y arises from the label of x by applying one of $(\neg E)$, $(\wedge E)$, $(\Box E)$ or (w), (iii) if x has two daughters y and z then the label of x is $Y;\neg(\varphi\wedge\psi)$ and the labels of y and z are $Y;\neg\varphi$ and $Y;\neg\psi$, respectively. There are possibly several tableaux for a given X. $(\vee E)$ introduces a *branching* into the tableau, and it is the only rule that does so. A branch of a tableau **closes** if it ends with $p;\neg p$ for some variable p. The tableau **closes** if all branches close.

Proposition 3.8.1. *If X has a $\mathcal{C}_\mathbf{K}$–tableau which closes then X is inconsistent in $\boldsymbol{K}_\kappa$.*

Proof. By induction on the length of the tableau. Clearly, at the leaves we have sets of the form $p;\neg p$, and they are all inconsistent. So, we have to prove for each rule that if the lower sets are inconsistent, so is the upper set. This is called the *correctness* of the rules. For example, if X is inconsistent, then $X;Y$ is inconsistent, so (w) is correct. The boolean rules are straightforward. For $(\Box E)$, assume that $X;\neg\varphi$ is inconsistent, that is, $X \vdash_{\mathbf{K}_\kappa} \varphi$. Then $\Box_j X \vdash_{\mathbf{K}_\kappa} \Box_j\varphi$, that is, $\Box_j X;\neg\Box_j\varphi$ is inconsistent. ∎

The calculus is also complete; that means, if no tableau closes, then X is consistent. We prove this by showing that whenever there is no closing tableau there is a model for X. Before we proceed, let us remark that one can remove weakening from the tableau calculus. However, this is possible only if $(\Box_j E)$ is replaced by the following rule.

$$(\Box_j E') \qquad \frac{X; \neg\Box_j \varphi}{X_\Box; \neg\varphi} \qquad\qquad X_\Box := \{\chi : \Box_j \chi \in X\}$$

Removing weakening is desirable from a computational point of view, because the rule of weakening introduces too many options in a search space. Given a set X we can apply $2^{\sharp X} - 2$ nontrivial weakenings but very few of them turn out to be sensible. Instead, if one intends to speed up the proof search one has to implement a different calculus. Moreover, one can try to close a branch as early as possible, e. g. when hitting a set containing φ and $\neg\varphi$, where φ can be any formula. One can also introduce priorities on the rules, such as to prefer applying $(\neg E)$ and $(\vee E)$ before any other rule, and to delay $(\vee E)$ (or $(\Box E')$). To show that the tableau calculus is complete we shall have to show that if no tableau for a set X closes then there is a model for X, which will be enough to show that X is consistent. To understand how such a model can be found we imagine the tableau as computing possible valuations at worlds in a model. We start somewhere and see whether we can fulfill X. The rules $(\neg E)$, $(\wedge E)$ and $(\vee E)$ are *local* rules. They allow us to derive something about what has to be true at the given world. By using the rule $(\Box E')$, however, we go into another world and investigate the valuation there. Thus, $(\Box E')$ is not local; we call it the *step rule*. Once we have made a step there is no returning back. This is why one should always apply local rules first. In fact, we can also *prove* that any closing tableau gives rise to another closing tableau where $(\Box E')$ has been applied only when nothing else was possible. Call a set **downward saturated** if it is closed under $(\neg E)$, $(\vee E)$ and $(\wedge E)$. Alternatively, X is downward saturated if (i) for every $\neg\neg\varphi \in X$ also $\varphi \in X$, (ii) for every $\varphi \wedge \psi \in X$ both $\varphi \in X$ and $\psi \in X$, and (iii) for every $\neg(\varphi \wedge \psi) \in X$ either $\neg\varphi \in X$ or $\neg\psi \in X$.

Lemma 3.8.2. *Suppose there is a closing tableau for X. Then there is a closing tableau for X where* $(\Box E')$ *is applied only to downward saturated sets.*

Proof. Consider an application of $(\Box E')$ where an application of $(\wedge E)$ has been possible instead:

$$\frac{X; \varphi \wedge \psi; \Box_j Y; \neg\Box_j \chi}{Y; \neg\chi}$$

Replace this derivation by

$$\cfrac{\cfrac{X; \varphi \wedge \psi; \Box_j Y; \neg\Box_j \chi}{X; \varphi; \psi; \Box_j Y; \neg\Box_j \chi}}{Y; \neg\chi}$$

By assumption, there is a closing tableau for $Y; \chi$, so the latter tableau can be brought to close as well. Similarly for $(\neg E)$. With $(\vee E)$ we get

$$\frac{X; \varphi \vee \psi; \Box_j Y; \neg\Box_j \chi}{Y; \neg\chi}$$

which we replace by

$$\begin{array}{c|c} \multicolumn{2}{c}{X; \varphi \vee \psi; \Box_j Y; \neg\Box_j \chi} \\ \hline X; \varphi; \Box_j Y; \neg\Box_j \chi & X; \psi; \Box_j Y; \neg\Box_j \chi \\ \hline Y; \chi & Y; \neg\chi \end{array}$$

Again, $Y; \chi$ has a closing tableau by assumption, so the latter derivation has a closing tableau. This process of swapping derivations yields less and less offending instances, so it terminates in a derivation where $(\Box E')$ is applied only when all formulae are either variables, negated variables or of the form $\neg\Box_j \chi$, $\Box_j \chi$. ∎

Call a $\mathfrak{C}_{\mathbf{K}}$–tableau **good** if the rule $(\Box E')$ is applied to downward saturated sets. Now we use good tableau–derivations to find our model. Assume that no closing tableau for X exists, hence also no closing good tableau. The frame will be based on worlds w_Z for downward saturated Z. Let S_X be the set of all sets in any tableau for X which is on a branch that does not close. By assumption, $X \in S_X$. Within S_X let Sat_X be the subcollection of downward saturated sets in S_X. By assumption, for each $Z \in S_X$ there exists a saturated $Z^\star \in Sat_X$ containing X. Let $Z, Y \in Sat_X$. Then put $Y \lhd_j Z$ iff Z is a saturation of a set U obtained by applying $(\Box_j E')$ to Y. This defines the frame $\mathfrak{Sat}_X := \langle Sat_X, \langle \lhd_j : j < \kappa \rangle \rangle$. Furthermore, let $Y \in \beta(p)$ iff $p \in Y$. By assumption, for no p we have both $p \in Y$ and $\neg p \in Y$ and so the definition is not contradictory. We will now show that if $\varphi \in Y$ then $\langle \mathfrak{Sat}_X, \beta, Y \rangle \vDash \varphi$. By definition of Sat_X we never have both $\varphi \in Y$ and $\neg\varphi \in Y$. Now, let $\varphi = \psi_1 \wedge \psi_2$. If $\psi_1 \wedge \psi_2 \in Y$ then both $\psi_1, \psi_2 \in Y$ and so by induction hypothesis the claim follows. Next, let $\varphi = \neg\neg\psi$. If $\varphi \in Y$, then also $\psi \in Y$ and applying the induction hypothesis, the claim follows again. Finally, if $\neg(\psi_1 \wedge \psi_2) \in Y$ then either $\neg\psi_1 \in Y$ or $\neg\psi_2 \in Y$. In either case we conclude $\langle \mathfrak{Sat}_X, \beta, Y \rangle \vDash \varphi$, using the induction hypothesis. Now assume $\varphi = \neg\Box_j \psi$ for some ψ. By construction there is a downward saturated Z such that $Y \lhd_j Z$ and $\neg\psi \in Z$. By induction hypothesis, $\langle \mathfrak{Sat}_X, \beta, Z \rangle \vDash \neg\psi$, showing that $\langle \mathfrak{Sat}_X, \beta, Y \rangle \vDash \varphi$. Now let $\varphi = \Box_j \psi$ and take a j–successor Z of Y. This successor is of the form $U^\star$ where $U^\star$ is the saturation of U, which in turn results from Y by applying $(\Box E')$. Thus $\langle \mathfrak{Sat}_X, \beta, U^\star \rangle \vDash \psi$, by induction hypothesis. Hence $\langle \mathfrak{Sat}_X, \beta, Y \rangle \vDash \varphi$.

Theorem 3.8.3. *A set X of modal formulae is $\boldsymbol{K}_\kappa$–consistent iff no $\boldsymbol{K}_\kappa$–tableau for X closes. $\boldsymbol{K}_\kappa$–satisfiability is in PSPACE.*

Proof. Only the last claim needs proof. It is enough to see that we can do the tableau algorithm keeping track only of a single branch, backtracking to a branching point whenever necessary. (This is not entirely straightforward but can be achieved with a certain amount of bookkeeping.) So, we need to show that a branch consumes only polynomial space. We see however that each branch of the tableau is of depth $\leq dp(\varphi)$ and that each node is of length $\leq |\varphi|^2$. ◻

We note that by a result LADNER [**137**], satisfiability is also PSPACE–complete. (In fact, LADNER proves in that paper that all logics in the interval $[\mathbf{K}, \mathbf{S4}]$ are PSPACE–hard.) With each tableau calculus we can associate a *dual* tableau calculus. Notice that for formulae a dual has been defined by reversing the roles of $\wedge$ and $\vee$ as well as $\Box_j$ and $\Diamond_j$. In particular, if φ is a formula, $\neg\varphi$ is the same as $\varphi^d[\neg p/p]$. The dual tableau calculus consists in reversing the roles of $\wedge$ and $\vee$, $\Box_j$ and $\Diamond_j$. Here a set X is read disjunctively, and we attempt to show that X is a theorem. For example, the following is a rule of the dual calculus

$$(\Diamond D) \qquad \frac{\Diamond_j X; \neg\Diamond_j\varphi}{X; \neg\varphi}$$

For suppose $\vdash_{\mathbf{K}} \bigvee X \vee \neg\varphi$. Then $\varphi \vdash_{\mathbf{K}} \bigvee X$ is valid, and so is $\Diamond_j\varphi \vdash_{\mathbf{K}} \Diamond_j \bigvee X$. Hence $\neg\Diamond_j\varphi \vee \bigvee \Diamond_j X$ is a theorem. It is the same to show that X has a closing tableau as it is to show that $\neg X$ has a closing dual tableau. Notice that the dual of $(\vee E)$ is $(\wedge E)$ and has no branching, while the dual of $(\wedge E)$ is $(\vee E)$ and does have branching. The dual of (w) is (w), and the dual of $(\neg E)$ is $(\neg E)$ as is easily checked.

THEOREM 3.8.4. *$\varphi \in \boldsymbol{K}_\kappa$ iff there is a closed dual tableau for $\{\varphi\}$.*

Let us now see in what ways a tableau calculus helps in finding interpolants. Consider $\varphi \vdash_{\mathbf{K}} \psi$, that is, $\varphi; \neg\psi$ is $\mathbf{K}$–inconsistent. Then we need to find a χ in the common variables of φ and ψ such that $\varphi; \neg\chi$ is inconsistent and $\chi; \neg\psi$ is inconsistent as well. The idea is now to start from a closing tableau for $\varphi; \neg\psi$ and define χ by induction on the tableau, starting with the leaves. To that end, let us introduce a marking or signing of the occurring formulae in a tableau. In the tableau we introduce the marks a and c, where a stands for *antecedent* and c for *consequent*. Each formula is thus signed, whereby we signal whether it is a part of φ or a part of ψ in the tableau. We start with the initial set $\varphi^a; (\neg\psi)^c$. The rules are then as follows. The tableau operates on sets of the $X; \varphi$, where X remains unchanged when passing from above the line to below (except for $(\Box E)$). In the marked rule, each formula in the set X inherits its previous mark, while the newly occurring formulae after applying the rule inherit their mark from φ, the formula on which the rule operates. (There is a slight twist here. A set may contain a formula φ that is marked φ^a as well as φ^c. In that case we will treat these two as distinct formulae. Alternatively, we may allow a formula to carry a double mark, i. e. in this case φ^{ac}.) For example, the marked rule $(\vee E)$ looks as follows.

$$\frac{X^a; Y^c; \neg(\varphi \wedge \psi)^c}{X^a; Y^c; (\neg\varphi)^c | X^a; Y^c; (\neg\psi)^c}$$

In $(\Box E)$, a formula $\Box_j\psi$ of $\Box_j X$ has the same mark as the corresponding ψ in X below. This concludes the marking procedure.

Now suppose that we have a closing tableau. We will now define *two* tableaux, one with formulae marked a and i and the other with formulae marked as i and c. The first will be a tableau for $\varphi; \neg\chi$ and the second a dual tableau for $\chi; \neg\psi$. Again, the same rules for marking apply. i stands here for *interpolant*. The construction will be by unzipping the original tableau as follows. Each time we have a set X in the tableau, we show that it is of the form $X = Y^a; Z^c$ and that there is a formula χ^i such that $Y^a; (\neg\chi)^i$ has a closing tableau and $\neg Z^c; (\neg\chi)^i$ has a closing dual tableau. This will yield the desired conclusion; for if $\varphi; \neg\chi$ closes, we get $\varphi \vdash \chi$, and if $\psi; \neg\chi$ dual closes, we get $\vdash \neg\chi \vee \psi$, that is, $\chi \vdash \psi$. In the actual proof we will construct $\neg\chi$ instead of χ, which makes the proof easier to read. So, we will construct a χ such that $\varphi; \chi$ has a closing tableau, and $\neg\psi; \chi$ a closing dual tableau.

A dual tableau for $\neg X$ is isomorphic to a tableau for X^d, the dual of X. So there is an inherent duality in the construction, which we will use extensively. The proof will be by induction. We begin with the leaves. There are four possibilities how a branch can close, either as (aa) $p^a; (\neg p)^a$ or as (ac) $p^a; (\neg p)^c$ or as (ca) $p^c; (\neg p)^a$ or as (cc) $p^c; (\neg p)^c$. In case (aa) we choose $\chi := \top$, in case (ac) we chose $\chi := \neg p$, in case (ca) we choose $\chi := p$, and in case (cc) we put $\chi := \bot$. In each case the claim is directly verified. Now suppose that (w) has been applied.

$$(w) \qquad \frac{X_1^a; X_2^c; Y_1^a; Y_2^c}{X_1^a; X_2^c}$$

By induction hypothesis, there is an interpolant χ for $X_1^a; X_2^c$. It is easy to verify that χ also is an interpolant for $X_1^a; X_2^c; Y_1^a; Y_2^c$. Next, suppose $(\Box E)$ has been applied, say on the antecedent formula $\neg\Box_j\varphi$.

$$(\Box E) \qquad \frac{\Box_j X^a; \Box_j Y^c; (\neg\Box_j\varphi)^a}{X^a; Y^c; (\neg\varphi)^a}$$

By induction hypothesis we have a closing tableau for $X^a; \neg\varphi^a; \chi^i$ and a dual closing tableau for $\chi^i; (\neg Y)^c$. Now consider the following rule applications.

$$\frac{\Box_j X^a; (\neg\Box_j\varphi)^a; (\Box_j\chi)^i}{X^a; (\neg\varphi)^a; \chi^i} \qquad \frac{\dfrac{\Diamond_j\neg Y^c; (\neg\Diamond_j\neg\chi)^i}{\neg Y^c; (\neg\neg\chi)^i}}{\neg Y^c; \chi^i}$$

By assumption, the lower line has a closing tableau; hence the upper line has a closing tableau as well. This shows that $(\Box_j\chi)$ is an interpolant for the premiss of $(\Box E)$. Now suppose that the rule has operated on a consequent formula, $(\neg\Box_j\psi)^c$. Then we have to put as the new interpolating formula the dual formula, $\Diamond_j\chi = \neg\Box_j\neg\chi$. The argumentation now is completely dual. We have

$$\frac{\dfrac{\Box_j X^a; (\neg\Box_j\neg\chi)^i}{X^a; \neg\neg\chi^i}}{X^a; \chi^i} \qquad \frac{\Diamond_j\neg X^c; (\neg\Box_j\psi)^c; (\Diamond_j\chi)^i}{\neg X^c; \neg\psi^c; \chi^i}$$

The lower lines have a closing tableau, and so do therefore the premisses. In the case where $(\neg E)$ and $(\wedge E)$ have applied, we choose for the new interpolant the old one. It is easy to check that this satisfies the requirements. In the first case, if $(\neg E)$ has operated on an a–formula we get

$$\frac{X^a; (\neg\neg\varphi)^a; \chi^i}{X^a; \varphi^a; \chi^i}$$

So, since $X^a; \varphi^a; \chi^i$ has a closing tableau, so has $X^a; (\neg\neg\varphi)^a; \chi^i$. On the dual side, X^c has remained unchanged, so nothing is to be proved. Dually if the rule $(\neg E)$ has applied to a c–formula. In case of $(\wedge E)$ and an a–formula we get

$$\frac{X^a; (\varphi \wedge \psi)^a; \chi^i}{X^a; \varphi^a; \psi^a; \chi^i}$$

By induction hypothesis, the lower set has a closing tableau; thus there is one for the upper set. Since X^c did not change, there is nothing to prove for X^c. If $(\wedge E)$ has applied to a c–formula $(\varphi \wedge \psi)$, the case is exactly dual to an application of $(\vee E)$ to an a–formula. Now, finally, the rule $(\vee E)$. Since we introduce a split, there are now two interpolation formulae, χ_1 and χ_2, one for each branch. Suppose first that an antecedent formula has been reduced. Then we put $\chi := \chi_1 \wedge \chi_2$.

$$\frac{X^a; (\neg(\varphi \wedge \psi))^a; (\chi_1 \wedge \chi)^i}{\dfrac{\dfrac{X^a; (\neg\varphi)^a; (\chi_1 \wedge \chi_2)^i}{X^a; (\neg\varphi)^a; \chi_1^i; \chi_2^i}}{X^a; (\neg\varphi)^a; \chi_1^i} \;\Bigg|\; \dfrac{\dfrac{X^a; (\neg\psi)^a; (\chi_1 \wedge \chi_2)^i}{X^a; (\neg\psi)^a; \chi_1^i; \chi_2^i}}{X^a; (\neg\psi)^a; \chi_2^i}} \qquad \frac{\neg X^c; (\chi_1 \wedge \chi_2)^i}{\neg X^c; \chi_1^i | \neg X^c; \chi_2^i}$$

Both tableaux can be brought to close. The first by the fact that we have chosen χ_i appropriately. The second by the fact that both $\chi_1^i; \neg X^c$ and $\chi_2^i; \neg X^c$ close by induction hypothesis. This concludes the inspection of all rules.

THEOREM 3.8.5. *$\boldsymbol{K}_\kappa$ has local interpolation. Moreover, an interpolant for φ and ψ can be constructed from a closing tableau for $\varphi; \neg\psi$.*

For special logics extending $\mathbf{K}_1$ there exist tableau calculi which allow to construct interpolants. We will display the relevant rules below.

$$(\mathbf{4}.)\quad \frac{\Box X; \neg\Box\varphi}{X; \Box X; \neg\varphi} \qquad\qquad (\mathbf{t}.)\quad \frac{X; \Box\varphi}{X; \varphi}$$

$$(\mathbf{g}.)\quad \frac{\Box X; \neg\Box\varphi}{X; \Box X; \neg\varphi; \Box\varphi} \qquad\qquad (\mathbf{grz}.)\quad \frac{\Box X; \neg\Box\varphi}{X; \Box X; \neg\varphi; \Box(\varphi \to \Box\varphi)}$$

$$(\mathbf{alt}_1.)\quad \frac{\Box X; \neg\Box Y; \neg\Box\varphi}{X; Y; \varphi}$$

It is an easy matter to verify that these rules are sound. Their completeness is harder to verify directly, but follows easily with the help of the reduction sets.

Now, finally, for the promised proof of interpolation for $\mathbf{K}_\kappa$ that does not make use of tableaux. In fact, it will show not only interpolation but a stronger property of $\mathbf{K}_\kappa$ called *uniform interpolation*.

Definition 3.8.6. *Let Λ be a modal logic. Λ has* **local uniform interpolation** *if (i.) given φ and variables $\vec{q}$, there exists a formula χ such that $var(\chi) \subseteq \vec{q}$ and for all formulae ψ such that $var(\varphi) \cap var(\psi) = \vec{q}$ we have $\varphi \vdash_\Lambda \chi \vdash_\Lambda \psi$, and (ii.) given ψ and variables $\vec{q}$, there exists a formula χ such that $var(\chi) \subseteq \vec{q}$ and for all formulae φ such that $var(\varphi) \cap var(\psi) = \vec{q}$ we have $\varphi \vdash_\Lambda \chi \vdash_\Lambda \psi$.*

The property (i.) alone is called **uniform preinterpolation** and the property (ii.) alone **uniform postinterpolation**. By definition, if a logic has uniform preinterpolation the interpolant does not depend on the actual shape of ψ but only on the set of shared variables. Since $\varphi \vdash_\Lambda \psi$ iff $\neg\psi \vdash_\Lambda \neg\varphi$ either of (i.) and (ii.) is sufficient for showing uniform interpolation. We will now show that $\mathbf{K}_\kappa$ has uniform preinterpolation. For simplicity we take the case of a single operator, that is, we prove the statement for $\mathbf{K}_1$. But the generalization is easy enough to make. The central idea is that when we have $\varphi \vdash \psi$, and we have a variable p that occurs in φ but not in ψ we want to simply erase the variable p in φ and define a formula $\varphi^{\top(p)}$ (or simply $\varphi^\top$) such that $\varphi^{\top(p)} \vdash \psi$. Define $\varphi^{\top(p)}$ as follows. An occurrence of the variable p is replaced by $\top$ if it is embedded in an even number negations, and by $\bot$ otherwise. Given this definition it is easily shown that $\varphi \vdash \varphi^\top$. This is one half of what we need; the other half is $\varphi^\top \vdash \psi$. Now here things can go wrong. Take, for example, $\varphi := p \wedge \neg p$ and $\psi := \bot$. Clearly, $\varphi \vdash \psi$. Given the current definition, $\varphi^\top = \top \wedge \neg\bot$. This formula is equivalent to $\top$. Hence $\varphi^\top \nvdash \bot$. To surround this problem, put φ into standard and explicit form. We will show in the next lemma that any model for $\varphi^\top$ based on an intransitive tree is the p–morphic image of a model for φ. Here, an intransitive tree is a frame $\langle f, \lhd\rangle$ such that it contains no cycles with respect to $\lhd$ and in which $x \lhd z$ and $y \lhd z$ implies $x = y$.

Lemma 3.8.7. *Suppose that φ is standard, explicit and clash free. Then for any finite model $\langle \mathfrak{f}, \beta, x\rangle$ for $\varphi^{\top(p)}$ based on an intransitive tree there exists a model $\langle \mathfrak{g}, \gamma, x'\rangle$ for φ such that*

- i $\langle g, \lhd\rangle$ *is an intransitive tree and there exists a contraction $c : \mathfrak{g} \twoheadrightarrow \mathfrak{f}$,*
- ii *for all $q \neq p$, $\beta(q) = c[\gamma(q)]$*
- iii $x = c(x')$.

Proof. By induction on the modal depth of φ. Assume that the depth is 0. Then $\varphi = \bigvee_{i<m} \nu_i$, each ν_i a conjunction of sentence letters or their negations. Then $\varphi^\top = \bigvee_{i<m} \nu_i^\top$. Suppose that $\langle \mathfrak{f}, \beta, w_0\rangle \vDash \varphi^\top$. We assume that β is defined only on the variables of $\varphi^\top$. Then there exists an $i < m$ such that $\langle \mathfrak{f}, \beta, x\rangle \vDash \nu_i^\top$. Now put $\gamma(q) := \beta(p)$ if $q \neq p$ and $\gamma(p) := \{w_0\}$ if p is a conjunct of ν_i and $\gamma(p) := \varnothing$, else. Since φ is clash free, this is well–defined. Then $\langle \mathfrak{f}, \gamma, w_0\rangle \vDash \nu_i$, and so $\langle \mathfrak{f}, \gamma, w_0\rangle \vDash \varphi$. Now let

$dp(\varphi) > 0$. Then φ is a disjunction of formulae φ_i of the form

$$\varphi_i = \mu \wedge \bigwedge_{j<n} \Diamond\psi_j \wedge \Box\chi.$$

Assume further that $\langle \mathfrak{f}, \beta, w_0 \rangle \vDash \varphi^\top$. β is defined only on the variables of $\varphi^\top$. Then for some i we have $\langle \mathfrak{f}, \beta, w_0 \rangle \vDash \varphi_i$. Furthermore,

$$\varphi_i^\top = \mu^\top \wedge \bigwedge_{j<n} \Diamond\psi_j^\top \wedge \Box\chi^\top .$$

Put $\mathfrak{f}^0 := \mathfrak{f}$, $\gamma^0(p) := \{w_0\}$ if p is a conjunct of μ, and $\gamma^0(p) := \varnothing$ otherwise. For $q \neq p$ put $\gamma^0(q) := \beta(q)$. Then $\langle \mathfrak{f}^0, \gamma^0, w_0 \rangle \vDash \mu$. Let the set of successsors of w_0 be $suc(w_0) = \{x_\alpha : \alpha < \lambda\}$. λ is finite. Inductively, for each $\alpha < \lambda$ we perform the following operation. **Case 1.** $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\alpha \rangle \vDash \psi_j^\top$ for some j. Then put $J := \{k : \psi_k \leftrightarrow \psi_j \in \mathbf{K}_\kappa\}$. Let $\mathfrak{g}$ be the transit of x_α in $\mathfrak{f}^\alpha$. By induction hypothesis, there exists a model $\langle \mathfrak{h}_k, \delta_k, y_k \rangle \vDash \psi_k$ for each $k \in J$ and a p–morphism $e_k : \mathfrak{h}_k \twoheadrightarrow \mathfrak{g}$ satisfying (ii) and (iii). Form $\mathfrak{f}^{\alpha+1}$ by blowing up the frame $\mathfrak{g}$ to $\bigoplus_{k\in J} \mathfrak{h}_k$. (See Theorem 3.3.3.) There exists a p–morphism $d^{\alpha+1} : \mathfrak{f}^{\alpha+1} \twoheadrightarrow \mathfrak{f}^\alpha$, obtained by extending $\bigoplus_{k\in J} e_k$ to $f^{\alpha+1}$. Moreover, $\gamma^{\alpha+1}(p) := \gamma^\alpha(p) \cup \bigcup_{k\in J}(d^{\alpha+1})^{-1}[\delta_k(p)]$, and for $q \neq p$, $\gamma^{\alpha+1}(q) := (d^{\alpha+1})^{-1}[\gamma^\alpha(q)]$. **Case 2.** $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\alpha \rangle \nvDash \psi_j^\top$ for all j. Then at least $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\alpha \rangle \vDash \chi^\top$, and we proceed as follows. We know that χ is a disjunction of simple standard, explicit and clash–free formulae τ_i. For some i we have $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\alpha \rangle \vDash (\tau_i)^\top$. Let $\mathfrak{g}$ be the transit of x_α in $\mathfrak{f}$. By induction hypothesis there exists a $\mathfrak{h}$ and a p–morphism $e : \mathfrak{h} \twoheadrightarrow \mathfrak{g}$, a δ^0 and y satisfying (ii) and (iii) such that $\langle \mathfrak{h}, \delta^0, y \rangle \vDash \tau_i$, and so $\langle \mathfrak{h}, \delta^0, y \rangle \vDash \chi$. Now blow up $\mathfrak{g}$ to $\mathfrak{h}$ in $\mathfrak{f}$. This defines $\mathfrak{f}^{\alpha+1}$. There exists a p–morphism $d^{\alpha+1} : \mathfrak{f}^{\alpha+1} \twoheadrightarrow \mathfrak{f}^\alpha$. Put $\gamma^{\alpha+1}(p) := \gamma^\alpha(p) \cup (d^{\alpha+1})^{-1}[\gamma^\alpha(p)]$, and let $\gamma^{\alpha+1}(q) := (d^{\alpha+1})^{-1}[\gamma^\alpha(q)]$ for $q \neq p$.

It is clear that $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\beta \rangle \vDash \psi_j$ iff $\langle \mathfrak{f}^{\alpha+1}, \gamma^{\alpha+1}, x_\beta \rangle \vDash \psi_j$ for all $\beta < \alpha$, and $\langle \mathfrak{f}^\alpha, \gamma^\alpha, x_\beta \rangle \vDash (\psi_j)^\top$ iff $\langle \mathfrak{f}^{\alpha+1}, \gamma^{\alpha+1}, x_\beta \rangle \vDash (\psi_j)^\top$ for $\beta > \alpha$. Furthermore, $\langle \mathfrak{f}^{\alpha+1}, \gamma^{\alpha+1}, w_0 \rangle \vDash \mu$. Put $\mathfrak{g} := \mathfrak{f}^\lambda$. Then the composition $c := d^1 \circ d^2 \circ \ldots \circ d^\lambda : \mathfrak{g} \twoheadrightarrow \mathfrak{f}$. For $q \neq p$, $\gamma(q) := \gamma^\lambda(q)$, is the result of blowing up β by c. Moreover, $c(w_0) = w_0$, and

$$\langle \mathfrak{g}, \gamma, w_0 \rangle \vDash \varphi_i ,$$

which had to be shown. ∎

Theorem 3.8.8. *$\mathbf{K}_\kappa$ has uniform local interpolation.*

Proof. Suppose that $\varphi(\vec{p}, \vec{q}) \nvdash \psi(\vec{q}, \vec{r})$. Let $\varphi_0(\vec{p}, \vec{q})$ be a standard, explicit and clash–free formula deductively equivalent to φ. Let $\vec{p}$ consist of the variables p_i, $i < n$. Now for $i < n$, let $\varphi_{i+1}(\vec{p}, \vec{q}) := \varphi_i(\vec{p}, \vec{q})^{\top(p_i)}$. $\chi := \varphi_n$. Then the variables p_i do not occur in χ. Moreover, $\varphi \vdash \chi$. Now assume that $\chi \nvdash \psi$. Then there exists a finite intransitive tree $\mathfrak{f}_0$, a valuation β_0 and a w_0 such that $\langle \mathfrak{f}_0, \beta_0, w_0 \rangle \vDash \chi; \neg\psi$. By the previous lemma, if $\langle \mathfrak{f}_i, \beta_i, w_0 \rangle \vDash \varphi_{n-i}; \neg\psi$ and $\mathfrak{f}_i$ is an intransitive tree then there exists a model $\langle \mathfrak{f}_{i+1}, \beta_{i+1}, w_0 \rangle \vDash \varphi_{n-i-1}; \neg\psi$ such that $\mathfrak{f}_{i+1}$ is an intransitive tree. Hence, by induction, we have $\langle \mathfrak{f}_n, \beta_n, w_0 \rangle \vDash \varphi_0; \neg\psi$. This means that $\varphi \nvdash \psi$, since φ_0 is deductively equivalent with φ. This contradicts the fact that $\varphi \vdash \psi$. ∎

It should be noted that the assumption that $\mathbf{K}_\kappa$ is complete with respect to finite intransitive trees is rather essential for the proof method. The method of reduction sets cannot be applied to show uniform interpolation for the standard systems. The notion of uniform interpolation has been introduced by Andrew Pitts [**161**]. It has been shown subsequently by Silvio Ghilardi and M. Zawadowski [**74**] and Albert Visser [**223**] that **K**, **Grz** and **G** have uniform interpolation, but that **S4** lacks uniform interpolation. Furthermore, Frank Wolter [**243**] proves that uniform interpolation is preserved under fusion. From the latter follows already that polymodal **K** has uniform interpolation if only **K** has uniform interpolation.

Notes on this section. The notion of a downward saturated set first appeared in Hintikka [**104**], who gave a tableau calculus for **K4** and other systems. Tableau calculi have attracted much interest in machine based theorem proving. The literature is too large to be adequately summarized here. Suffice to mention here the work by Melvin Fitting [**70**], Rajeev Goré [**88**] and Martin Amerbauer [**1**]. Furthermore, [**224**] contains an overview of proof theory in modal logic. Tableau calculi are closely connected to Gentzen–calculi. A Gentzen calculus operates on pairs of sets of formulae, $\langle \Gamma, \Delta \rangle$, written $\Gamma \vdash \Delta$. It is possible to reformulate a Gentzen calculus as a tableau calculus. The method for proving interpolation employed above is quite similar to the one introduced by S. Maehara in [**144**].

Exercise 122. Show that the rule (**4**.) is sound and complete for **K4**. *Hint.* Use the reduction sets for **K4**.

Exercise 123. Show that the rule (**t**.) is sound and complete for **K.T**, and that the rules (**t**.) and (**4**.) together are sound and complete for **S4**.

Exercise 124. Show that (**g**.) is sound and complete for **G**, and that (**grz**.) is sound and complete for **Grz**. *Hint.* You have to use two reductions in succession here, first one to **K4** and then one to **K**.

Exercise 125. Show that ($\mathbf{alt}_1$.) is sound and complete for $\mathbf{K}.\mathbf{alt}_1$.

Exercise 126. Show that $\varphi \vdash_{\mathbf{K}} \varphi^{\top(p)}$.

Exercise 127. (Wolter [**243**].) Show that if $\mathsf{Alg}\,\Lambda$ is a locally finite variety and Λ has interpolation, then Λ has uniform interpolation. It follows, for example, that **S5** has uniform interpolation. *Hint.* The notion of a locally finite variety is defined in Section 4.8. You may work with the following characterization: $\mathsf{Alg}\,\Lambda$ is locally finite iff $\mathrm{Can}_\Lambda(n)$ is finite for every $n \in \omega$.

Exercise 128. Show that **G** and **Grz** are in PSPACE. *Hint.* Show that the length of a branch for in a tableau for φ is bounded by the number of subformulae of φ.

Exercise 129. Here is a tableau calculus for determining whether or not $\Delta \Vdash_{\mathbf{K}_\kappa} \varphi$. Introduce a formula diacritic $^\spadesuit$. Now replace the rule $(\Box_j E)$ by

$$\frac{\neg\Box_j\varphi; \Box_j X; Y^\spadesuit}{\neg\varphi; X; Y^\spadesuit}$$

Furthermore, add the rule

$$\frac{X; Y^\spadesuit}{X; Y^\spadesuit; Y}$$

Show that $\Delta \Vdash_{\mathbf{K}_\kappa} \varphi$ iff $\Delta^\spadesuit; \varphi$ has a closing tableau. Show then that the problem '$\psi \Vdash_{\mathbf{K}_\kappa} \varphi$?' is in EXPTIME.

3.9. Modal Consequence Relations

In this section we will study the lattice of all consequence relations. Thereby we will also elucidate the role of the consequence relations $\vdash_\Lambda$ and $\Vdash_\Lambda$. We start by defining the notion of a *modal consequence relation* and give some alternative characterizations.

Definition 3.9.1. *A* ***modal consequence relation*** *is a consequence relation over $\mathcal{P}_\kappa$ which contains at least the rule (mp.) and whose set of tautologies is a modal logic. If $\vdash$ is a modal consequence relation and $\Lambda := \mathsf{Taut}(\vdash)$ then $\vdash$ is a* ***modal consequence relation for*** *Λ. $\vdash$ is* ***normal*** *(****quasi–normal****,* ***classical****,* ***monotone****) if Λ is.*

In sequel we will deal exclusively with normal consequence relations. Notice that if $\vdash \subseteq \wp(\mathcal{P}_\kappa) \times \mathcal{P}_\kappa$ is a consequence relation, then we may define $\vdash^\lambda := \vdash \cap (\wp(\mathcal{P}_\lambda) \times \mathcal{P}_\lambda)$ if $\lambda \leq \kappa$, and if $\lambda > \kappa$ let $\vdash^\lambda$ denote the least consequence relation over $\mathcal{P}_\lambda$ containing $\vdash$. A special case is $\lambda = 0$.

Proposition 3.9.2. *Let $\vdash$ be a consistent modal consequence relation. Then its reduct to the language $\top$, $\neg$ and $\wedge$ is the consequence relation of boolean logic.*

Proof. $\mathsf{Taut}(\vdash^0)$ contains all boolean tautologies. For we have

$$\mathsf{Taut}(\vdash^0) = \mathsf{Taut}(\vdash) \cap \mathcal{P}_0 \ .$$

The rule of modus ponens is contained in $\vdash^0$. Therefore, $\vdash^0$ contains the consequence relation $\vdash_\mathbf{2}$. If $\vdash$ is consistent, we have $p \nvdash q$, so $p \nvdash^0 q$, and so $\vdash^0$ is consistent too. However, $\vdash_\mathbf{2}$ is maximal, and so equal to $\vdash^0$. ◻

If $\vdash$ is a normal consequence relation, we denote by $\mathcal{Q}(\vdash)$ the lattice of extensions of $\vdash$ and by $\mathcal{E}(\vdash)$ the lattice of normal extensions. We will show below that this is a complete and algebraic lattice. Now let Λ be a modal logic. Then define

$$\mathit{CRel}(\Lambda) := \{\vdash : \mathsf{Taut}(\vdash) = \Lambda\}$$

As we have seen in Section 1.4 of Chapter 1, $\mathit{CRel}(\vdash)$ is an interval with respect to inclusion. The smallest member is in fact $\vdash_\Lambda$, as follows immediately from the definition. The largest member will be denoted by $\vdash_\Lambda^m$; it is structurally complete. As is clear, for a given set of tautologies there exists exactly one structurally complete logic with the given tautologies, and at most one logic with a deduction theorem for $\rightarrow$. $\vdash^m_\Lambda$ is the structurally complete logic with respect to Λ and $\vdash_\Lambda$ is the logic with a deduction theorem for $\rightarrow$.

Proposition 3.9.3. *Let Λ be a modal logic. Then*

$$CRel(\Lambda) = \{\vdash : \vdash_\Lambda \subseteq \vdash \subseteq \vdash^m_\Lambda\}$$

Moreover, $\vdash_\Lambda$ is the unique member of $CRel(\Lambda)$ having a deduction theorem for $\rightarrow$ and $\vdash^m_\Lambda$ the unique member which is structurally complete.

To see some more examples, consider the rule $\langle\{\Box p\}, p\rangle$. It is admissible in **K**. For assume that $\varphi := p^\sigma$ is not a theorem. Then there exists a model $\langle \mathfrak{f}, \beta, x\rangle \vDash \neg\varphi$. Consider the frame $\mathfrak{g}$ based on $f \cup \{z\}$, where $x \notin f$, and the relation $\blacktriangleleft := \lhd \cup \{\langle z, y\rangle : y \in f\}$. Take $\gamma(p) := \beta(p)$. Then $\langle \mathfrak{g}, \gamma, z\rangle \vDash \neg\Box\varphi$. We warn the reader here that even though for any modal consequence relation, $\Box p \vdash p$ is equivalent to $p \vdash \Diamond p$, the rule $\langle\{p\}, \Diamond p\rangle$ is not admissible in **K** despite the admissibility of $\langle\{\Box p\}, p\rangle$. Take $p := \top$. $\Diamond\top$ is not a theorem of **K**. Similarly, the so–called MacIntosh rule $\langle\{p \rightarrow \Box p\}, \Diamond p \rightarrow p\rangle$ is not admissible for **K**. Namely, put $p := \Box\bot$. $\Box\bot \rightarrow \Box\Box\bot$ is a theorem but $\Diamond\Box\bot \rightarrow \Box\bot$ is not. Notice also that if a rule ρ is admissible in a logic Θ we may not conclude that ρ is admissible in every extension of Θ. A case in point is the rule $\langle\{\Box p\}, p\rangle$, which is not admissible in $\mathbf{K} \oplus \Box\bot$.

Recall the notation $\vdash^R$. This denotes the consequence relation generated by the rules R. At present we may tacitly assume that R contains (mp.). Equivalently, $\vdash^R$ is the least modal consequence relation containing R. Notice that for every modal consequence relation $\vdash$ there exists an R with $\vdash \ = \ \vdash^R$ (for example the set of all finitary rules of $\vdash$ itself).

Proposition 3.9.4. *The set of modal consequence relations over $\mathcal{P}_\kappa$ forms an algebraic lattice. The compact elements are exactly the finitely axiomatizable consequence relations. The lattice of quasi–normal consequence relations is the sublattice of consequence relations containing $\vdash_{\mathbf{K}_\kappa}$.*

Proof. Clearly, the operation $\bigsqcap$ is set intersection, and $\vdash_1 \sqcup \vdash_2$ is the smallest consequence relation containing both $\vdash_1$ and $\vdash_2$. If $\vdash_1 \ = \ \vdash^{R_1}$ and $\vdash_2 \ = \ \vdash^{R_2}$ then $\vdash_1 \sqcup \vdash_2 \ = \ \vdash^{R_1 \cup R_2}$. With this latter characterization it is easy to define the infinite union. Namely, if $\vdash_i \ = \ \vdash_{R_i}$ for $i \in I$, put $\bigsqcup_I \vdash_i := \vdash^S$, where $S := \bigcup_I R_i$. All rules are finitary by definition. Therefore, if a rule is derivable in $\vdash^S$, then it is derivable already from a finite union of the $\vdash_i$. It follows that a finitely axiomatizable consequence relation is compact, and that a compact consequence relation is finitely axiomatizable. Moreover, the lattice is algebraic, since $\vdash^R = \bigsqcup_{\rho\in R} \vdash^\rho$. The last claim is a consequence of the fact that $\vdash'$ is quasi–normal iff $\mathsf{Taut}(\vdash')$ is quasi–normal iff $\mathsf{Taut}(\vdash')$ contains $\mathbf{K}_\kappa$. ∎

Proposition 3.9.5. *For each quasi–normal logic* Λ *and each quasi–normal consequence relation* $\vdash'$,

$$\vdash_\Lambda \subseteq \vdash' \qquad \Leftrightarrow \qquad \Lambda \subseteq \mathsf{Taut}(\vdash')$$

$\mathsf{Taut}(-)$ *commutes with infinite intersections,* $\vdash_\Lambda$ *with infinite intersections and joins.*

Proposition 3.9.6. $\mathcal{E}(\vdash_{\mathbf{K}_\kappa})$ *is a complete sublattice of* $\mathcal{Q}(\vdash_{\mathbf{K}_\kappa})$.

Proposition 3.9.7. *In monomodal logic,* $\vdash_\Lambda$ *is maximal iff* Λ *is a coatom.*

Proof. Clearly, if $\vdash_\Lambda$ is maximal in $\mathcal{E}(\vdash_{\mathbf{K}})$, Λ must be a coatom. To show the converse, we need to show is that for a maximal consistent normal logic Λ, $\vdash_\Lambda$ is structurally complete. (It will follow that *CRel*(Λ) has exactly one element.) Now, Λ is Post–complete iff it contains either the formula $\Box\top$ or the formula $p \leftrightarrow \Box p$. Assume that $\vdash_\Lambda$ can be expanded by a rule $\rho = \langle\Delta, \varphi\rangle$. Then, by using the axioms ρ can be transformed into a rule $\rho' = \langle\Delta', \varphi'\rangle$ in which the formulae are nonmodal. (Namely, any formula in a rule may be exchanged by a deductively equivalent formula. Either $\Box\top \in \Lambda$ and any subformula $\Box\chi$ may be replaced by $\top$, or $p \leftrightarrow \Box p \in \Lambda$ and then $\Box\chi$ may be replaced by χ.) A nonmodal rule not derivable in $\vdash_\Lambda$ is also not derivable in its boolean fragment, $\vdash^0_\Lambda$. By the maximality of the latter, adding ρ' yields the inconsistent logic. ∎

In polymodal logics matters are a bit more complicated. We will see that there exist in fact $2^{\aleph_0}$ logics which are coatoms in $\mathcal{E}\,\mathbf{K}_2$ without their consequence relation being maximal. Moreover, we will see that even in monomodal logics there exist $2^{\aleph_0}$ maximal consequence relations, which are therefore not of the form $\vdash_\Lambda$ (except for the two abovementioned consequence relations). Notice that even though a consequence is maximal iff it is structurally complete and Post–complete, Post–completeness is relative to the derivable rules. Therefore, this does *not* mean that the tautologies are a maximally consistent modal logic. We define the **T–spectrum** of Λ, *TSp*(Λ), to be the cardinality of the set of consequence relations whose set of tautologies is Λ.

$$TSp(\Lambda) := card\, CRel(\Lambda)$$

To characterize the choices for the T–spectrum we will first deal with a seemingly different question.

Proposition 3.9.8. *Let* Λ *be a normal logic. Then the following are equivalent.*

(1) $\vdash_\Lambda = \Vdash_\Lambda$.
(2) $\Vdash_\Lambda$ *admits a deduction theorem for* $\rightarrow$.
(3) $\Lambda \supseteq \boldsymbol{K}_\kappa \oplus \{p \rightarrow \Box_j p : j < \kappa\}$.
(4) Λ *is the logic of some set of Kripke–frames containing only one world.*

Proof. Clearly, if (1.) holds, then (2.) holds as well. Now let (2.) be the case. Then since $p \Vdash_\Lambda \Box_j p$, by the deduction theorem, $\Vdash_\Lambda p \rightarrow \Box_j p$, which gives (3.). From (3.) we deduce (1.) as follows. Since $p; p \rightarrow \Box_j p \vdash_\Lambda \Box_j p$, and $p \rightarrow \Box_j p \in \Lambda$, we get $p \vdash_\Lambda \Box_j p$. Therefore, $\Vdash_\Lambda = \vdash_\Lambda$. Finally we establish the equivalence of (3.) and

(4.). Assume (4.). Then clearly the formulae $p \to \Box_j p$ are axioms, since they hold on any one–point frame. Now assume that (4.) fails. Consider a rooted generated subframe $\mathfrak{F}$ of $\mathfrak{Can}_\Lambda(var)$ consisting of more than one point. Such a frame must exist, by assumption. Let X be the root and Y a j–successor. Then there exists a set $\widehat{\varphi}$ such that $X \in \widehat{\varphi}$ but $Y \notin \widehat{\varphi}$. Now put $\beta(p) := \widehat{\varphi}$. It follows that $\langle \mathfrak{F}, \beta, X\rangle \vDash p; \neg\Box_j p$. Hence (3.) fails as well. ∎

Now let us return to the question of T–spectra. Clearly, if $\vdash_\Lambda \neq \Vdash_\Lambda$ then the T–spectrum of Λ cannot be 1. We will show now that the converse almost holds.

Proposition 3.9.9. *Let Λ be a modal logic. Then the following are equivalent.*

(1) *The T–spectrum of Λ is 1.*
(2) *$\vdash_\Lambda$ is structurally complete.*
(3) *Λ is the logic of a single Kripke–frame containing a single world.*
(4) *Λ is a fusion of monomodal logics of the frames $\boxed{\bullet}$ or $\boxed{\circ}$.*

Proof. The equivalence between (1.) and (2.) is immediate. The equivalence of (3.) and (4.) is also not hard to show. If Λ is a fusion of logics for one–point frames it contains for each operator either the axiom $\Box_j\top$ or $p \leftrightarrow \Box_j p$. It means that the relation $\lhd_j$ is on all frames empty or on all frames the diagonal. Hence the generated subframes of the canonical frame are one–point frames and they are all isomorphic. Finally, we show (2.) $\Leftrightarrow$ (3.). Assume (3.). Then by the fact that the $\vdash_\Lambda$ is the logic of a single algebra based on two elements, and has all constants, it is structurally complete. Now let (3.) fail. There are basically two cases. If Λ is not the logic of one–point frames, then $\vdash_\Lambda$ is anyway not structurally complete by the previous theorem. Otherwise, it is the intersection of logics determined by matrices of the form $\langle\mathfrak{A}, D\rangle$, D an open filter, $\mathfrak{A}$ the free algebra in $\aleph_0$ generators. (In fact, the freely 0–generated algebra is enough.) $\mathfrak{A}$ contains a constant c such that $0 < c < 1$. Namely, take two different one point frames. Then, say, $\Box_0$ is the diagonal on one frame and empty on the other. Then $c := \blacklozenge_0 1$ is a constant of the required form. The rule $\langle\{\Diamond_0\top\}, p\rangle$ is admissible but not derivable. ∎

The method of the last proof can be used in many different ways.

Lemma 3.9.10. *Let Λ be a logic and χ a constant formula such that neither χ not $\neg\chi$ are inconsistent. Then the rule $\rho[\chi] := \langle\{\chi\}, \bot\rangle$ is admissible for Λ but not derivable in $\vdash_\Lambda$.*

Proof. Since $\chi \notin \Lambda$ and $var(\chi) = \varnothing$, for no substitution σ, $\chi^\sigma \in \Lambda$. Hence the rule $\rho[\chi]$ is admissible. If it is derivable in $\vdash_\Lambda$ then $\vdash_\Lambda \chi \to \bot$, by the deduction theorem. So $\neg\chi \in \Lambda$, which is not the case. So, $\rho[\chi]$ is not derivable. ∎

Theorem 3.9.11. *Let Λ be a logic such that $\mathfrak{Fr}_\Lambda(0)$ has infinitely many elements. Then $TSp(\Lambda) = 2^{\aleph_0}$.*

Proof. Let $A \subseteq Fr_\Lambda(0)$. Call A a **block** if $0, 1 \notin A$ and for every $a, b \in A$, if $a \neq b$ then $a \cap b = 0$. For each $a \in A$ there exists a constant formula χ_a whose value in $\mathfrak{Fr}_\Lambda(0)$ is a. So, for a subset $S \subseteq A$, put

$$\vdash^S_\Lambda := \vdash_\Lambda \sqcup \bigsqcup_{a \in S} \vdash^{\rho[\chi_a]}$$

We claim that (1) $\vdash^S_\Lambda \in T(\vdash_\Lambda)$, (2) if $S \neq T$ then $\vdash^S_\Lambda = \vdash^T_\Lambda$, and (3) $\mathfrak{Fr}_\Lambda(0)$ contains a block of size $\aleph_0$. **Ad (1).** By the previous lemma, all the rules $\rho[\chi_a]$ are admissible, by the requirement that $a \in A$ and A is a block. **Ad (2).** Suppose that $a \in S - T$. Then let U be the closure under (mp.) of χ_a in Λ. We claim that U is a theory of $\vdash^T_\Lambda$ but not of $\vdash^S_\Lambda$. To be a theory of $\vdash^R_\Lambda$ for some set R no more is required than it be closed under (mp.) and consistent if it is does not contain any χ_a, $a \in R$, or else be inconsistent (and contain all formulae). Now since $a \notin T$ and χ_a is consistent, the (mp.) closure does not contain any χ_b, $b \in T$, since $a \cap b = 0$ (which means that $\chi_a \vdash_\Lambda \neg\chi_b$). Since $a \in S$, $\chi_a \vdash^S_\Lambda \bot$. **Ad (3).** We distinguish three cases. **Case A.** $\mathfrak{Fr}_\Lambda(0)$ has infinitely many atoms. Then the atoms form a block of size at least $\aleph_0$. **Case B.** $\mathfrak{Fr}_\Lambda(0)$ has no atoms. Then there exists a sequence $\langle c_i : i \in \omega \rangle$ such that $0 < c_{i+1} < c_i < 1$ for all $i < \omega$. Put $a_i := c_i - c_{i+1}$. Then $0 < a_i$ since $c_i > c_{i+1}$ and $a_i < 1$ since $c_i < 1$. Furthermore, let $i < j$. Then $a_i \cap a_j = (c_i \cap -c_{i+1}) \cap (c_j \cap -c_{j+1}) = c_j \cap -c_{i+1} \leq c_{i+1} \cap -c_{i+1} = 0$. So there exists an infinite block. **Case C.** There exist finitely many atoms. Then let $\mathfrak{C}$ be the boolean algebra underlying $\mathfrak{Fr}_\Lambda(0)$. We claim that $\mathfrak{C} \cong \mathfrak{A} \times \mathfrak{B}$, where $\mathfrak{A}$ is finite and $\mathfrak{B}$ is atomless. (This is left as an exercise.) Now $\mathfrak{B}$ contains an infinite block, $\langle b_i : i \in \omega \rangle$. Put $a_i := \langle 1, b_i \rangle$. The set $A := \{a_i : i \in \omega\}$ is a block in $\mathfrak{A} \times \mathfrak{B}$. ◻

Corollary 3.9.12. *Let Λ be a monomodal logic and $\Lambda \subseteq$ **G.3**. Then $TSp(\Lambda) = 2^{\aleph_0}$.*

Proof. **G.3** has infinitely many distinct constants, namely $\Box^n\bot$, $n \in \omega$. This applies as well to Λ. ◻

Let us remain with **G.3** a little bit. Consider the consequence $\vdash^m_{\mathbf{G.3}}$. We claim that it is maximal. To see this we need to show that it is Post–complete. This follows from a general fact that we will establish here.

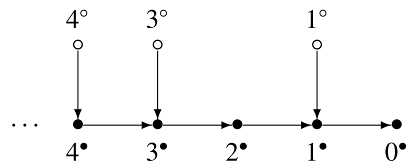
Theorem 3.9.13. *Let Λ be 0–characterized. Then $\vdash^m_\Lambda$ is maximal.*

Proof. Let $\Vdash \supsetneq \vdash^m_\Lambda$. Then $\mathsf{Taut}(\Vdash) \supsetneq \Lambda$. Since Λ is 0–characterized, there is a constant χ such that $\Lambda \subsetneq \Lambda \oplus \chi \subseteq \mathsf{Taut}(\Vdash)$. Therefore, $\chi \notin \Lambda$. Two cases arise. **Case 1.** $\neg\chi \notin \Lambda$. Then the rule $\rho[\chi]$ is admissible in Λ and so derivable in $\vdash^m_\Lambda$. Therefore $\rho[\chi] \in \Vdash$, and so since $\Vdash \chi$, also $\Vdash \bot$. So, $\Vdash$ is inconsistent. **Case 2.** $\neg\chi \in \Lambda$. Then $\mathsf{Taut}(\Vdash)$ is inconsistent. So $\Vdash$ is inconsistent as well. ◻

This theorem makes the search for maximal consequence relations quite easy. Let us note in passing that there are consequences relations $\vdash_1$ and $\vdash_2$ such that

$$\mathsf{Taut}(\vdash_1 \sqcup \vdash_2) \neq \mathsf{Taut}(\vdash_1) \sqcup \mathsf{Taut}(\vdash_2) .$$

FIGURE 3.2. $\mathfrak{T}_M$, $M = \{1, 3, 4, \ldots\}$



Namely, let $\vdash_1 := \vdash^m_{\mathbf{G.3}}$ and $\vdash_2 := \vdash_{\mathbf{K}\oplus\Box\bot}$. Then $\mathsf{Taut}(\vdash_1 \sqcup \vdash_2) = \mathbf{K} \oplus \bot$, but $\mathbf{G.3} \sqcup \mathbf{K} \oplus \Box\bot = \mathbf{K} \oplus \Box\bot$.

We will now investigate the cardinality of the set of coatoms in $\mathcal{E}(\vdash_{\mathbf{K}})$. We will show that there are exactly $2^{\aleph_0}$. Of course, showing that there are at least that many is enough; there cannot be more. In the light of the previous theorem, we are done if we can find $2^{\aleph_0}$ distinct logics which are 0–characterized. Let $M \subseteq \omega$. Put $T_M := \{n^\bullet : n \in \omega\} \cup \{n^\circ : n \in M\}$. Put

$$x \lhd y \quad \Leftrightarrow \quad \begin{cases} \quad\;\; (1.) & x = m^\bullet, y = n^\bullet \text{ and } m > n \\ \text{or } (2.) & x = m^\circ, y = n^\bullet \text{ and } m \geq n \\ \text{or } (3.) & x = m^\circ, y = n^\circ \text{ and } m = n \end{cases}$$

Let $\mathbb{T}_M$ be the algebra of 0–definable sets. Put $\mathfrak{T}_M := \langle T_M, \lhd, \mathbb{T}_M\rangle$. We will show now that if $M \neq N$ then $\mathsf{Th}\,\mathfrak{T}_M \neq \mathsf{Th}\,\mathfrak{T}_N$. To see this, we show that every one–element set $\{n^\circ\}$ in T_M is definable by a formula $\chi(n)$ that depends only on n, not on M. First, take the formula

$$\delta(n) := \Box^{n+1}\bot \wedge \neg\Box^n\bot$$

$\delta(n)$ defines the set $\{n^\bullet\}$. Now put

$$\chi(n) := \Diamond\delta(n) \wedge \neg\delta(n+1) \wedge \neg\Diamond\delta(n+1)$$

It is easily checked that $\chi(n)$ defines $\{n^\circ\}$. Hence, if $n \notin M$, $\neg\chi(n) \in \mathsf{Th}\,\mathfrak{T}_M$. So, $\neg\chi(n) \in \mathsf{Th}\,\mathfrak{T}_M$ iff $n \notin M$. This establishes that if $M \neq N$, $\mathsf{Th}\,\mathfrak{T}_M \neq \mathsf{Th}\,\mathfrak{T}_N$.

THEOREM 3.9.14. *The lattice of normal monomodal consequence relations contains* $2^{\aleph_0}$ *many coatoms.*

Notes on this section. In contrast to the theory of modal logics, the theory of modal consequence relations is not very well developed. Nevertheless, there has been significant progress in the understanding of consequence relations, notably through the work of VLADIMIR RYBAKOV. In a series of papers (see [**178**], [**179**], [**180**], [**181**] as well as the book [**182**] and references therein) he has investigated the question of axiomatizing $\vdash^m_\Theta$ by means of a rule basis. A major result was the solution of a problem by HARVEY FRIEDMAN to axiomatize the calculus of admissible rules for intuitionistic logic. In the more philosophically oriented literature, certain special rules have been in the focus of attention. Typically, the notion of a rule is somewhat more general. It is a pair $\langle \Delta, \Gamma \rangle$, where Δ and Γ are sets of formulae. It is **admissible** for a logic Λ, if for every substitution σ such that $\Delta^\sigma \subseteq \Lambda$ we have $\Gamma^\sigma \cap \Lambda \neq \varnothing$. Examples are the rule of margins

$$\langle \{p \to \Box p\}, \{p, \neg p\} \rangle$$

the MacIntosh rule (which is of course also a rule in our sense), the (strong) rules of disjunction

$$\langle \{\bigvee_{i<n} \Box p_i\}, \{p_i : i < n\} \rangle$$

and the weak rules of disjunction

$$\langle \{\bigvee_{i<n} \boxplus p_i : \boxplus \text{ compound}\}, \{p_i : i < n\} \rangle$$

The latter are nonfinitary rules. See for example work by Brian Chellas and Krister Segerberg [**46**] and Timothy Williamson [**227**], [**226**] and [**228**].

Exercise 130. Show Proposition 3.9.5.

Exercise 131. Let $\Lambda \supseteq \mathbf{K}_\kappa \oplus \{p \to \Box_j p : j < \kappa\}$. Show that Λ is the logic of its 0–generated algebra. It follows that it has all constants. However, $\vdash_\Lambda$ is not necessarily structurally complete. Can you explain why?

Exercise 132. Let $\mathfrak{A}$ be a boolean algebra with finitely many atoms. Show that $\mathfrak{A} \cong \mathfrak{B} \times \mathfrak{C}$, where $\mathfrak{B}$ is finite and $\mathfrak{C}$ is atomless.

Exercise 133. Show that the rule(s) of disjunction are admissible for **K**. *Hint.* Start with models refuting $\varphi_i := p_i^\sigma$. Now build a model refuting $\bigvee_{i<n} \Box\varphi_i$.

Part 2

The General Theory of Modal Logic

CHAPTER 4

Universal Algebra and Duality Theory

4.1. More on Products

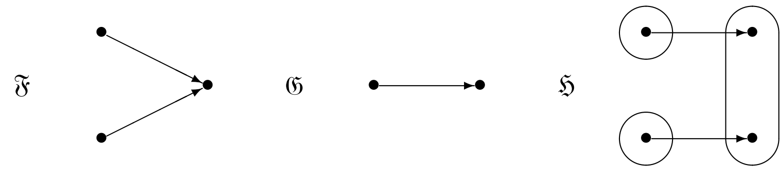
In this chapter we will develop the algebraic theory of modal algebras, taking advantage of some strong theorems of universal algebra. First, we know from a theorem by Garreth Birkhoff that equational classes correspond one–to–one to varieties, and that the lattice of modal logics is dual to the lattice of varieties. Second, by using the representation theory of boolean algebras by Marshall H. Stone we can derive many useful results about general frames, in particular deriving a theorem about modally definable classes of Kripke–frames. Fuller expositions on universal algebra can be found in [**37**], [**89**].

We have to begin by talking more about *products*. A generalization of the (direct) product of algebras is the so–called *subdirect product*. We call $\mathfrak{A}$ a **subdirect product** of the $\mathfrak{B}_i$, $i \in I$, if $\mathfrak{A}$ is a subalgebra of the direct product $\prod_{i\in I} \mathfrak{B}_i$ such that for each projection $\pi_i : \prod_{i\in I} \mathfrak{B}_i \twoheadrightarrow \mathfrak{B}_i$ we have $\pi_i[A] = B_i$. In other words, if $\mathfrak{A}$ is projected onto any factor of the product, we get the full factor rather than a proper subalgebra. Moreover, also every algebra isomorphic to $\mathfrak{A}$ will be called a subdirect product of the $\mathfrak{B}_i$. An alternative characterization of this latter, broader notion is the following. $\mathfrak{C}$ is a **subdirect product** of the $\mathfrak{B}_i$, $i \in I$, if there exists an embedding $j : \mathfrak{C} \rightarrowtail \prod_{i\in I} \mathfrak{B}_i$ such that for every $i \in I$, $\pi_i \circ j : \mathfrak{C} \twoheadrightarrow \mathfrak{B}_i$. To see a nontrivial example of a subdirect product, take the algebra $\mathfrak{A}$ of the frame $\mathfrak{F}$ in Figure 4.1. It is a subdirect product of the algebra $\mathfrak{B}$ of the frame $\mathfrak{G}$; simply take the direct product $\mathfrak{B} \times \mathfrak{B}$, which is isomorphic to the algebra over $\mathfrak{G} \oplus \mathfrak{G}$. Now take the subalgebra $\mathfrak{C}$ generated by the encircled sets. It is isomorphic to the algebra $\mathfrak{A}$. (The sets define the frame $\mathfrak{H}$.) An algebra is called **subdirectly irreducible** (**s.i.**) if for every subdirect product $h : \mathfrak{A} \rightarrowtail \prod_{i\in I} \mathfrak{B}_i$ we have that $\pi_i \circ h : \mathfrak{A} \twoheadrightarrow \mathfrak{B}_i$ is an isomorphism for some $i \in I$. There are some useful theorems on subdirect products and subdirect irreducibility. Recall that there is a smallest congruence on $\mathfrak{A}$, $\Delta_A = \{\langle a, a\rangle : a \in A\}$, also denoted by Δ, and a largest congruence $\nabla_A = A \times A$, denoted by ∇ when no confusion arises. The congruences of an algebra form an algebraic lattice.

Proposition 4.1.1. *Let $\mathfrak{A}$ be a subdirect product of the algebras $\mathfrak{B}_i$, $i \in I$. Let $\pi_i : \prod_{i\in I} \mathfrak{B}_i \to \mathfrak{B}_i$ be the projections. Then we have*

$$\bigcap_{i\in I} ker(\pi_i \restriction A) = \Delta_A \ .$$

FIGURE 4.1. A subdirect product



Conversely, if $\Delta_A = \bigcap_{i\in I} \Theta_i$ *then* $\mathfrak{A}$ *is a subdirect product of the algebras* $\mathfrak{A}/\Theta_i$, $i \in I$. *The embedding of* $\mathfrak{A}$ *into* $\prod_{i\in I} \mathfrak{A}/\Theta_i$ *is defined by the map*

$$h(a) := \langle [a]\Theta_i : i \in I\rangle \ .$$

PROOF. An element of the direct product is a sequence $a = \langle a(i) : i \in I\rangle$. For two such elements we deduce from $\langle a, b\rangle \in \bigcap_{i\in I} ker(\pi_i \restriction A)$ that $\langle a, b\rangle \in ker(\pi_i \restriction A)$ for all $i \in I$, that is, $a(i) = b(i)$ for all i. Hence $a = b$. This proves the first claim. For the second observe first that h is a homomorphism. Now $h(a) = h(b)$ implies $[a]\Theta_i = [b]\Theta_i$ for all $i \in I$, that is $\langle a, b\rangle \in \Theta_i$ for all i. Hence $a = b$, since $\bigcap_{i\in I} \Theta_i = \Delta_A$. Now $\pi_i \circ h[A] = A/\Theta_i$, because if $[a]\Theta_i$ is given, then $\pi_i \circ h(a) = \pi_i(\langle [a]\Theta_i : i \in I\rangle) = [a]\Theta_i$. ◻

THEOREM 4.1.2. $\mathfrak{A}$ *is subdirectly irreducible iff there exists a congruence* Θ *such that every congruence* $\neq \Delta$ *contains* Θ. *Equivalently,* $\mathfrak{A}$ *is subdirectly irreducible iff* Δ *is* $\bigcap$*–irreducible in the lattice of congruences of* $\mathfrak{A}$.

PROOF. We show the second claim first. Let $\bigcap_{i\in I} \Theta_i = \Delta$ for some Θ_i which are all different from Δ. Then consider the subdirect representation $h : \mathfrak{A} \rightarrowtail \prod_{i\in I}(\mathfrak{A}/\Theta_i)$. Then none of the maps $\pi_i \circ h$ is injective since their kernel is exactly Θ_i. So, $\mathfrak{A}$ is subdirectly irreducible. Let on the other hand $\bigcap_{i\in I} \Theta_i \neq \Delta$ for all $\Theta_i \neq \Delta$. Then in a subdirect representation $h : \mathfrak{A} \rightarrowtail \prod_{i\in I} \mathfrak{B}_i$ we have $\bigcap_{i\in I} ker(\pi_i \circ h) = \Delta$ and thus $\Theta_i = \Delta$ for some $i \in I$, showing $\mathfrak{A}$ to be subdirectly irreducible. Now for the first claim. Let $\mathfrak{A}$ be subdirectly irreducible. Then let $\Theta := \bigcap_{\Phi\neq\Delta} \Phi$. Since Δ is $\bigcap$–irreducible, $\Theta \neq \Delta$. Conversely, if there exists a congruence $\Theta \neq \Delta$ such that every congruence $\neq \Delta$ contains Θ, then $\Theta = \bigcap_{\Phi\neq\Delta} \Phi$, so Δ is $\bigcap$–irreducible. ◻

In case that $\mathfrak{A}$ is subdirectly irreducible, the smallest congruence above Δ is called the **monolith** of $\mathfrak{A}$. Now recall the notation $\Theta(E)$ for the smallest congruence containing E. We say E **generates** $\Theta(E)$. A congruence is called **principal** if there is a one–membered E generating it. If $E = \{\langle a, b\rangle\}$ we write $\Theta(a, b)$ for $\Theta(E)$. It is easy to see that the monolith of a subdirectly irreducible algebra is principal. The following is due to [**17**].

THEOREM 4.1.3 (Birkhoff). *Every algebra is the subdirect product of subdirectly irreducible algebras.*

Proof. For each pair $a, b \in A$ of elements consider the set $\mathcal{E}(a, b)$ of congruences such that $\langle a, b\rangle \notin \mathcal{E}(a, b)$. Suppose we can show that each $\mathcal{E}(a, b)$ has a maximal element $\Psi(a, b)$ if $a \neq b$. Then let $\Theta(a, b)$ be the least congruence containing $\langle a, b\rangle$. $\Theta(a, b) \notin \mathcal{E}(a, b)$, by definition. Consider now any congruence $\Theta' \supsetneq \Psi(a, b)$. By the fact that $\Psi(a, b)$ is maximal for not being above $\Theta(a, b)$, we must have $\Theta' \supseteq \Psi(a, b) \sqcup \Theta(a, b)$. Thus, $\mathfrak{A}/\Psi(a, b)$ is subdirectly irreducible by Theorem 4.1.2. Moreover, we have $\Delta = \bigcap_{a \neq b} \Psi(a, b)$, hence $\mathfrak{A}$ is a subdirect product of the $\mathfrak{A}/\Psi(a, b)$.

But now for the promised proof of the existence of $\Psi(a, b)$. We consider the following property $\mathcal{P}$ of subsets of $A \times A$. S has $\mathcal{P}$ iff $\Theta(S)$ does not contain $\langle a, b\rangle$. All we have to show is that $\mathcal{P}$ is of finite character. For then by Tukey's Lemma, maximal sets exist. By Proposition 1.2.6,

$$\Theta(S) = \bigcup \langle \Theta(S_0) : S_0 \subseteq S, S_0 \text{ finite } \rangle ,$$

and so the claim is obvious. ◻

We will close this discussion of decomposition by a criterion of decomposability into a direct product.

Definition 4.1.4. *Let $\mathfrak{A}$ be an algebra. $\mathfrak{A}$ is called* ***directly reducible*** *if there exist algebras $\mathfrak{B}$ and $\mathfrak{C}$ such that $\sharp B \neq 1$ and $\sharp C \neq 1$ and $\mathfrak{A} \cong \mathfrak{B} \times \mathfrak{C}$. If $\mathfrak{A}$ is not directly reducible it is called* ***directly irreducible****.*

Let $\mathfrak{A}$ be an algebra and $\Theta, \Psi \in Con(\mathfrak{A})$. Θ and Ψ are said to **permute** if $\Theta \circ \Psi = \Psi \circ \Theta$. In this case, $\Theta \sqcup \Psi = \Theta \circ \Psi$.

Theorem 4.1.5. *An algebra $\mathfrak{A}$ is directly reducible iff there exist congruences Θ and Ψ, both different from ∇_A and Δ_A, such that (i.) $\Theta \sqcup \Psi = \nabla_A$, (ii.) $\Theta \sqcap \Psi = \Delta_A$, (iii.) Θ and Ψ permute.*

Proof. Suppose $\mathfrak{A}$ is directly reducible. Then there exist algebras $\mathfrak{B}$ and $\mathfrak{C}$, such that $\sharp B > 1$ and $\sharp C > 1$, and an isomorphism $h : \mathfrak{A} \to \mathfrak{B} \times \mathfrak{C}$. We may actually assume that $\mathfrak{A} = \mathfrak{B} \times \mathfrak{C}$. Let p_1 and p_2 be the canonical projections from $\mathfrak{B} \times \mathfrak{C}$ onto $\mathfrak{B}$ and $\mathfrak{C}$. These are surjective homomorphisms. Let $\Theta := ker(p_1)$ and $\Psi := ker(p_2)$. Then $\Theta \neq \nabla_A$ since $\mathfrak{B}$ is not isomorphic to $\mathbf{1}$. Also, $\Theta \neq \Delta_A$ since $\mathfrak{C}$ is not isomorphic to $\mathbf{1}$. Likewise it holds that Ψ is different from ∇_A and Δ_A. Now $\langle x_0, x_1\rangle \,\Theta\, \langle y_0, y_1\rangle$ iff $x_0 = y_0$ and $\langle x_0, x_1\rangle \,\Psi\, \langle y_0, y_1\rangle$ iff $x_1 = y_1$. It is easy to see that $\Theta \circ \Psi = \Psi \circ \Theta = \nabla_A$, showing (i.) and (iii.), and that $\Theta \cap \Psi = \Delta_A$, showing (ii.). Now assume conversely that Θ and Ψ are nontrivial congruences of $\mathfrak{A}$ satisfying (i.), (ii.) and (iii.). Let $h_\Theta : \mathfrak{A} \twoheadrightarrow \mathfrak{A}/\Theta$ and $h_\Psi : \mathfrak{A} \twoheadrightarrow \mathfrak{A}/\Psi$ be the natural maps with kernel Θ and Ψ, and let $h := \langle h_\Theta, h_\Psi\rangle : \mathfrak{A} \to \mathfrak{A}/\Theta \times \mathfrak{A}/\Psi$ be the map defined by $h(a) := \langle [a]\Theta, [a]\Psi\rangle$. This map is a well–defined homomorphism. We have to show that h is bijective. (1.) h is injective. Let $h(a) = h(b)$. Then $[a]\Theta = [b]\Theta$ and $[a]\Psi = [b]\Psi$, hence $[a](\Theta \cap \Psi) = [b](\Theta \cap \Psi)$. By (ii.), $a = b$. (2.) h is surjective. Let $\langle [a]\Theta, [b]\Psi\rangle \in A/\Theta \times A/\Psi$. Since $\Theta \circ \Psi = \nabla_A$, by (i.) and (iii.), there exists a c such that $a \,\Theta\, c \,\Psi\, b$. Then $[c]\Theta = [a]\Theta$ and $[c]\Psi = [b]\Psi$ and so $h(c) = \langle [a]\Theta, [b]\Psi\rangle$, as required. ◻

An algebra has **permuting congruences** if all nontrivial congruences permute pairwise. A class of algebras is **congruence permutable** if all its members have permuting congruences. The following is due to A. I. Malcev [**155**].

Theorem 4.1.6 (Malcev). *Let $\mathcal{V}$ be a variety of algebras. $\mathcal{V}$ is congruence permutable iff there exists a term $p(x, y, z)$ such that for all $\mathfrak{A} \in \mathcal{V}$ and all $a, b \in A$*

$$p^{\mathfrak{A}}(a, a, b) = b, \qquad p^{\mathfrak{A}}(a, b, b) = a\ .$$

Proof. Assume that there exists a term $p(x, y, z)$ with the properties given above. Then let a and b be elements such that $a\,\Theta \circ \Psi\,b$. Then there exists a c such that $a\,\Theta\,c\,\Psi\,b$. Hence

$$a = p^{\mathfrak{A}}(a, b, b)\,\Psi\,p^{\mathfrak{A}}(a, c, b)\,\Theta\,p^{\mathfrak{A}}(c, c, b) = b\ .$$

So $a\,\Psi \circ \Theta\,b$. Now assume that all members of $\mathcal{V}$ have permuting congruences. Then in particular the algebra $\mathfrak{A} := \mathfrak{Fr}_{\mathcal{V}}(\{x, y, z\})$ freely generated by x, y and z has permuting congruences. Let $\Theta := \Theta(\langle x, y\rangle)$ and $\Psi := \Theta(\langle y, z\rangle)$. Then $\langle x, z\rangle \in \Theta \circ \Psi$, and so $\langle x, z\rangle \in \Psi \circ \Theta$. Hence there exists an element u such that $x\,\Psi\,u\,\Theta\,z$. This element is of the form $p^{\mathfrak{A}}(x, y, z)$ for some ternary termfunction p. We claim that $x = p^{\mathfrak{A}}(x, y, y)$ and $y = p^{\mathfrak{A}}(x, x, y)$. This is enough to show the theorem. To that end, consider the canonical homomorphism h_Ψ. Modulo an isomorphism, $h_\Psi : \mathfrak{Fr}_{\mathcal{V}}(\{x, y, z\}) \twoheadrightarrow \mathfrak{Fr}_{\mathcal{V}}(\{x, y\}) : x \mapsto x, y \mapsto y, z \mapsto y$. Put $\mathfrak{B} := \mathfrak{F}_{\mathcal{V}}(\{x, y)\}$. $h_\Psi(p^{\mathfrak{A}}(x, y, z)) = p^{\mathfrak{B}}(x, y, y) = h_\Psi(x) = x$, since $h_\Psi(u) = h_\Psi(x)$. So, $p^{\mathfrak{B}}(x, y, y) = x$ holds in the algebra freely generated by x and y. To see that the equation holds in any algebra, let $\mathfrak{D}$ be any algebra in $\mathcal{V}$, and let $a, b \in D$ be elements. Let $j : \mathfrak{B} \to \mathfrak{D}$ be the unique homomorphism satisfying $j(x) = a$ and $j(y) = b$. Then $p^{\mathfrak{D}}(a, b, b) = j(p^{\mathfrak{B}}(x, y, y)) = j(x) = a$. Hence the first equation holds in all algebras. Analogously for the second equation, using the congruence Θ instead. □

An algebra is said to be **congruence distributive** if the lattice of congruences is distributive.

Theorem 4.1.7. *An algebra is congruence distributive if there is a ternary termfunction $m(x, y, z)$ such that for all $a, b \in A$:*

$$m(a, a, b) = m(a, b, a) = m(b, a, a) = a.$$

Proof. First of all, from lattice theory we get $\Theta \sqcap \Phi \subseteq \Theta \sqcap (\Phi \sqcup \Psi)$ and $\Theta \sqcap \Psi \subseteq \Theta \sqcap (\Phi \sqcup \Psi)$, so that $(\Theta \sqcap \Phi) \sqcup (\Theta \sqcap \Psi) \subseteq \Theta \sqcap (\Phi \sqcup \Psi)$. For the converse inclusion assume $\langle a, b\rangle \in \Theta \sqcap (\Phi \sqcup \Psi)$. Then $a\,\Theta\,b$ and there is a sequence c_i, $i < n + 1$, such that

$$a = c_0\,\Phi\,c_1\,\Psi\,c_2\,\Phi\,c_3 \ldots \Phi\,c_{n-1}\,\Psi\,c_n = b$$

We then have $m(a, c_i, b)\,\Theta\,m(a, c_i, a) = a$ and $m(a, c_i, b)\,\Theta\,m(b, c_i, b) = b$ for all i so that

$$m(a, c_i, b)\,\Theta\,a\,\Theta\,b\,\Theta\,m(a, c_{i+1}, b)$$

and $m(a, c_i, b)\,\Phi\, m(a, c_{i+1}, b)$ for even i. Similarly for odd i it is shown that $m(a, c_i, b)\,\Psi\, m(a, c_{i+1}, b)$. Therefore $m(a, c_i, b)\,(\Theta\sqcap\Phi)\, m(a, c_{i+1}, b)$ for even i and $m(a, c_i, b)\,(\Theta\sqcap\Psi)\, m(a, c_{i+1}, b)$ for odd i. Thus $\langle a, b\rangle \in (\Theta \sqcap \Phi) \sqcup (\Theta \sqcap \Psi)$. ◻

For example, let $\mathfrak{A}$ be an algebra in which there are termfunctions $\sqcap$, $\sqcup$ such that $\langle A, \sqcap, \sqcup\rangle$ is a lattice. Then $\mathfrak{A}$ is congruence distributive. Namely, take

$$m(x, y, z) := (x \sqcup y) \sqcap (y \sqcup z) \sqcap (z \sqcup x).$$

This termfunction satisfies the above criterion. If in addition there is a termfunction $'$ such that $\langle A, \sqcap, \sqcup, '\rangle$ is a boolean algebra, then $\mathfrak{A}$ has permuting congruences. For take

$$p(x, y, z) := (x \sqcap z) \sqcup (x \sqcap y' \sqcap z') \sqcup (x' \sqcap y' \sqcap z)\ .$$

If $y = z$, this reduces to $p(x, y, y) = (x\sqcap y)\sqcup(x\sqcap y')\sqcup(x'\sqcap y'\sqcap y) = (x\sqcap y)\sqcup(x\sqcap y') = x$; if $x = y$ this gives $p(x, x, z) = (x\sqcap z)\sqcup(x\sqcap x'\sqcap z')\sqcup(x'\sqcap x'\sqcap z) = (x\sqcap z)\sqcup(x'\sqcap z) = z$. It also follows that subalgebras, homomorphic images and products of similar congruence distributive algebras are again congruence distributive, if the same term can be chosen in all algebras. This is the case with modal algebras.

Corollary 4.1.8. *The variety of modal algebras has permuting congruences and is congruence distributive.*

Congruence distributivity is important in connection with *reduced products*. Let I be an index set, $\prod_{i\in I} \mathfrak{B}_i$ be a product. Let F be a filter on 2^I and let Θ_F be the set of all pairs $\langle a, b\rangle$ such that $\{i : a(i) = b(i)\} \in F$. Θ_F is a congruence. This is left as an exercise. Congruences of this form are called **filtral**. We define the F–reduced product of the $\mathfrak{B}_i$ by

$$\prod_F \mathfrak{B}_i := (\prod_{i\in I} \mathfrak{B}_i)/\Theta_F.$$

If F is an ultrafilter, we speak of an **ultraproduct**. Given a class $\mathcal{K}$ of algebras, $\mathsf{Up}(\mathcal{K})$ denotes the closure of $\mathcal{K}$ under ultraproducts. Let us note that if $F \subseteq G$ then there is a surjective homomorphism $\prod_F \mathfrak{B}_i \twoheadrightarrow \prod_G \mathfrak{B}_i$. The following theorem is due to Bjarni Jónsson [**111**], also known as *Jónsson's Lemma*.

Theorem 4.1.9 (Jónsson). *Let $\mathcal{K}$ be a class of algebras and $\mathcal{V} = \mathsf{HSP}(\mathcal{K})$ be a congruence distributive variety. If $\mathfrak{A} \in \mathcal{V}$ is subdirectly irreducible then $\mathfrak{A} \in \mathsf{HSUp}(\mathcal{K})$.*

Proof. Let $h : \mathfrak{B} \twoheadrightarrow \mathfrak{A}$ for $\mathfrak{B} \rightarrowtail \prod_{i\in I} \mathfrak{C}_i$. (This is not necessarily a subdirect product.) Then put $\Phi := h^{-1}[\Delta_A]$; this is a congruence. Moreover, Φ is $\sqcap$–irreducible in $\mathfrak{B}$. For if $\Theta_1 \sqcap \Theta_2 = \Phi$ we have

$$(\Theta_1 \sqcap \Theta_2)/\Phi = \Theta_1/\Phi \sqcap \Theta_2/\Phi = \Delta_A$$

in $Con(\mathfrak{A})$. This implies $\Theta_1/\Phi = \Delta_A$ or $\Theta_2/\Phi = \Delta_A$, which is nothing but $\Theta_1 = \Phi$ or $\Theta_2 = \Phi$.

For subsets S of I let $\overline{\Theta}_S$ denote the congruence induced on $\prod_{i\in I} \mathfrak{C}_i$ by the principal filter $\uparrow S = \{T : S \subseteq T \subseteq I\}$. Furthermore, let D be the set of all $S \subseteq I$ such that

$\overline{\Theta}_S \restriction B \subseteq \Phi$, that is, $\Phi = \Phi \sqcup (\overline{\Theta}_S \restriction B)$. Choose U to be a maximal filter contained in D. Then $\Theta_U = \bigcup \langle \overline{\Theta}_S : S \in U \rangle$, and so $\Theta_U \restriction B \subseteq \Phi$. All there is to be done is to show that U is an ultrafilter over I. For then the map $\mathfrak{B} \to \prod_U \mathfrak{C}_i$ has the kernel $\Theta_U \restriction B$. So this induces an embedding $\mathfrak{B}/(\Theta_U \restriction B) \rightarrowtail \prod_U \mathfrak{C}_i$. Since $\Theta_U \restriction B \subseteq \Phi$, there is a homomorphism $\mathfrak{B}/(\Theta_U \restriction B) \twoheadrightarrow \mathfrak{A}$.

Now observe that if $S, T \subseteq I$ then

$$\overline{\Theta}_{S \cup T} \restriction B = (\overline{\Theta}_S \restriction B) \sqcap (\overline{\Theta}_T \restriction B).$$

Therefore we have for $S, T \in D$

$$\Phi = \Phi \sqcup (\overline{\Theta}_{S \cup T} \restriction B) = (\Phi \sqcup (\overline{\Theta}_S \restriction B)) \sqcap (\Phi \sqcup (\overline{\Theta}_T \restriction B)).$$

Since Φ is $\sqcap$–irreducible in $Con(\mathfrak{B})$ we have either $\Phi = \Phi \sqcup (\overline{\Theta}_S \restriction B)$ or $\Phi = \Phi \sqcup (\overline{\Theta}_T \restriction B)$. And so we conclude that

$$\text{If } S \cup T \in D \text{ then } S \in D \text{ or } T \in D.$$

Furthermore

$$\text{If } S \in D \text{ and } S \subseteq T \text{ then } T \in D.$$

From these properties we can derive that U is an ultrafilter. For if not, there is a set S such that neither $S \in U$ nor $I - S \in U$. Thus there are sets $K, L \in U$ such that $S \cap K \notin D$ and $(I - S) \cap L \notin D$. We show the existence of K; the existence of L is proved analogously. Suppose that $\varnothing \in D$. Then $D = 2^I$ and U is an ultrafilter by construction. Now assume $\varnothing \notin D$. Consider the system of sets $V := \{T : T \supseteq S \cap K, K \in U\}$. If all $S \cap K \in D$, we have a system of sets which is a filter containing U and fully contained in D. In particular, $S \in V$, contradicting the maximality of U. So there is a $K \in U$ such that $S \cap K \notin D$. Likewise, we have an $L \in U$ such that $(I - S) \cap L \notin D$. Put $M := K \cap L$. U is a filter, so $M \in U$. Thus also $M \in D$. Now $M = (S \cap M) \cup ((I - S) \cap M)$ but neither set is in D. This is a contradiction. ◻

By the fact that modal algebras are congruence distributive we can now infer that a subdirectly irreducible algebra in the variety generated by some class of modal algebras $\mathcal{K}$ is an image of a subalgebra of an ultraproduct from $\mathcal{K}$. Let us return, however, to the criterion of subdirect irreducibility. First of all, a congruence Θ on a boolean algebra $\mathfrak{A}$ defines a homomorphism $h_\Theta : \mathfrak{A} \twoheadrightarrow \mathfrak{A}/\Theta$ with kernel Θ. Put $F_\Theta := h_\Theta^{-1}(1)$. This is a filter, as is easily checked. And we have $a \in F$ iff $a \,\Theta\, 1$. Moreover, suppose that $a \,\Theta\, b$. Then $h_\Theta(a) = h_\Theta(b)$ and so $h_\Theta(a \leftrightarrow b) = 1$, whence $a \leftrightarrow b \in F$. In other words, filters are the congruence classes of the top elements and they are in one–to–one correspondence with congruences on the algebra. Let $\mathfrak{A}$ be a boolean algebra. We have seen in Section 1.7 that the map $f : \Theta \mapsto F_\Theta = \{a : a \,\Theta\, 1\}$ is a one–to–one map from the lattice of congruences of $\mathfrak{A}$ onto the lattice of filters on $\mathfrak{A}$. Furthermore, if F is a filter, then Θ_F defined by $a \,\Theta_F\, b$ iff $a \leftrightarrow b \in F$ is the inverse under f. In modal algebras, we have to take open filters rather than filters. Recall from Section 3.1 that a filter is open if it is closed under (fi$\blacksquare$.) If $a \in F$ then $\blacksquare_j a \in F$

We have shown in Lemma 3.1.5 that open filters indeed correspond to homomorphisms. Any intersection of open filters is again an open filter, so the open filters form a complete lattice, with the join $F \sqcup G$ defined as the smallest open filter containing both F and G.

Theorem 4.1.10. *The map $f : \Theta \mapsto F_\Theta := \{a : a \,\Theta\, 1\}$ is an isomorphism from the lattice $Con(\mathfrak{A})$ of a modal algebra $\mathfrak{A}$ onto the lattice of open filters on $\mathfrak{A}$, with inverse $F \mapsto \Theta_F := \{\langle a, b\rangle : a \leftrightarrow b \in F\}$.*

In particular, if $C \subseteq A$ then the smallest open filter containing C, denoted by $\langle C\rangle$, can be obtained as follows.

$$\langle C\rangle = \{d : (\exists \boxplus)(\exists C_0 \subseteq_{fin} C)(d \geq \boxplus \bigcap C_0)\}.$$

If C is finite, then put $c := \bigcap C$. The above condition then reduces to the condition that $d \geq \boxplus c$ for some compound modality $\boxplus$.

We now wish to characterize subdirectly irreducible modal algebras. We know that $\mathfrak{A}$ is subdirectly irreducible iff it has a monolith Θ. Θ is principal, say $\Theta = \Theta(a, b)$. Now, in terms of filters this means that there exists a unique minimal open filter in the set of open filters $\neq \{1\}$ and this filter is generated by a single element o (for example $o = a \leftrightarrow b$). To say that the filter generated by o is minimal is to say that every other filter generated by an element $a \neq 1$ must contain o, which in turn means that there must be a compound modality $\boxplus$ such that $o \geq \boxplus a$. Thus we obtain the following criterion of [**170**].

Theorem 4.1.11 (Rautenberg). *A modal algebra is subdirectly irreducible iff there exists an element $o \neq 1$ such that for every $a \in A - \{1\}$ there is a compound modality $\boxplus$ such that $o \geq \boxplus a$. Such an o is called an* **opremum** *of $\mathfrak{A}$.*

For a Kripke–frame $\mathfrak{f}$ the question of subdirect irreducibility of the algebra $\mathfrak{Ma}(\mathfrak{f})$ of all subsets of f has a rather straightforward answer. We warn the reader that this theorem fails in general. This is discussed in Section 4.8.

Theorem 4.1.12. *For a Kripke–frame $\mathfrak{f}$ the algebra $\mathfrak{Ma}(\mathfrak{f})$ of all subsets of f is subdirectly irreducible iff $\mathfrak{f}$ is rooted.*

Proof. Suppose that $\mathfrak{f}$ is rooted at x. Then take $o := f - \{x\}$. We claim that o is an opremum. So take a set $a \subseteq f$ such that $a \neq f$. Then there exists a $y \in f$ such that $y \notin a$. By assumption, there is a $k \in \omega$ such that a finite path exists from x to y. Then $x \notin \boxplus a$ for some compound modality $\boxplus$. Thus $\boxplus a \subseteq o$. By Theorem 4.1.11, $\mathfrak{Ma}(\mathfrak{f})$ is subdirectly irreducible. Now assume that there is no single point x which can generate $\mathfrak{f}$. Then there exist two nonempty sets S and T such that for $X := Tr(S)$ and $Y := Tr(T)$, $f = X \cup Y$, but $X \neq f$ and $Y \neq f$. Suppose there exists an opremum, o. Since $\blacksquare X \supseteq X$, we must have $o \supseteq X$. Similarly $o \supseteq Y$, and so $o \supseteq X \cup Y = f$, a contradiction. $\Box$

Exercise 134. Show that the filtral congruence Θ_F defined above is indeed a congruence.

Exercise 135. Show that a boolean algebra is subdirectly irreducible iff it is isomorphic to **2**. Hence $\mathsf{BA} = \mathsf{HSP}(\mathbf{2})$.

Exercise 136. Suppose that $\mathfrak{G} = \langle G, 1, ^{-1}, \cdot \rangle$ is a group and Θ a congruence on G. Show that the congruence class of the unit 1 must be a normal subgroup. Furthermore, suppose that $[a]\Theta$ is given. Then $[1]\Theta = a^{-1} \cdot [a]\Theta = \{a^{-1} \cdot b : b \in [a]\Theta\}$. Thus, show that there is a one–to–one correspondence between congruences on $\mathfrak{G}$ and normal subgroups.

Exercise 137. Let $\mathbb{Z}_2$ be the additive group of integers modulo 2. That is, we have

$\cdot$	0	1
0	0	1
1	1	0

	$^{-1}$
0	0
1	1

Show that the lattice of congruences of the group $\mathbb{Z}_2 \times \mathbb{Z}_2$ is not distributive. *Hint.* Use the previous exercise.

Exercise 138. A vector space $\mathfrak{V}$ over a field $\mathbb{F}$ can be made into an algebra by adding a unary function θ_r for each $r \in \mathbb{F}$. Its action is defined by $\theta_r(v) = r \cdot v$, where $r \cdot v$ is the usual scalar multiplication. (Why is this complicated definition necessary?) Again, a congruence defines a normal subgroup. Moreover, this subgroup must be closed under all θ_r. Show that there is a one–to–one correspondence between congruences on $\mathfrak{V}$ and subspaces.

Exercise 139. Continuing the previous exercise, show that a vector space is subdirectly irreducible iff it is one–dimensional.

4.2. Varieties, Logics and Equationally Definable Classes

Two of the most fundamental theorems of universal algebra are due to Birkhoff of which the first states that a class of algebras is definable by means of equations iff it is a variety, that is, closed under H, S and P. The second gives an explicit characterization of all equations that hold in a variety which is defined by some given set of equations. We will prove both theorems in their full generality and derive some important consequences for modal logics. To start, let Ω be a signature, $\mathcal{L}$ a language for Ω and $t(\vec{x}), s(\vec{x})$ two terms based on the variables x_i, $i < n$. An Ω–algebra $\mathfrak{A}$ satisfies the equation $s(\vec{x}) \approx t(\vec{x})$, written $\mathfrak{A} \vDash s(\vec{x}) \approx t(\vec{x})$, if for all $\vec{a} \subseteq A$, $s^{\mathfrak{A}}(\vec{a}) = t^{\mathfrak{A}}(\vec{a})$. Furthermore, $\mathcal{V} \vDash s(\vec{x}) \approx t(\vec{x})$ if for all $\mathfrak{A} \in \mathcal{V}$, $\mathfrak{A} \vDash s(\vec{x}) \approx t(\vec{x})$. Often we write $s \approx t$, dropping the variables.

Proposition 4.2.1. *The class of all algebras which satisfy an equation $s \approx t$ is closed under taking products, subalgebras and homomorphic images.*

The proof of this theorem is routine and left as an exercise. We will prove here that the converse is true as well. Let $\mathcal{V}$ be a variety, and and X a set of variables. Then put

$$\mathsf{Eq}_X(\mathcal{V}) := \{\langle s(\vec{x}), t(\vec{x})\rangle : \mathcal{V} \vDash s(\vec{x}) \approx t(\vec{x}), \vec{x} \subseteq X\}$$

In case that $X = \{x_i : i \in \omega\}$ put $\mathsf{Eq}(\mathcal{V}) := \mathsf{Eq}_X(\mathcal{V})$. Let E be a set of equations over X. Define $\mathsf{Alg}(E)$ to be the class of algebras satisfying E. By Proposition 4.2.1, $\mathsf{Alg}(E)$ is a variety. Moreover, for any class $\mathcal{K}$ of Ω–algebras we always have $\mathcal{K} \subseteq \mathsf{Alg}\,\mathsf{Eq}(\mathcal{K})$. What we have to show is that for a given variety $\mathcal{V}$ we have $\mathcal{V} = \mathsf{Alg}\,\mathsf{Eq}(\mathcal{V})$. There is a way to restate this using free algebras. Recall from Section 1.3 that an algebra $\mathfrak{Fr}_{\mathcal{V}}(Y)$ is said to be a *freely Y–generated algebra* if for every algebra $\mathfrak{A} \in \mathcal{V}$ and every map $v : Y \to A$ there is a homomorphism $\widehat{v} : \mathfrak{Fr}_{\mathcal{V}}(Y) \to \mathfrak{A}$ such that $\widehat{v} \upharpoonright Y = h$. We have seen in Theorem 1.3.5 that a variety has free algebras for all sets Y. Moreover, $\mathsf{Alg}\,\mathsf{Eq}(\mathcal{V})$ has free algebras because they can be obtained from the term algebras. Namely, if $\mathfrak{Tm}_{\Omega}(Y)$ is the algebra of Y–terms over the language $\mathcal{L}$ with signature Ω, define a congruence Θ by $s(\vec{x})\,\Theta\,t(\vec{x})$ iff for all $\mathfrak{A} \in \mathcal{V}$ and all $\vec{a} \subseteq A$ we have $s^{\mathfrak{A}}(\vec{a}) = t^{\mathfrak{A}}(\vec{a})$. Then $\mathfrak{Tm}_{\Omega}(Y)/\Theta$ is freely generated by Y in $\mathsf{Alg}\,\mathsf{Eq}(\mathcal{V})$. For let $v : Y \to A$ be a map, $\mathfrak{A}$ an algebra over A. Then there is a unique extension $\overline{v} : \mathfrak{Tm}_{\Omega}(Y) \to \mathfrak{A}$. Let $s(\vec{x})\,\Theta\,t(\vec{x})$. Then $\overline{v}(s(\vec{x})) = s^{\mathfrak{A}}(v(\vec{x})) = t^{\mathfrak{A}}(v(\vec{x})) = \overline{v}(t(\vec{x}))$, by definition of Θ. Hence there exists a unique homomorphism $\widehat{v} : \mathfrak{Tm}_{\Omega}(Y)/\Theta \to \mathfrak{A}$.

Proposition 4.2.2. *For any variety $\mathcal{V}$, $\mathfrak{Fr}_{\mathcal{V}}(X)$ is a subdirect product of the $\mathfrak{Fr}_{\mathcal{V}}(E)$, E a finite subset of X.*

Proof. Let E be a finite subset of X. We let $\mathfrak{A}(E)$ be the subalgebra of $\mathfrak{Fr}_{\mathcal{V}}(X)$ generated by the terms x_i, $x_i \in E$. It is easy to see that $\mathfrak{A}(E)$ is isomorphic to $\mathfrak{Fr}_{\mathcal{V}}(E)$. Let κ_E be a map $\kappa_E : X \to E$ such that $\kappa_E \upharpoonright E = id_E$. This map can be extended to a homomorphism $\overline{\kappa}_E : \mathfrak{Fr}_{\mathcal{V}}(X) \twoheadrightarrow \mathfrak{Fr}_{\mathcal{V}}(E)$. (That this map is onto follows from Theorem 1.3.6.) Let F be the collection of all finite subsets of X. We have

$$\bigcap_{E \in F} ker(\overline{\kappa}_E) = \Delta\ .$$

For if $\mathfrak{Fr}_{\mathcal{V}}(X) \nvDash s(\vec{x}) \approx t(\vec{x})$, then let E consist of the variables in $\vec{x}$. Now $\mathfrak{Fr}_{\mathcal{V}}(E) \nvDash s(\vec{x}) \approx t(\vec{x})$ and thus $\overline{\kappa}_E(s(\vec{x})) = s(\vec{x}) \not\approx t(\vec{x}) = \overline{\kappa}_E(t(\vec{x}))$. (Here we write $s(\vec{x})$ also for the equivalence class of the term $s(\vec{x})$ in the free algebras.) By Proposition 4.1.1, $\mathfrak{Fr}_{\mathcal{V}}(X)$ is a subdirect product of the $\mathfrak{Fr}_{\mathcal{V}}(E)$. ◻

Corollary 4.2.3. *For every variety $\mathcal{V}$, $\mathcal{V} = \mathsf{HSP}(\mathfrak{Fr}_{\mathcal{V}}(\aleph_0))$.*

Now, if we are able to show that $\mathfrak{Tm}_{\Omega}(X)/\Theta$ is also the freely X–generated algebra of $\mathcal{V}$, we are obviously done. For then the classes $\mathcal{V}$ and $\mathsf{Alg}\,\mathsf{Eq}(\mathcal{V})$ contain the same countably generated free algebras. Now recall the construction of a free algebra from the algebras of $\mathcal{V}$ from Section 1.2.

Proposition 4.2.4. *For a class $\mathcal{K}$, and a set X*

$$\mathfrak{Tm}_\Omega(X)/\mathsf{Eq}_X(\mathcal{K}) \in \mathsf{SP}(\mathcal{K}) .$$

Proof. Let $Q := \{\langle s(\vec{x}), t(\vec{x})\rangle : \vec{x} \subseteq X, \mathcal{K} \nvDash s \approx t\}$. For each $\epsilon \in Q$ we pick a witness algebra $\mathfrak{A}_\epsilon$. This means that there is a sequence $\vec{a}$ in $\mathfrak{A}_\epsilon$ such that $s^{\mathfrak{A}_\epsilon}(\vec{a}) \neq t^{\mathfrak{A}_\epsilon}(\vec{a})$. Thus we have a map $v_\epsilon : X \to A$ such that $v_\epsilon : x_i \mapsto a_i$, $i < n$. This extends to a homomorphism $\overline{v}_\epsilon : \mathfrak{Tm}_\Omega(X) \to \mathfrak{A}_\epsilon$. Now let $h : \mathfrak{Tm}_\Omega(X) \to \prod_{\epsilon\in Q} \mathfrak{A}_\epsilon$ be the canonical homomorphism defined by all the v_ϵ. We show now that $ker(h)$ is exactly $\mathsf{Eq}_X(\mathcal{K})$. For if $\epsilon = s(\vec{x}) \approx t(\vec{x})$ holds in all algebras, it holds in their product as well, and so $\epsilon \in ker(h)$. However, if it does not hold in all algebras of $\mathcal{K}$, then $h(s(\vec{x})) \neq h(t(\vec{x}))$, for $\overline{v}_\epsilon(s(\vec{x})) = s^{\mathfrak{A}_\epsilon}(\vec{a}) \neq t^{\mathfrak{A}_\epsilon}(\vec{a}) = \overline{v}_\epsilon(t(\vec{x}))$, as had to be shown. □

Theorem 4.2.5 (Birkhoff). *A class of Ω–algebras is definable by means of equations over a language $\mathcal{L}$ of signature Ω exactly if it is a variety.*

In addition, an equational theory $\mathsf{Eq}(\mathcal{K})$ can actually be identified with a particular kind of congruence on $\mathfrak{Tm}_\Omega(X)$, the so–called *fully invariant* congruence.

Definition 4.2.6. *A congruence Θ on an algebra $\mathfrak{A}$ is called **fully invariant** if it is compatible with all endomorphisms of $\mathfrak{A}$. That is, Θ is fully invariant if whenever $h : \mathfrak{A} \to \mathfrak{A}$ is an endomorphism and $a\,\Theta\, b$ then also $h(a)\,\Theta\, h(b)$.*

Our aim is to show that all modal logics are theories of certain classes of algebras. In order to do this, we need to be explicit about what equations can be derived from other equations. The following axiomatization is due to Birkhoff. Unlike in propositional calculi, we derive equations from sets of equations and not terms from sets of terms. We write $\Gamma \vdash_V t \approx s$ if $t \approx s$ can be derived in finitely many steps by applying one of the following rules in addition to (ext.), (mon.) and (trs.) of Section 1.3.

(V1) $\vdash_V s \approx s$
(V2) $s \approx t \vdash_V t \approx s$
(V3) $s \approx t; t \approx u \vdash_V s \approx u$
(V4) $s_0 \approx t_0; \dots; s_{n-1} \approx t_{n-1} \vdash_V f(s_0, \dots, s_{n-1}) \approx f(t_0, \dots, t_{n-1})$
(V5) $s(x_0, \dots, x_{n-1}) \approx t(x_0, \dots, x_{n-1}) \vdash_V$
$s(u_0, \dots, u_{n-1}) \approx t(u_0, \dots, u_{n-1})$

(V1), (V2) and (V3) are axioms of pure equality. (V4) is known as the **replacement rule**, (V5) as the **substitution rule**. Notice that this calculus also satisfies (sub.) and (cmp.), the latter by the fact that it is a calculus defined by finitary rules, so it is finitary by the nature of a proof. In (V4), f is any n–ary term function. However, it can be shown that it is enough to require the validity of (V4) only for the basic functions, f_i, $i \in I$. Now take a set Γ of equations. Consider the congruence on $\mathfrak{Tm}_\Omega(X)$ induced by Γ. The condition of reflexivity of Θ corresponds to (V1), the symmetry to (V2) and the transitivity to (V3). Finally, (V4) corresponds to the compatibility with all functions. Hence, the calculus without substitution defines the smallest congruence induced by Γ. Now a substitution is nothing but an endomorphism of the term algebra so (V5) enshrines the requirement that the congruence derivable from Γ is fully invariant.

Theorem 4.2.7. *Let Γ be a set of equations over $\mathcal{L}$ of signature Ω. The smallest fully invariant congruence on $\mathfrak{Tm}_\Omega(X)$ containing Γ is the set of all $s \approx t$ such that $\Gamma \vdash_V s \approx t$.*

Let us call a set of the form $\mathsf{Eq}(\mathcal{K})$ an **equational theory**. Then we have

Corollary 4.2.8 (Birkhoff). *A set of equations is an equational theory iff it is closed under the rules of $\vdash_V$.*

Proof. First of all, the rules (V1) – (V5) are correct. That is, given an algebra $\mathfrak{A}$ and given a rule, if $\mathfrak{A}$ satisfies every premiss of a rule then $\mathfrak{A}$ also satisfies the conclusion. **(V1).** For all $v : X \to A$ we have $\overline{v}(s) = \overline{v}(s)$ for any term s. **(V2).** Assume that for all $v : X \to A$, $\overline{v}(s) = \overline{v}(t)$. Then for all $v : X \to A$ also $\overline{v}(t) = \overline{v}(s)$, showing $\mathfrak{A} \vDash t \approx s$. **(V3).** Assume that $\mathfrak{A} \vDash s \approx t; t \approx u$. Take a map $v : X \to A$. Then $\overline{v}(s) = \overline{v}(t)$ as well as $\overline{v}(t) = \overline{v}(u)$, from which $\overline{v}(s) = \overline{v}(u)$. Thus $\mathfrak{A} \vDash s \approx u$. **(V4).** Assume that $\mathfrak{A} \vDash s_i \approx t_i$ for all $i < n$. Take $v : X \to A$. Then $\overline{v}(f(s_0, \dots, s_{n-1})) = f^{\mathfrak{A}}(\overline{v}(s_0), \dots, \overline{v}(s_{n-1})) = f^{\mathfrak{A}}(\overline{v}(t_0), \dots, \overline{v}(t_{n-1})) = \overline{v}(f(t_0, \dots, t_{n-1}))$. Hence $\mathfrak{A} \vDash f(s_0, \dots, s_{n-1}) \approx f(t_0, \dots, f_{n-1})$. **(V5).** Assume $\mathfrak{A} \vDash s \approx t$. Define a substitution σ by $\sigma : x_i \mapsto u_i$, $i < n$, $\sigma : x_i \mapsto x_i$ for $i \geq n$. Then $\overline{v} \circ \sigma : X \to A$, and the homomorphism extending the map is just $\overline{v} \circ \overline{\sigma}$, since it coincides on X with $\overline{v} \circ \sigma$. Now let $v : X \to A$ be given. Then $\overline{v}(s(u_0, \dots, u_{n-1})) = \overline{v}(\overline{\sigma}(s)) = \overline{v} \circ \overline{\sigma}(s) = \overline{v} \circ \overline{\sigma}(t) = \overline{v}(t(u_0, \dots, u_{n-1}))$. Thus $\mathfrak{A} \vDash s(u_0, \dots, u_{n-1}) \approx t(u_0, \dots, u_{n-1})$.

Now let Γ be any set of equations and Θ its closure under (V1) to (V5). Let $\mathfrak{A} := \mathfrak{Tm}_\Omega(X)/\Gamma$. Then if $s \approx t \notin \Gamma$ we have $\mathfrak{A} \nvDash s \approx t$. For just take the canonical homomorphism $h_\Theta : \mathfrak{Tm}_\Omega(X) \to \mathfrak{A}$ with kernel Θ. Since Θ is a congruence, this is well–defined and we have $h_\Theta(s) \neq h_\Theta(t)$, as required. Next we have to show that $\mathfrak{A} \vDash \Gamma$. To see that take any equation $s \approx t \in \Gamma$ and $v : X \to A$. Since A is generated by terms over X modulo Θ, we can define a substitution σ such that $\overline{v} = h_\Theta \circ \overline{\sigma}$. Namely, put $\sigma(x) := t(\vec{y})$, where $t(\vec{y}) \in h_\Theta^{-1}(h(x))$ is freely chosen. Then $h_\Theta(\sigma(x)) = v(x)$, as required. Now $\overline{\sigma}(s) \,\Theta\, \overline{\sigma}(t)$, by closure under substitution, so that $\overline{v}(s) = h_\Theta(\overline{\sigma}(s)) = h_\Theta(\overline{\sigma}(t)) = \overline{v}(t)$. Thus $\mathfrak{A} \vDash \Gamma$. ◻

Theorem 4.2.9. *There is a one–to–one correspondence between varieties of Ω–algebras and fully invariant congruences on the freely countably generated algebra. Moreover, $\mathcal{V}_1 \subseteq \mathcal{V}_2$ iff for the corresponding congruences $\Theta_1 \supseteq \Theta_2$.*

One half of this theorem is actually Theorem 2.2.9. The converse direction was actually much harder to prove but makes the result all the more useful.

Consider now what this means for modal logic. (We will henceforth write again p and q for variables instead of x and y, as well as φ and ψ for formulae instead of

s and t.) We have seen earlier for propositional formulae φ and ψ that $\mathfrak{A} \vDash \varphi \approx \psi$ iff $\mathfrak{A} \vDash \varphi \leftrightarrow \psi \approx \top$. The latter is nothing but $\mathfrak{A} \vDash \varphi \leftrightarrow \psi$, now read in the standard sense. Thus the equational theory of a class $\mathcal{K}$ of modal algebras can be interpreted as the logical theory of a class. We will show that if we have a class of modal algebras, then the set of equations $\varphi \approx \top$ in that class is a modal logic, and conversely. So the two alternative ways to specify classes of algebras — namely via equations and via logical axioms — coincide. The equational theory of polymodal algebras is as follows. The primitive function symbols are here taken to be $\top$, $\bot$, $\neg$, $\wedge$, $\vee$ and the modal operators $\Box_j$. The following equations must hold.

$$\begin{array}{lcllcl}
p \wedge p & \approx & p & p \vee p & \approx & p \\
p \wedge q & \approx & q \wedge p & p \vee q & \approx & q \vee p \\
p \wedge (q \wedge r) & \approx & (p \wedge q) \wedge r & p \vee (q \vee r) & \approx & (p \vee q) \vee r \\
p \wedge (q \vee p) & \approx & p & p \vee (q \wedge p) & \approx & p \\
p \wedge (q \vee r) & \approx & (p \wedge q) \vee (p \wedge r) & p \vee (q \wedge r) & \approx & (p \vee q) \wedge (p \vee r) \\
p \wedge \top & \approx & p & p \vee \bot & \approx & p \\
\neg(p \wedge q) & \approx & (\neg p) \vee (\neg q) & \neg(p \vee q) & \approx & (\neg p) \wedge (\neg q) \\
p \wedge (\neg p) & \approx & \bot & p \vee (\neg p) & \approx & \top \\
\neg(\neg p) & \approx & p & \neg\top & \approx & \bot \\
\Box_j(p \wedge q) & \approx & (\Box_j p) \wedge (\Box_j q) & \Box_j \top & \approx & \top
\end{array}$$

We call this set of equations Mal. The first five rows specify that the algebras are distributive lattices; then follow laws to the effect that there is a top and a bottom element, and that there is a negation. Finally, there are two laws concerning the box operators. We define $\varphi \to \chi$ by $(\neg\varphi) \vee \chi$, $\varphi \leftrightarrow \chi$ by $(\varphi \wedge \chi) \vee ((\neg\varphi) \wedge (\neg\chi))$ and $\Diamond_j \varphi$ by $\neg\Box_j\neg\varphi$.

Proposition 4.2.10. *(i)* $\mathsf{Mal}; \varphi \approx \chi \vdash_V \varphi \leftrightarrow \chi \approx \top$. *(ii)* $\mathsf{Mal}; \varphi \leftrightarrow \chi \approx \top \vdash_V \varphi \approx \chi$.

Proof. The proof will be a somewhat reduced sketch. A full proof would consume too much space and is not revealing. The reader is asked to fill in the exact details. We perform the proof only to show how the calculus works in practice. (i) $\mathsf{Mal}; \varphi \approx \chi \vdash_V \varphi \vee (\neg\chi) \approx \chi \vee (\neg\chi); \varphi \vee (\neg\varphi) \approx \chi \vee (\neg\varphi)$, by applying (V4). Furthermore, $\mathsf{Mal} \vdash_V \varphi \vee (\neg\varphi) \approx \top; \chi \vee (\neg\chi) \approx \top$. Applying (V2) and (V3) we get $\mathsf{Mal}; \varphi \approx \chi \vdash_V \varphi \vee (\neg\chi) \approx \top; \chi \vee (\neg\varphi) \approx \top$. So, $\mathsf{Mal}; \varphi \approx \chi \vdash_V (\varphi \vee (\neg\chi)) \wedge (\chi \vee (\neg\varphi)) \approx \top \wedge \top$, by (V4). Now by (V5) we have $\mathsf{Mal} \vdash_V \top \wedge \top \approx \top$; thus $\mathsf{Mal}; \varphi \approx \chi \vdash_V (\varphi \vee (\neg\chi)) \wedge (\chi \vee (\neg\varphi)) \approx \top$. Applying the distributivity law twice we arrive at

$$\mathsf{Mal}; \varphi \approx \chi \vdash_V (\varphi \wedge \chi) \vee (\varphi \wedge (\neg\varphi)) \vee ((\neg\chi) \wedge \chi) \vee ((\neg\chi) \wedge (\neg\varphi)) \approx \top$$

We can replace $\varphi \wedge \neg\varphi$ as well as $(\neg\chi) \wedge \chi$ by $\bot$ and drop both occurrences of $\bot$ from the disjunction. Commutativity of $\wedge$ yields $\mathsf{Mal}; \varphi \approx \chi \vdash_V (\varphi \wedge \chi) \vee ((\neg\varphi) \wedge (\neg\chi)) \approx \top$, the desired result, i. e. $\varphi \leftrightarrow \chi \approx \top$. (ii) (We will now write $\varphi \approx \chi$ rather than $\mathsf{Mal} \vdash_V \varphi \approx \chi$.) Assume $(\varphi \wedge \chi) \vee ((\neg\varphi) \wedge (\neg\chi)) \approx \top$. Then $\varphi \wedge ((\varphi \wedge \chi) \vee ((\neg\varphi) \wedge (\neg\chi))) \approx \varphi \wedge \top$.

Now $\varphi \wedge \top \approx \varphi$ and so we have $\varphi \wedge ((\varphi \wedge \chi) \vee ((\neg\varphi) \wedge (\neg\chi))) \approx \varphi$. Distributing φ we get $(\varphi \wedge (\varphi \wedge \chi)) \vee (\varphi \wedge ((\neg\varphi) \wedge (\neg\chi))) \approx \varphi$. From there with associativity $((\varphi \wedge \varphi) \wedge \chi) \vee ((\varphi \wedge (\neg\varphi)) \wedge (\neg\chi)) \approx \varphi$ which results in $(\varphi \wedge \chi) \vee (\bot \wedge (\neg\chi)) \approx \chi$, by (V4) with $\varphi \wedge \varphi \approx \varphi$ and $\varphi \wedge (\neg\varphi) \approx \bot$. This gets reduced to $\varphi \wedge \chi \approx \varphi$. Similarly, one can derive $\varphi \wedge \chi \approx \chi$. So, $\mathsf{Mal}; \varphi \leftrightarrow \psi \approx \top \vdash_V \varphi \approx \chi$. ◻

Let Γ be a set of equations. We put $\mathsf{Th}(\Gamma) := \{\varphi : \Gamma \vdash_V \varphi \approx \top\}$. Now let Λ be a modal logic. Then we define $\mathsf{Eq}(\Lambda) := \{\varphi \approx \psi : \varphi \leftrightarrow \psi \in \Lambda\}$. The proof of the next theorem is left as an exercise.

Theorem 4.2.11. *Let Γ be a set of equations, Δ be a set of formulae. Then the following holds.*

(1) $\mathsf{Th}(\Gamma)$ *is a normal modal logic.*
(2) $\mathsf{Eq}(\Delta)$ *is an equational theory of modal algebras.*
(3) $\mathsf{Th}\,\mathsf{Eq}\,\mathsf{Th}(\Gamma) = \mathsf{Th}(\Gamma)$.
(4) $\mathsf{Eq}\,\mathsf{Th}\,\mathsf{Eq}(\Lambda) = \mathsf{Eq}(\Lambda)$.

Corollary 4.2.12. *There is a dual isomorphism between the lattice of normal κ–modal logics and the lattice of varieties of normal modal algebras with κ operators.*

This is a considerable strengthening of Proposition 2.2.7, Lemma 2.2.8 and Theorem 2.2.9. For now we do not only know that different logics have different varieties associated with them, we also know that different varieties have different varieties associated with them. The relation between equational calculi and deductive calculi has been a topic of great interest in the study of general logical calculi, see [**29**]. Wim Blok and Don Pigozzi have tried to isolate the conditions under which a mutual translation is possible between these two deductive formulations of a logic. It would take us too far afield to discuss these developments, however.

Notes on this section. In a series of papers Wolfgang Rautenberg partly together with Burghard Herrmann have studied the possibility to export an axiomatization of a variety in the Birkhoff–calculus to a Hilbert–style proof system for the logic determined by some unital semantics over that variety, see [**102**], [**174**] and [**173**]. The equational rules present no problem, likewise the rule of substitution. However, the rule of replacement is not straightforward; it may lead to an infinite axiomatization. (See next section on that theme.) Therefore, the so–called *finite replacement property* was defined. It guarantees that adding a finite set of instances of the rule of replacement will be sufficient for validity of all rule instances. It is a theorem by Robert C. Lyndon [**142**] that the equational theory of any 2–element algebra is finitely axiomatizable. In [**102**] it has been shown that all varieties of 2–element algebras have the finite replacement property. It follows that the logic of any 2–element matrix is finitely axiomatizable. For 3–element algebras both theorems are false.

Exercise 140. Prove Proposition 4.2.1

Exercise 141. Prove Theorem 4.2.9

4.3. Weakly Transitive Logics II

In this section we will prove some connections between purely algebraic notions and properties of logics. Again, weakly transitive logics play a fundamental role. The following definition is due to T. Prucnal and A. Wroński [**166**].

Definition 4.3.1. *Let $\vdash$ be a consequence relation. A set $\Delta(p,q) := \{\delta_i(p,q) : i \in I\}$ is called a* ***set of equivalential terms for*** *$\vdash$ if the following holds*

($eq1$) $\vdash \Delta(p,p)$
($eq2$) $\Delta(p,q) \vdash \Delta(q,p)$
($eq3$) $\Delta(p,q); \Delta(q,r) \vdash \Delta(p,r)$
($eq4$) $\bigcup_{i<\Omega(f)} \Delta(p_i,q_i) \vdash \Delta(f(\vec{p}), f(\vec{q}))$
($eq5$) $p; \Delta(p,q) \vdash q$

$\vdash$ is called ***equivalential*** *if it has a set of equivalential terms, and* ***finitely equivalential*** *if it has a finite set of equivalential terms. If $\Delta(p,q) = \{\delta(p,q)\}$ is a set of equivalential terms for $\vdash$ then $\delta(p,q)$ is called an* ***equivalential term*** *for $\vdash$.*

Let us investigate the notion of an equivalential logic for modal logics. Clearly, for any modal logic Λ, $\vdash_\Lambda$ is always equivalential; a set of equivalential terms is the following.

$$\Delta(p,q) := \{\boxplus(p \leftrightarrow q) : \boxplus \text{ a compound modality}\} .$$

Moreover, $\Vdash_\Lambda$ is always finitely equivalential; $p \leftrightarrow q$ is an equivalential term for $\Vdash_\Lambda$. Thus, the only remaining question is whether $\vdash_\Lambda$ is finitely equivalential. Note that if $\vdash_\Lambda$ is finitely equivalential it also has an equivalential term. For if $\Delta(p,q) = \{\delta_i(p,q) : i < n\}$ is a finite set of equivalential terms for $\vdash_\Lambda$ then $\delta(p,q) := \bigwedge_{i<n} \delta_i(p,q)$ is an equivalential term.

Proposition 4.3.2. *Let Λ be a modal logic and $\Delta(p,q)$ a set of equivalential terms for Λ. Then the following holds.*

(1) $\Delta(p,q) \vdash_\Lambda p \leftrightarrow q$
(2) $p \leftrightarrow q \Vdash_\Lambda \Delta(p,q)$
(3) $\Delta(p,q) \Vdash_\Lambda \Delta(p \leftrightarrow q, \top)$
(4) $\Delta(p \leftrightarrow q, \top) \Vdash_\Lambda \Delta(p,q)$

Proof. (1.) follows from (eq5) and (eq2) with the deduction theorem. For (2.) note that $p \leftrightarrow q \Vdash_\Lambda \Delta(p,p) \leftrightarrow \Delta(p,q)$ (by Proposition 3.1.7). By (eq1) we get $p \leftrightarrow q \Vdash_\Lambda \Delta(p,q)$, as desired. To prove (3.) observe that $p \leftrightarrow q \Vdash_\Lambda \Delta(p \leftrightarrow q, \top)$, again by Proposition 3.1.7. Since we have established that $\Delta(p,q) \Vdash_\Lambda p \leftrightarrow q$, the third claim follows. For (4.) observe that, by (1.), $\Delta(p \leftrightarrow q, \top) \Vdash_\Lambda p \leftrightarrow q$ and that $p \leftrightarrow q \Vdash_\Lambda \Delta(p,q)$ (by (2.)). ∎

Now let Λ be a modal logic and Σ a set of formulae. Let $\Sigma^{\vdash} := \{\psi : \Sigma \vdash_\Lambda \psi\}$ and $\Sigma^{\Vdash} := \{\psi : \Sigma \Vdash_\Lambda \psi\}$. Consider the set $\Delta(\Sigma, \top)^{\vdash}$ where

$$\Delta(\Sigma, \top) := \{\delta(\varphi, \top) : \delta \in \Delta, \varphi \in \Sigma\} .$$

This set is closed under (mp.). We show that it is closed under (mn.). Consider first $\psi \in \Sigma$. We clearly have $\Delta(\Sigma, \top) \vdash_\Lambda \Delta(\psi, \top)$ and so by (eq4) also $\Delta(\Sigma, \top) \vdash_\Lambda \Delta(\boxplus\psi, \boxplus\top)$ for all compound modalities. Since $\boxplus\top$ is a theorem, it can be substituted by $\top$. By (eq5), $\Delta(\boxplus\psi, \top) \vdash \boxplus\psi$. So, $\Delta(\Sigma, \top) \vdash \boxplus\psi$ for all $\boxplus$ and $\psi \in \Sigma$. This shows that $\Delta(\Sigma, \top) \vdash_\Lambda \boxtimes^{\omega}\Sigma$. From this we get $\Delta(\Sigma, \top) \Vdash_\Lambda \Sigma$. Hence, $\Delta(\Sigma, \top) \vdash_\Lambda \boxtimes^{\omega}\Delta(\Sigma, \top)$, by (2) of the previous theorem, since $\Sigma \Vdash_\Lambda \Delta(\Sigma, \top)$. Now let $\Delta(\Sigma, \top) \vdash_\Lambda \varphi$ for some φ. Then $\Box_j\Delta(\Sigma, \top) \vdash_\Lambda \Box_j\varphi$. Since $\Delta(\Sigma, \top) \vdash_\Lambda \Box_j\Delta(\Sigma, \top)$, we have succeeded to show that $\Box_j\varphi \in \Delta(\Sigma, \top)^{\vdash}$. Hence $\Delta(\Sigma, \top)^{\vdash} \supseteq \Sigma^{\Vdash}$. The converse inclusion is a consequence of (2) of Proposition 4.3.2.

Proposition 4.3.3. *Let Λ be a modal logic and $\Delta(p, q)$ a set of equivalential terms. Then for any set Σ*

$$\Delta(\Sigma, \top)^{\vdash} = \Sigma^{\Vdash} .$$

Definition 4.3.4. *A variety $\mathcal{V}$ has* ***equationally definable principal congruences (EDPC)*** *if there exists a number $n \in \omega$ and terms $s_i(w, x, y, z)$, $t_i(w, x, y, z)$, $i < n$, such that for every $\mathfrak{A} \in \mathcal{V}$ and $a, b, c, d \in A$*

$$\langle c, d\rangle \in \Theta(a, b) \text{ iff } s_i^{\mathfrak{A}}(a, b, c, d) = t_i^{\mathfrak{A}}(a, b, c, d) \text{ for all } i < n .$$

Proposition 4.3.5. *A variety $\mathcal{V}$ of modal algebras has EDPC iff there exists a term $u(x, y)$ such that for all $\mathfrak{A} \in \mathcal{V}$ and elements $a, b \in A$, b is in the open filter generated by a iff $u^{\mathfrak{A}}(a, b) = 1$.*

Proof. Suppose that $\mathcal{V}$ has EDPC, and let $s_i(w, x, y, z)$ and $t_i(w, x, y, z)$, $i < n$, be terms defining principal congruences in $\mathcal{V}$. Then put

$$u(x, y) := \bigwedge_{i<n} s_i(x, \top, y, \top) \leftrightarrow t_i(x, \top, y, \top) .$$

Then $u^{\mathfrak{A}}(a, b) = 1$ iff for all $i < n$ we have $s_i^{\mathfrak{A}}(a, \top, b, \top) = t_i^{\mathfrak{A}}(a, \top, b, \top)$ iff $\langle b, \top\rangle \in \Theta(a, \top)$ iff b is in the open filter generated by a. Conversely, let $u(x, y)$ be a term such that for all algebras $\mathfrak{A} \in \mathcal{V}$ b is in the open filter generated by a iff $u^{\mathfrak{A}}(a, b) = 1$. Then let $n = 1$, $s_0(w, x, y, z) := u(w \leftrightarrow x, y \leftrightarrow z)$ and $t_0(w, x, y, z) := \top$. Let $a, b, c, d \in A$, $\mathfrak{A} \in \mathcal{V}$. Then $s_0^{\mathfrak{A}}(a, b, c, d) = t_0^{\mathfrak{A}}(a, b, c, d)$ iff $u^{\mathfrak{A}}(a \leftrightarrow b, c \leftrightarrow d) = 1$ iff $c \leftrightarrow d$ is in the filter generated by $a \leftrightarrow b$ iff $\langle c, d\rangle \in \Theta(a, b)$. ◻

We say that $u(x, y)$ **defines principal open filters** in $\mathcal{V}$ if for all $\mathfrak{A} \in \mathcal{V}$ and $a, b \in A$ we have $u^{\mathfrak{A}}(a, b) = 1$ iff b is in the open filter generated by a. $\mathcal{V}$ has **definable principal open filters** (**DPOF**) iff there exists a $u(x, y)$ defining open filters. By the previous theorem, a variety of modal algebras has EDPC iff it has definable open filters. The following theorem has been obtained in [**26**].

Theorem 4.3.6 (Blok & Pigozzi & Köhler). *For any normal modal logic Λ the following are equivalent:*

(1) $\vdash_\Lambda$ *is finitely equivalential.*
(2) $\mathsf{Alg}\,\Lambda$ *has DPOF.*
(3) $\mathsf{Alg}\,\Lambda$ *has EDPC.*
(4) $\Vdash_\Lambda$ *admits a deduction theorem.*
(5) Λ *is weakly transitive.*

Proof. We have shown earlier that (4.) $\Leftrightarrow$ (5.) and we have shown in Proposition 4.3.5 that (2.) $\Leftrightarrow$ (3.). We show that (1.) $\Rightarrow$ (2.) $\Rightarrow$ (4.) and (5.) $\Rightarrow$ (1.). Assume (1.). Then there exists an equivalential term $\delta(p,q)$ for $\vdash_\Lambda$. Now put $u(p,q) := \delta(p,\top) \to q$. Let $\mathfrak{A} \in \mathsf{Alg}\,\Lambda$ and $a \in A$. Then by Proposition 4.3.3 the set $F := \{b : b \geq \delta^{\mathfrak{A}}(a,\top)\}$ is the open filter generated by a. So $b \in F$ iff $u^{\mathfrak{A}}(a,b) = 1$. Hence $u(p,q)$ defines principal open filters. Hence (2.) is proved. Now assume (2.). Suppose that $u(p,q)$ defines principal open filters in $\mathsf{Alg}\,\Lambda$. We claim that $u(p,q)$ satisfies a deduction theorem for $\Vdash_\Lambda$. Namely, let Δ be a set of formulae, and let φ and ψ be formulae. Let $\mathfrak{A}$ be a Λ–algebra, F an open filter in $\mathfrak{A}$. We can actually assume that $F = \{1\}$. Then $\Delta \vdash_{\langle\mathfrak{A},F\rangle} u(\varphi,\psi)$ iff for every valuation β such that $\overline{\beta}[\Delta] \subseteq \{1\}$ we have $u^{\mathfrak{A}}(\overline{\beta}(\varphi),\overline{\beta}(\psi)) = 1$ iff for all valuations β such that $\overline{\beta}[\Delta] \subseteq \{1\}$, $\overline{\beta}(\psi)$ is in the open filter generated by $\overline{\beta}(\varphi)$ iff for every valuation β such that $\overline{\beta}[\Delta;\varphi] \subseteq \{1\}$ we also have $\overline{\beta}(\psi) = 1$ iff $\Delta;\varphi \vdash_{\langle\mathfrak{A},F\rangle} \psi$. Thus $\Delta;\varphi \Vdash_\Lambda \psi$ iff $\Delta \Vdash_\Lambda u(\varphi,\psi)$. This shows (4.). Finally, assume (5.). Let Λ be weakly transitive with master modality $\boxplus$. Then $\boxplus(p \leftrightarrow q)$ is an equivalential term for $\vdash_\Lambda$. ◻

Definition 4.3.7. *Let $\mathfrak{A}$ be an algebra. A ternary termfunction $t(x,y,z)$ is called a* ***ternary discriminator term for*** *$\mathfrak{A}$ if the following holds*

$$t^{\mathfrak{A}}(a,b,c) = \begin{cases} c & \text{if } a = b \\ a & \text{otherwise} \end{cases}$$

Let $\mathcal{V}$ be a variety. $\mathcal{V}$ is called a ***discriminator variety*** *if there exists a class $\mathcal{K}$ of algebras such that $\mathcal{V}$ is the least variety containing $\mathcal{K}$ and there exists a term $t(x,y,z)$ which is a discriminator term for all $\mathfrak{A} \in \mathcal{K}$.*

We remark here that except in trivial cases a discriminator for an algebra $\mathfrak{A}$ cannot be a discriminator for $\mathfrak{A} \times \mathfrak{A}$. This is why the definition of a discriminator variety is somewhat roundabout.

Proposition 4.3.8. *Let $\mathfrak{A}$ be an algebra, and $t(x,y,z)$ a ternary discriminator for $\mathfrak{A}$. Then $\mathfrak{A}$ is simple.*

Proof. Let $\Theta \neq \Delta_A$ be a congruence. Then there exist $a,b \in A$ such that $a \neq b$ and $a\,\Theta\,b$. Then

$$a = t^{\mathfrak{A}}(a,b,c)\ \Theta\ t^{\mathfrak{A}}(a,a,c) = c$$

Hence $\Theta = \nabla_A$, and so $\mathfrak{A}$ is simple. ◻

Proposition 4.3.9. *Let $\mathcal{K}$ be a class of algebras, and assume that $t(x, y, z)$ is a discriminator term for all algebras of $\mathcal{K}$. Then it is a discriminator term for all members of* $\mathsf{HSUp}\,\mathcal{K}$.

Proof. Let $t(x, y, z)$ be a discriminator term for $\mathfrak{A}$; then it is obviously a discriminator term for every subalgebra of $\mathfrak{A}$. It is also a discriminator term for every homomorphic image of $\mathfrak{A}$, for the only images up to isomorphism are $\mathfrak{A}$ and the trivial algebra. Finally, let $\mathfrak{B}$ be an ultraproduct of $\mathfrak{A}_i$, $i \in I$, with ultrafilter U over I. Then the congruence Θ_U is defined as in Section 4.1. We write $\overline{a}_U$ instead of $[\overline{a}]\Theta_U$. Let $\overline{a}, \overline{b}, \overline{c} \in \prod_{i \in I} A_i$. Assume that the set D defined by $D := \{i : a_i = b_i\}$ is in U. Then $\overline{a}_U = \overline{b}_U$ and moreover $D \subseteq \{i : t^{\mathfrak{A}_i}(a_i, b_i, c_i) = c_i\}$, whence $t^{\mathfrak{B}}(\overline{a}_U, \overline{b}_U, \overline{c}_U) = \overline{c}_U$. Now assume that $D \notin U$. Then $\overline{a}_U \neq \overline{b}_U$, and $D \subseteq \{i : t^{\mathfrak{A}_i}(a_i, b_i, c_i) = a_i\}$. Thus $t^{\mathfrak{B}}(\overline{a}_U, \overline{b}_U, \overline{c}_U) = \overline{a}_U$. □

In the remaining part of this section we shall be concerned with the relationship between three properties of a variety of modal algebras: being a discriminator variety, being semisimple and being weakly transitive and cyclic. Semisimplicity is defined as follows.

Definition 4.3.10. *An algebra is called* **semisimple** *if it is a subdirect product of simple algebras. A variety is called* **semisimple** *if it consists entirely of semisimple algebras.*

Proposition 4.3.11. *Let $\mathcal{V}$ be a congruence distributive variety. Then if $\mathcal{V}$ is a discriminator variety, $\mathcal{V}$ is semisimple.*

Proof. Suppose $\mathcal{V}$ is a discriminator variety. Then it is generated by a class $\mathcal{K}$ of simple algebras. If $\mathfrak{B}$ is subdirectly irreducible, it is by Jónsson's Theorem contained in $\mathsf{HSUp}\,\mathcal{K}$. By Propositions 4.3.9 and 4.3.8, $\mathfrak{B}$ is simple. □

It can be shown in general that a discriminator variety is congruence distributive, so that the previous theorem actually holds without assuming the congruence distributivity of $\mathcal{V}$. (See exercises below.) Varieties of modal algebras have however already been shown to be congruence distributive, so we do not need to work harder here.

We shall now prove that a semisimple variety of modal algebras is weakly transitive on condition that it has only finitely many operators. Let $\mathcal{V}$ be a semisimple variety of κ–modal algebras, κ finite. We assume that there is a modality $\Box_i$ such that $\mathcal{V} \vDash \Box_i p \leftrightarrow p$. This makes life a little bit easier. Obviously, given this assumption $\mathcal{V}$ is weakly transitive iff it satisfies the equation $\boxtimes^k x = \boxtimes^{k+1} x$ for some $k \in \omega$. For a simple $\mathfrak{A} \in \mathcal{V}$ we put

$$\Omega_{\mathfrak{A}} := \{c \in A : c \neq 1 \text{ and } \boxtimes -c = 0\} .$$

We call c **dense in** $\mathfrak{A}$ if it is in $\Omega_{\mathfrak{A}}$. For $- \boxtimes -c$ is the closure of c. Denote by $\mathcal{V}_S$ the class of simple algebras from $\mathcal{V}$. Obviously, one of the following must hold for our variety $\mathcal{V}$:

(A) $(\forall n \in \omega)(\exists \mathfrak{A} \in \mathcal{V}_S)(\exists c \in \Omega_{\mathfrak{A}})(\boxtimes^n c > 0 \text{ and } \boxtimes^n (c \to \boxtimes c) \nleq c)$,

(B) $(\exists n \in \omega)(\forall \mathfrak{A} \in \mathcal{V}_S)(\forall c \in \Omega_{\mathfrak{A}})(\boxtimes^n c > 0$ implies $\boxtimes^n (c \to \boxtimes c) \leq c)$.

We will refer to $\mathcal{V}$ as being *of type A* or *of type B* depending on which of the above holds. In case B obtains, we shall denote the least number such that B holds for $\mathcal{V}$ by N.

Definition 4.3.12. *Let $\mathfrak{A}$ be a modal algebra, $\kappa < \aleph_0$. Call $c \in A$* **deep in** *$\mathfrak{A}$ if for all $m \in \omega$ we have $\boxtimes^{\leq m+1} c < \boxtimes^{\leq m} c$.*

Since $\boxtimes c \leq c$ by our assumptions, c is deep if $\boxtimes^{m+1} c < \boxtimes^m c$ for all m. Obviously, it is enough to require this to hold for almost all m. For if $\boxtimes^{m+1} c = \boxtimes^m c$ for some m, then equality holds for almost all m. Now, using the ultraproduct construction we can easily show the following:

Lemma 4.3.13. *Suppose that $\mathcal{V}$ is not weakly transitive. Then there exists an algebra in $\mathcal{V}$ containing a deep element.*

Lemma 4.3.14. *For every $c \in \mathfrak{A}$ such that $\boxtimes^k c = 0$ there is a dense $b \geq c$.*

Proof. To see this, let $b := - \boxtimes^m -c$, where m is maximal with the property $\boxtimes^m - c > 0$. Then, $\boxtimes - b = \boxtimes^{m+1} c = 0$, and $c \leq - \boxtimes^m -c = b$, as required. □

Lemma 4.3.15. *Let $k \in \omega$. If for all $\mathfrak{A} \in \mathcal{V}_S$ and all $c \in \Omega_{\mathfrak{A}}$ we have $\boxtimes^k c = 0$, then $\mathcal{V}$ satisfies $\boxtimes^{k+1} x = \boxtimes^k x$.*

Proof. Let $\mathfrak{A} \in \mathcal{V}_S$. Then by the previous lemma, for any non–unit element c of A there is a dense b such that $c \leq b$. It follows that $0 = \boxtimes^k b \geq \boxtimes^k c$. Hence we have $\boxtimes^k c = 0$ for all $c \in A - \{1\}$. Thus, $\mathfrak{A} \vDash \boxtimes^{k+1} x = \boxtimes^k x$. □

Lemma 4.3.16. *If $\mathfrak{A} \in \mathcal{V}_S$, and $\mathcal{V}$ is semisimple of type B. If c is deep in $\mathfrak{A}$ then*

$$\boxtimes^N((\boxtimes^m c \to \boxtimes^{m+1} c) \to \boxtimes(\boxtimes^m c \to \boxtimes^{m+1} c)) \leq c$$

Proof. Suppose that c is deep. Then $\boxtimes^k c > \boxtimes^{k+1} c$ for all k. Now let m be given. Then $\boxtimes^m c \to \boxtimes^{m+1} c$ belongs to $\Omega_{\mathfrak{A}}$, for $\boxtimes - (\boxtimes^m c \to \boxtimes^{m+1} c) = \boxtimes^m c \cap \boxtimes - \boxtimes^{m+1} c = \boxtimes^{m+1} c \cap \boxtimes - \boxtimes^{m+1} c \leq \boxtimes^{m+1} c \cap - \boxtimes^{m+1} c = 0$. Moreover, $\boxtimes^N(\boxtimes^m \to \boxtimes^{m+1} c) \geq \boxtimes^N \boxtimes^{m+1} c > \boxtimes^{N+m+2} c$, since c is deep. Thus, since $\mathcal{V}$ is of type B,

$$\boxtimes^N((\boxtimes^m c \to \boxtimes^{m+1} c) \to \boxtimes(\boxtimes^m c \to \boxtimes^{m+1} c)) \leq c\ .$$

□

Theorem 4.3.17 (Kowalski). *If $\mathcal{V}$ is semisimple, then $\mathcal{V}$ is weakly transitive.*

Suppose for contradiction that $\mathcal{V}$ is semisimple and not weakly transitive. Without loss of generality we may assume that it does not satisfy $\boxtimes^{n+1} x = \boxtimes^n x$, for any given $n \in \omega$. Now, for any n, we take a simple algebra $\mathfrak{A}_n$ falsifying $\boxtimes^{n+1} x = \boxtimes^n x$. Then, by Lemma 4.3.15 there is a $c_n \in A_n$ such that $\boxtimes^n c_n > 0$ and $\boxtimes - c_n = 0$, that is, $c_n \in \Omega_{\mathfrak{A}_n}$.

Now, let U be a nonprincipal ultrafilter over ω. Put $\mathfrak{A} := \prod_{n\in\omega} \mathfrak{A}_n$, and $c := \langle c_n : n \in \omega\rangle/U$. Then for all $n \in \omega$ we get $\boxtimes^n c > \boxtimes^{n+1} c > 0$ and $\boxtimes - c = 0$. So, c is both deep an dense.

Obviously, $\mathfrak{A}$ is a subdirect product of subdirectly irreducible algebras, which, by our assumption, are also simple. We will derive a contradiction from this. More precisely, we will derive a contradiction from the assumption that all subdirectly irreducible members of $\mathsf{H}(\mathfrak{A})$ are simple.

Consider the congruence $\Theta = \Theta(c, 1)$ on $\mathfrak{A}$. By the choice of c, our Θ is neither the diagonal nor the full congruence. As Θ is principal, there must be a congruence Π covered by Θ. With the choice of Π, our reasoning splits into two cases, one for either of the two types.

Type A. If $\mathcal{V}$ is of type A, then, in $\mathfrak{A}$ we have $\boxtimes^n(c \to \boxtimes c) \nleq c$ for all $n \in \omega$; hence $\Theta(c \to \boxtimes c, 1)$ is strictly below Θ. We choose a $\Pi \prec \Theta$ from the interval $I[\Theta(c \to \boxtimes c, 1), \Theta]$. Since the lattices of congruences are algebraic there always is one.

Type B. If $\mathcal{V}$ is of type B, then we just choose any $\Pi \prec \Theta$. This case has one feature that deserves to be spelled out as a separate fact.

Lemma 4.3.18. *Let $\mathcal{V}$ be of type B, and $m \in \omega$. If $\Phi \in Con(\mathfrak{A})$ is a congruence satisfying $\boxtimes^m c\ \Phi\ \boxtimes^{m+1} c$ then $\Phi \geq \Theta$.*

Proof. By the construction, c is deep in $\mathfrak{A}$. It follows from Lemma 4.3.16 that

$$\boxtimes^N((\boxtimes^m c \to \boxtimes^{m+1} c) \to \boxtimes(\boxtimes^m c \to \boxtimes^{m+1} c)) \leq c\ .$$

Now, if $\Phi \in Con(\mathfrak{A})$ satisfies $\boxtimes^m c\ \Phi\ \boxtimes^{m+1} c$, then $\boxtimes^N((\boxtimes^m c \to \boxtimes^{m+1} c) \to \boxtimes(\boxtimes^m c \to \boxtimes^{m+1} c))\ \Phi\ 1$, which implies $c\ \Phi\ 1$. So, $\Phi \geq \Theta$. □

Now we return to the main argument. We shall develop it for both types together, splitting the reasoning only when necessary. Take the set of congruences

$$\Gamma := \{\Phi \in Con(A) : \Phi \geq \Pi \text{ and } \Phi \ngeq \Theta\}\ .$$

Let $\Psi := \bigsqcup \Gamma$; by congruence distributivity, $\Psi \in \Gamma$. Then $\mathfrak{A}/\Psi$ is subdirectly irreducible. Observe that we cannot have $\Gamma = \{\Pi\}$, for in such a case $\mathfrak{A}/\Psi = \mathfrak{A}/\Pi$, and this would be subdirectly irreducible but non–simple, since $\Theta \neq \nabla$. Thus, since $\Psi \nleq \Theta$, we obtain that $\Theta \sqcup \Psi$ is full (which by congruence permutability equals $\Theta \circ \Psi$). For otherwise $\mathfrak{A}/\Psi$ would be a non–simple subdirectly irreducible algebra in $\mathcal{V}$. Now, as $\mathfrak{A}/\Psi = (\mathfrak{A}/\Pi)/\Psi$, and since principal congruences remain principal in homomorphic images, we can shift the whole argument to $\mathfrak{A}/\Pi$.

Let $\mathfrak{B} := \mathfrak{A}/\Pi$. As Θ and Ψ are above Π in $Con(\mathfrak{A})$ we shall write Θ and Ψ in place of Θ/Π and Ψ/Π from $Con(\mathfrak{B})$. Thus, for instance, we will say that Θ (and not Θ/Π) is an atom of $Con(\mathfrak{B})$. We shall also write c instead of $[c]_\Pi$.

Now we have an algebra $\mathfrak{B}$; a principal congruence $\Theta \in Con(\mathfrak{B})$ such that $\Theta \succ 0$; further, there is a congruence Ψ — which is non–principal in general — which is the largest congruence not containing Θ, and $\Theta \circ \Psi = 1$. Hence, there is an element

$d \in B - \{1\}$ with $(d, 1) \in \Theta$ and $(d, 0) \in \Psi$. The latter is equivalent to $(-d, 1) \in \Psi$. It follows that $d \geq \boxtimes^m c$ for some $m \in \omega$. Moreover, as $\Omega_{\mathfrak{A}}/\Pi \subseteq \Omega_{\mathfrak{B}}$, we obtain that $c(= [c]_\Pi) \in \Omega_{\mathfrak{B}}$. Thus, $c \in \Omega_{\mathfrak{B}} - \{1\}$.

The statement $(-d, 1) \in \Psi$ can be further broken up as follows: first,

$$\Psi = \bigsqcup\{\Phi \in Cp(\mathfrak{A}) : 0 < \Phi \leq \Psi\} ,$$

where $Cp(\mathfrak{A})$ stands for the set of all compact congruences of $\mathfrak{A}$ (which are also the principal congruences of $\mathfrak{A}$). Thus, each Φ above is of the form $\Phi = \Theta(b, 1)$ for some $b \neq 1$. Let $C \subset B$ be the set of all such $b \in B$. Passing from congruences to open filters we obtain: $-d \in [\bigcup_{b \in C}[\boxtimes^n b : n \in \omega))$.

Secondly, C is a downward directed set (in fact, a filter, but we do not need that). For take $b_0, \dots, b_{k-1} \in C$, and let $b = b_0 \cap \dots \cap b_{k-1}$. Then, $\Theta(b, 1) = \Theta(b_0, 1) \vee \dots \vee \Theta(b_{k-1}, 1)$, and all the congruences on the right–hand side of the equation are below Ψ by definition. Thus, so is $\Theta(b, 1)$; hence $b \in C$, as needed. Moreover, all $b \in C$ satisfy $\forall n \in \omega : \boxtimes^n b \nleq a$; otherwise some congruence below Ψ would contain Θ, which is impossible.

Thirdly, we have: $-d \in [\bigcup_{b \in C}[\boxtimes^n b : n \in \omega))$ iff $-d \geq d_0 \cap \dots \cap d_{k-1}$, with $d_i \in [\boxtimes^n b_i : n \in \omega)$ $(0 \leq i \leq k - 1)$. Since $\boxtimes$ distributes over meet, this gives

$$-d \in [\boxtimes^n(b_0 \cap \dots \cap b_{k-1}) : n \in \omega) ,$$

and by the previous argument $b_0 \cap \cdots \cap b_{k-1} \in C$.

Gathering all this together, we get $-d \in [\boxtimes^n b : n \in \omega)$ for some $b \in B$ such that $\forall n \in \omega : \boxtimes^n b \nleq c$. On the other hand, $d \in [\boxtimes^n c : n \in \omega)$. Therefore, $d \geq \boxtimes^m c$ for some m, and $-d \geq \boxtimes^k b$, for some k. Thus, $\boxtimes^m c \leq d \leq - \boxtimes^k b$; in particular, $\boxtimes^m c \leq - \boxtimes^k b$.

Consider $q := - \boxtimes^k b \to \boxtimes^m c = \boxtimes^k b \vee \boxtimes^m c$. As $q \geq \boxtimes^m c$ we have:

$$\Theta(q, 1) \leq \Theta(\boxtimes^m c, 1) = \Theta(c, 1) = \Theta .$$

Since Θ is an atom, $\Theta(q, 1)$ is either 0 or Θ.

Let us first deal with the case $\Theta(q, 1) = \Theta = \Theta(c, 1)$. We then have: $\boxtimes^r q \leq c$, for some r. This yields: $c \geq \boxtimes^r(\boxtimes^k b \vee \boxtimes^m c) \geq \boxtimes^{r+k} b$. Thus, $c \geq \boxtimes^{r+k} b$, which is a contradiction.

The remaining possibility is $\Theta(q, 1) = 0$. Then we have $q = 1$, and that means $- \boxtimes^k b \leq \boxtimes^m c$. Together with the inequality from the previous paragraph, this implies $- \boxtimes^k b = \boxtimes^m c$. Two cases arise.

Type A. By the choice of Π we have that $c \to \boxtimes c \;\Pi\; 1$ in $\mathfrak{A}$, and so $c \;\Pi\; \boxtimes c$. Hence $c = \boxtimes c$ in $\mathfrak{B}$, and we get $\boxtimes^k b = -c$. Further, since c is dense, we have $\boxtimes - c = 0$, from which we get $\boxtimes^{k+1} b = 0$. However, $(b, 1) \in \Psi$, and so $0 = \boxtimes^{k+1} b \;\Psi\; 1$. Hence Ψ is full. Contradiction.

Type B. If either $\boxtimes - \boxtimes^k b = \boxtimes^{m+1} c < \boxtimes^m c$, or $\boxtimes - \boxtimes^m c = \boxtimes^{k+1} b < \boxtimes^k b$, then the congruences $\Theta(b, 1)$ and $\Theta(c, 1)$ have a non–trivial intersection; namely, both contain the pair $(\boxtimes^{m+1} c \vee \boxtimes^{k+1} b, 1)$. This pair is not in Δ. This cannot happen, for then

$\Theta \sqcap \Psi > 0$, which contradicts the definition of Ψ. So, we obtain that $\boxtimes^{m+1} c = \boxtimes^m c$ and $\boxtimes - \boxtimes^m c = - \boxtimes^m c$ in $\mathfrak{B}$. Hence, $\boxtimes^{m+1} c \ \Pi \ \boxtimes^m c$ in $\mathfrak{A}$, and since $\mathcal{V}$ is of type B, we can apply Lemma 4.3.18 to get that $\Pi \geq \Theta$. This, however, cannot happen either, since $\Pi < \Theta$, by its definition. This completes the proof of Theorem 4.3.17.

Recall from Section 2.5 that Λ is called cyclic if for every basic modal operator $\Box_i$ there exists a compound modality $\boxminus$ such that $p \to \Box_i \Diamond p \in \Lambda$. If Λ is weakly transitive with master modality $\boxplus$ this is equivalent to the requirement that $\boxplus$ satisfies **S5**.

Lemma 4.3.19. ($\kappa < \aleph_0$.) *Let Λ be a modal logic. Then if Λ is weakly transitive and cyclic,* $\mathsf{Alg}\,\Lambda$ *is a discriminator variety and semisimple.*

Proof. Assume that Λ is weakly transitive with master modality $\boxplus$. First we show that $\mathsf{Alg}\,\Lambda$ is semisimple. Let $\mathfrak{A}$ be an algebra and let $a \in A$. Assume that a is open, that is, $a = \boxplus a$. We claim that $-a$ is also open. Namely, from $a \leq \boxplus a$ we conclude that $- \boxplus -(-a) \leq -a$. Therefore

$$-a \leq \boxplus(- \boxplus -(-a)) \leq \boxplus - a \leq -a\,.$$

Hence $\boxplus - a = -a$. Thus $-a$ is open. Now assume that $\mathfrak{A}$ is not simple. Then there exists a proper open filter F such that Θ_F is not the monolith. It is easy to see that F can be assumed to be generated by a single element a. By weak transitivity, we can also assume that a is open, and that $F = \{b : b \geq a\}$. Then $G := \{b : b \geq -a\}$ is also an open filter. Moreover, $F \cap G = \{1\}$. Hence $\Theta_G \cap \Theta_F = \Delta_A$. Hence $\mathfrak{A}$ is directly decomposable. Thus, $\mathsf{Alg}\,\Lambda$ is semisimple.

Now put $t(x, y, z) := \boxplus(x \leftrightarrow y) \wedge z. \vee .\neg \boxplus (x \leftrightarrow y) \wedge x$. We claim that $t(x, y, z)$ is a discriminator term. To that effect, take a subdirectly irreducible algebra. It is simple, by since $\mathsf{Alg}\,\Lambda$ is semisimple. So, 0 is an opremum and $\boxplus a = 0$ iff $a \neq 1$. Let $a = b$, and c be any element. Then $t^{\mathfrak{A}}(a, b, c) = \boxplus(a \leftrightarrow b) \cap c. \cup . - \boxplus(a \leftrightarrow b) \cap a = c$. Let now $a \neq b$. Then $a \leftrightarrow b \neq 1$ and $\boxplus(a \leftrightarrow b) = 0$. Thus $t^{\mathfrak{A}}(a, b, c) = a$. This shows that $\mathsf{Alg}\,\Lambda$ is a discriminator variety. ◻

Theorem 4.3.20 (Kowalski). ($\kappa < \aleph_0$.) *Let Λ be a modal logic. Then the following are equivalent:*

(1) Λ *cyclic and is weakly transitive.*
(2) $\mathsf{Alg}\,\Lambda$ *is semisimple.*
(3) $\mathsf{Alg}\,\Lambda$ *is a discriminator variety.*

Proof. By Lemma 4.3.19 (1) implies both (2) and (3), and by Proposition 4.3.11 (3) implies (2). It remains to be shown that (2) implies (1). So, suppose that $\mathsf{Alg}\,\Lambda$ is semisimple. By Theorem 4.3.17 Λ is weakly transitive. What remains to be shown is that if $\mathsf{Alg}\,\Lambda$ is semisimple, Λ must be cyclic. Since Λ is weakly transitive and κ finite, there is a maximal compound modality, $\boxplus$. Suppose that Λ is not cyclic. Then $p \to \boxplus\neg \boxplus \neg p \notin \Lambda$. Hence there exists a subdirectly irreducible algebra $\mathfrak{A}$ such that $\mathfrak{A} \nvDash p \to \boxplus\neg \boxplus \neg p$. So there exists a $b \in A$ such that $b \cap - \boxplus - \boxplus -b \neq 0$. Consider the open filter F generated by $-b$. $F = \{c : c \geq \boxplus - b\}$, by the assumption that $\boxplus$ is a strongest compound modality. Suppose that $\boxplus{-b} = 0$. Then $-\boxplus-\boxplus-b = -\boxplus-0 = 0$, a contradiction. So, $F \neq A$. Suppose that $\boxplus - b = 1$. Then, since $\boxplus - b \leq -b$, $-b = 1$. It follows that $b = 0$, again a contradiction. So, $F \neq \{1\}$. We conclude that $\mathfrak{A}$ is not simple. This contradicts our assumption. So, Λ must be cyclic. □

Notes on this section. The complex of ideas surrounding the property of equationally definable principal congruences has been studied in the papers [**28**], [**26**], [**30**] and [**27**] by Wim Blok, Don Pigozzi and Peter Köhler. In [**26**] it is proved that if a congruence–permutable variety is semisimple it has EDPC iff it is a discriminator variety.

Exercise 142. Let $\mathfrak{A}$ be a modal algebra. Denote by $\mathfrak{Cp}(\mathfrak{A})$ the semilattice of compact congruences. Show that if $\mathsf{Th}\,\mathfrak{A}$ is weakly transitive and cyclic then $\mathfrak{Cp}(\mathfrak{A})$ is the semilattice reduct of a boolean algebra.

Exercise 143. Show that a finite tense algebra is semisimple.

Exercise 144. An algebra $\mathfrak{A}$ is called **hereditarily simple** if every subalgebra of $\mathfrak{A}$ is simple. Show that every simple modal algebra is hereditarily simple.

Exercise 145. Show that a discriminator variety is congruence distributive. *Hint.* First show that it has permuting congruences. Then show that the variety is congruence distributive.

4.4. Stone Representation and Duality

This section provides the rudiments of representation theory and duality theory. For a proper understanding of the methods it is useful to learn a bit about category theory. In this section, we provide the reader with the essentials. More can be found in [**143**], [**101**] or [**80**]. The basic notion is that of a *category*. A **category** is a structure $\mathfrak{C} = \langle \mathsf{Ob}, \mathsf{Mor}, dom, cod, \circ, id \rangle$ where Ob is a class, called the class of **objects**, Mor another class, the class of **morphisms**, $dom, cod : \mathsf{Mor} \to \mathsf{Ob}$ two functions assigning to each morphism a **domain** and a **codomain**, $\circ : \mathsf{Mor} \times \mathsf{Mor} \to \mathsf{Mor}$ a *partial* function, assigning to suitable pairs of morphisms their **composition**, and $id : \mathsf{Ob} \to \mathsf{Mor}$ a function assigning to each object a morphism, the **identity** on that object. We write $f : A \to B$ or $A \xrightarrow{f} B$ to state that f is a morphism with domain A and codomain B. We also say that f is an **arrow** from A to B. We require the following.

(1.) For morphisms f, g the composition $f \circ g$ is defined iff $cod(g) = dom(f)$.
(2.) For every object A, $dom(id(A)) = cod(id(A)) = A$.
(3.) For every morphism $f : A \to B$ we have $f \circ id(A) = f$ and $id(B) \circ f = f$.
(4.) If $f : A \to B$, $g : B \to C$ and $h : C \to D$ then $h \circ (g \circ f) = (h \circ g) \circ f$.

For example, take the class of sets with the functions defined as usual, this gives rise to a category called Set. To give another example, let $\mathcal{L}$ be a language of signature Ω and T an equational theory over Ω, then the class $\mathsf{Alg}\,T$ of Ω–algebras for T form a category with the morphisms being the Ω–homomorphisms. An arrow $f : A \to B$ is an **isomorphism** if there is a $g : B \to A$ such that $g \circ f = id(A)$ and $f \circ g = id(B)$. If $f : A \to B$ is an isomorphism, A and B are said to be **isomorphic**. A pair $\Delta = \langle O, M\rangle$, where O is a class of objects and M a class of morphisms of $\mathcal{C}$ such that for each $p \in M$ we have $dom(p), cod(p) \in O$ is called a **diagram**. A diagram **commutes** if for any three arrows $A \stackrel{f}{\to} B$, $B \stackrel{g}{\to} C$ and $A \stackrel{h}{\to} C$ of M we have $h = g \circ f$.

$$\begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, "h = g \circ f"'] & & B \arrow[dl, "g"] \\ & C & \end{tikzcd}$$

Definition 4.4.1. *Let $\mathcal{C}$ be a category. Put $\mathsf{Ob}^{op} := \mathsf{Ob}$, $\mathsf{Mor}^{op} := \mathsf{Mor}$, $id^{op} := id$; moreover, put $dom^{op} := cod$ and $cod^{op} := dom$, and finally $f \circ^{op} g := g \circ f$. Then $\mathcal{C}^{op}$ defined as $\mathcal{C}^{op} := \langle \mathsf{Ob}^{op}, \mathsf{Mor}^{op}, \mathsf{dom}^{op}, cod^{op}, \circ^{op}, id^{op}\rangle$ is called the **dual** or **opposite category**.*

The dual category arises by reversing the direction of the arrows. So if for example $f : A \to B$ is a map, then there is a dual map from B to A in $\mathcal{C}^{op}$, which is usually also called f. The notation is rather unfortunate here since it uses the same name for the objects and for the arrows and only assigns a different direction to the arrow in the opposite category. So, the arrow exists in the category as well as in the dual category and which way it goes is determined by the context. To remove this ambiguity we will write $f^{op} : B \to A$ to distinguish it from the corresponding arrow $f : A \to B$ of $\mathcal{C}$.

Definition 4.4.2. *Let $\mathcal{C}$ and $\mathcal{D}$ be categories, F both a map from the objects of $\mathcal{C}$ to the objects of $\mathcal{D}$ and from the morphisms of $\mathcal{C}$ to the morphisms of $\mathcal{D}$. Then F is called a **covariant functor** if*

1. *$F(dom(f)) = dom(F(f))$, $F(cod(f)) = cod(F(f))$,*
2. *$F(g \circ f) = F(g) \circ F(f)$ and*
3. *$F(id(A)) = id(F(A))$.*

*F is called a **contravariant functor** if*

1. *$F(cod(f)) = dom(F(f))$, $F(dom(f)) = cod(F(f))$,*
2. *$F(g \circ f) = F(f) \circ F(g)$ and*
3. *$F(id(A)) = id(F(A))$.*

Thus, a covariant functor maps $f : A \to B$ into $F(f) : F(A) \to F(B)$ while a contravariant functor maps f into $F(f) : F(B) \to F(A)$. Therefore, as the direction of the arrows is reversed, the composition must also work in reverse order. Now, an essential feature of modal duality theory is to find well–behaved functors between the various categories that arise in modal logic. Typically, in the standard algebraic tradition one would be content to define just representation theories for the objects of that category (e. g. of finite boolean algebras as algebras of subsets of a set). However, from the categorial perspective the most natural representation is that which is in addition also functorial. The advantage is for example that the characterization of modally definable classes of algebras (namely varieties) is transferred to a characterization of modally definable classes of frames. For we do not only know something about the objects, we also learn something about the maps between them. An instructive example is the representation theorem for modal algebras. There is a function mapping the modal algebras to descriptive frames and functions between modal algebras to p–morphisms; in other words we have a *functor* between the respective categories. We will show that this functor is contravariant and has an inverse. So, every p–morphism between frames is the image of a homomorphism between the corresponding algebras under the functor. This has far reaching consequences. A similar example is Theorem 4.4.10 of this section. If F is a functor from $\mathcal{C}$ to $\mathcal{D}$ and G a functor from $\mathcal{D}$ to $\mathcal{E}$ then $G \circ F$ is a functor from $\mathcal{C}$ to $\mathcal{E}$. $G \circ F$ is covariant if F and G are both covariant or both contravariant; $G \circ F$ is contravariant if exactly one of F and G is contravariant.

The following construction will be a major tool in representation theory, both in this section and in Section 7.4. For the category $\mathcal{C}$ let $hom_{\mathcal{C}}(A, B)$ denote the class of arrows from A to B. A category $\mathcal{C}$ is called **locally small** if $hom_{\mathcal{C}}(A, B)$ is a *set* for all objects A, B. All categories considered in this book will be locally small. If $\mathcal{C}$ is locally small then for every object A the map $hom_{\mathcal{C}}(-, A)$ can be turned into a contravariant functor H_A from $\mathcal{C}$ into the category Set of sets and functions. Namely, for an object B we put $H_A(B) := hom_{\mathcal{C}}(B, A)$ and for a function $f : C \to B$ we put $H_A(f) : H_A(B) \to H_A(C) : g \mapsto g \circ f$. This is well defined. It is a functor; for let $e : D \to C$, $f : C \to B$ and $g : B \to A$. Then $H_A(f \circ e) : g \mapsto g \circ (f \circ e)$ and $H_A(e) \circ H_A(f) : g \mapsto (g \circ f) \circ e$. These two functions are the same, by definition of a category. Furthermore, if $f = id(B)$ then $H_A(f) = id(H_A(B))$, as is easily verified. This shows that we have a contravariant functor.

Proposition 4.4.3. *Let $\mathcal{C}$ be a locally small category and A an object in $\mathcal{C}$. Define H^A by $H^A(B) := hom_{\mathcal{C}}(A, B)$ and $H^A(f) : H^A(B) \to H^A(C) : g \mapsto f \circ g$, where $f : B \to C$ is an arrow in $\mathcal{C}$. Also, define H_A by putting $H_A(B) := hom_{\mathcal{C}}(B, A)$ and $H_A(f) : H_A(C) \to H_A(B) : g \mapsto g \circ f$. Then H^A is a covariant functor from $\mathcal{C}$ into* Set *and H_A a contravariant functor from $\mathcal{C}$ into* Set.

Usually, H^A is called the **covariant hom–functor** and H_A the **contravariant hom–functor**. Let us apply this to boolean algebras. By BA we denote the category of boolean algebras with boolean homomorphisms. It is a locally small category as is easily verified. By Proposition 1.7.10 an ultrafilter U on an algebra $\mathfrak{A}$ defines a homomorphism $f_U : \mathfrak{A} \twoheadrightarrow \mathbf{2}$ by putting $f_U(a) = 1$ iff $a \in U$. Conversely, for every map $f : \mathfrak{A} \twoheadrightarrow \mathbf{2}$ the set $f^{-1}(1)$ is an ultrafilter on $\mathfrak{A}$. Finally, every map $f : \mathfrak{A} \to \mathbf{2}$ is surjective, since $f(1^{\mathfrak{A}}) = 1$ and $f(0^{\mathfrak{A}}) = 0$. We will henceforth call maps $\mathfrak{A} \twoheadrightarrow \mathbf{2}$ **points** of the algebra $\mathfrak{A}$. The contravariant hom–functor $H_{\mathbf{2}}$ will now be denoted by $(-)_*$. To give the reader a feeling for the construction we will repeat it in this concrete case. We denote the set of points of a boolean algebra $\mathfrak{A}$ by $\mathfrak{A}_*$. Moreover, for any algebra $\mathfrak{B}$ and boolean homomorphism $f : \mathfrak{B} \to \mathfrak{A}$ the map $f_* : \mathfrak{A}_* \to \mathfrak{B}_*$ is defined by $f_*(g) := g \circ f$.

$$\begin{tikzcd}
\mathfrak{B} \arrow[rr, "f"] \arrow[dr, "f_*(g) := g \circ f"'] & & \mathfrak{A} \arrow[dl, "g"] \\
& \mathbf{2} &
\end{tikzcd}$$

Let us investigate this a bit closer. We will show that f is surjective iff f_* is injective; and that f is injective iff f_* is surjective. Assume f is surjective. Let g and h be two two different points of $\mathfrak{A}$. Then there is an a such that $g(a) \neq h(a)$. Therefore, there is an element $b \in f^{-1}(a)$ with $g \circ f(b) \neq h \circ f(b)$. Hence $f_* : \mathfrak{A}_* \rightarrowtail \mathfrak{B}_*$. Now assume that f is not surjective. Then let $im[f]$ be the direct image of f in $\mathfrak{A}$. Since this is not the whole algebra, there is an element a such that neither a nor $-a$ is in $im[f]$. Now take an ultrafilter U on $\mathfrak{A}$. Let $V := U \cap im[f]$ be the restriction of U to $im[\mathfrak{B}]$. Then V is an ultrafilter on $im[f]$ (can you see why this is so?). Now both the set $V \cup \{a\}$ and $V \cup \{-a\}$ have the finite intersection property in $\mathfrak{A}$, hence can be extended to ultrafilters V^1 and V^2. Then $f^{-1}[U] = f^{-1}[V^1] = f^{-1}[V^2]$, and so f_* is not injective since $V_1 \neq V_2$. This concludes the case of surjectivity of f. Now assume that f is injective. We will show that f_* is surjective. So let $g : \mathfrak{B} \to \mathbf{2}$. Put $F := f[g^{-1}(1)]$, the direct image of the ultrafilter defined by g. This generates a filter in $\mathfrak{A}$ and can be extended to an ultrafilter. Hence f_* is surjective. Now let f be not injective. Then there exist elements a and b such that $a \neq b$ but $f(a) = f(b)$. This means $f(a \leftrightarrow b) = 0$, but $a \leftrightarrow b \neq 0$. Then either $a \cap -b \neq 0$ or $b \cap -a \neq 0$. Hence there is an ultrafilter containing one but not the other. This ultrafilter is not of the form $f^{-1}[U]$ for any ultrafilter on $\mathfrak{A}$.

Theorem 4.4.4. $H_{\mathbf{2}} : \mathfrak{A} \mapsto \mathfrak{A}_*$ *is a contravariant functor from the category of boolean algebras to the category of sets. Moreover,* f_* *is injective iff* f *is injective, and* f_* *is surjective iff* f *is injective.*

Now let us try to define an inverse functor from Set to BA. Take the set $2 = \{0, 1\}$ and look at $hom_{\mathsf{Set}}(X, 2)$. It is not difficult to turn this into a boolean algebra. Namely, observe that for a function $f : X \to 2$ each fibre $f^{-1}(1)$ is a subset of X, and every subset $A \subseteq X$ defines a function χ_A by $\chi_A(x) = 1$ iff $x \in A$. This function is known as the **characteristic function** of A. Now define

$$\begin{array}{lll} 0 & := & \chi_\varnothing \\ 1 & := & \chi_X \\ -\chi_A & := & \chi_{X-A} \\ \chi_A \cap \chi_B & := & \chi_{A\cap B} \\ \chi_A \cup \chi_B & := & \chi_{A\cup B} \end{array}$$

So, for X we let X^* be the boolean algebra defined on the set $hom_{\mathsf{Set}}(X, 2)$. And for $f : Y \to X$ let $f^*(g) := g \circ f$.

$$\begin{tikzcd} Y \arrow[rr, "f"] \arrow[dr, "f^*(g) := g \circ f"'] & & X \arrow[dl, "g"] \\ & 2 & \end{tikzcd}$$

Theorem 4.4.5. *The map $X \mapsto X^*$ is a contravariant functor from the category of sets into the category of boolean algebras. Moreover, f^* is injective iff f is surjective and f^* is surjective iff f is injective. f^* is called the* ***powerset functor****.*

The proof is left to the reader. Now, we have a contravariant functor from BA to Set and a contravariant functor from Set to BA. Combining these functors we get a covariant functor from BA to BA and from Set to Set. For example, starting with a boolean algebra $\mathfrak{A}$ we can form the set $\mathfrak{A}_*$ and then turn this into a boolean algebra $(\mathfrak{A}_*)^*$. (The latter operation we sometimes refer to as *raising* a set into a boolean algebra.) Likewise, we can start with the set X, raise this to a boolean algebra X^* and form the point set $(X^*)_*$. Unfortunately, in neither case we can expect to have an isomorphism. This is analogous to the case of vector spaces and their biduals. In the finite dimensional case there is an isomorphism between a vector space and its bidual, but in the infinite dimensional case there is only an embedding of the former into the latter. The same holds here. The map $x \mapsto U_x = \{Y \subseteq X : x \in Y\}$ embeds X into the point set of X^*. The map $a \mapsto \widehat{a} = \{f : \mathfrak{A} \to \mathbf{2} : f(a) = 1\}$ embeds $\mathfrak{A}$ into the powerset–algebra over the points of $\mathfrak{A}$.

Thus the situation is not optimal. Nevertheless, let us proceed further along this line. First of all, if intuitively boolean logic is the logic of sets, then in a way we have succeeded, because we have shown that anything that satisfies the laws of boolean algebras is in effect an algebra of sets. We cash out on this as follows. Let $\mathfrak{A}$ be an algebra, X a set. We call a boolean homomorphism $f : \mathfrak{A} \rightarrowtail X^*$ a **realization** of $\mathfrak{A}$. A realization turns the elements of the algebra into subsets of X, and interprets the operations on $\mathfrak{A}$ as the natural ones on sets. (Actually, X^* was construed via the hom–functor, so we have on the right hand side the characteristic functions rather than the sets, but this is inessential for the argument here, since the two algebras are isomorphic.) What we have proved so far is that every boolean algebra can be realized. Now, in order to recover $\mathfrak{A}$ in X^* we do not need the operations — because they are standard. All we need is the collection $f[A] = \{f(a) : a \in A\}$. So, we can represent $\mathfrak{A}$ by the pair $\langle X, f[A]\rangle$. There is a structure in mathematics which is almost of that form, namely a *topological space*. Recall that a **topological space** is a pair $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ where $\mathbb{X}$ is a collection of subsets of X, called the set of **open sets**, which contains $\varnothing$, X and which is closed under finite intersections and arbitrary unions. Maps between topological spaces are the **continuous functions**, where a function $f : X \to Y$ is a continuous function from $\langle X, \mathbb{X}\rangle$ to $\langle Y, \mathbb{Y}\rangle$ if for every $A \in \mathbb{Y}$, $f^{-1}[A] \in \mathbb{X}$. Alternatively, since the open sets of $\mathfrak{X}$ form a locale $\Omega(\mathfrak{X}) := \langle \mathbb{X}, \cap, \bigcup\rangle$ with finite meets and infinite joins, we can say that a function is continuous if the function $\Omega(f) : \Omega(\mathfrak{Y}) \to \Omega(\mathfrak{X})$ defined by $\Omega(f)(A) := f^{-1}[A]$ is a homomorphism preserving finite meets and infinite joins. Let $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ be a topological space. A set $A \subseteq X$ is **clopen** in the space $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ if both A and $X - A$ are open. $\mathfrak{X}$ is **discrete** if every subset is open. $\mathfrak{X}$ is discrete iff for every $x \in X$ the singleton set $\{x\}$ is open. $\mathfrak{X}$ is called **compact** if for every union $\bigcup_I x_i = X$ of open sets x_i we have a finite subset $J \subseteq I$ such that $\bigcup_J x_i = X$. Finally, a subset $\mathbb{B}$ of $\mathbb{X}$ is called a **basis** of the topology if every open set is the (possibly infinite) union of members of $\mathbb{B}$.

Definition 4.4.6. *A topological space is* ***zero–dimensional*** *if the clopen sets are a basis for the topology.*

Let $\mathfrak{A}$ be a boolean algebra. Put $X := pt(\mathfrak{A})$. For $a \in A$ put

$$\widehat{a} := \{p \in X : p(a) = 1\} .$$

Let $\mathbb{A} := \{\widehat{a} : a \in A\}$. $\mathbb{A}$ is closed under all boolean operations. Now let $\mathbb{X}$ be the set of all unions of members of $\mathbb{A}$. Alternatively, $\mathbb{X}$ is the smallest topology induced by $\mathbb{A}$ on X. Put $\mathfrak{A}_o := \langle X, \mathbb{X}\rangle$. $\mathfrak{A}_o$ is a zero–dimensional topological space. Before we set out to study the functorial properties of this map, let us turn to a fundamental problem of this construction, namely how to recover the set $\mathbb{A}$ when given the topology $\mathbb{X}$. That we can succeed is not at all obvious. So far we have a set X and a collection of clopen subsets forming a boolean algebra which is a basis for the topology. To show that the clopen sets are not always reconstructible take $X = \omega$, B the collection of finite and cofinite sets and C the collection of all subsets of X. Both B and C are a basis of the same topology, namely the discrete topology. In this topology, every subset is clopen.

Proposition 4.4.7. *Let $\langle X, \mathbb{X}\rangle$ be a compact topological space. Assume $\mathbb{X}_0$ is a basis for the topology and that $\mathbb{X}_0$ is closed under complements and finite unions. Then for $x \in \mathbb{X}$ we have $x \in \mathbb{X}_0$ iff x is clopen.*

Proof. If x is in $\mathbb{X}_0$ then it is clearly clopen, since $X - x \in \mathbb{X}_0$ as well. Now assume that x is clopen. Let $x = \bigcup_I y_i$, $X - x = \bigcup_J z_j$ be two representations such that $y_i, z_j \in \mathbb{X}_0$ for all $i \in I$ and $j \in J$. Now $X = x \cup (X - x) = (\bigcup_I y_i) \cup (\bigcup_J z_j)$. Thus there is a finite $K \subseteq I$ and a finite $L \subseteq J$ such that $X = (\bigcup_K y_i) \cup (\bigcup_L z_j)$. Then $x = x \cap ((\bigcup_K y_i) \cup (\bigcup_L z_j)) = \bigcup_K (x \cap y_i) \cup \bigcup_L (x \cap z_j) = \bigcup_K x \cap y_i = \bigcup_K y_i \in \mathbb{X}_0$. □

As before, a map $f : \mathfrak{B} \to \mathfrak{A}$ induces a continuous function $f_o : \mathfrak{A}_o \to \mathfrak{B}_o$ via $f_o(g) := g \circ f$. For first of all we have the set–map $f_* : \mathfrak{A}_* \to \mathfrak{B}_*$ and we simply define $f_o(\bigcup_I x_i) := \bigcup_I f_o(x_i)$. (The reader is asked to verify that this definition is consistent. This is not entirely harmless.) This defines the first functor. For the opposite direction, take a topological space $\mathfrak{X}$ and define the set of points by $hom_{\mathsf{Top}}(\mathfrak{X}, \mathbf{2})$, where $\mathbf{2}$ is the discrete topological space with 2 elements. (So, $\mathbf{2} = \langle\{0, 1\}, \{\varnothing, \{0\}, \{1\}, \{0, 1\}\}\rangle$.) Such functions are uniquely characterized by the set on which they give the value 1. So they are of the form χ_x for a set x. x must be open, being of the form $f^{-1}(1)$, and closed, being the complement of $f^{-1}(0)$. Thus, we get as before a boolean algebra of functions χ_x, namely for all clopen elements. We call this algebra $\mathfrak{X}^o$.

Now, do we have $\mathfrak{A} \cong (\mathfrak{A}_o)^o$ as well as $(\mathfrak{X}^o)_o$? Let us begin with the first question. We have defined $\mathfrak{A}_o$ on the set of points, or — equivalently — on the set of ultrafilters. We will show that this space is compact, allowing us to recover the original boolean algebra as the algebra of clopen sets. Since this is the algebra we will get when raising via ^{o}, we do in fact have the desired isomorphism. A space is compact if for any intersection $\bigcap_I x_i$ of closed sets there is a finite $J \subseteq I$ such that $\bigcap_J x_i = \varnothing$. Alternatively, consider a family of sets $\langle x_i : i \in I\rangle$ such that any finite subfamily has non–empty intersection. This is the finite intersection property defined in Proposition 1.7.11. If a space is compact, such a family must have non–empty intersection. Now consider a set $\{S_i : i \in I\}$ of closed sets in $\mathfrak{A}_o$ with the finite intersection property. Each S_i is an intersection of sets of the form $\widehat{a}$. Without loss of generality we may therefore assume that we have a family $\langle\widehat{a}_j : j \in J\rangle$ of elements $\widehat{a}_j$ with the finite intersection property. Then there is an ultrafilter U containing that family. There is a function f_U such that $f^{-1}(1) = U$. Hence, $f_U \in \bigcap_J \widehat{a}_j$ and so the intersection is not empty. So, $\mathfrak{A}_o$ is a compact space. Thus, by Proposition 4.4.7, $\mathfrak{A} \cong (\mathfrak{A}_o)^o$. In general, for a compact topological space $\mathfrak{X} \cong (\mathfrak{X}^o)_o$ does not hold. Thus, we must restrict the class of topological spaces under consideration. This leads to the following definition.

Definition 4.4.8. *A topological space is called a* ***Stone space*** *if it is compact and for two different points there is a clopen set containing one but not the other. The category of Stone spaces and continuous maps between them is denoted by* StoneSp.

Consider this in contrast to the separation axioms T_0 and T_2. A topological space is a **Hausdorff space** or $\mathbf{T_2}$**–space** if whenever x and y are distinct there are disjoint open sets U and V such that $x \in U$ and $y \in V$. A space is a T_0**–space** if for every given pair x, y of points there exists an open set A such that $\sharp(A \cap \{x, y\}) = 1$. (More about T_0–spaces in Section 7.4.) These two conditions are not the same. For example, the topological space over $0, 1$ with the sets $\varnothing$, $\{0\}$ and $\{0, 1\}$ satisfies the condition that there is an open set containing 0 but not 1; but there is no open set containing 1 and not 0. So it is a T_0–space but not a T_2–space. (This space is known as the **Sierpiński–space**.) Now consider requirement in the above definition. It is almost like the T_0–axiom but requires not just an open set but a clopen set. This however means that it is the same as a T_2–axiom with respect to the clopen sets. For if a contains x and not y and a is clopen, then $X - a$ is clopen as well, and it contains y but not x.

Definition 4.4.9. *Two categories $\mathcal{C}$ and $\mathcal{D}$ are* ***equivalent*** *if there exist covariant functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ such that for each $A \in Ob(\mathcal{C})$ we have $A \cong G(F(A))$ and for each $B \in Ob(\mathcal{D})$ we have $B \cong F(G(B))$.*

Theorem 4.4.10 (Stone). *The category* BA *of boolean algebras is equivalent to the category* $\mathsf{StoneSp}^{op}$*, the dual category of the category of the Stone–spaces.*

Proof. We know that we have a contravariant functor $(-)_o : \mathsf{BA} \to \mathsf{StoneSp}$ and a contravariant functor $(-)^o : \mathsf{StoneSp} \to \mathsf{BA}$. These can be made into covariant functors by switching to the opposite category $\mathsf{StoneSp}^{op}$. The remaining bit is to show that $\mathfrak{X} \cong (\mathfrak{X}^o)_o$. Now, for a point $x \in X$ put $U_x := \{a \in \mathbb{X} : a \text{ clopen}, x \in a\}$. This is an ultrafilter on the algebra of clopen sets. We show that the map $u : x \mapsto U_x$ is a topological isomorphism. It is injective by the definition of a Stone space; for if $x \neq y$ then there is a clopen set a such that $x \in a$ but $y \notin a$. The map is also surjective, by construction. Finally, a clopen set of $(\mathfrak{X}^o)_o$ is a set of the form $\widehat{a} := \{U : a \in U\}$, where U ranges over the ultrafilters of $\mathfrak{X}_o$ and a is clopen in $\mathfrak{X}$. Now let a be a clopen subset of $\mathfrak{X}$. Then

$$\begin{aligned} u[a] &= \{U_x : x \in a\} &&= \{U_x : a \in U_x\} \\ &= \{U : a \in U\} &&= \widehat{a} \end{aligned}$$

(Here we use the fact that every ultrafilter is of the form U_x for some x.) Hence, u induces a bijection of the clopen sets. Therefore, it induces a bijection between the open sets. Hence u is a topological isomorphism. □

Exercise 146. Show that the dual category $\mathcal{C}^{op}$ of a category $\mathcal{C}$ is a category. Show also that if $F : \mathcal{C} \to \mathcal{D}$ is a contravariant functor, then $F^{op} : \mathcal{C} \to \mathcal{D}^{op}$ is a covariant functor, where $F^{op}(A) := F(A)$ and $F^{op}(f) := F(f)^{op}$.

Exercise 147. Show Theorem 4.4.5.

Exercise 148. Show that a topological space is a Stone space iff it is compact and zero–dimensional.

Exercise 149. Let $J \subseteq \mathbb{R}$ be a subset of $\mathbb{R}$, and $\mathbb{J}$ the set of sets of the form $O \cap J$, O open in $\mathbb{R}$. Now let $\mathfrak{J}$ be the topological space $\langle J, \mathbb{J}\rangle$. Show that *points* in $\mathfrak{J}$ are nothing but *Dedekind cuts*.

Exercise 150. Show that intervals $I = [x, y] \subset \mathbb{R}$ endowed with the relative topology of $\mathbb{R}$ have no points. Show that $\mathbb{Q}$ is not a Stone–space.

Exercise 151. Let the real numbers between 0 and 1 be presented as 3–adic numbers, that is, as sequences $\langle a_i : 0 < i \in \omega \rangle$ such that $a \in \{0, 1, 2\}$. Such a sequence corresponds to the real number $\sum_{i=1} a_i \cdot 3^{-i}$. To make this correspondence one–to–one, we require that there is no i_0 such that $a_j = 2$ for all $j \geq i_0$. Now let $C \subset [0, 1]$ be the set of reals corresponding to sequences in which no a_i is equal to 1. This set is known as the *Cantor–Set*. Show that this set, endowed with the relative topology of the real line, is a Stone–space.

***Exercise 152.** Let X be countably infinite. Construct a compact T_0–space over X. Moreover, show that no T_2–space over X can be compact.

4.5. Adjoint Functors and Natural Transformations

In this section we will prove a representation theorem for frames that makes use of the topological representation developed by Marshall Stone.

Definition 4.5.1. *Let $\mathcal{C}$ and $\mathcal{D}$ be categories and $F, G : \mathcal{C} \to \mathcal{D}$ functors. Let $\eta : \mathsf{Ob}(\mathcal{C}) \to \mathsf{Mor}(\mathcal{D})$ be a map such that for all $\mathcal{C}$–objects A we have $G(A) = \eta(F(A))$ and that for all $\mathcal{C}$–arrows $f : A \to B$ we have $G(f) \circ \eta(A) = \eta(B) \circ F(f)$. Then η is called a* **natural transformation** *from F to G.*

The last condition can be presented in the form of a commutative diagram.

$$\begin{tikzcd} F(A) \arrow[r, "\eta(A)"] \arrow[d, "F(f)"'] & G(A) \arrow[d, "G(f)"] \\ F(B) \arrow[r, "\eta(B)"'] & G(B) \end{tikzcd}$$

It is important in category theory that everything is defined not only with respect to objects but also with respect to morphisms. The latter requirement is often called *naturalness*. Thus, a natural transformation is natural because it conforms with the arrows in the way indicated by the picture. Another example of this naturalness condition is in the definition of *adjoint functors* given below.

Let $\mathcal{C}$ and $\mathcal{D}$ be categories, and $F, G, H : \mathcal{C} \to \mathcal{D}$ be functors, η a natural transformation from F to G and θ a natural transformation from G to H. Define a map $\theta \bullet \eta : Ob(\mathcal{C}) \to Mor(\mathcal{D})$ by $(\theta \bullet \eta)(A) := \theta(A) \circ \eta(A)$. This is a natural transformation. For let $A \xrightarrow{f} B$. Then $\eta(B) \circ F(f) = G(f) \circ \eta(A)$ and $\theta(B) \circ G(f) = H(f) \circ \theta(A)$ by assumption that η and θ are natural transformations. Then

$$\begin{aligned} (\theta \bullet \eta)(B) \circ F(f) &= \theta(B) \circ \eta(B) \circ F(f) \\ &= \theta(B) \circ G(f) \circ \eta(A) \\ &= H(f) \circ \theta(A) \circ \eta(A) \\ &= H(f) \circ (\theta \bullet \eta)(A) \end{aligned}$$

Proposition 4.5.2. *Let $\mathcal{C}$ and $\mathcal{D}$ be categories. Then the functors from $\mathcal{C}$ to $\mathcal{D}$ form the objects of a category with arrows being the natural transformations.*

Definition 4.5.3. *Let $\mathcal{C}$ and $\mathcal{D}$ be categories, and $F : \mathcal{C} \to \mathcal{D}$ and let $G : \mathcal{D} \to \mathcal{C}$ be functors. F is called* **left adjoined** *to G, in symbols $F \dashv G$, if there exists for every $\mathcal{C}$–object A of $\mathcal{C}$ and every $\mathcal{D}$–object B a bijection $\beta_{AB} : hom_{\mathcal{D}}(F(A), B) \to hom_{\mathcal{C}}(A, G(B))$. Moreover, β_{AB} must be natural in both arguments; this means that for arrows $f : A \to A'$ and $g : B \to B'$ we have*

$$\begin{array}{rcl} \beta_{AB} \circ H^{GB}(f) & = & H^{B}(F(f)) \circ \beta_{A'B} \\ \beta_{AB'} \circ H_{A}(G(g)) & = & H_{FA}(g) \circ \beta_{AB} \end{array}$$

The definition of naturalness is best understood if presented in the form of a picture. Here is the picture corresponding to the first of these conditions.

$$\begin{tikzcd} hom_{\mathcal{D}}(A', GB) \arrow[r, "\beta_{A'B}"] \arrow[d, "H^{GB}(f)"'] & hom_{\mathcal{C}}(FA', B) \arrow[d, "H^{B}(F(f))"] \\ hom_{\mathcal{D}}(A, GB) \arrow[r, "\beta_{AB}"'] & hom_{\mathcal{C}}(FA, B) \end{tikzcd}$$

Assume now that F is left adjoined to G. Then there exists a bijection $\beta_{A,FA}$ from $hom(FA, FA)$ to $hom(A, GF(A))$. In particular, $\beta_{A,FA}(id(A)) : A \to GF(A)$. The map $\eta : A \mapsto \beta_{A,FA}(id(A))$ is a natural transformation from the identity functor on $\mathcal{C}$ to the functor GF. Similarly, starting with a bijection between $hom(GB, GB)$ and $hom(FG(B), B)$ we obtain a natural transformation from the functor FG to the identity functor on $\mathcal{D}$. These two transformations are called the **unit** ($\eta : 1_{\mathcal{C}} \to GF$) and the **counit** ($\theta : FG \to 1_{\mathcal{D}}$) of the adjunction. Moreover, the following so–called **triangular identities** hold for all objects A of $\mathcal{C}$ and B of $\mathcal{D}$:

$$\begin{array}{rcl} \theta(FA) \circ F(\eta(A)) & = & id(FA) \\ G(\theta(B)) \circ \eta(GB) & = & id(GB) \end{array}$$

It can be shown that the existence of a natural transformation $\eta : 1_{\mathcal{C}} \to GF$ (the unit) and a natural transformation $\theta : FG \to 1_{\mathcal{D}}$ (the counit) is enough to ensure that two functors are adjoint. Namely, consider the following schemata.

$$\begin{array}{rc} g & A \to G(B) \\ \cline{2-2} F(g) & F(A) \to FG(B) \\ \cline{2-2} \theta(B) \circ F(g) & F(A) \to B \end{array} \qquad\qquad \begin{array}{rc} f & F(A) \to B \\ \cline{2-2} G(f) & GF(A) \to G(B) \\ \cline{2-2} G(f) \circ \eta(A) & A \to G(B) \end{array}$$

So, $\beta_{AB} : g \mapsto F(g) \circ \eta(A)$ is bijective. It is not hard to show that if the triangular identities hold then these bijections are natural in both arguments. The existence of a unit and a counit satisfying the triangular identities is typically somewhat easier to verify.

To understand this terminology, let us look at a particular case. Let $\mathcal{V}$ be a variety of Ω–algebras for a given signature Ω. They form a category, which we also denote by $\mathcal{V}$. $F : \mathcal{V} \to \mathsf{Set} : \mathfrak{A} \mapsto A$ the functor sending an algebra to its underlying set, and a homomorphism between algebras to the corresponding set–map. F is called the **forgetful functor**. Consider the functor $G : \mathsf{Set} \to \mathcal{V} : X \mapsto \mathfrak{Fr}_{\mathcal{V}}(X)$, sending a set X to the algebra freely generated by X in $\mathcal{V}$. We call it the **free functor**. We claim that $G \dashv F$. For let $f : X \to F(\mathfrak{A})$ be a map. Then, by definition of a free algebra there exists an extension $\overline{f} : \mathfrak{Fr}_{\mathcal{V}}(X) \to \mathfrak{A}$; the correspondence between f and $\overline{f}$ is bijective. It is a matter of straightforward but tedious calculations to verify that this bijection is natural. We conclude the following theorem.

Theorem 4.5.4. *Let $\mathcal{V}$ be a variety of Ω–algebras. Then the forgetful functor has a right adjoint, the free functor.*

The next example appears also in many guises throughout this book. A poset $\langle P, \leq\rangle$ can be regarded as a category as follows. Put $Ob(\mathcal{C}) := P$, $Mor(\mathcal{C}) := \leq$, $id(x) := \langle x, x\rangle$, and $\langle x, y\rangle \circ \langle y, z\rangle := \langle x, z\rangle$. The domain of $\langle x, y\rangle$ is x, and the codomain is y. We call a category a **poset category** if for every pair A, B of objects there exists at most one arrow from A to B. In a poset category it is unnecessary to name arrows. So we simply write $A \to B$ to denote the unique arrow from A to B; and we write $A \leq B$ to state that there exists an arrow from A to B. Let now $\mathcal{C}$ and $\mathcal{D}$ be poset categories. A functor $F : \mathcal{C} \to \mathcal{D}$ is uniquely determined by its action on the objects. For if $f : A \to B$ then $F(f) : F(A) \to F(B)$ is an arrow and so is uniquely determined. The requirement that F is a functor is equivalent to the condition that F be **isotonic**, that is, if $A \leq B$ then $F(A) \leq F(B)$. Suppose now that $\mathfrak{P}$ and $\mathfrak{Q}$ are posets and $f : P \to Q$, $g : Q \to P$ be isotonic maps. We may think of the maps as functors between the corresponding poset categories. We claim that f is left–adjoint to g iff the following holds for all $x \in P$ and all $y \in Q$:

$$\frac{x \leq g(y)}{\overline{\underline{\;f(x) \leq y\;}}}$$

(This is read as follows. The situation above the line obtains iff the situation below the line obtains.) The proof is straightforward. There exists a bijection between the hom–sets, and this bijection is uniquely defined by the fact that the hom–sets contain only one member. The naturalness is also immediately verified. Now let $x \in P$, $y \in Q$. Then from $g(y) \leq g(y)$ we get $fg(y) \leq y$, and from $f(x) \leq f(x)$ we get $x \leq gf(x)$. Assume now that $f : \mathfrak{P} \to \mathfrak{Q}$ and $g : \mathfrak{Q} \to \mathfrak{P}$ such that $fg(y) \leq y$ for all $y \in Q$ and $x \leq gf(x)$ for all $x \in P$. Then from $y \leq f(x)$ we deduce $g(y) \leq gf(x)$. Since $gf(x) \leq x$ we have $g(y) \leq x$. If $g(y) \leq x$ then $fg(y) \leq f(x)$. Together with $y \leq fg(y)$ we get $y \leq f(x)$.

We will now extend the results of duality theory to modal algebras. This can be done by pushing the topological duality further, as outlined in Giovanni Sambin and Virginia Vaccaro [**186**]. We will sketch this approach, proving only part of the results. Some technical details have to be adapted when lifting this approach to polymodal algebras. This is the reason why we do not simplify the exposition to monomodal algebras; otherwise it looks overly complicated. The key is to regard polymodal algebras as functors from a special diagram into the category of boolean algebras with hemimorphisms as functions. Recall that a hemimorphism from a boolean algebra $\mathfrak{A}$ to a boolean algebra $\mathfrak{B}$ is a map $\tau : A \to B$ such that $\tau(1) = 1$ and $\tau(a \cap b) = \tau(a) \cap \tau(b)$ for all $a, b \in A$. We write $\tau : \mathfrak{A} \rightharpoonup \mathfrak{B}$ for the fact that τ is a hemimorphism. A **co–hemimorphism** is a map $\sigma : A \to B$ such that $\sigma(0) = 0$ and $\sigma(a \cup b) = \sigma(a) \cup \sigma(b)$ for all $a, b \in A$. If τ is a hemimorphism, $\sigma(a) := -\tau - (a)$ is a co–hemimorphism, and if σ is a co–hemimorphism then $\tau(a) := -\tau - (a)$ is a hemimorphism. The category of boolean algebras as objects and hemimorphisms as arrows is denoted by Bal. The dual of a boolean algebra is a Stone–space; the dual of a hemimorphism turns out to be a relation between the points of the spaces. Namely, if $\tau : \mathfrak{A} \rightharpoonup \mathfrak{B}$ and $f : \mathfrak{A} \to \mathbf{2}$, $g : \mathfrak{B} \to \mathbf{2}$ are points, then put $g \lhd f$ if for all $a \in A$, $g(\tau(a)) = 1$ implies $f(a) = 1$.

Consider a relation $\lhd \subseteq X \times Y$. Given a set $S \subseteq X$ and a set $T \subseteq Y$, write

$$\begin{array}{lll}
\Diamond T & := & \{x \in X : (\exists y)(x \lhd y \text{ and } y \in T)\} \\
\Box T & := & \{x \in X : (\forall y)(\text{if } x \lhd y \text{ then } y \in T)\} \\
\Diamond S & := & \{y \in Y : (\exists x)(x \lhd y \text{ and } x \in S)\} \\
\Box S & := & \{y \in Y : (\forall x)(\text{if } x \lhd y \text{ then } x \in S)\}
\end{array}$$

Then $\Diamond, \Box : \wp(X) \to \wp(Y)$ and $\Diamond, \Box : \wp(Y) \to \wp(X)$. Moreover, the following laws of adjunction hold.

$$\frac{S \subseteq \Box T}{\overline{\Diamond S \subseteq T}} \qquad \frac{\Box S \supseteq T}{\overline{S \supseteq \Diamond T}}$$

These laws are reflected in the postulates $p \to \Box\Diamond p$ and $p \to \Box\Diamond p$ of tense logic. (In fact, the latter encode that $\Box\Diamond$ and $\Box\Diamond$ are the unit and counit of this adjunction.) Let $\mathfrak{X}$ and $\mathfrak{Y}$ be topological spaces and $f : \mathbb{X} \to \mathbb{Y}$ be a function. f is called **open** if for every open set $S \subseteq X$, $f[S]$ is open in f. Likewise, f is called **closed** (**clopen**) if the direct image of a closed (clopen) set is closed (clopen). In general a map that is both open and closed is also clopen. The converse is generally false.

Definition 4.5.5. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be topological spaces, and $\lhd \subseteq X \times Y$ be a relation. $\lhd$ is called a* ***continuous relation*** *from $\mathfrak{X}$ to $\mathfrak{Y}$ if $\Diamond$ is a clopen map from $\mathbb{Y}$ to $\mathbb{X}$. The category of topological spaces and continuous relations as maps is denoted by* Spa. *$\lhd$ is called* ***closed*** *if $\Diamond$ is a closed map.*

A relation is continuous iff $\Box$ is clopen. The reader is warned that $\Box H$ is not the inverse image of H under $\lhd$. The latter is $\Diamond H$. The two coincide just in case $\lhd$ is a surjective function. Let $\mathfrak{X}^\diamond$ denote the boolean algebra of clopen subsets of $\mathfrak{X}$.

$$\mathfrak{X}^\diamond := \langle \{H : H \text{ clopen in } \mathfrak{X}\}, 1, -, \cap \rangle$$

Let $\lhd \subseteq X \times Y$ be a continuous relation. Then $\Diamond$ commutes with arbitrary unions. So $\Diamond : \mathfrak{Y}^\diamond \to \mathfrak{X}^\diamond$ is a co–hemimorphism and $\Box : \mathfrak{Y}^\diamond \to \mathfrak{X}^\diamond$ is a hemimorphism. Now let $\blacktriangleleft \subseteq Y \times Z$ be a continuous relation from $\mathfrak{Y}$ to $\mathfrak{Z}$. Then $\lhd \circ \blacktriangleleft \subseteq X \times Z$ is a continuous relation from X to Z.

THEOREM 4.5.6. $(-)^\diamond$ *is a contravariant functor from* Spa *to* Bal.

Now return to the case of a hemimorphism $\tau : \mathfrak{A} \rightharpoonup \mathfrak{B}$. We define $\mathfrak{A}^\diamond$ to be simply the space of points endowed with the topology generated by the sets $\widehat{a} = \{U \in pt(\mathfrak{A}) : a \in U\}$. Its sets can be characterized as follows.

PROPOSITION 4.5.7. *Let $\mathfrak{X}$ be the Stone space of $\mathfrak{A}$. A set S is closed in $\mathfrak{X}$ iff it is of the form $\{U : U \supseteq F\}$, where F is a filter in $\mathfrak{A}$. A set is clopen iff it is of the form $\{U : U \supseteq F\}$, where F is a principal filter.*

PROOF. We prove the second claim first. Assume S is clopen. Then $S = \widehat{a} = \{U : a \in U\}$ for some $a \in A$. Now put $F = \{b : b \geq a\}$. Then $S = \{U : U \supseteq F\}$. F is principal. Conversely, if F is principal, say $F = \{b : b \geq a\}$ for some a, then $S = \widehat{a}$, hence S is clopen. Now assume that S is closed. Then $S = \bigcap_{i \in I} R_i$ for some family of clopen sets R_i. Let $R_i = \{U : U \supseteq F_i\}$, $i \in I$, where each F_i is a principal filter. Let G be the filter generated by the F_i. We have $S = \{U : U \supseteq G\}$. Conversely, assume that there exists a filter G such that $S = \{U : U \supseteq G\}$. Then G is generated by a family $\langle F_i : i \in I\rangle$ of principal filters. It follows that $S = \bigcap_{i \in I} R_i$, where $R_i = \{U : U \supseteq F_i\}$. Each R_i is clopen, and so S is closed. ◻

Now for $U \in pt(\mathfrak{B})$ and $V \in pt(\mathfrak{A})$ we put $U \,\tau_\diamond\, V$ iff for every $a \in U$, $\tau a \in U$ implies $a \in V$. The latter is the same as: $a \in V$ implies $\sigma(a) \in U$, where $\sigma(a) := -\tau - a \in U$. With $\lhd$ understood to be $\tau_\diamond$, $\diamondsuit$, $\diamondsuit$ etc. are properly defined. Therefore, by definition

$$\begin{array}{rcl} U\tau_\diamond V & \Leftrightarrow & V \subseteq \tau[U] \\ U\tau_\diamond V & \Leftrightarrow & U \subseteq \sigma[V] \end{array}$$

DEFINITION 4.5.8. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be topological spaces. A relation $\lhd \subseteq X \times Y$ is called* ***point closed****, if for every $x \in X$, $\diamondsuit\{x\}$ is closed in $\mathfrak{Y}$.*

PROPOSITION 4.5.9. *Suppose $\tau : \mathfrak{A} \rightharpoonup \mathfrak{B}$. Then $\tau_\diamond$ is a continuous, point closed relation from $\mathfrak{B}_\diamond$ to $\mathfrak{A}_\diamond$.*

PROOF. It is not hard to see that $\tau_\diamond$ is point closed. For let $U \in pt(\mathfrak{B})$. Then $\tau^{-1}[U]$ is a filter of $\mathfrak{A}$; hence $H := \{T \in pt(\mathfrak{A}) : T \supseteq \tau^{-1}[U]\}$ is closed. Moreover, $H = \{T : T\tau_\diamond U\}$. Now for the first claim. Let C be a clopen set, $C = \widehat{a}$. Then $S \in \diamondsuit C$ iff $S \,\tau_\diamond\, T$ for some $T \in \widehat{a}$ iff $S \,\tau_\diamond\, T$ for some $T \ni a$ iff $-\tau - a \in S$. The latter is a clopen set. Hence, $\diamondsuit$ maps clopen sets onto clopen sets. ◻

THEOREM 4.5.10. $(-)_\diamond$ *is a contravariant functor from* Bal *into* Spa.

PROOF. Let $\tau : \mathfrak{A} \rightharpoonup \mathfrak{B}$ and $\upsilon : \mathfrak{B} \rightharpoonup \mathfrak{C}$. We have to show that

$$(\upsilon \circ \tau)_\diamond = \tau_\diamond \circ \upsilon_\diamond \ .$$

Claim. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be Stone spaces. If $\lhd$ is a continuous and point closed relation from $\mathfrak{X}$ to $\mathfrak{Y}$ then $\diamondsuit$ is closed.* For a proof let D be a closed set of $\mathfrak{X}$. We show that if $y \notin \diamondsuit D$ then there exists a clopen $C \subseteq Y$ such that $\diamondsuit D \subseteq C$ but $y \notin C$. This suffices for a proof. So let $y \notin \diamondsuit D$. For every $x \in D$, $y \notin \diamondsuit\{x\}$, and since $\diamondsuit\{x\}$ is closed there exists a clopen set C_x such that $y \notin C_x$ but $\Diamond\{x\} \subseteq C_x$. (This follows from the fact that each closed set of $\mathfrak{Y}$ is the intersection of clopen sets.) Hence, $x \in \Box C_x$. Therefore $D \subseteq \bigcup_{x \in D} \Box C_x$. This is an open cover of D, and by compactness of $\mathfrak{X}$ there exists a finite set S such that $D \subseteq \bigcup_{x \in S} \Box C_x$. Then also $\Diamond D \subseteq \bigcup_{x \in S} C_x$. Put $C := \bigcup_{x \in S} C_x$. C is clopen and does not contain y. This proves the claim.

From this fact we deduce that $\tau_\Diamond \circ \upsilon_\Diamond$ is point–closed. For, being the composition of closed maps, it is closed, and a fortiori point–closed. We finally need to verify that this map is based on the same continuous relation as $(\upsilon \circ \tau)_\Diamond$. The proof is essentially the same as in Theorem 3.2.8; if $U(\tau \circ \upsilon)_\Diamond V$ then one has to find a point Z such that $U\, \tau_\Diamond\, Z$ and $Z\, \upsilon_\Diamond\, V$. Alternatively, two maps between Stone spaces are identical iff they are identical on the clopen sets. We leave the verification of that fact to the reader. ☐

Finally, consider having κ–many modal operators. Let $\mathfrak{J}(\kappa)$ be the category consisting of a single object, denoted here by $\oslash$. The set of morphisms consists in finite sequences from κ. If σ and τ are such sequences, then $\sigma(\tau(\oslash)) = (\sigma^\frown\tau)(\oslash)$. The identity on $\oslash$ is the map ϵ, where ϵ is the empty sequence. (Moreover, $dom(f) = cod(f) = \oslash$ for every arrow f.) This defines the category $\mathfrak{J}(\kappa)$. Consider a functor $F : \mathfrak{J}(\kappa) \to \mathsf{Bal}$. Then $F(\oslash) = \mathfrak{A}$ for some boolean algebra $\mathfrak{A}$ and for each $j < \kappa$, $F(j) : \mathfrak{A} \rightharpoonup \mathfrak{A}$. For each sequence σ, $F(\sigma) : \mathfrak{A} \rightharpoonup \mathfrak{A}$ is uniquely determined by the $F(j)$, $j < \kappa$. Hence, the functor can be viewed as a κ–modal algebra. Next, consider another functor $G : \mathfrak{J}(\kappa) \to \mathsf{Bal}$. Suppose that η is a natural transformation from F to G. This means that $\eta(\oslash) : \mathfrak{A} \rightharpoonup \mathfrak{B}$ and that

$$\eta(\oslash) \circ F(j) = G(j) \circ \eta(\oslash)$$

A natural transformation η is **boolean** if $\eta(\oslash)$ is a boolean homomorphism. In that case, the conditions on η being a natural transformation are equivalent to $\eta(\oslash)$ being a homomorphism.

Theorem 4.5.11 (Sambin & Vaccaro). *The category of κ–modal algebras is equivalent to the category of functors from $\mathfrak{J}(\kappa)$ to* Bal*, with arrows being the boolean natural transformations. This category is denoted by* Mal_κ.

In a similar way we can introduce functors from $\mathfrak{J}(\kappa)$ to Spa. They correspond to κ–modal frames. In other words, a κ–modal frame is a topological space $\mathfrak{X}$ endowed with a family of continuous relations from $\mathfrak{X}$ to $\mathfrak{X}$ indexed by κ. For the moment, however, frames are functors. Let F and G be two such functors. A **weak contraction** from F to G is a continuous relation c from $\mathfrak{X} := F(\oslash)$ to $\mathfrak{Y} := G(\oslash)$ such that (1.) For every clopen set H of $\mathfrak{Y}$, $c^{-1}[H]$ is clopen in $\mathfrak{X}$, (2.) For every clopen set H of $\mathfrak{Y}$, $\Diamond c^{-1}[H] = c^{-1}[\Diamond H]$. A **contraction** is a function (!) and a weak contraction that satisfies (2.) for all sets $\{x\}$ (and so for all subsets of Y). Hence a contraction is a weak contraction. It is tempting to conclude that the category of frames is the same as the category of functors from $\mathfrak{J}(\kappa)$ into the category of topological spaces together with weak contractions. This is not so, however. Rather, this result can be true only if we take zero–dimensional topological spaces.

THEOREM 4.5.12 (Sambin & Vaccaro). *The functors from $\mathfrak{J}(\kappa)$ to the category of zero–dimensional compact spaces with contractions as arrows form a category denoted by* Fra_κ. Fra_κ *is equivalent to the category of κ–modal frames and p–morphisms.*

The proof of this theorem is left as an exercise. The reader should convince himself that this theorem is essentially a reformulation of the notion of a general frame and a p–morphism between such frames into category theory. Now consider the category of Stone spaces, denoted by StSpa_κ, and the category of functors from $\mathfrak{J}(\kappa)$ with point closed contractions as arrows.

THEOREM 4.5.13. StSpa_κ *is dual to the category* Mal_κ.

Exercise 153. Show Theorem 4.5.12.

Exercise 154. Show that if $f : X \to Y$ is a function, the map $f : \wp(X) \to \wp(Y) : A \mapsto f[A]$ has both a left adjoint and a right adjoint (if viewed as a map between posets).

4.6. Generalized Frames and Modal Duality Theory

We have seen that a boolean algebra can be realized as a set algebra i. e. as a subalgebra of a powerset–algebra. Also, we have seen that boolean algebras can be repesented by certain topological spaces. So, we can either choose a topological representation or a representation of boolean algebras by so–called *boolean spaces* (which are pairs $\langle X, \mathbb{X}\rangle$ where $\mathbb{X}$ is closed under complement and union). In a subsequent section we have developed the topological representation of modal algebras. In a second step we restrict the topological space to the clopen sets. In this way we get the standard representation of modal algebras as certain general frames.

DEFINITION 4.6.1. *Let $\mathfrak{A}$ be a κ–modal algebra. Then the generalized frame $\mathfrak{A}^+$ is defined as follows. The* **worlds** *are the* **ultrafilters** *of $\mathfrak{A}$, and $U \lhd_j V$ iff for all $b \in V$ we have $\blacklozenge_j b \in U$. Furthermore, the* **internal sets** *are the sets of the form $\widehat{b} = \{U : b \in U\}$. For a homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ we let $h^+ : \mathfrak{B}^+ \to \mathfrak{A}^+$ be defined by $h^+(U) := h^{-1}[U] = \{b : h(b) \in U\}$.*

The converse direction is harmless as well.

DEFINITION 4.6.2. *Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ be a generalized frame. Then the modal algebra $\mathfrak{F}_+$ is defined as follows. The* **elements** *are the* **internal sets**, *and the* **operations** *are intersection, complement, and $\blacklozenge_j$, $j < \kappa$, defined by*

$$\blacklozenge_j b := \{w : (\exists x)(w \lhd_j x \text{ and } x \in b)\}.$$

If $p : \mathfrak{F} \to \mathfrak{G}$ is a p–morphism, then $p_+ : \mathfrak{G}_+ \to \mathfrak{F}_+$ is defined by $p_+(b) := p^{-1}[b]$.

Theorem 4.6.3. $(-)^+$ *is a contravariant functor from the category* Mal *of modal algebras to the category* Frm *of general frames;* $(-)_+$ *is a contravariant functor from* Frm *to* Mal. *Moreover, for every modal algebra* ${\mathfrak{A}^+}_+ \cong \mathfrak{A}$.

We entrust the proof of the fact that $(-)^+$ and $(-)_+$ are functors onto the reader. Again we are faced with the problem of the converse, namely, to say when for a general frame ${\mathfrak{F}_+}^+ \cong \mathfrak{F}$. In order to state this condition, let us give another definition.

Definition 4.6.4. *A frame is* ***differentiated*** *if for every pair x and y of different worlds there is an internal set containing x but not y. A frame is* ***tight*** *if for every pair x and y of worlds such that* $x \ntriangleleft_j y$ *there is an internal set b such that* $x \in \blacksquare_j b$ *but* $y \notin b$. *A frame is* ***compact*** *if every family of internal sets with the finite intersection property has a nonempty intersection. A frame is* ***refined*** *if it is both differentiated and tight, and it is* ***descriptive*** *if it is refined and compact. A frame is* ***atomic*** *if for every world x the set* $\{x\}$ *is internal and* ***full****, if every subset is internal.*

We introduce also abbreviations for classes. $\mathfrak{Krp}$ denotes the class of Kripke–frames, $\mathfrak{G}$ the class of generalized frames, $\mathfrak{Df}$ the class of differentiated frames, $\mathfrak{Ti}$ the class of tight frames, $\mathfrak{R}$ the class of refined frames, $\mathfrak{Cmp}$ the class of compact frames, $\mathfrak{D}$ the class of descriptive frames and $\mathfrak{C}$ the class of canonical frames. Atomic frames will play only a marginal role, although atomicity is a rather desirable property of frames, playing a fundamental role in the completeness proofs for fusions (see Chapter 6). The definition of compactness is equivalent to the definition of compactness of the space generated by the sets of the algebra. Moreover, the postulate of differentiatedness is the separation postulate for Stone–spaces. We may say informally, that a compact frame has *enough worlds*, a differentiated frame has *enough internal sets for identity* and a refined frame has *enough sets for the basic relations*.

The properties of frames have a topological counterpart.

Proposition 4.6.5. *Let* $\mathfrak{F}$ *be a* κ*–modal frame. (i.)* $\mathfrak{F}$ *is differentiated iff* $\{x\}$ *is closed for all* $x \in f$ *iff the corresponding topological space is Hausdorff. (ii.)* $\mathfrak{F}$ *is refined iff the space is Hausdorff and* $\triangleleft_j$ *is point closed for all* $j < \kappa$. *(iii.)* $\mathfrak{F}$ *is compact iff the corresponding topological space is compact.*

We remark here that the property of tightness also is a topological separation property. It says, namely, that for every x and $j < \kappa$ the set $suc_j(x) = \{y : x \triangleleft_j y\}$ can be separated by a clopen neighbourhood from any point not contained in it.

Theorem 4.6.6. *For a descriptive frame,* $\mathfrak{F} \cong {\mathfrak{F}_+}^+$.

Proof. Consider the map $x \mapsto U_x := \{b : x \in b\}$. This map is injective, since the frame is differentiated. It is surjective, since the frame is compact. Now assume that $x \triangleleft_j y$. Then if $b \in U_y$, $\blacklozenge_j b \in U_x$. Hence $U_x \triangleleft_j U_y$, by definition of the relation $\triangleleft_j$. Now assume $x \ntriangleleft_j y$. Then, since the frame is tight, there is a $b \in U_y$ such that $\blacksquare - b \in U_x$, that is, $\blacklozenge b \notin U_x$. Hence $U_x \ntriangleleft_j U_y$. ◻

Corollary 4.6.7 (Duality Theorem). *The categories* Mal_κ *of (κ–)modal algebras and* DFrm_κ *of descriptive κ–frames are dually equivalent. Moreover, a homomorphism of modal algebras is surjective iff its dual map is an embedding; it is injective iff its dual map is a contraction.*

The duality theorem distinguishes the descriptive frames from the other frames and will be used quite effectively. From a theoretical point this is a very satisfactory result, allowing us to transfer algebraic proofs into geometric proofs and conversely. However, as it turns out, descriptive frames are hard to construct. We do not have such a good intuition about them. Typically, it is much easier to construct refined frames than to construct descriptive frames. So, we will quite often work with refined frames instead. As a last point notice that there exists a modal algebra based on a single element. This algebra is denoted by **1**. Applying the representation theory we obtain that $\mathbf{1}^+$ is the empty frame. It is this fact to allow frames to have no worlds at all.

In addition to (general) frames we also have Kripke–frames and there is a rather easy way to turn a frame into a Kripke–frame, namely by just forgetting the internal sets. So, given a frame $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ we put $\mathfrak{F}_\sharp := \mathfrak{f}$, and given a p–morphism $\pi : \mathfrak{F} \to \mathfrak{G}$ we put $\pi_\sharp := \pi$. Then $\pi_\sharp : \mathfrak{F}_\sharp \to \mathfrak{G}_\sharp$. This is a functor, as is easily checked. This is a kind of forgetful functor. Conversely, given a Kripke frame $\mathfrak{f}$, let $\mathfrak{f}^\sharp := \langle \mathfrak{f}, 2^f\rangle$. In analogy, we dub this the **recovery functor**. Actually, the same terminology can be applied to the previous case. To pass from a general frame to a modal algebra is practically a forgetful operation, and to get a frame from the algebra is a recovery process. There is a complete analogy here, because forgetting what has been reconstructed results in the same structure; we have both $\mathfrak{A}^+{}_+ \cong \mathfrak{A}$ and $\mathfrak{f}^\sharp{}_\sharp \cong \mathfrak{f}$. But the converse need not hold; we are not sure to be able to reconstruct what we have forgotten.

Theorem 4.6.8. *The map* $(-)_\sharp$ *is a covariant functor from the category* Frm *of frames into the category* Krp *of Kripke–frames. The map* $(-)^\sharp$ *is a covariant functor from* Krp *into* Frm*. Moreover, for a Kripke–frame* $\mathfrak{f}^\sharp{}_\sharp \cong \mathfrak{f}$.

Evidently, $\mathfrak{F}$ is full iff $\mathfrak{F} = \mathfrak{F}_\sharp{}^\sharp$. So, the category of full frames is equivalent to the category of Kripke–frames. (This is not as difficult as it sounds.) A nice consequence is the following construction, originally due to Bjarni Jónsson and Alfred Tarski. Take a κ–modal algebra $\mathfrak{A}$. It is called **complete** if the lattice reduct is complete. The **completion** of $\mathfrak{A}$, $\mathfrak{Em}(\mathfrak{A})$, is a modal algebra together with an embedding $c : \mathfrak{A} \rightarrowtail \mathfrak{Em}(\mathfrak{A})$, which is complete and satisfies that for every complete $\mathfrak{B}$ and homomorphism $h : \mathfrak{A} \to \mathfrak{B}$ there exists a $k : \mathfrak{Em}(\mathfrak{A}) \to \mathfrak{B}$ with $k \circ c = h$. It is easy to see that a complete modal algebra is isomorphic to an algebra of the form $\mathfrak{Ma}(\mathfrak{f})$ for some Kripke–frame $\mathfrak{f}$ (where $\mathfrak{Ma}(\mathfrak{f})$ as defined earlier can now be redefined as $\mathfrak{f}^\sharp{}_+$). Simply take as elements of $\mathfrak{f}$ the atoms of $\mathfrak{A}$, and put $b \lhd_j c$ for atoms b and c, iff $b \leq \blacklozenge_j c$. Now if $\mathfrak{A}$ is a modal algebra, then let $\mathfrak{Em}(\mathfrak{A}) := (\mathfrak{A}^+{}_\sharp)^\sharp{}_+$. This means the following. We pass from $\mathfrak{A}$ to the dual general frame. Next we pass to the corresponding full frame by taking all sets of the frame as internal sets (this is the actual completion). Finally, we take the algebra of sets of this full frame. The identity on $pt(\mathfrak{A})$ is a p–morphism $i : (\mathfrak{A}^+)_\sharp{}^\sharp \rightarrowtail \mathfrak{A}^+$. It is surjective, and so $i^+ : \mathfrak{A} \rightarrowtail \mathfrak{Em}(\mathfrak{A})$. We show that $c := i^+$ has the required property. To that end, let $h : \mathfrak{A} \to \mathfrak{B}$ be a homomorphism and $\mathfrak{B}$ complete. Then $h^+ : \mathfrak{B}^+ \to \mathfrak{A}^+$ is a p–morphism of frames. Since $\mathfrak{B}^+$ is complete, we can actually factor h^+ through i and a function $j : \mathfrak{B}^+ \to (\mathfrak{A}^+)_\sharp{}^\sharp$, by extending h^+ to all subsets of the frame. We have $h^+ = i \circ j$. Now switch back to the modal algebras. By duality, this gives $h = j^+ \circ i^+ = j^+ \circ c$. Put $k := j^+$. This concludes the proof.

Theorem 4.6.9. *For every modal algebra, the natural embedding map* $\mathfrak{A} \rightarrowtail \mathfrak{Em}(\mathfrak{A}) := (\mathfrak{A}^+{}_\sharp)^\sharp{}_+$ *is a completion.*

Now, what are the relationships between all these classes? First of all, a full frame is both differentiated and refined. A full frame is compact iff it is finite. For consider the family of sets $f - \{x\}$ for $x \in f$. If f is infinite, this family has the finite intersection property. Yet the intersection of all these sets is empty. A differentiated compact frame is also tight and hence descriptive. (See exercises; a proof using quite different methods will be given in the next chapter.) There are tight, compact frames which are not differentiated. For example, take a descriptive frame $\mathfrak{F}$. Form a new frame $\mathfrak{F}^\delta$ from $\mathfrak{F}$ by taking twins w^1 and w^2 for each $w \in f$. Put $x^i \lhd_j y^k$ iff $i = k$ and $x \lhd_j y$; finally, the internal sets are the sets of the form $a^1 \cup a^2 = \{x^1 : x \in a\} \cup \{x^2 : x \in a\}$. This is a frame; it is tight, and compact. But it is not differentiated.

Now for the difference between tightness and differentiatedness. We have seen already that there are tight frames which are not differentiated. For the converse we have to work harder. A finite frame will not do here. Define a general frame $\mathfrak{R} := \langle \mathfrak{r}, \mathbb{R} \rangle$, where $\mathfrak{r} := \langle \omega, \leq \rangle$ and $\mathbb{R}$ is the set of all finite unions of sets of the form

$$r(i, j) = \{k : k \equiv j \pmod{i}\}$$

where $0 \leq j < i$. Since $-r(i, j) = \bigcup_{j' \neq j} r(i, j')$, $\mathbb{R}$ is closed under complements. $\mathbb{R}$ closed under intersection, too, by the Chinese Remainder Theorem. Finally, $\blacklozenge a = \omega$ iff $a \neq \emptyset$, and $\blacklozenge \emptyset = \emptyset$, so $\mathfrak{R}$ is a frame. It is differentiated. For let $i \neq j$, say $j < i$. Then $i \in r(i + 1, -1)$ but $j \notin r(i + 1, -1)$. Now $\mathfrak{R}$ is not tight. For $i \ntriangleleft j$ iff $j < i$. But there is no set b such that $i \in \blacksquare b$ but $j \notin b$. For either $b \neq 1$ and then $i \notin \blacksquare b (= \emptyset)$, or $b = \omega$ and then $j \in b$.

This is actually an instructive frame and we will prove something more than necessary right now.

Theorem 4.6.10. $\mathsf{Th}(\mathfrak{R}) = \boldsymbol{S5}$. *Moreover, every finite Kripke–frame for* $\boldsymbol{S5}$ *is a p–morphic image of* $\mathfrak{R}$.

Proof. We prove the second claim first. Consider an **S5**–frame with i elements. Then the map $p : j \mapsto j \ (mod\ i)$ is a p–morphism. We take as $\lhd$ the direct image of $\leq$ under p. Now if $k < i$ then for some $r \in \omega$ we have $j \leq r \cdot i + k$, so that $p(j) \lhd p(k)$. Thus

$\lhd = i \times i$, and so we have a p–morphism of Kripke–frames. Next, take any subset $J \subseteq \{0, \dots, i-1\}$. Then $p^{-1}[J] = \bigcup_{j \in J} r(i, j)$, which is internal. Finally, by the fact that **S5** has the finite model property we have $\mathsf{Th}(\mathfrak{R}) \subseteq$ **S5**. However, $\mathfrak{R} \vDash p \to \Box\Diamond p$. For if $\beta(p) = \varnothing$ then $\beta(p) \subseteq \overline{\beta}(\Box\Diamond p)$. And if $\beta(p) \neq \varnothing$ then $\overline{\beta}(\Box\Diamond p) = \blacksquare 1 = 1$. ◻

It seems at first sight that a frame which is not differentiated can be made into a differentiated frame by taking the map defined by $x \mapsto U_x$. We call this the **refinement map**. It maps two points x and x' onto the same target point if $U_x = U_{x'}$. If $U_x = U_{x'}$ we also write $x \sim x'$. Unfortunately, this generally is not a p–morphism. We need an extra condition to ensure this. One possibility is to require tightness. For then, if $p(x) \lhd p(y)$, that is, $U_x \lhd U_y$ then for every $x' \sim x$ there is a y' such that $x' \lhd y'$ and $y' \sim y$. In the case of tightness we can even show something stronger, namely that $x' \lhd y$. Namely, let $y \in b$, that is, $b \in U_y$. Then $\blacklozenge b \in U_x = U_{x'}$, that is, $x' \in \blacklozenge b$. Hence, by tightness, $x' \lhd y$. However, tightness is stronger than necessary.

Proposition 4.6.11. *If $\mathfrak{F}$ is tight, the refinement map $p : x \mapsto U_x$ is a p–morphism. Moreover, if $x \lhd y$, $x \sim x'$ and $y \sim y'$ then $x' \lhd y'$.*

We will now turn to some important questions concerning the relationship between geometrical properties of a frame $\mathfrak{F}$ and properties of the algebra of sets. In particular, we will be concerned with the question of whether $\mathfrak{F}$ *is rooted* corresponds to $\mathfrak{F}_+$ *is subdirectly irreducible*. The material presented here is based on [**185**]. We will give two examples, showing that in fact neither implication holds.

Example. Consider the frame $\mathfrak{F} = \langle \mathbb{Z}, \prec, \mathbb{F} \rangle$ where $x \prec y$ iff $|x-y| = 1$ and $\mathbb{F}$ is the set of finite and cofinite subsets of $\mathbb{Z}$. This is well–defined. The frame is connected and the only open filters are the trivial filters. (Let F be an open filter. If $F \neq \mathbb{F}$, then F does not contain a finite set. For if a is finite, there is an $n \in \omega$ such that $\blacksquare^n a = \varnothing$. Now, assume that $F \neq \{1\}$. Then for some z, $Y := \mathbb{Z} - \{z\} \in F$. Then if $|z' - z| \leq n$, $z' \notin \blacksquare^n Y \in F$, whence $\mathbb{Z} - \{z'\} \in F$ for all z'. Any cofinite set is an intersection of such sets. Hence, F is the filter containing all cofinite sets.) Thus, $Con(\mathfrak{F}_+) \cong \mathbf{3}$. This shows that the algebra of $\mathfrak{F}$ is subdirectly irreducible. Now consider the bidual $\mathfrak{F}_+{}^+$. A point of $\mathfrak{F}_+{}^+$ is an ultrafilter of $\mathfrak{F}_+$. If U is not principal then it contains only infinite sets, hence only cofinite sets. Moreover, it must contain all cofinite sets. So, there exists exactly one nonprincipal ultrafilter, which we denote by V. Also, $U_x := \{a \in \mathbb{F} : x \in a\}$. Then for all x, $V \ntriangleleft U_x$, since $\blacklozenge\{x\}$ is finite and so not in V. Likewise $U_x \ntriangleleft V$. For let $b := \{y : |x - y| > 1\}$. Then $b \in V$ but $x \notin \blacklozenge b$. It is not hard to verify that $V \lhd V$. So, the bidual of $\mathfrak{F}$ has as its underlying frame the disjoint union of $\langle \mathbb{Z}, \prec \rangle$ and $\boxed{\circ}$. It is therefore not rooted, while its algebra is subdirectly irreducible. This example is quite similar to the algebra of finite and cofinite subsets of the infinite garland in Section 7.9.

Example. Consider the frame $\Omega := \langle \omega, >, \mathbb{O} \rangle$, where $\mathbb{O}$ is the set of finite and cofinite sets. Here, Ω_+ is not subdirectly irreducible, but $\Omega_+{}^+$ is rooted. To verify the first claim, notice that Ω_+ has an infinite descending chain of congruences whose intersection is the diagonal. These congruences correspond to the finite generated subframes of Ω. Hence, Ω_+ is not subdirectly irreducible. On the other hand, ${\Omega_+}^+$ has one more element than Ω, consisting of the ultrafilter V of cofinite sets. We show that for all ultrafilters U, $V \lhd U$. For let $b \in U$. Then $n \in b$ for some natural number n, and so $\blacklozenge b \supseteq \{y : y > n\}$, which is cofinite. Thus $\blacklozenge b \in V$, and therefore $V \lhd U$, by definition of $\lhd$.

Theorem 4.6.12 (Sambin). *There exist descriptive frames $\mathfrak{F}$ which are rooted such that $\mathfrak{F}_+$ is not subdirectly irreducible. There exist algebras $\mathfrak{A}$ which are subdirectly irreducible such that $\mathfrak{A}^+$ is not rooted.*

Let us stay a little bit with the theory of descriptive frames. Johan van Benthem [**12**] has investigated the structure of some ultrafilter extensions of frames. We will show that though the structure of ultrafilter extensions gives some evidence for the structure of biduals, it is by no means complete. Before we give examples we will develop some terminology. We will simplify the discussion by restricting ourselves to a single operator, $\Box$. Notice, however, that when we have a Kripke–frame $\langle f, \lhd\rangle$, then automatically we have two operators on the algebra of sets, namely $-\!\blacksquare$ and $\blacksquare\!-$. We will use this notation throughout this section. Notice on the other hand that a (general) frame for the monomodal language need not be a frame for the bimodal language, since the algebra of internal sets need not be closed under $-\!\blacksquare$. Nevertheless, we will use $-\!\blacksquare$, keeping in mind that it may result in noninternal sets. Moreover, we will reserve $\blacksquare$ for the operator on $\mathfrak{A}$, and use $\blacksquare\!-$ and $-\!\blacksquare$ for the operations on $\mathfrak{A}^+$. Finally, in a frame $\mathfrak{F}$, a set is denoted by a lower case letter only if it is an internal set. Upper case letters stand for sets which may also be external. A reminder on the use of topology: internal sets are also clopen sets. We will switch freely between these two characterizations.

Definition 4.6.13. *An element of a modal algebra $\mathfrak{A}$ is called* **essential** *if the open filter generated by it is a minimal open filter of $\mathfrak{A}$ distinct from $\{1\}$. $E_{\mathfrak{A}}$ denotes the set of essential elements of $\mathfrak{A}$.*

The reader may verify that an element is essential in $\mathfrak{A}$ iff it is an opremum.

Lemma 4.6.14. *Suppose that $E_{\mathfrak{A}}$ is not empty. Then $E_{\mathfrak{A}} \cup \{1\}$ is the smallest open filter. So, $E_{\mathfrak{A}}$ is nonempty iff $\mathfrak{A}$ is subdirectly irreducible.*

Now we start to investigate the nature of $\mathfrak{A}^+$. Let us call a set X in a frame $\mathfrak{F}$ a **transit** if it is successor closed. Furthermore, X is closed if it is an intersection of internal sets (iff it is a closed set of the topology induced by $\mathbb{F}$ on f). It is not hard to see that the closed transits are closed under arbitrary intersection and union, and that they form a locale.

Theorem 4.6.15 (Sambin & Vaccaro). *The locale of open filters of $\mathfrak{A}$ is anti–isomorphic to the dual locale of closed transits of $\mathfrak{A}^+$.*

Proof. Consider the map C defined by

$$F \mapsto C(F) := \{U \in pt(\mathfrak{A}) : F \subseteq U\} = \bigcap\{\widehat{a} : a \in F\}$$

This map is a bijection between closed sets of $\mathfrak{A}^+$ and filters of the boolean reduct of $\mathfrak{A}$. We need to show that F is open iff $C(F)$ is a transit.

$$
\begin{array}{lll}
 & (\forall a)(a \in F \Rightarrow \blacksquare a \in F) & \\
\Leftrightarrow & (\forall a)(C(F) \subseteq \widehat{a} \Rightarrow C(F) \subseteq \widehat{\blacksquare a}) & \text{since } a \in F \text{ iff } C(F) \subseteq \widehat{a} \\
\Leftrightarrow & (\forall a)(C(F) \subseteq \widehat{a} \Rightarrow C(F) \subseteq \blacksquare\!-\widehat{a}) & \text{since } \widehat{\blacksquare a} = \blacksquare\!-\widehat{a} \\
\Leftrightarrow & (\forall a)(C(F) \subseteq \widehat{a} \Rightarrow -\!\blacklozenge\,\widehat{a} \subseteq \widehat{a}) & \text{by adjunction} \\
\Leftrightarrow & -\!\blacklozenge\, C(F) \subseteq C(F) & \text{since } C(F) \text{ and } -\!\blacklozenge\, C(F) \text{ are closed} \\
\Leftrightarrow & C(F) \text{ is a transit} &
\end{array}
$$

□

Corollary 4.6.16. *$\mathfrak{A}$ is subdirectly irreducible iff $\mathfrak{A}^+$ has a greatest closed transit.*

For frames, a similar terminology can be set up. Notice first of all the following.

Proposition 4.6.17. *Let $\mathfrak{F}$ be a frame and $C \subseteq f$ a set. Then the smallest transit containing C is the set $T(C) := \bigcup_{k\in\omega} -\!\blacklozenge^{k} C$. It is open if C is internal. The largest transit contained in C is the set $K(C) := \bigcap_{k\in\omega} \blacksquare\!-^{k} C$. It is closed if C is internal.*

Definition 4.6.18. *Let $\mathfrak{F}$ be a frame. Let $I_{\mathfrak{F}}$ denote the set of all points x such that the transit of x is $\mathfrak{F}$. Furthermore, put $H_{\mathfrak{F}} := f - I_{\mathfrak{F}}$.*

Lemma 4.6.19. *For every frame, $H_{\mathfrak{F}}$ is a transit. Moreover, if $\mathfrak{F}$ has a greatest nontrivial transit C then $C = H_{\mathfrak{F}}$. Otherwise, $f = H_{\mathfrak{F}}$. So $\mathfrak{F}$ is rooted iff it has a greatest nontrivial transit.*

Lemma 4.6.20. *For every frame $\mathfrak{F}$ and every set C, if $H_{\mathfrak{F}} \subseteq C$ and $C \neq f$ then $H_{\mathfrak{F}} = K(C)$.*

Recall that in a topological space $\mathfrak{X}$, a set S is called **dense** if the closure in $\mathfrak{X}$ is the entire space. A set is of **measure zero** if it does not contain any open set iff its open kernel is empty iff its complement is dense.

Lemma 4.6.21. *For every frame, $H_{\mathfrak{F}}$ is either dense or closed.*

Proof. Suppose $H_{\mathfrak{F}}$ is not dense. Then there exists a clopen set C such that $H_{\mathfrak{F}} \subseteq C \subseteq f$. (For every closed set is the intersection of clopen subsets, and the closure of $H_{\mathfrak{F}}$ is not f.) Therefore, by Lemma 4.6.20, $H_{\mathfrak{F}} = K(C)$. By Proposition 4.6.17, $H_{\mathfrak{F}}$ is closed. □

Proposition 4.6.22. *For every algebra $\mathfrak{A}$, if $I_{\mathfrak{A}^+}$ is not of measure zero, then $\mathfrak{A}$ is subdirectly irreducible.*

Proof. Let $I_{\mathfrak{A}^+}$ be not of measure zero. Then $H_{\mathfrak{A}^+}$ is not dense. Therefore it is closed, by Lemma 4.6.21. By Corollary 4.6.16, $\mathfrak{A}$ is subdirectly irreducible. □

It might be deemed unnecessary to invoke the topology when we want to speak about Kripke–frames. For example, we know that for a Kripke–frame $\mathfrak{f}$, $\mathfrak{Ma}(\mathfrak{f})$ is subdirectly irreducible iff $\mathfrak{f}$ is rooted. But this is almost the only result that does not make use of the topological methods. So, let us deal first with the question of open filters in the algebra $\mathfrak{Ma}(\mathfrak{f})$ of a Kripke–frame. At first blush they seem to correspond simply to the transits of $\mathfrak{f}$. As a particular corollary, if $\mathfrak{f}$ is connected, $\mathfrak{Ma}(\mathfrak{f})$ should be simple. But this is wrong. For let $\mathfrak{f}$ be countably infinite. Then $\mathfrak{Ma}(\mathfrak{f})$ has at least $2^{\aleph_0}$ elements. It is easy to show, however, that such an algebra cannot be simple. For take any element $b \neq 1$. The least open filter containing b is countable. Hence it is not $\{1\}$ and not the entire algebra. Thus, the conjecture that $\mathfrak{Ma}(\mathfrak{f})$ is simple if every world of $\mathfrak{f}$ is a root is easily refuted. Nevertheless, it is instructive to see a concrete counterexample.

Example. Let $\mathfrak{z} = \langle Z, \lhd\rangle$ be the frame with Z the set of integers and $x \lhd y$ iff $|x - y| = 1$. Every world of $\mathbb{Z}$ is a root. So, $H_{\mathfrak{Z}} = \varnothing$, which is a clopen set. So, the algebra of sets is subdirectly irreducible (which also follows from the fact that $\mathbb{Z}$ is rooted). Therefore, $E_{\mathfrak{Z}}$ is not empty. That is, we should have essential elements.

Lemma 4.6.23. *$E_{\mathfrak{z}}$ consists exactly of the cofinite sets $\neq Z$.*

Proof. Suppose that A is cofinite. Then there exists a $k \in \omega$ such that $Z - A \subseteq [-k, k]$. Let $B \subsetneq \omega$. Then for some $m \in Z$, $m \notin B$. So, $[m - n, m + n] \subseteq \blacksquare^n B$. Hence, $[-k, k] \supseteq \blacksquare^{m+k} B$, from which follows that $A \supseteq \blacksquare^{m+k} B$. So, A is essential. Now let A not be cofinite and let $B := \omega - \{0\}$. Then for no $k \in \omega$, $\blacksquare^k B \subseteq A$, since $\blacksquare^k B$ is cofinite. □

We remark here that the converse of Proposition 4.6.22 is false.

Example. Let $\mathfrak{Z} := \langle \mathfrak{z}, \mathbb{O}\rangle$ where $\mathbb{O}$ consists of all finite unions of sets of the form $o(k, a) := \{n \cdot 2^k + a : n \in Z\}$, where k is a natural number and a an integer. The algebra $\mathfrak{A} := \mathfrak{Z}_+$ is simple. For any set of $\mathbb{O}$ is of the form $b = \bigcup_{i<p} o(k, a_i)$ for some p and k and some a_i. Let $\{b_j : j < q\}$ be the set of numbers such that for every $j < q$ there exist $i, i' < p$ such that $b_j \equiv a_i + 1 \pmod{2^k}$ and $b \equiv a_{i'} - 1 \pmod{2^k}$. Then $\blacksquare b = \bigcup_{j<q} o(k, b_j)$. It follows that $\blacksquare^{2^k} b = \varnothing$ iff $b \neq 1$. By the criterion for subdirect irreducibility, $\mathfrak{A}$ is subdirectly irreducible with 0 an opremum. Hence $\mathfrak{A}$ is simple. We will show that the frame underlying $\mathfrak{A}^+$ is decomposable into a disjoint union of at least two frames, from which follows that $I_{\mathfrak{A}^+} = \varnothing$. To see that, take an ultrafilter U in $\mathfrak{A}$. We define a sequence $J(U) = \langle j_k : k \in \omega\rangle$ as follows. For each $k \in \omega$ let j_k be the (unique) remainder $\pmod{2^k}$ such that $o(k, j_k) \in U$. $J(U)$ is a sequence satisfying

$$(\ddagger) \qquad j_{k+1} = j_k \text{ or } j_{k+1} = j_k + 2^1 k$$

for all $k \in \omega$. Conversely, let $S = \langle s_k : k \in \omega\rangle$ be a sequence satisfying $(\ddagger)$. Let $U(S)$ be the ultrafilter containing $o(k, s_k)$. It is easy to see that this ultrafilter exists (this collection has the finite intersection property, by $(\ddagger)$), and that $U(S)$ is indeed uniquely defined. Now, there are evidently $2^{\aleph_0}$ many such sequences. Hence, $\mathfrak{A}^+$ has uncountably many worlds.

$\mathfrak{Z}$ satisfies $\mathbf{alt}_2$ and $\mathbf{B}$ and so $\mathfrak{A}^+$ satisfies $\mathbf{alt}_2$ and $\mathbf{B}$ as well. Hence, the underlying frame has the property that (a) each point sees at most two points, (b) if x sees y, y sees x. (This follows from the results of Chapter 5, but can also be established directly.) It follows that a connected component is generated by any of its points, and that it is countable. Since the frame is uncountable, it is has more than two connected components. Consequently, $I_{\mathfrak{A}^+} = \varnothing$.

Exercise 155. Show that a finite differentiated frame is full.

Exercise 156. Any finite frame is compact. Hence a finite differentiated frame is descriptive.

Exercise 157. Prove Proposition 4.6.5.

Exercise 158. (This example is due to Frank Wolter. It is very closely related to the frame considered in Section 3.5.) Take the set $\omega + 1$ and put $i \lhd j$ iff $i < j < \omega$ or $i = j = \omega$. The internal sets are all sets which are finite and do not contain ω or else are cofinite and contain ω. Show that this frame is compact, differentiated but not tight.

Exercise 159. Let $\mathcal{V}$ be a variety, $n \in \omega$ and $\mathfrak{A}_i \in \mathcal{V}$, $i < n$. A subalgebra $\mathfrak{B} \rightarrowtail \prod_{i<n} \mathfrak{A}_i$ is called **skew–free** if for every congruence Θ on $\mathfrak{B}$ there exist $\Psi_i \in Con(\mathfrak{A}_i)$ such that $\Theta = (\mathsf{X}_{i<n}\Psi) \cap B^2$ (see Section 1.3 for definitions). Show that every subalgebra of a direct product of modal algebras is skew–free. *Hint.* Use duality.

Exercise 160. Show that there exist simple modal algebras which do not generate semisimple varieties.

4.7. Frame Constructions III

Our aim in this section is to translate Birkhoff's Theorem on varieties into general frames. This will allow us to say which classes of frames are classes of frames for a logic. In the form that this theorem takes here, however, it is not very surprising, but we will transform it later into stronger variants that allow deep insights into the structure of such classes. Recall that Birkhoff's Theorem says that a class is equationally definable iff it is closed under products, subalgebras and homomorphic images. In the context of modal algebras equational classes coincide with modally definable classes, where a class $\mathcal{K}$ of algebras is called **modally definable** if there is a set Φ of modal formulae such that $\mathcal{K}$ contains exactly the algebras satisfying Φ. Likewise, a class $\mathcal{K}$ of frames is **modally definable** if $\mathcal{K}$ is the class of frames for some Φ. This can be relativized to some class, e. g. the class of refined frames, descriptive, full frames etc. Of particular interest is the class of modally definable Kripke–frames.

One immediate problem is that we have to translate the notions of product, subalgebra etc. into frames. Here, category theory is of good use since it provides canonical definitions of these things. A good illustration is the notion of a *product*.

Definition 4.7.1. *Let $\mathfrak{C}$ be a category, and B_i, $i \in I$, an indexed collection of $\mathfrak{C}$–objects. A pair $\langle A, \{p_i : i \in I\}\rangle$, where A is a $\mathfrak{C}$–object and $p_i : A \to B_i$ $\mathfrak{C}$–arrows, is called a* **product** *of the B_i if for each object C and arrows $q_i : C \to B_i$ there is a unique morphism $m : C \to A$ such that $q_i = m \circ p_i$ for all $i \in I$. The maps p_i are called* **projections***. $\langle A, \{p_i : i \in I\}\rangle$ is called a* **coproduct** *of the B_i if $\langle A, \{(p_i)^{op} : i \in I\}\rangle$ is a product in the dual category, $\mathfrak{C}^{op}$.*

$$\begin{tikzcd}
 & & B_2 \\
C \arrow[r, dashed, "!"] \arrow[urr, "q_2"] \arrow[drr, "q_1"'] & A \arrow[ur, "p_2"] \arrow[dr, "p_1"'] & \\
 & & B_1
\end{tikzcd}$$

Usually, only the object A in the pair $\langle A, \{p_i : i \in I\}\rangle$ is referred to as the product. (This at least is common usage in algebra.) So, a product is an object for which projections p_i that make $\langle A, \{p_i : i \in I\}\rangle$ a product in the sense of the definition. However, notice that the map denoted by '!' in the picture above is not uniquely defined by A and C alone but only given the pairs $\langle A, \{p_1, p_2\}\rangle$ and $\langle C, \{q_1, q_2\}\rangle$. We can view a product as a solution to a special diagram (consisting of two objects and identity arrows). This solution consists in an object and arrows from that object into the objects of the diagram. What makes such a solution a product is a condition that is usually referred to as the *universal property*. (See the exercises.) The reader is advised to spell out the definition of a coproduct in detail. Before we proceed to examples, let us note one important fact about products and coproducts.

Theorem 4.7.2. *Let C and D be products of the B_i, $i \in I$. Then C is isomorphic to D. Moreover, if C is a product and isomorphic to D, then D also is a product of the B_i, $i \in I$.*

Proof. By assumption there are maps $p_i : C \to B_i$ and $q_i : D \to B_i$ such that for any E with maps $r_i : E \to B_i$ we have $m : E \to C$ and $n : E \to D$ such that $r_i = p_i \circ m$ and $r_i = q_i \circ n$ for all $i \in I$. We apply the universal property for C to D and get a map $f : D \to C$ such that $q_i = p_i \circ f$ for all i. We apply the universal property of D to C and obtain likewise a map $g : C \to D$ such that $p_i = q_i \circ g$ for all i. Then we have $q_i = p_i \circ f = (q_i \circ g) \circ f = q_i \circ (g \circ f)$ as well as $p_i = q_i \circ g = (p_i \circ f) \circ g = p_i \circ (f \circ g)$, again for all i. Since C is a product, there is only one map $i : C \to C$ satisfying $p_i = p_i \circ i$.

Since the identity $id(C)$ on C has this property, we conclude that $f \circ g = id(C)$. And analogously we get that $g \circ f = id(D)$. So C and D are isomorphic.

For the second claim let C be a product with projections p_i, $i \in I$. Let $h : D \to C$ be an isomorphism with inverse $k : C \to D$. Then we claim that D is a product with projections $q_i := p_i \circ h$. For let E be given and $r_i : E \to B_i$, $i \in I$. Then there exists a unique map $f : E \to C$ such that $r_i = p_i \circ f$ for all $i \in I$. Let $g := k \circ f$. Then $r_i = p_i \circ f = q_i \circ k \circ f = q_i \circ g$ for all $i \in I$. Furthermore, g is unique. For if g' also satisfies $r_i = q_i \circ g'$ for all i, then $r_i = p_i \circ h \circ g'$ for all i, from which $h \circ g' = h \circ g$. This implies $k \circ h \circ g' = k \circ h \circ g$, which is the same as $g = g'$. ◻

The product of two algebras, defined earlier, is a product in the categorial sense. In general, let Ω be a signature, T an equational theory in Ω, and let $\mathsf{Alg}\, T$ be the category of Ω–algebras for T. This category has products. $\langle \prod_{i \in I} \mathfrak{A}_i, p_i \rangle$, where p_i is the projection onto the ith component, is a product of the family $\{\mathfrak{A}_i : i \in I\}$. This fact is used later. We perform the argument with $I := \{1, 2\}$. Let $\mathfrak{B}_1$, $\mathfrak{B}_2$ be given. Put $\mathfrak{A} := \mathfrak{B}_1 \times \mathfrak{B}_2$, and let the projections be $p_1 : \langle b_1, b_2 \rangle \mapsto b_1$ and $p_2 : \langle b_1, b_2 \rangle \mapsto b_2$. Take any $\mathfrak{C}$ with homomorphisms $q_i : \mathfrak{C} \to \mathfrak{B}_i$. Put $m : \mathfrak{C} \to \mathfrak{B}_1 \times \mathfrak{B}_2 : c \mapsto \langle q_1(c), q_2(c) \rangle$. There is no other choice; for assuming $m(c) = \langle x_1, x_2 \rangle$ we get

$$\begin{array}{llll} q_1(c) & = (p_1 \circ m)(c) & = p_1(\langle x_1, x_2 \rangle) & = x_1 \\ q_2(c) & = (p_1 \circ m)(c) & = p_2(\langle x_1, x_2 \rangle) & = x_2 \end{array}$$

So m is unique, and we only have to show that it is a homomorphism. We trust that the reader can fill in the proof. From the previous theorem we get that products are unique up to isomorphism. Let us cash out on this immediately. Say that a category **has products** (**has coproducts**) if for any family of objects the product (coproduct) of that family exists. The category of algebras in a variety has products.

Theorem 4.7.3. *Let Λ be a polymodal logic. The category of descriptive Λ–frames has coproducts.*

Proof. The proof is by duality of the category of descriptive frames and the category of modal algebras. Suppose that $\mathfrak{F}_i$, $i \in I$, is a family of frames; put

$$\coprod_{i \in I} \mathfrak{F}_i := (\prod_{i \in I} (\mathfrak{F}_i)_+)^+ \ .$$

We claim that this a coproduct. By the Duality Theorem it is enough to show that $(\coprod_{i \in I} \mathfrak{F}_i)_+$ is a product. However, it is isomorphic to $\prod_{i \in I} (\mathfrak{F}_i)_+$. The latter is a product of the $(\mathfrak{F}_i)_+$. ◻

There also is a notion of a coproduct of frames, called the *disjoint union*. Let $\mathfrak{F}_i = \langle \mathfrak{f}_i, \mathbb{F}_i \rangle$ be frames. The disjoint union, $\bigoplus_{i \in I} \mathfrak{F}_i$, is defined over the disjoint union of the sets f_i with relations being the disjoint union of the respective relations. The sets are of the form $b = \bigcup_{i \in I} b_i$ where $b_i \in \mathbb{F}_i$. Since the f_i are disjoint, so are then the b_i, and it turns out that we get

$$(\bigoplus_{i\in I}\mathfrak{F}_i)_+ \cong \prod_{i\in I}(\mathfrak{F}_i)_+ \ .$$

The projections are $p_i : a \mapsto a \cap f_i$. The so defined sets form a modal algebra. To see this, consider the map $\pi : b \mapsto \bigcup_{i\in I} b_i$. This is readily checked to be a homomorphism and bijective. The following theorem has been shown in the special case $I = \{1, 2\}$ and Kripke–frames in Theorem 2.4.4.

Theorem 4.7.4. *$\bigoplus_{i\in I}\mathfrak{F}_i$ is a coproduct of the frames $\mathfrak{F}_i$ in* Frm.

Proof. Put $\mathfrak{G} := \bigoplus_{i\in I}\mathfrak{F}_i$. Let $\mathfrak{H}$ be a frame and $h_i : \mathfrak{F}_i \to \mathfrak{H}$. We define $m : \bigoplus_{i\in I}\mathfrak{F}_i \to \mathfrak{H}$ by $m(x) := h_i(x)$ if $x \in f_i$. Since the f_i are all disjoint, this definition by cases is unambigous. m is a p–morphism, for if $x \lhd_j y$ in $\mathfrak{G}$ and $x \in f_i$ then also y in f_i and $x \lhd_j y$ already in f_i. Since h_i is a p–morphism, $m(x) = h_i(x) \lhd_j h_i(y) = m(y)$. Now assume $m(x) \lhd_j u$. Let $x \in f_i$. Then also $h_i(x) \lhd_j u$, and we get an $y \in f_i$ such that $h_i(y) = u$ and $x \lhd_j y$. But then $m(y) = h_i(y) = u$, and this proves the second condition. Finally, let $b \in \mathbb{H}$ be a set. We have $c_i := h_i^{-1}[b] \in \mathbb{F}_i$. Put $c := \bigcup_{i\in I} c_i$. c is an internal set of $\mathfrak{G}$ and $m[c] = b$. ☐

Now we have two definitions of coproducts, one for descriptive frames and one for arbitrary frames. Since the category DFrm is a subcategory of Frm in the sense that it contains a subclass of the objects of Frm but all Frm–arrows between them, it can be shown that for descriptive frames $\mathfrak{F}_i$, $i \in I$, there exists a map $\bigoplus_{i\in I}\mathfrak{F}_i \rightarrowtail \coprod_{i\in I}\mathfrak{F}_i$. As it turns out, the two coproducts are not always the same. For example, take the full frame $\mathfrak{ch}_n^\sharp$, $\mathfrak{ch}_n := \langle n, >\rangle$. These frames are descriptive. Put $\mathfrak{D} := \coprod_{n\in\omega}\mathfrak{ch}_n^\sharp$, and $\mathfrak{K} := \bigoplus_{n\in\omega}\mathfrak{ch}_n^\sharp$. There exists an arrow $\mathfrak{K} \rightarrowtail \mathfrak{D}$. Furthermore, the algebras $\mathfrak{K}_+$ and $\mathfrak{D}_+$ are isomorphic, since they are (isomorphic to the) the product of $(\mathfrak{ch}_n^\sharp)_+$. Nevertheless, there is no isomorphism between these frames. One way to see this is as follows. The frame $\mathfrak{D}$ contains a world which has no finite depth (or no depth at all), while each point in $\mathfrak{K}$ has a depth. Another argument is the following. $\mathfrak{K}$ has countably many worlds, the set of worlds being a countable union of finite sets. But $\mathfrak{D}$ has uncountably many points, as there are uncountably many ultrafilters in the algebra of sets. (This is given as an exercise below.) It appears to be paradoxical that we should end up with several versions of a coproduct, but this is due to the fact that a coproduct is a notion that makes sense only within a category; it has no absolute meaning. Since the category Frm has more objects, we end up with different coproducts. Finally, it is clear that we shall end up with copoducts of frames rather than products, by duality, or simply the fact that arrows go the opposite way. The reader may however also check that the categories of frames and descriptive frames have no products. Namely, take $\mathfrak{F}_1$ to be the one point reflexive frame ($\kappa = 1$), $\mathfrak{F}_2$ the one point irreflexive frame. Then there is no product of these two in either category.

Figure 4.2.

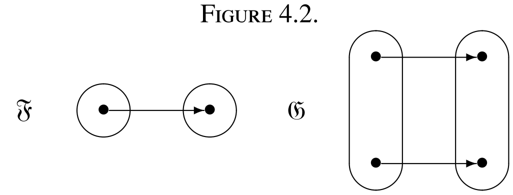
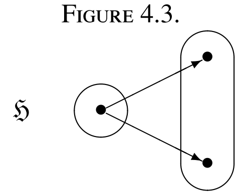


Figure 4.3.



Different problems arise with subalgebras and homomorphic images. It appears at first sight that subalgebras translate into p–morphic images and homomorphic images into generated subframes. But the terminology on the frames is more subtle here. We know that a subalgebra–map $h : \mathfrak{A} \rightarrowtail \mathfrak{B}$ defines a unique $h^+ : \mathfrak{B}^+ \twoheadrightarrow \mathfrak{A}^+$ and conversely. But we write $f : \mathfrak{F} \twoheadrightarrow \mathfrak{G}$ for frames if f is surjective on the worlds only. Now consider case of Figure 4.2. We have $\mathfrak{F} \rightarrowtail \mathfrak{G}$, and we have an isomorphism between the algebras of internal sets. We expect therefore that there is a surjective p–morphism from $\mathfrak{F}$ to $\mathfrak{G}$. But there is none. The duality theory cannot be used in this case. A similar counterexample can be constructed for generated subframes. Consider namely the frame $\mathfrak{H}$. Although the algebras of sets is isomorphic to that of $\mathfrak{G}$, none is a generated subframe of the other.

Moreover, notice that there is a distinction between *embeddings* and *subframe embeddings*. An *embedding* is just an injective p–morphism, while a map $i : \mathfrak{F} \rightarrowtail \mathfrak{G}$ is a *subframe embedding* if $i^{-1} : \mathbb{G} \to \mathbb{F}$ is surjective. In the latter case $\mathfrak{F}$ is required to have no more sets than necessary to make i a p–morphism. If $e : \mathfrak{F} \to \mathfrak{G}$ is an embedding, take $\mathbb{F}_2 = \{e^{-1}[b] : b \in \mathfrak{G}\}$. Then the identity is an embedding: $id : \langle \mathfrak{f}, \mathbb{F}_2 \rangle \rightarrowtail \langle \mathfrak{f}, \mathbb{F} \rangle$ and the image of $\langle f, \mathbb{F}_2 \rangle$ under e is a generated subframe.

Now let us proceed to the promised characterization of modally definable classes of frames. The easiest case is that of classes of descriptive frames. In descriptive frames, the correspondence is as exact as possible.

Theorem 4.7.5. *A class of descriptive frames is modally definable iff it is closed under coproducts, p–morphic images and generated subframes.*

Proof. This follows from the Duality Theorem as follows. Suppose that $\mathfrak{X}$ is modally definable. Then $\mathfrak{X}_+$ is equationally definable, hence a variety. Therefore $\mathfrak{X}$ is closed under coproducts. For if $\mathfrak{F}_i$, $i \in I$, is a family of descriptive frames from

$\mathfrak{X}$ then $(\bigoplus_{i\in I}\mathfrak{F}_i)_+ \cong \prod_{i\in I}(\mathfrak{F}_i)_+$. Since $\mathfrak{X}_+$ is closed under products the latter is in $\mathfrak{X}_+$, and thus is the former, by closure under isomorphisms. But since $(\mathfrak{X}_+)^+ \cong \mathfrak{X}$ we also have $\bigoplus_{i\in I}\mathfrak{F}_i \in \mathfrak{X}$. Analogously it is shown that $\mathfrak{X}$ is closed under generated subframes since $\mathfrak{X}_+$ is closed under homomorphic images, and that $\mathfrak{X}$ is closed under contractions since $\mathfrak{X}_+$ is closed under subalgebras. Conversely, if a class $\mathfrak{X}$ of descriptive frames is closed under coproducts, $\mathfrak{X}_+$ is closed under products. If $\mathfrak{X}$ is closed under generated subframes, $\mathfrak{X}_+$ is closed under homomorphic images, and if $\mathfrak{X}$ is closed under contractions, $\mathfrak{X}_+$ is closed under subalgebras. ◻

We know that two frames $\mathfrak{F}$ and $\mathfrak{G}$ whose algebras are isomorphic have the same modal theory. Hence, if we want to see what classes of frames are modally definable, we just have to close it under the operation $\mathsf{B} : \mathfrak{F} \mapsto (\mathfrak{F}_+)^+$ as well as B^{-1}. $(\mathfrak{F}_+)^+$ is known as the **bidual** of $\mathfrak{F}$.

Theorem 4.7.6. *A class of frames is modally definable iff it is closed under disjoint unions, p–morphic images, generated subframes, and if it and its complement are closed under biduals.*

Proof. Suppose that $\mathfrak{X}$ is modally definable. Then it is closed under biduals, and its complement is closed under biduals as well. By the previous theorem, $\mathfrak{X}$ is closed under p–morphic images and generated subframes. Furthermore, the bidual of $\bigoplus_{i\in I}\mathfrak{F}_i$ is nothing but the coproduct in DFrm. Since $\mathfrak{X}$ is closed under coproducts and inverse biduals, $\mathfrak{X}$ is closed under disjoint unions. Suppose then that $\mathfrak{X}$ has all the required closure properties. Then consider $\mathfrak{X}_d := \mathfrak{X} \cap \mathfrak{D}$, the class of descriptive frames contained in $\mathfrak{X}$. We have $\mathfrak{X} = \mathsf{B}^{-1}[\mathfrak{X}_d]$, by closure and inverse closure under B. Therefore, the modal theory of $\mathfrak{X}$ is the same as the modal theory of $\mathfrak{X}_d$, and $\mathfrak{X}$ is the class of frames satisfying the modal description of $\mathfrak{X}_d$. So, we only need to show that $\mathfrak{X}_d$ is modally definable. $\mathfrak{X}_d$ is closed under coproducts since the coproduct in DFrm is the bidual of the disjoint union. $\mathfrak{X}_d$ is closed under p–morphic images and generated subframes by assumption on $\mathfrak{X}$. Thus $\mathfrak{X}_d$ is modally definable. ◻

Exercise 161. Recall the notion of a net extension of Chapter 2.4. Prove that given $\mathfrak{f}$, $\mathfrak{g}$ and $\mathfrak{h}$, an embedding $e : \mathfrak{f} \rightarrowtail \mathfrak{g}$ and a contraction $c : \mathfrak{f} \twoheadrightarrow \mathfrak{h}$, there exists a $\mathfrak{e}$ and maps $p : \mathfrak{g} \to \mathfrak{e}$, $q : \mathfrak{h} \to \mathfrak{e}$ such that (i) $p \circ e = q \circ c$ and (ii) for all $\mathfrak{e}'$, $p' : \mathfrak{g} \to \mathfrak{e}'$ and $q' : \mathfrak{h} \to \mathfrak{e}'$ satisfying (i) there exists a unique $i : \mathfrak{e} \rightarrowtail \mathfrak{e}'$ such that $p' = i \circ p$, and $q' = i \circ q$. Prove also that p is a contraction and q an embedding.

Exercise 162. Generalize the setting of the previous exercise as follows. Let A, B and C be objects and morphisms p, q as in the picture below to the left. This is a special diagram. Say that D together with morphisms $B \to D$, $C \to D$ is a **pushout** of this diagram in a category $\mathfrak{C}$ if for any other D' with maps $B \to D'$, $C \to D'$ there exists a unique map $D \to D'$ making the entire diagram to the right commute. D together with the maps into D is called a *co–cone* of the diagram to the left. The dual notion is that of a **pullback**. It is based on the opposite or dual of the diagram to the left. Try to formulate the notion of a pullback. Show that any two pushouts of a given diagram are isomorphic.

$$\begin{tikzcd} A \arrow[r, "p"] \arrow[d, "q"'] & B \\ C & \end{tikzcd} \qquad \begin{tikzcd} A \arrow[r, "p"] \arrow[d, "q"'] & B \arrow[d] \arrow[ddr] & \\ C \arrow[r] \arrow[drr] & D \arrow[dr, "!"] & \\ & & D' \end{tikzcd}$$

Exercise 163. Let $\Delta = \langle Ob, Mor\rangle$ be a diagram in a category $\mathcal{C}$. A **cone** for Δ is a pair $\mathfrak{L} = \langle L, \langle p_C : C \in Ob\rangle\rangle$ where L is an object and $p_C : L \to C$ such that for any pair $C, D \in Ob$ and $C \xrightarrow{f} D \in Mor$ we have $p_D = f \circ p_C$. A **limit** for Δ is a cone $\mathfrak{L}$ such that for any cone $\mathfrak{L}'$ for Δ we have a uniquely defined map $i : L' \to L$ such that $p'_C = p_C \circ i$. Now show that products and pullbacks are limits for special kinds of diagrams. Dualize to define the notion **co–cone** and **co–limit** and see where you can find instances of **co–limits**.

Exercise 164. Show that there exist $2^{\aleph_0}$ ultrafilters on $\wp(\omega)$. (In fact, there are $2^{2^{\aleph_0}}$ many.) *Hint.* Consider the sets $I_n := \{i : i \equiv 0 \pmod{p_n}\}$, where p_n is the nth prime number. For each $M \subseteq \omega$ let U_M be an ultrafilter containing I_n iff $n \in M$.

4.8. Free Algebras, Canonical Frames and Descriptive Frames

In this section we will discuss the difference between canonical frames and descriptive frames. Before we do so, let us reflect on duality in connection with canonical frames. We have shown in Section 1.3 that in any given variety $\mathcal{V}$, for any cardinality α, $\mathcal{V}$ contains freely α–generated algebras. Furthermore, if $v : \alpha \rightarrowtail \beta$ then $\overline{v} : \mathfrak{Fr}_{\mathcal{V}}(\alpha) \rightarrowtail \mathfrak{Fr}_{\mathcal{V}}(\beta)$ and if $w : \alpha \twoheadrightarrow \beta$ then $\overline{w} : \mathfrak{Fr}_{\mathcal{V}}(\alpha) \twoheadrightarrow \mathfrak{Fr}_{\mathcal{V}}(\beta)$. Now let $\mathcal{V}$ be the variety of Θ–algebras. Then $\mathfrak{Can}_\Theta(\alpha)$ turns out to be the dual of $\mathfrak{Fr}_{\mathcal{V}}(\alpha)$. Moreover, $\overline{v}^+ : \mathfrak{Can}_\Theta(\beta) \twoheadrightarrow \mathfrak{Can}_\Theta(\alpha)$ and also $\overline{w}^+ : \mathfrak{Can}_\Theta(\beta) \rightarrowtail \mathfrak{Can}_\Theta(\alpha)$. The difference is that while the worlds of $\mathfrak{Can}_\Theta(\alpha)$ are maximally consistent sets the worlds of the dual of $\mathfrak{Fr}_\Theta(\alpha)$ are ultrafilters. The difference, however, is negligeable. Let us elaborate on this a little bit. For given cardinal number α let $V_\alpha := \{p_\beta : \beta < \alpha\}$. Now let $\varphi \equiv \chi$ iff $\varphi \leftrightarrow \chi \in \Theta$. This is a congruence. Moreover, $\mathfrak{Fr}_{\mathcal{V}}(\alpha) = \mathfrak{Tm}(V_\alpha)/\equiv$. The homomorphism corresponding to $\equiv$ is denoted by $h_\equiv$ and the congruence class of φ is denoted by $[\varphi]$.

Proposition 4.8.1. *A set Δ is a maximally consistent set of V_α–terms iff there exists an ultrafilter U in $\mathfrak{Fr}_{\mathcal{V}}(\alpha)$ such that $\Delta = h_\equiv^{-1}[U]$.*

Proof. Let Δ be maximally consistent. Then it is deductively closed, as is easily seen. In particular, if $\varphi \in \Delta$ and $\varphi \equiv \chi$, then $\chi \in \Delta$. So we have $\Delta = h_\equiv^{-1}[h_\equiv[\Delta]]$.

Furthermore, the image U of Δ under $h_\equiv$ is deductively closed and so it is a filter, by Proposition 1.7.8. To show that U is an ultrafilter, take $b \notin U$. Then $b = [\varphi]$ for some $\varphi \notin \Delta$. By maximal consistency of Δ, $\neg\varphi \in \Delta$ (Lemma 2.8.2). Moreover, $[\neg\varphi] = -[\varphi] = -b$. Hence $-b \in U$. If $-b \in U$, then $b \notin U$, otherwise Δ is inconsistent. So, U is an ultrafilter. Conversely, assume that U is an ultrafilter, and put $\Delta := h_\equiv^{-1}[U]$. Δ is deductively closed. For let $\varphi \in \Delta$ and $\varphi \to \chi \in \Delta$. Then $[\varphi] \in U$ and $[\varphi] \to [\chi] = [\varphi \to \chi] \in U$. U is deductively closed, so $[\chi] \in U$, from which $\chi \in \Delta$. Δ is maximal. For given φ, $[\varphi] \in U$ or $[\neg\varphi] = -[\varphi] \in U$. Hence $\varphi \in \Delta$ or $\neg\varphi \in \Delta$. But not both, since not both $[\varphi] \in U$ and $-[\varphi] \in U$. ◻

Proposition 4.8.2. *The map $h_\equiv^{-1}$ is an isomorphism from $\mathfrak{Fr}_\Theta(\alpha)^+$ onto $\mathfrak{Can}_\Theta(\alpha)$.*

The proof of this proposition is left as an exercise. Now let $v : \alpha \to \beta$ be a function. Denote by v also the function $p_\mu \mapsto p_{v(\mu)}$, $\mu < \alpha$. Then $\overline{v} : \mathfrak{Fr}_\Theta(\alpha) \to \mathfrak{Fr}_\Theta(\beta)$. Furthermore, if v is injective iff $\overline{v}$ is, and v is surjective iff $\overline{v}$ is. Then $\overline{v}^+ : \mathfrak{Fr}_\Theta(\beta)^+ \to \mathfrak{Fr}_\Theta(\alpha)^+$, by duality. Furthermore, $\overline{v}^+$ is injective iff $\overline{v}$ is surjective, and surjective iff $\overline{v}$ is injective. Recall also how $\overline{v}^+$ is defined; given an ultrafilter U of $\mathfrak{Fr}_\Theta(\beta)$, $\overline{v}^+(U) := \overline{v}^{-1}[U]$. Now recall from Section 2.8 the map $X_v : \mathfrak{Tm}(V_\beta) \to \mathfrak{Tm}(V_\beta)$, taking a β–term to its preimage under v. This map can be shown to map maximally consistent sets onto maximally consistent sets. Namely, given Δ, put

$$Y_v(\Delta) := h_\equiv^{-1} \circ \overline{v} \circ h_\equiv[\Delta]$$

Lemma 4.8.3. *For a maximally consistent set Δ, $Y_v(\Delta) = X_v(\Delta)$.*

The proof is straightforward but rather unrevealing. The Theorem 2.8.11 follows in this way from the duality theory developed in this chapter. On the one hand duality theory is more general, since there are descriptive frames which are not canonical for any logic. On the other hand, we will show below that descriptive frames are generated subframes of canonical frames. This result of duality theory therefore follows already from Theorem 2.8.11.

Definition 4.8.4. *Let Λ be a modal logic and α a cardinal number. A frame $\mathfrak{F}$ is called a* ***canonical frame for*** *Λ if $\mathfrak{F}$ is an α–canonical frame for some α; and $\mathfrak{F}$ is* ***canonical*** *simpliciter if it is a canonical frame for some logic.*

Recall that the α–canonical frame for Λ is isomorphic to $\mathfrak{Fr}_\Lambda(\alpha)^+$. It may appear at first blush that canonical frames are the same as descriptive frames. This is not so as we will show below. However, every descriptive frame is a generated subframe of a canonical frame. Namely, fix any logic Λ and let $\mathfrak{D}$ be a descriptive frame for Λ. Then $\mathfrak{D}_+$ is a Λ–algebra. Every Λ–algebra is the image of a free Λ–algebra by Theorem 1.3.5. Consequently, $\mathfrak{D} \rightarrowtail \mathfrak{Fr}_\Lambda(\alpha)$ for some α. We will use this fact to derive a useful characterization of canonicity of logics first shown in [**186**]

Definition 4.8.5. *A logic Λ is called* ***d–persistent*** *if for every descriptive frame $\mathfrak{D}$ for Λ the underlying Kripke–frame, $\mathfrak{D}_\sharp$, is a Λ–frame as well.*

THEOREM 4.8.6 (Sambin & Vaccaro). *A logic is canonical iff it is d–persistent.*

PROOF. Since every canonical frame is descriptive, d–persistence implies canonicity. So let us assume that Λ is canonical and $\mathfrak{D}$ a descriptive frame such that $\mathfrak{D} \vDash \Lambda$. Then $\mathfrak{D} \rightarrowtail \mathfrak{C}$ for some canonical $\mathfrak{C}$ frame for Λ. Because $\mathfrak{C} \vDash \Lambda$, we have $\mathfrak{C}_\sharp \vDash \Lambda$, by the fact that Λ is canonical. But then $\mathfrak{D}_\sharp \vDash \Lambda$, since $\mathfrak{D}_\sharp \rightarrowtail \mathfrak{C}_\sharp$. ◻

We can cash out here a nice result concerning finite model property. A variety is said to be **locally finite** if every finitely generated algebra is finite. Obviously, a variety is locally finite if all finitely generated free algebras are finite. By duality, they are isomorphic to the full algebra of sets over a finite Kripke–frame. Now say that a logic is **locally finite** if its corresponding variety is. Say that a logic Λ has a property $\mathfrak{P}$ **essentially** if every extension of Λ has $\mathfrak{P}$.

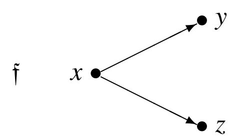
THEOREM 4.8.7. *If Θ is locally finite then every extension of Θ is weakly canonical and has the finite model property essentially.*

Now let us return to the question whether descriptive frames are also canonical. The question is whether modal algebras are free algebras in some variety. For example, vector spaces are freely generated by their basis. However, not all boolean algebras are free, for example the algebra 2^3. (A finite boolean algebra is freely generated by n elements iff it has 2^{2^n} elements.) Nevertheless, it is interesting to approach the question to see clearly what the relation between the two notions is. Let $\mathfrak{D}$ be a descriptive frame and put $\Theta := \mathsf{Th}(\mathfrak{D})$. First, if $\mathfrak{D} \cong \mathfrak{Can}_\Lambda(\alpha)$ for some α and some Λ, then $\mathfrak{D} \cong \mathfrak{Can}_\Theta(\alpha)$. In other words, $\mathfrak{D}$ is α–canonical in its own variety iff it is α–canonical. For by duality, if an algebra $\mathfrak{A}$ is freely α–generated in a variety $\mathcal{V}$, we have $\mathfrak{A} \in \mathcal{V}$ and so the variety generated by $\mathfrak{A}$ is included in $\mathcal{V}$. However, $\mathfrak{A}$ is then freely α–generated in any smaller variety containing it, and so freely α–generated in the least such variety.

PROPOSITION 4.8.8. *Let $\mathfrak{F}$ be a frame. $\mathfrak{F}$ is α–canonical for some logic iff it is α–canonical for* $\mathsf{Th}\,\mathfrak{F}$.

We will need the distinction between an algebra generating a variety and an algebra free in that variety later on in the connection with splittings. It is necessary for the understanding of the results proved there to have seen an explicit construction of α–free algebras and we will provide such an example now. Consider the frame of Figure 4.4 This frame is not 0–generated; but it is 1–generated, for example, by the set $\{y\}$. Therefore it is a canonical frame iff it is also freely 1–generated in its own variety. Now, how does the 1–generated canonical frame look like? To that effect recall that the freely n–generated algebra in a variety $\mathcal{V}$ is a subalgebra of the direct product of the members of $\mathcal{V}$, indexed by n–tuples of elements in the algebras. In the present context, where $\mathcal{K} = \{\mathfrak{A}\}$, this reduces to the direct product $\prod_{b \in A} \mathfrak{A}$. The freely one–generated algebra is computed as the subalgebra generated by the function which picks b in the component indexed by a. (Recall that the product is indexed over A, so that each factor takes an index $b \in A$. In each factor, s takes a

Figure 4.4.



value, in this case $s(b) = b$.) The reason why this is so lies in the following. Let $\mathcal{V}$ be generated by $\mathcal{K}$. Let us call $\mathfrak{A}$ α–**free** for $\mathcal{K}$ if there is a subset $X \subseteq A$ of cardinality α such that for every $\mathfrak{B} \in \mathcal{K}$ and maps $v : X \to B$ there is a homomorphism $\overline{v} : \mathfrak{A} \to \mathfrak{B}$. The difference with the concept of a freely generated algebra is that $\mathfrak{A} \in \mathcal{K}$ is not required; moreover, $\mathfrak{A}$ need not be generated by X. But the following holds.

Proposition 4.8.9. *Let $\mathfrak{A}$ be α–free for $\mathcal{K}$. Then $\mathfrak{A}$ is α–free for* $\mathsf{HSP}(\mathcal{K})$.

Proof. Let $\mathfrak{A}$ be α–free for $\mathcal{K}$. We show that then $\mathfrak{A}$ is α–free for the classes $\mathsf{H}(\mathcal{K})$, $\mathsf{S}(\mathcal{K})$ and $\mathsf{P}(\mathcal{K})$. To simplify the argumentation, let us first remark that if $\mathfrak{A}$ is α–free for $\mathcal{K}$, then it is α–free for $\mathsf{I}(\mathcal{K})$, the closure of $\mathcal{K}$ under taking isomorphic copies.
(1.) Let $\mathfrak{C} \in \mathsf{H}(\mathcal{K})$. Then there is a $\mathfrak{B} \in \mathcal{K}$ and a homomorphism $h : \mathfrak{B} \twoheadrightarrow \mathfrak{C}$. Now take a map $m : X \to C$. There exists a map $n : X \to B$ such that $m = h \circ n$. (Just let $n(x) := a$ for some $a \in h^{-1}(m(x))$.) By assumption there exists a homomorphism $\overline{n} : \mathfrak{A} \to \mathfrak{B}$ extending n. Then $h \circ \overline{n} : \mathfrak{A} \to \mathfrak{C}$ is a homomorphism extending m.
(2.) Let $\mathfrak{C} \in \mathsf{S}(\mathcal{K})$. Then there is a $\mathfrak{B} \in \mathcal{K}$ such that $\mathfrak{C} \leq \mathfrak{B}$ and so $C \subseteq B$. Let $i : C \to B$ be the inclusion map. Let $m : X \to C$ be a map. Then $i \circ m : X \to B$ and by assumption there is an extension $\overline{i \circ m} : \mathfrak{A} \to \mathfrak{B}$. However, the image of this map is contained in C, and so restricting the target algebra to $\mathfrak{C}$ we get the desired homomorphism $\overline{m} : \mathfrak{A} \to \mathfrak{C}$.
(3.) Let $\mathfrak{C} \in \mathsf{P}(\mathcal{K})$. Then there are $\mathfrak{B}_i$, $i \in I$, such that $\mathfrak{C} \cong \prod_{i \in I} \mathfrak{B}_i$. We may assume $\mathfrak{C} = \prod_{i \in I} \mathfrak{B}_i$. Now take $m : X \to C$. Then for the projections p_i we have $p_i \circ m : X \to B_i$, and by assumption there are homomorphisms $\overline{p_i \circ m} : \mathfrak{A} \to \mathfrak{B}_i$. By the fact that the algebraic product is a product in the categorial sense there is a unique $f : \mathfrak{A} \to \prod_{i \in I} \mathfrak{B}_i$ such that $p_i \circ f = \overline{p_i \circ m}$. Then f extends m. ∎

The present algebra, the algebra of sets over the frame $\mathfrak{f}$, has eight elements. Thus, the freely one–generated algebra is a subalgebra of $\mathfrak{A}^8$. The eight choices are diagrammed in Figure 4.5 below; in each copy the set of elements which are values of p are put into a box. All we have to do is to calculate the algebra generated in this complicated frame. However, we are helped by a number of facts. First, $\mathfrak{f}$ admits an automorphism, namely $x \mapsto x$, $y \mapsto z$, $z \mapsto y$. By this automorphism, iii is mapped into iv and v into vi. All other models are mapped onto themselves. This fact has as

FIGURE 4.5.

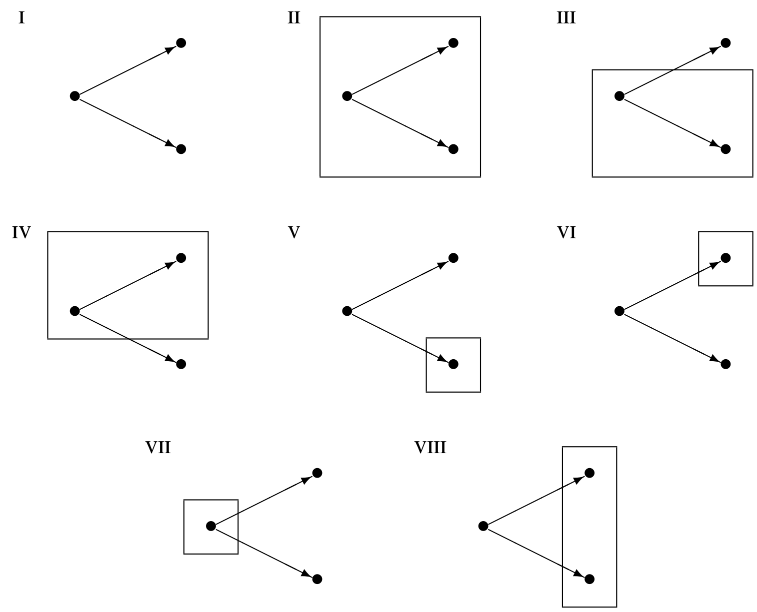
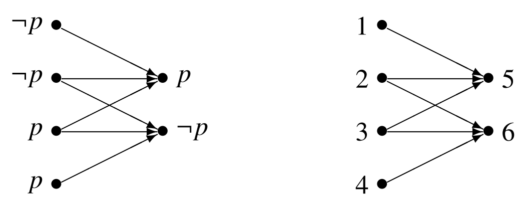


FIGURE 4.6. $\mathfrak{Can}_{\Theta}(1)$



a consequence that the algebra induced on iii and iv jointly (on the underlying frame $\mathfrak{f} \oplus \mathfrak{f}$) is isomorphic to the one iii induces on $\mathfrak{f}$ (and isomorphic to the one induced on iv on its copy of $\mathfrak{f}$). Hence, we can drop iii and v in the direct sum. Next, the frames induced on i, ii, vii and viii are not refined. i and ii are actually isomorphic to a one–element frame, vii and viii to a two–element chain. This gives a reduced representation of the underlying frame as the direct sum of two one–element chains, two two–element chains and two copies of $\mathfrak{f}$. The general frame is still not refined. Its refinement is the frame shown in Figure 4.6. (The frame is shown to the left. To the right we repeat the frame, with worlds being numbered from 1 to 6.) It might be surprising that the frame just constructed should indeed be canonical, as it contains more points. But notice that this frame is rather regular, having 8 automorphisms, and that $\mathfrak{Aut}(\mathfrak{Can}_\Theta(1)) \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$, generated by the permutations (1 4)(2)(3)(5)(6), (1)(2 3)(4)(5)(6) and (1)(2)(3)(4)(5 6). The orbits are $\{1, 4\}$, $\{2, 3\}$ and $\{5, 6\}$. There is up to automorphisms of the frame only one generating set, containing one world from each orbit of the group.

Notes on this section. Canonical frames have been heavily used for obtaining results, such as completeness results. Yet, their structure is not well–understood. Some new results can be found in the thesis by Timothy Surendonk ([**204**]). It is not known, for example, whether there exists a cardinal number α such that if a modal logic is α–canonical it is also β–canonical for any β.

Exercise 165. Prove Proposition 4.8.2.

Exercise 166. Show Theorem 4.8.7.

*****Exercise 167.** Let S be a set, and let $\mathfrak{G}$ be a group of permutations of S. We say that $\mathfrak{G}$ is **transitive** on S, if for every given pair of points $\langle x, y\rangle \in S^2$ there exists a $g \in G$ such that $g(x) = y$. We say that $\mathfrak{G}$ is **sharply transitive** if at most such g exists given $\langle x, y\rangle$. Now let $\mathfrak{Fr}_\Lambda(\lambda)$ be the algebra freely generated by λ many elements. Call a function $f : \kappa \to Fr_\Lambda(\lambda)$ an ***M*–system** if $f[\kappa]$ is a maximal independent subset of $\mathfrak{Fr}_\Lambda(\lambda)$. (A set $H \subseteq A$ is called **independent** in an algebra $\mathfrak{A}$ if for all $a \in H$, a is not contained in the subalgebra generated by $H - \{a\}$ in $\mathfrak{A}$.) Show that $\mathfrak{Aut}(\mathfrak{Fr}_\Lambda(\lambda))$ is sharply transitive on the set of M–systems of $\mathfrak{Fr}_\Lambda(\lambda)$.

Exercise 168. Let Θ be a consistent modal logic and n a natural number. Show that $\mathfrak{Aut}(\mathfrak{Fr}_\Theta(n))$ has a subgroup of size $n! \cdot 2^n$.

4.9. Algebraic Characterizations of Interpolation

Categories of frames have coproducts since they are dual to the algebras. However, to have coproducts is a rather rare property. This accounts in a way for the fact that properties such as interpolation and Halldén–completeness are rather rare. We will prove in this section two standard results on interpolation, both shown by Larisa Maksimova in a series of papers, and then derive some useful characterizations of Halldén–completeness.

Definition 4.9.1. *A variety $\mathcal{V}$ of polymodal algebras is said to have the* ***amalgamation property*** *if for any triple $\mathfrak{A}_0$, $\mathfrak{A}_1$ and $\mathfrak{A}_2$ of algebras in $\mathcal{V}$ and embeddings $i_1 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_1$ and $i_2 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_2$ there exists an algebra $\mathfrak{A}_3 \in \mathcal{V}$ and maps $e_1 : \mathfrak{A}_1 \rightarrowtail \mathfrak{A}_3$, $e_2 : \mathfrak{A}_2 \rightarrowtail \mathfrak{A}_3$ such that $e_1 \circ i_1 = e_2 \circ i_2$. $\mathcal{V}$ is said to be have the* ***superamalgamation property*** *if in addition the e_i can be required to have the property that whenever $e_1(a_1) \leq e_2(a_2)$ for $a_1 \in A_1$, $a_2 \in A_2$ there exists an $a_0 \in A_0$ such that $a_1 \leq i_1(a_0)$ and $i_2(a_0) \leq a_2$.*

$$
\begin{tikzcd}
& \mathfrak{A}_1 \arrow[dr, dotted, "e_1"] & \\
\mathfrak{A}_0 \arrow[ur, "i_1"] \arrow[dr, "i_2"'] & & \mathfrak{A}_3 \\
& \mathfrak{A}_2 \arrow[ur, dotted, "e_2"'] &
\end{tikzcd}
$$

Theorem 4.9.2 (Maksimova). *Let Λ be a polymodal logic. Then the following are equivalent.*

(1) *Λ has local interpolation.*
(2) *The variety of Λ–algebras has the superamalgamation property.*

Proof. Assume that Λ has local interpolation. Let $\mathfrak{A}_0$, $\mathfrak{A}_1$ and $\mathfrak{A}_2$ be Λ–algebras and $i_j : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_j$ ($j \in \{1, 2\}$) be embeddings. Without loss of generality we can assume that $A_0 \subseteq A_1 \cap A_2$. For each element $a \in A_1 \cup A_2$ fix a variable x_a. We assume all these variables are distinct for distinct elements. Denote by $\mathfrak{F}_i$, $i \in \{0, 1, 2\}$, the Λ–algebra freely generated by $\{x_a : a \in A_i\}$. Denote by $\mathfrak{F}_3$ the algebra generated by $\{x_a : a \in A_1 \cup A_2\}$. There are natural embeddings $\mathfrak{F}_0 \rightarrowtail \mathfrak{F}_i \rightarrowtail \mathfrak{F}_3$, $i \in \{1, 2\}$. Also there are homomorphisms $b_i : \mathfrak{F}_i \twoheadrightarrow \mathfrak{A}_i$, $i \in \{1, 2\}$ defined by $b_1 : x_a \mapsto a$ and $b_2 : x_b \mapsto b$. The maps assign the same value to each x_a where $a \in A_0$. Now let $T_1 := \{\varphi : b_1(\varphi) = 1\}$ and $T_2 := \{\varphi : b_2(\varphi) = 1\}$. Put

$$T := \{\chi : T_1 \cup T_2 \Vdash_\Lambda \chi\}.$$

Now we show that the following holds for $\{i, j\} = \{1, 2\}$

$$T \Vdash_\Lambda \varphi \to \psi \qquad \Leftrightarrow \qquad (\exists \chi \in F_0)(\varphi \to \chi \in T_i \;\&\; \chi \to \psi \in T_j)\ .$$

From right to left is clear. Now assume that $T \Vdash_\Lambda \varphi \to \psi$. Then for some finite set $\Gamma_1 \subseteq T_1$ and some finite set $\Gamma_2 \subseteq T_2$ we have $\Gamma_1; \Gamma_2 \Vdash_\Lambda \varphi \to \psi$ and so for some compound modality $\boxplus$ we get $\boxplus\Gamma_1; \boxplus\Gamma_2 \vdash_\Lambda \varphi \to \psi$. Thus

$$\vdash_\Lambda \varphi \wedge \boxplus\Gamma_1. \to . \boxplus\,\Gamma_2 \to \psi.$$

There exists now a local interpolant χ, that is, a formula using the common variables of φ and ψ and $\vdash_\Lambda (\boxplus\Gamma_1 \wedge \varphi) \to \chi$ as well as $\vdash_\Lambda \chi \to (\boxplus\Gamma_2 \to \psi)$. With this χ we have $\Gamma_1 \Vdash_\Lambda \varphi \to \chi$ and $\Gamma_2 \Vdash_\Lambda \chi \to \psi$. Thus we have $\varphi \to \chi \in T_1$ and $\chi \to \psi \in T_2$.

$$\begin{tikzcd}
& \mathfrak{F}_2 \arrow[drr] \arrow[dd] & & \\
\mathfrak{F}_0 \arrow[ur] \arrow[drr] \arrow[dd] & & & \mathfrak{F}_3 \arrow[dd] \\
& \mathfrak{A}_2 \arrow[drr] & \mathfrak{F}_1 \arrow[ur] \arrow[dd] & \\
\mathfrak{A}_0 \arrow[ur] \arrow[drr] & & & \mathfrak{A}_3 \\
& & \mathfrak{A}_1 \arrow[ur] &
\end{tikzcd}$$

Now define $\varphi\,\Theta\,\psi$ by $T \Vdash_\Lambda \varphi \leftrightarrow \psi$. This defines a congruence on $\mathfrak{F}_3$, since it defines an open filter. Now put $\mathfrak{A}_3 := \mathfrak{F}_3/\Theta$. There is a homomorphism $b_3 : \mathfrak{F}_3 \to \mathfrak{A}_3$. We will show that the filters defined by Θ on the algebras $\mathfrak{F}_1$ and $\mathfrak{F}_2$ are exactly T_1 and T_2. (See picture above.) This means that $ker(b_3) \restriction F_i = ker(b_i)$ $(i \in \{1, 2\})$. Namely, for $\varphi \in F_1$ we have $\varphi\,\Theta\,1$ iff $T \Vdash_\Lambda \top \to \varphi$. This means that for some χ in the common variables we have $\top \to \chi \in T_2$ and $\chi \to \varphi \in T_1$. Then χ is constant and since $\top \in T_1$ we also have $\top \to \chi \in T_1$ showing $\top \to \varphi \in T_1$, that is, $\varphi \in T_1$. Likewise we show that $\psi\,\Theta\,1$ implies $\psi \in T_2$ for $\psi \in F_2$. Denote by c_i the restriction of b_3 to $\mathfrak{F}_i$, $i = 1, 2$. So $c_i : \mathfrak{F}_i \to \mathfrak{A}_3$. We have shown that $ker(c_i) = ker(b_i)$. There now exist maps $e_i : \mathfrak{A}_i \rightarrowtail \mathfrak{A}_3$ $(i \in \{1, 2\})$ such that c_i factors through e_i, and $e_1 \circ b_1 = c_1$ as well as $e_2 \circ b_2 = c_2$. Finally, let $e_1(a) \leq e_2(b)$. Then $e_1(a) \to e_2(b) = 1$ and so $x_a \to x_b\,\Theta\,1$ which means that $T \Vdash_\Lambda x_a \to x_b$. From this we get a $\chi \in F_0$ such that $x_a \to \chi \in T_1$ and $\chi \to x_b \in T_2$, and so $a = b_1(x_a) \leq b_1(\chi)$ as well as $b_2(\chi) \leq b_2(x_b) = b$. Moreover, $b_1(\chi) = b_2(\chi)$, since b_1 and b_2 assign the same value in A_0 to χ, which is a common subalgebra of both $\mathfrak{A}_1$ and $\mathfrak{A}_2$. This shows that the superamalgamation property holds for $\mathcal{V}$.

Now assume that the variety of Λ–algebras has the superamalgamation property. Let $\varphi = \varphi(\vec{p}, \vec{r})$ and $\psi = \psi(\vec{r}, \vec{q})$ be formulae such that for no χ based on the variables $\vec{r}$ we have $\vdash_\Lambda \varphi \to \chi; \chi \to \psi$. Let $\mathfrak{F}_0$ be the algebra freely generated by $\vec{r}$, $\mathfrak{F}_1$ the algebra freely generated by $\vec{p}$ and $\vec{r}$, $\mathfrak{F}_2$ the algebra freely generated by $\vec{q}$ and $\vec{r}$ and $\mathfrak{F}_3$ the algebra freely generated by $\vec{p}$, $\vec{q}$ and $\vec{r}$. Let U_1 be an ultrafilter on $\mathfrak{F}_3$ containing φ, and let V be the ultrafilter induced by U_1 on the subalgebra $\mathfrak{F}_0$. Then $V \cup \{\neg\psi\}$ has the finite intersection property in $\mathfrak{F}_3$ and so there exists an ultrafilter U_2 containing it. We then have $U_2 \cap F_0 = U_1 \cap F_0$. Let Θ_1 be the largest congruence contained in U_1, and Θ_2 be the largest congruence contained in U_2. Put $\Phi := \Theta_1 \sqcup \Theta_2$. (The largest congruence induced by a filter F is the same as the congruence induced by the largest open filter contained in F; and this is equal to the set of elements a such that $\boxplus a \in F$ for all compound modalities $\boxplus$.) For elements of F_0, $u\,\Theta_1\,v$ iff for all compound modalities $\boxplus(u \leftrightarrow v) \in U_1$ iff for all compound modalities $\boxplus(u \leftrightarrow v) \in V$ iff for all compound modalities $\boxplus(u \leftrightarrow v) \in U_2$ iff $u\,\Theta_2\,v$. Put $\Theta_0 := \Phi \cap (F_0 \times F_0)$ and $\mathfrak{A}_0 := \mathfrak{F}_0/\Theta_0$. Then we get an embedding $i_1 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_1$

as well as $i_2 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_2$. Assume now that we have an element $\chi \in F_0$ such that $\varphi \to \chi \,\Theta_1\, 1$ and $\chi \to \psi \,\Theta_2\, 1$. Then $\chi \in U_1$ since $\varphi \in U_1$, and then also $\psi \in U_2$. But this contradicts our choice of U_1 and U_2. Hence no such element exists. By superamalgamation, however, we get an algebra $\mathfrak{B}$ and maps $e_i : \mathfrak{A}_i \rightarrowtail \mathfrak{B}$ such that $e_1 \circ i_1 = e_2 \circ i_2$ and $e_1([\varphi]\Theta_1) \nleq e_2([\psi]\Theta_2)$. Now define a map $b : \mathfrak{F}_3 \to \mathfrak{B}$ by $b(p_i) := e_1([p_i]\Theta_1)$, $b(q_j) := e_2([q_j]\Theta_2)$ as well as $b(r_k) = \epsilon_1([r_k]\Theta_1) = e_2([r_k]\Theta_2)$. This is uniquely defined and we have $b(\varphi) \nleq b(\psi)$ from which $\nvdash_\Lambda \varphi \to \psi$. ◻

Theorem 4.9.3 (Maksimova). *Let Λ be a polymodal logic. Then the following are equivalent.*

(1) *Λ has global interpolation.*
(2) *The variety of Λ–algebras has the amalgamation property.*

Proof. The proof of amalgamation from global interpolation is actually analogous to the previous one. So let us prove that amalgamability implies global interpolation. We assume that $\varphi \Vdash_\Lambda \psi$ but no interpolant exists. Define $\mathfrak{F}_0$ to be algebra freely generated by the common variables of φ and ψ, and $\mathfrak{F}_3$ the algebra generated by all the variables of φ and ψ. Let O_1 be the open filter generated by φ, and O_2 an open filter containing $\neg\psi$ and $O_1 \cap F_0$. Such a filter exists. Then $O_1 \cap F_0 = O_2 \cap F_0$. Let Θ_i be the congruence associated with O_i. Now put $\mathfrak{A}_1 := \mathfrak{F}_3/\Theta_1$ and $\mathfrak{A}_2 := \mathfrak{F}_3/\Theta_2$. Then as before for elements of F_0, $u \,\Theta_1\, v$ iff $u \,\Theta_2\, v$, and hence Θ_1 and Θ_2 induce the same congruence on $\mathfrak{F}_0$. Therefore, we have an embedding $i_1 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_1$ and $i_2 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_2$. Thus, assuming the amalgamation property, there is an algebra $\mathfrak{B}$ and morphisms e_i, $i = 1, 2$, satisfying $e_1 \circ i_1 = e_2 \circ i_2$. Define v by $v(p) := e_1([p]\Theta_1)$ if $p \in var(\varphi)$ and $v(p) := e_2([p]\Theta_2)$ if $p \in var(\psi)$. This is noncontradictory. Since we have $\varphi \in O_1$ we also have $\overline{v}(\boxplus\varphi) = 1$ for all compound modalities. Since $\neg\psi \in O_2$ we have $\overline{v}(\psi) \neq 1$, and so $\varphi \nVdash_\Lambda \psi$. ◻

The results on amalgamation property can be improved by showing that $\mathfrak{A}_3$ enjoys a so–called universal property. This means that given $\mathfrak{A}_0$, $\mathfrak{A}_1$ and $\mathfrak{A}_2$, embeddings $i_j : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_i$ ($j \in \{1, 2\}$) there exists an $\mathfrak{A}_3$ and maps $e_j : \mathfrak{A}_j \to \mathfrak{A}_3$ such that $e_1 \circ i_1 = e_2 \circ i_2$ and for every algebra $\mathfrak{B}$ together with maps $d_1 : \mathfrak{A}_1 \to \mathfrak{B}$, $d_2 : \mathfrak{A}_2 \to \mathfrak{B}$ with $d_1 \circ i_1 = d_2 \circ i_2$, then there exists a unique homomorphism $h : \mathfrak{A}_3 \to \mathfrak{B}$ such that $d_1 = h \circ e_1$ and $d_2 = h \circ e_2$. Namely, consider the maps $v_1 : x_a \mapsto d_1 \circ b_1(x_a) = d_1(a)$, $v_2 : x_a \mapsto d_2 \circ b_2(x_a) = d_2(a)$. Then v_1 and v_2 agree on the elements in A_0, and so $v := v_1 \cup v_2$ is well–defined. It extends to a unique homomorphism $\overline{v} : \mathfrak{F}_3 \to \mathfrak{B}$. We have that $\Theta_1 := ker(\overline{v}) \restriction \mathfrak{F}_1 = ker(d_1 \circ b_1)$ and $\Theta_2 := ker(\overline{v}) \restriction \mathfrak{F}_2 = ker(d_2 \circ b_2)$. Put $\Theta := \Theta_1 \sqcup \Theta_2$. As in the previous proof it is shown that $\Theta \restriction F_1 = \Theta_1$ and $\Theta \restriction F_2 = \Theta_2$. Moreover, Θ is the kernel of the map $\overline{v}$ and includes $ker(b_3)$. Thus it can be factored uniquely through b_3, yielding a homomorphism $h : \mathfrak{A}_3 \to \mathfrak{B}$ with the desired properties. In the sense of the definitions in the exercises of Chapter 4.7 what we have shown is that a variety of modal algebras that has amalgamation also has pushouts for those diagrams in which both arrows are injective.

Now we turn to local and global Halldén–completeness. Recall that if a logic is locally or globally Halldén–complete it has trivial constants; another way of saying this is that the freely zero–generated algebra contains exactly two elements (if Λ is consistent). Then for every pair of nontrivial algebras $\mathfrak{A}_1$ and $\mathfrak{A}_2$ we can find an $\mathfrak{A}_0$ and injections $i_0 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_1$ and $i_1 : \mathfrak{A}_0 \rightarrowtail \mathfrak{A}_2$. Simply take the zero–generated subalgebra; if the algebras have more than one element, this algebra is isomorphic to $\mathfrak{Fr}_\Lambda(0)$.

Definition 4.9.4. *A variety $\mathcal{V}$ has **fusion** if for every pair $\mathfrak{A}_1, \mathfrak{A}_2 \in \mathcal{V}$ of nontrivial algebras there exists an algebra $\mathfrak{B}$ and embeddings $e_1 : \mathfrak{A}_1 \rightarrowtail \mathfrak{B}$ and $e_2 : \mathfrak{A}_2 \rightarrowtail \mathfrak{B}$. $\mathcal{V}$ has **superfusion** if for every pair $\mathfrak{A}_1, \mathfrak{A}_2 \in \mathcal{V}$ of nontrivial algebras there exists an algebra $\mathfrak{B}$ and embeddings $e_i : \mathfrak{A}_i \rightarrowtail \mathfrak{B}$ such that for every $a \in A_1 - \{0\}$ and every $b \in A_2 - \{1\}$ the inequation $i_1(a) \leq i_2(b)$ does not hold.*

The following theorem can be found in a slightly different form in [**153**].

Theorem 4.9.5 (Maksimova). *Let Λ be a polymodal logic. Then the following are equivalent*

(1) *Λ is locally Halldén–complete.*
(2) *The variety of Λ–algebras has superfusion, and the zero–generated algebra contains at most two elements.*

The global version is as follows:

Theorem 4.9.6 (Maksimova). *Let Λ be a polymodal logic. Then the following are equivalent*

(1) *Λ is globally Halldén–complete.*
(2) *The variety of Λ–algebras has fusion, and the zero–generated algebra contains at most two elements.*

For a proof, let Λ be (locally/globally) Halldén–complete. We may assume that Λ is consistent; otherwise the equivalence is clearly valid. Take two algebras $\mathfrak{A}_1$ and $\mathfrak{A}_2$. We can embed $\mathfrak{Fr}_\Lambda(0)$ into both of these algebras. Now follow the proofs of the Theorems 4.9.2 and 4.9.3. We obtain an algebra $\mathfrak{B}$ and embeddings $e_i : \mathfrak{A}_i \rightarrowtail \mathfrak{B}$. The superfusion condition is verified for the local case. For the converse, let the variety of Λ–algebras have (super)fusions, and let the zero–generated algebra have two elements. Then enter the second half of the proof of Theorem 4.9.2 with φ and ψ assuming that for no constant proposition both $\varphi \vdash_\Lambda \chi$ and $\chi \vdash_\Lambda \psi$. But either $\chi = \top$ and then $\nvdash_\Lambda \psi$, or $\chi = \bot$ and then $\varphi \nvdash_\Lambda \bot$. Performing the same argument as in the proof we get that $\varphi \nvdash_\Lambda \psi$. As a consequence we get the following theorem of [**15**].

Theorem 4.9.7 (van Benthem & Humberstone). *Let Λ be a logic such that for every pair $\langle \mathfrak{F}_1, x_1 \rangle$ and $\langle \mathfrak{F}_2, x_2 \rangle$ of pointed frames there exists a pointed frame $\langle \mathfrak{G}, y \rangle$ and two contractions $p_1 : \mathfrak{G} \twoheadrightarrow \mathfrak{F}_1$ $p_2 : \mathfrak{G} \twoheadrightarrow \mathfrak{F}_2$ such that $p_1(x_1) = p_2(x_2) = y$. Then Λ is locally Halldén–complete.*

Definition 4.9.8. *A variety $\mathcal{V}$ has* ***finite coproducts*** *if for every pair $\mathfrak{A}_1$, $\mathfrak{A}_2$ of algebras in $\mathcal{V}$ there exists a third algebra $\mathfrak{B}$ and maps $i_1 : \mathfrak{A}_1 \to \mathfrak{B}$, $i_2 : \mathfrak{A}_2 \to \mathfrak{B}$ such that for every algebra $\mathfrak{C}$ and every pair of maps $e_1 : \mathfrak{A}_1 \to \mathfrak{C}$ and $e_2 : \mathfrak{A}_2 \to \mathfrak{C}$ a unique $h : \mathfrak{B} \to \mathfrak{C}$ exists satisfying $e_1 = h \circ i_1$ and $e_2 = h \circ i_2$. We denote $\mathfrak{B}$ by* $\mathfrak{A}_1 \oplus \mathfrak{A}_2$.

Theorem 4.9.9. *A logic is globally Halldén–complete iff it has trivial constants and the corresponding variety has finite coproducts.*

Exercise 169. Give a characterization of those logics whose variety has coproducts, without any restriction on constant formulae.

Exercise 170. Let Λ_i, $i < \omega$, be logics which have (local/global) interpolation. Suppose that $\Lambda_i \subseteq \Lambda_j$ if $i \leq j$. Show that $\bigsqcup_{i\in\omega}\Lambda_i$ has (local/global) interpolation.

Exercise 171. Give an example of logics Λ_i which have local interpolation, such that $\Lambda_i \geq \Lambda_j$ for $i \leq j$, such that $\bigcap_{i\in\omega} \Lambda_i$ fails to have interpolation.

CHAPTER 5

Definability and Correspondence

5.1. Motivation

Correspondence theory developed from a number of insights about the possibility of defining certain elementary properties of frames via modal axioms. For example, transitivity of Kripke–frames may either be characterized by the first–order formula $(\forall xyz)(x \lhd y \lhd z. \rightarrow .x \lhd z)$ or by the modal axiom $\Diamond\Diamond p \rightarrow \Diamond p$. We therefore say that the axiom **4** *corresponds* to transitivity on Kripke–frames. These insights have sparked off the search for the exact limit of this correspondence. In particular, the following two questions have been raised by JOHAN VAN BENTHEM in [**10**], who has also done much to give complete answers to them.

- $*$ Which elementary properties of Kripke–frames can be characterized by modal axioms?
- $*$ Which modal axioms determine an elementary property on Kripke–frames?

Both questions were known to have nontrivial answers. Irreflexivity cannot be characterized modally, so not all first–order properties are modally characterizable. On the other hand, some modal axioms like the **G**–axiom determine a non–elementary property of frames. Many people have contributed to the area of correspondence theory, which is perhaps the best worked out subtheory of modal logic, e. g. JOHAN VAN BENTHEM [**8**], ROBERT GOLDBLATT [**77**], [**79**]. With HENDRIK SAHLQVIST's classical paper [**183**] the theory reached a certain climax. There have been attempts to strengthen this theorem, but without success. It still stands out as *the* result in correspondence theory. Nevertheless, there has been a lot of improvement in understanding it. The original proof was rather arcane and methods have been found to prove it in a more systematical way (see VAN BENTHEM [**10**], SAMBIN and VACCARO [**187**] and also KRACHT [**121**] and [**124**]).

Meanwhile, the direction of the research has also changed somewhat. Correspondence theory as defined above is just part of a general discipline which has emerged lately, namely *definability theory*. Definability theory is the abstract study of definability of sets and relations in general frames. There are a number of reasons to shift to this more abstract investigation. First, as has been observed already in SAMBIN and VACCARO [**187**], there is a real benefit in raising the above questions not for Kripke–frames but for frames in general and suitable classes thereof. For suppose that (as in SAHLQVIST's original proof) correspondence of modal axioms of a certain form with first–order conditions is established not only for Kripke–frames but also for descriptive frames. Then one can immediately derive that logics axiomatized by such formulae must be complete. First–orderness on Kripke–frames is not enough for that. Second, the success of *canonical formulae* of Michael Zakharyaschev for logics containing **K4** (see Chapter 8) shows that there can be useful geometric characterizations of axioms which are not first–order in general. Even though first–order properties are well–understood and there is a powerful machinery for first–order logic, there is nevertheless much to be said in favour of an attempt to characterize in *whatever terms possible* the geometric condition imposed on frames by an axiom. They cannot all be first–order as we know, but they may in many cases be simple, as the notorious example of **G** shows. The third generalization concerns the use of *relations* rather than *properties*. That is, we ask which relations are characterizable modally in a given class of frames. This move has many advantages, although we need to clarify what we mean by a modal characterization of relations since modal formulae are invariably properties of worlds rather than sequences of worlds. Nevertheless, we will develop such a theory here and it will be possible to say for example that a frame is differentiated iff equality is modally definable. In this way natural classes of frames are motivated by the possibility to define certain relations.

Definability theory is thus the study of the following questions.

* Given a class $\mathfrak{X}$ of frames, which relations between worlds are characterizable modally in $\mathfrak{X}$?
* Given a class $\mathfrak{X}$ of frames, what geometric properties of frames do modal axioms impose on the frames of $\mathfrak{X}$?

From there many questions arise which have been treated in the literature with sometimes surprising answers. These questions are for example the following.

* Which classes of frames can be characterized by modal axioms?
* What is the relation between closure properties of classes and the syntactic form of definable properties or relations?

5.2. The Languages of Description

Before we can enter the discussion on definability we will fix a suitable language within which we derive formal results. Such a language is traditionally seen to be *monadic second–order logic*. Here, however, we will use a notational variant, namely *two–sorted first–order predicate logic*. The special language will be called the *external language* and denoted by $\mathcal{L}^e$. It is *two–sorted*; that means, there are two sorts of well–formed expressions, namely *modal propositions* (called *i*–formulae) and *formulae*. Moreover, $\mathcal{L}^e$ has two sublanguages, $\mathcal{L}^m$, the *modal language*, for talking about modal propositions and $\mathcal{L}^f$, the *frame language*, for talking about worlds. $\mathcal{L}^m$ is in fact isomorphic to the basic modal language we are operating in. Remember that there is not just a single modal language, but a whole family of them with varying number of modal operators and varying number of variables and constants. For the moment we assume that we have countably many propositional variables and no constants, though nothing depends on that. Of proposition variables and constants there will otherwise be as many as needed. Similarly, depending on the cardinality κ of basic modal operators $\mathcal{L}^m$ will have the following symbols

* proposition variables $p\text{-}var = \{p_i : i \in \omega\}$,
* boolean connectives $\top$, $\neg$, $\wedge$,
* modal connectives $\Box_j$, $j < \kappa$.

All symbols are standard, and their interpretation will be as well. Notice that $\Diamond_j$ is not a primitive symbol. This is done to make the subsequent discussion simpler. Nothing hinges essentially on that. Now the frame language $\mathcal{L}^f$ has

* world–variables $w\text{-}var := \{w_i : i \in \omega\}$,
* equality $\doteq$, and relations $\lhd_i$ for $i < \kappa$,
* logical junctors t, $\neg$, $\wedge$ and $\rightarrow$,
* quantifiers $\forall$ and $\exists$.

(Here t is a constant, which receives the value *true*.) Finally, $\mathcal{L}^e$ has two more ingredients, namely

* a membership predicate ϵ,
* proposition–quantifiers $\forall$ and $\exists$.

For a rigorous definition of formulae we need two sorts, *propositions* and *worlds*, over which the two sorts of quantifiers may range. Moreover, we define two sorts of well–formed expressions, **internal formulae (i–formulae)** and **external formulae (e–fomulae)**. They are composed as follows

(1) p_i, $i < \omega$, $\top$, $\bot$ are i–formulae.
(2) If φ and ψ are i–formulae, so are $\neg\varphi$, $\varphi \wedge \psi$, $\Box_j\varphi$ ($j < \kappa$).
(3) If φ is an i–formula and $i < \omega$ then $w_i \in \varphi$ is an e–formula.
(4) t is an e–formula.
(5) $w_i \doteq w_j$ and $w_i \lhd_k w_j$ are e–formulae for all $i, j < \omega$, $k < \kappa$.
(6) If ζ and η are e–formulae then so are $\neg\zeta$, $\zeta \wedge \eta$ and $\zeta \rightarrow \eta$.
(7) If ζ is an e–formula and $i < \omega$ then $(\forall w_i)\zeta$ and $(\exists w_i)\zeta$ are e–formulae.
(8) If ζ is an e–formula and $i < \omega$ then $(\forall p_i)\zeta$ and $(\exists p_i)\zeta$ are e–formulae.

An **e–model** is a triple $\langle \mathfrak{F}, \beta, \iota \rangle$ with $\mathfrak{F}$ a frame, and with $\beta : p\text{-}var \rightarrow \mathbb{G}$ and $\iota : w\text{-}var \rightarrow g$. Given an e–model $\langle \mathfrak{G}, \beta, \iota \rangle$ and an e–formula ζ, the relation $\langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta$ is defined inductively as follows. (Here, given two functions f and g, $f \sim_a g$

abbreviates the fact that $f(x) \neq g(x)$ only if $x = a$.)

$$\begin{array}{lcll}
\langle \mathfrak{G}, \beta, \iota \rangle \vDash w_i \epsilon \varphi & \Leftrightarrow & \iota(w_i) \in \overline{\beta}(\varphi) & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash w_i \doteq w_j & \Leftrightarrow & \iota(w_i) = \iota(w_j) & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash w_i \lhd_k w_j & \Leftrightarrow & \iota(w_i) \lhd_k \iota(w_j) & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash \neg \zeta & \Leftrightarrow & \langle \mathfrak{G}, \beta, \iota \rangle \nvDash \zeta & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta \wedge \eta & \Leftrightarrow & \langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta \text{ and } \langle \mathfrak{G}, \beta, \iota \rangle \vDash \eta & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta \rightarrow \eta & \Leftrightarrow & \text{from } \langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta \text{ follows } \langle \mathfrak{G}, \beta, \iota \rangle \vDash \eta & \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash (\forall w_i)\zeta & \Leftrightarrow & \text{for all } \iota' \sim_{w_i} \iota & \langle \mathfrak{G}, \beta, \iota' \rangle \vDash \zeta \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash (\exists w_i)\zeta & \Leftrightarrow & \text{for some } \iota' \sim_{w_i} \iota & \langle \mathfrak{G}, \beta, \iota' \rangle \vDash \zeta \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash (\forall p_i)\zeta & \Leftrightarrow & \text{for all } \beta' \sim_{p_i} \beta & \langle \mathfrak{G}, \beta', \iota \rangle \vDash \zeta \\
\langle \mathfrak{G}, \beta, \iota \rangle \vDash (\exists p_i)\zeta & \Leftrightarrow & \text{for some } \beta' \sim_{p_i} \beta & \langle \mathfrak{G}, \beta', \iota \rangle \vDash \zeta
\end{array}$$

Further, $\mathfrak{G} \vDash \zeta$ iff for all β and all ι we have $\langle \mathfrak{G}, \beta, \iota \rangle \vDash \zeta$. Notice that we have used $\overline{\beta}(\varphi)$ in the first clause. This is strictly speaking yet to be defined. However we assume that $\overline{\beta}(\varphi)$ is computed as before by induction on φ. Notice that in addition to equality there is one symbol whose interpretation is rather special, namely ϵ. It must always be interpreted as membership. The following sentences are theorems of the $\mathcal{L}^e$–logic of generalized frames. They show that the $\mathcal{L}^i$–connectives are in principle dispensable. (Here $\zeta \equiv \eta$ abbreviates $\zeta \rightarrow \eta. \wedge .\eta \rightarrow \zeta$. Open formulae are as usual treated as if all free variables were universally quantified.)

$$\begin{array}{lcl}
w_i \epsilon \neg\varphi & \equiv & \neg(w_i \epsilon \varphi) \\
w_i \epsilon \varphi \wedge \psi & \equiv & w_i \epsilon \varphi \ \wedge \ w_i \epsilon \psi \\
w_i \epsilon \varphi \rightarrow \psi & \equiv & w_i \epsilon \varphi. \rightarrow .w_i \epsilon \psi \\
w_i \epsilon \Box_j \varphi & \equiv & (\forall w_k)(w_i \lhd_j w_k. \rightarrow .w_k \epsilon \varphi)
\end{array}$$

$\mathcal{L}^e$ is not interesting for us because it defines a logic of structures, but because it is a rather strong language within which we can express (almost) everything we wish to say. For example, it contains both first–order properties for frames and properties expressed by modal axioms. With respect to the latter only the notation has changed somewhat. We can no longer write $\mathfrak{F} \vDash \varphi$ but must instead write

$$\mathfrak{F} \vDash (\forall w_0)(w_0 \epsilon \varphi).$$

Also, the so–called **standard translation** of a modal formula is defined as follows.

$$\begin{array}{lcl}
ST(p, x) & := & x \epsilon p \\
ST(\neg\varphi, x) & := & \neg ST(\varphi, x) \\
ST(\varphi \wedge \psi, x) & := & ST(\varphi, x) \ \wedge \ ST(\psi, x) \\
ST(\Box_j \varphi, x) & := & (\forall y)(x \lhd_j y. \rightarrow .ST(\varphi, y))
\end{array}$$

(In the last clause, y must be a variable not already occurring in $ST(\varphi, x)$. By construction, x is always the unique free variable. Clearly, $ST(\varphi, y)$ is then the same as $ST(\varphi, x)[y/x]$.) Obviously we have in all frames

$$(\forall x)(x \epsilon \varphi \equiv ST(\varphi, x)) .$$

To conclude this part we will introduce the language $\mathcal{R}^e$. In this language the quantifiers over worlds are replaced by so–called *restricted quantifiers*. They are defined as follows.

$$\begin{array}{lcl}(\forall w_j \rhd_k w_i)\zeta & := & (\forall w_j)(w_i \lhd_k w_j. \rightarrow .\zeta) \\ (\exists w_j \rhd_k w_i)\zeta & := & (\exists w_j)(w_i \lhd_k w_j. \wedge .\zeta)\end{array}$$

The restricted quantifiers can be defined from the unrestricted quantifiers as shown, but the converse does not hold in absence of weak transitivity. Syntactically, we want to construe the restricted quantifier as follows. It takes an e–formula ζ and two variables w_i, w_j and returns an e–formula. Hence, for each $i < \kappa$ there is a distinct restricted quantifier. Notice that this quantifier is said to *bind* only w_j, and that w_i is called a **restrictor**. Let $\mathcal{R}^e$ denote the language $\mathcal{L}^e$ with restricted world quantifiers instead of unrestricted ones. Likewise, $\mathcal{R}^f$ denotes the language obtained from $\mathcal{L}^f$ by replacing the unrestricted quantifiers by restricted quantifiers. $\mathcal{R}^e$ can be construed as a sublanguage of $\mathcal{L}^e$, and $\mathcal{R}^f$ as a sublanguage of $\mathcal{L}^f$. The restricted languages are expressively weaker, but the difference turns out to be inessential in the connection with modal logic. On the other hand, $\mathcal{R}^e$ and its frame counterpart $\mathcal{R}^f$ have several advantages, as will be seen shortly. With the restricted quantifiers we will define the following shorthand notation, corresponding to compound modalities. Let σ range over sequences of numbers $< \kappa$, and s, t over sets of such sequences. (If $s = \{\sigma\}$, we omit the brackets.)

$$\begin{array}{lcl}(\forall x \rhd^{\epsilon} w)\zeta & := & \zeta[w/x] \\ (\exists x \rhd^{\epsilon} w)\zeta & := & \zeta[w/x] \\ (\forall x \rhd^{\sigma^\frown i} w)\zeta & := & (\forall y \rhd^{i} w)(\forall x \rhd^{\sigma} y)\zeta \\ (\forall x \rhd^{s \cup t} w)\zeta & := & (\forall x \rhd^{s} w)\zeta \ \wedge \ (\forall x \rhd^{t} w)\zeta \\ (\exists x \rhd^{\sigma^\frown i} w)\zeta & := & (\exists y \rhd^{i} w)(\exists x \rhd^{\sigma} y)\zeta \\ (\exists x \rhd^{s \cup t} w)\zeta & := & (\exists x \rhd^{s} w)\zeta \vee (\exists x \rhd^{t} w)\zeta\end{array}$$

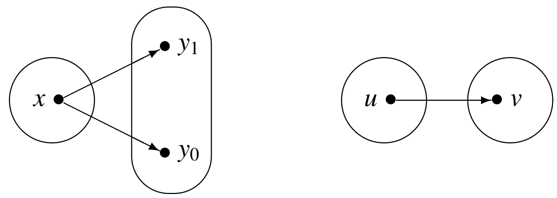
Here, it is assumed that y is not free in ζ. Similarly the shorthand notation $x \lhd^s y$ is defined. It corresponds to the compound modality $\Box^s$. Sets of the form $\{x : w \lhd^s x\}$ are called **cones**. The restricted quantifiers range over cones. The last language introduced is $\mathcal{S}^f$. It has in addition to the symbols of $\mathcal{R}^f$ all $\lhd^s$ as primitive symbols, where s is a finite union of finite sequences over κ. Also, it has the following additional axioms

$$\begin{array}{lcl}x \lhd^{j^\frown\sigma} y & \equiv & (\exists z \rhd_j x)(z \lhd^{\sigma} y) \\ x \lhd^{\langle\rangle} y & \equiv & x \doteq y \\ x \lhd^{\varnothing} y & \equiv & \neg(x \doteq x) \\ x \lhd^{s \cup t} y & \equiv & x \lhd^{s} y. \ \vee \ .x \lhd^{t} y\end{array}$$

Exercise 172. Show that no $\mathcal{R}^e$ and $\mathcal{R}^f$ formula can be a sentence, i. e. have no free w–variables.

Exercise 173. Show that $\mathcal{R}^f$ has nontrivial constant formulae, in contrast to $\mathcal{L}^f$.

FIGURE 5.1. $\mathfrak{F}$ (left) and $\mathfrak{G}$ (right)



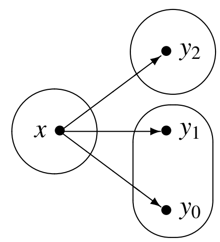
5.3. Frame Correspondence — An Example

Before we start with the general theory we will look at an instructive example. Let there be just one operator, for simplicity. Consider the axiom $\mathbf{alt}_1 := \Diamond p \wedge \Diamond q. \rightarrow .\Diamond(p \wedge q)$. It is quite easy to show that a Kripke–frame satisfies this axiom iff the frame is quasi–functional, that is, satisfies the first–order axiom $(\forall x)(\forall y_0 \rhd x)(\forall y_1 \rhd x)(y_0 \doteq y_1)$. Namely, if the frame $\mathfrak{f} = \langle f, \lhd \rangle$ is quasi–functional, and $\langle \mathfrak{f}, \beta, x \rangle \vDash \Diamond p \wedge \Diamond q$, then there is a successor $y_0 \vDash p$ and a successor $y_1 \vDash q$. But $y_0 = y_1$, and so $y_0 \vDash p \wedge q$ from which we get $x \vDash \Diamond(p \wedge q)$. On the other hand, if $\mathfrak{f}$ is *not* quasi–functional there is an $x \in f$ and $x \lhd y_0, y_1$ for distinct y_0, y_1. Put $\beta(p) := \{y_0\}$ and $\beta(q) := \{y_1\}$. Then we have $y_0 \vDash p; \neg q$, $y_1 \vDash q; \neg p$ and for all other points z we have $z \vDash \neg p; \neg q$. Whence $x \vDash \Diamond p; \Diamond q; \neg\Diamond(p \wedge q)$.

Now lets suppose we have a general frame $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$. Does it still hold that it satisfies $\mathbf{alt}_1$ iff it is quasi–functional? Well, one direction is uncomplicated. If the underlying Kripke–frame is quasi–functional then $\mathfrak{F} \vDash \mathbf{alt}_1$. Just copy the proof given above. However, in the converse direction we encounter problems. When x has two different successors we took two special sets for $\beta(p)$ and $\beta(q)$ and there is no guarantee that we may still be able to do so. In fact, the following frame shows that the converse direction is false in general. Namely, let $\mathfrak{f}$ be based on three points x, y_0, y_1 with $x \lhd y_0, y_1$. Define $\mathbb{F} := \{\varnothing, \{x\}, \{y_0, y_1\}, \{x, y_0, y_1\}\}$. Then $\mathfrak{F} := \langle \mathfrak{f}, \mathbb{F} \rangle$ is a frame since $\mathbb{F}$ is closed under boolean operations and under $\blacklozenge$, as is quickly computed. (See Figure 5.1.) The valuation that we used to show that $\mathbf{alt}_1$ can be refuted is now no longer available. Moreover, the map $p : x \mapsto u, y_0 \mapsto v, y_1 \mapsto v$ is a p–morphism onto the quasi–functional frame $\mathfrak{g}$. Moreover, $\mathbb{F}$ is the p–preimage of the powerset of $\{u, v\}$. Thus, we actually have $\mathfrak{F} \vDash \mathbf{alt}_1$.

We see that the correspondence between first–order properties and modal properties may break down if we pass to a larger class of frames. This is an important point. In most of the literature on correspondence theory only the problem of correspondence with respect to Kripke–frames is discussed. Although this is by itself a legitimate problem, it is for reasons discussed earlier advisable to broaden the class of frames to be looked at. We can also put the question on its head and ask how large the class is for which the correspondence between $\mathbf{alt}_1$ and quasi–functionality can be shown. The way to approach that question is to look for sets which can serve as

FIGURE 5.2.



values for valuations falsifying a given formula once the corresponding first–order condition is not met. For example, if we have a violation of quasi–functionality because $x \lhd y_0, y_1$ for $y_0 \neq y_1$ we want to be able to produce a valuation β such that $\langle \mathfrak{F}, \beta, x\rangle \nvDash \mathbf{alt}_1$. Above we have chosen $\beta(p) = \{y_0\}$ and $\beta(q) = \{y_1\}$, so we conclude that in *atomic frames* the correspondence will still hold. But this is not such a good result. Consider the frame $\mathfrak{H} = \langle \mathfrak{h}, \mathbb{H}\rangle$ where $h = \{x\} \cup \{y_i : i \in \omega\}$ and $x \lhd y_i$, but no other relations hold. Finally, $\mathbb{H}$ is the boolean algebra generated by the sets of the form $r(i,k) := \{y_n : (\exists \ell \in \omega)(n = i \cdot \ell + k)\}$, where $k < i$. Thus $\mathbb{H}$ consists of finite unions of such sets possibly with $\{x\}$. $\mathbb{H}$ as defined is closed under complements and unions, as an analogous frame constructed in Section 4.6. The point about this frame is that it is not atomic but nevertheless the correspondence holds. For pick any y_i and y_j with $i \neq j$. What we need is sets a_0, a_1 such that $y_i \in a_0$, $y_j \in a_1$, and $y_k \notin a_0 \cap a_1$ for all k. Assume $i < j$. Then put $a_1 := r(j,0) - r(j,i)$ and $a_0 := r(j,i)$. We have $a_1 \cap a_0 = \varnothing$ by construction and $y_i \in a_0$, $y_j \in a_1$. So, such sets exist for all choices for offending triples x, y_0, y_1.

In general it is sufficient that $\mathfrak{F}$ be differentiated. For let us assume that x sees two distinct points y_0 and y_1. Then differentiatedness guarantees the existence of a set a such that $y_0 \in a$ but $y_1 \notin a$. Putting $\beta(p) := a$, $\beta(q) := -a$ we get the desired valuation proving $\mathfrak{F} \nvDash \mathbf{alt}_1$.

PROPOSITION 5.3.1. *A differentiated frame is a frame for* $\boldsymbol{K.alt}_1$ *iff the underlying Kripke–frame is quasi–functional.*

Now consider the frame $\mathfrak{E}$ in Figure 5.2. This frame is not quasi–functional, it is not differentiated, but it also does not satisfy $\mathbf{alt}_1$. Thus, the result above is still not optimal. The reason for this failure is easy to spot. On the one hand we have $x \lhd y_1$ and $x \lhd y_2$ and there is a set a such that $y_1 \in a$ but $y_2 \notin a$. This alone suffices to establish the equivalence between failing $\mathbf{alt}_1$ and not being quasi–functional. However, on the other hand we also have $x \lhd y_0$ and $x \lhd y_1$ and y_0 and y_1 are not separable by a set. So, what we have shown in the cases above is that no matter what triple of points x, y, z we choose such that $x \lhd y, z$ there always is a set a such that $y \in a$ but $z \notin a$.

The frame here does not have this property. To distinguish these two properties let us say that quasi–functionality corresponds **casewise** to $\mathbf{alt}_1$ if for every choice of points $x \lhd y, z$ such that $y \neq z$ there is a set a with $y \in a$ but $z \notin a$. And let us say that quasi–functionality corresponds *simply* if from the failure of quasi–functionality somewhere we can deduce the existence of triple $x \lhd y, z$ and a set a such that $y \in a$ but $z \notin a$. Simple correspondence is clearly weaker.

Proposition 5.3.2. *In the class of differentiated frames the property of being quasi–functional corresponds casewise to the property of being a frame for* $\boldsymbol{K.alt}_1$.

Casewise correspondence is not actually the same as local correspondence, the notion we are ultimately interested in. Local correspondence is defined only with reference to the root x. Namely, quasi–functionality locally corresponds to $\mathbf{alt}_1$ in a frame $\mathfrak{G}$ if

$$\mathfrak{G} \vDash (\forall u)[(\forall y_0 \rhd u)(\forall y_1 \rhd u)(y_0 \doteq y_1) \quad \equiv \quad (\forall p)(\forall q)(u \epsilon \Diamond p \wedge \Diamond q \to \Diamond(p \wedge q))]$$

The frame $\mathfrak{C}$ above satisfies this correspondence as well, showing that the casewise correspondence as just defined is really weaker. Nevertheless, it seems a rather artificial concept to begin with. The problem is, however, that the schema above can be written down rather nicely. On the left hand side we have a first–order formula satisfied at u, and on the right hand side we have a modal formula satisfied at u. Notice that the $\vDash$ is ambigous here. On the left we would have to construe the statement as being expressed in $\mathcal{R}^f$, the frame part of the external language, whereas on the right hand side we have the ordinary $\vDash$ from the internal language modal logic, and not from the modal fragment of the external language. It is now in principle possible to rephrase the left hand side to state that we have a concrete triple violating functionality.

$$\langle \mathfrak{F}, \iota \rangle \vDash x \lhd y_0 \ \wedge\ x \lhd y_1 \ \wedge\ y_0 \not\doteq y_1$$

On the left hand side we have *several* such statements:

$$\langle \mathfrak{F}, \beta, \iota(y_0) \rangle \vDash p, \qquad \langle \mathfrak{F}, \beta, \iota(y_1) \rangle \vDash q, \qquad \langle \mathfrak{F}, \beta, \iota(x) \rangle \vDash \neg\Diamond(p \wedge q)$$

Thus, we can define (strong) local correspondence by requiring that the violation of quasi–functionality is equivalent to the simultaneous satisfaction of three internal formulae, one at each of the worlds which constitute the triple violating quasi–functionality. It is this latter formulation of correspondence that we will take as the key definition. If we use $\mathcal{L}^e$ here, we can write this sequence of conditions as

$$\langle \mathfrak{F}, \beta, \iota \rangle \vDash x \epsilon \neg\Diamond(p \wedge q) \ \wedge\ y_0 \epsilon p \ \wedge\ y_1 \epsilon q$$

Furthermore, we can abstract from the valuation β:

$$\langle \mathfrak{F}, \iota \rangle \vDash (\exists p)(\exists q)(x \epsilon \neg\Diamond(p \wedge q) \ \wedge\ y_0 \epsilon p \ \wedge\ y_1 \epsilon q)$$

Exercise 174. Show that $\mathfrak{H}$ is a frame. In particular, show that the intersection of two sets $r(i_0, k_0)$, $r(i_1, k_1)$ is a finite union of sets of the form $r(i, k)$.

Exercise 175. Show that for tight frames $\mathfrak{F}$ and points x, y, z such that $\langle \mathfrak{F}, \iota \rangle \vDash x \lhd y \lhd z; x \not\lhd z$ iff for some valuation $x \vDash \neg\Diamond p$, $y \vDash \top$ and $z \vDash p$. Hence, transitivity corresponds locally to **.4** in the class of tight frames.

5.4. The Basic Calculus of Internal Descriptions

Informally, we will say that a (first–order) relation $\zeta(x_0, \ldots, x_{n-1})$ can be *internally described* in a given class $\mathfrak{X}$ of frames if we can find a sequence $\langle \varphi_0, \ldots, \varphi_{n-1} \rangle$ of modal formulae such that for every frame $\mathfrak{F} \in \mathfrak{X}$ we have

$$\langle \mathfrak{F}, \iota \rangle \vDash \zeta(x_0, \ldots, x_{n-1}) \quad \Leftrightarrow \quad \text{for some } \beta \ \langle \mathfrak{F}, \beta, \iota(x_0) \rangle \vDash \varphi_0, \ldots, \langle \mathfrak{F}, \beta, \iota(x_{n-1}) \rangle \vDash \varphi_{n-1}$$

We will rewrite the right–hand side into

$$\langle \mathfrak{F}, \beta, \iota \rangle \vDash x_0 \epsilon \varphi_0 \ \wedge \ldots \wedge \ x_{n-1} \epsilon \varphi_{n-1}$$

We use overstrike arrows in the following way. Let $\vec{x}$ and $\vec{\varphi}$ be n–long sequences. Then

$$\vec{x} \epsilon \vec{\varphi} := \bigwedge_{i<n} x_i \epsilon \varphi_i$$

Notice that the number i does double duty in x_i by both identifying x_i and assigning to it a modal formula φ. This will be rather cumbersome. Thus, to make the association of the variables with the modal formulae independent of the index i on the variables x_i, we use the following notational device. We write $\zeta[\vec{x}]$ to denote (a pair consisting of) the formula ζ and a sequence $\langle x_i : i < n \rangle$ such that every free variable of ζ is identical to some x_i, $i < n$. (It is not required that x_i is distinct from x_j if $i \neq j$.) The numbers $i < n$ of an n–long sequence are also called *slots*. Given an n–long sequence $\vec{\varphi}$ of modal formulae, the slots of $\zeta[\vec{x}]$ are in one to one correspondence with the slots of $\vec{\varphi}$. Notice that writing $\zeta[\vec{x}]$ we may nevertheless have $\vec{x} \nsubseteq \mathit{fvar}(\zeta)$, a fact which we will make use of. We call $\zeta[\vec{x}]$ a **slotted formula**. We emphasize that this is just a piece of notation, nothing more. If the association between variables and slots is clear (especially when there is just one variable), we may drop the sequence. Occasionally we will also use subscripts 0, 1 etc. rather than the sequence of slots.

Definition 5.4.1. *Let $\mathfrak{X}$ be a class of frames, $\zeta[x_0, \ldots, x_{n-1}]$ a slotted $\mathfrak{L}^f$–formula and $\vec{\varphi} = \langle \varphi_0, \ldots, \varphi_{n-1} \rangle$ a sequence of length n. We say that $\vec{\varphi}$* ***internally describes*** *ζ in $\mathfrak{X}$ if for all $\mathfrak{F} \in \mathfrak{X}$ and all ι, $\langle \mathfrak{F}, \iota \rangle \vDash \zeta(\vec{x})$ iff $\langle \mathfrak{F}, \iota \rangle \vDash \vec{x} \epsilon \vec{\varphi}$. Symbolically, we write $\zeta[\vec{x}] \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi}$, or simply $\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}$. Given $\mathfrak{X}$ and $\vec{\varphi}$ we say that $\vec{\varphi}$ is* ***elementary in*** *$\mathfrak{X}$ if an $\zeta \in \mathfrak{L}^f$ exists which is described by $\vec{\varphi}$ in $\mathfrak{X}$.*

Notice first of all that internal describability of ζ itself is $\mathfrak{L}^e$–definable. Namely, $\langle \mathfrak{F}, \iota \rangle \vDash \vec{x} \epsilon \vec{\varphi}$ is just a shorthand for the conjunction of $x_i \epsilon \varphi_i$. Now we have the following equivalence

$$\zeta \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi} \quad \Leftrightarrow \quad \mathfrak{X} \vDash (\forall \vec{x})(\zeta[\vec{x}]. \equiv .(\exists \vec{p})(\vec{x} \epsilon \vec{\varphi}))$$

Here $\vec{p}$ collects all variables of $\vec{\varphi}$. A particularly interesting example of an elementary sequence are those of length 1. In that case it is of the form φ_0, i. e. an ordinary proposition. φ_0 is elementary in $\mathfrak{X}$ iff there is an $\zeta(x_0) \in \mathcal{L}^f$ such that

$$\mathfrak{X} \vDash (\forall x_0)(\zeta(x_0) \equiv (\exists \vec{p})(x_0 \in \varphi_0)) .$$

Alternatively,

$$\mathfrak{X} \vDash (\forall x_0)(\neg\zeta(x_0) \equiv .(\forall \vec{p})(x_0 \in \neg\varphi_0)) .$$

If the latter holds we say that $\neg\varphi_0$ **defines** $\neg\zeta[x_0]$ and is **elementary**, and that ζ is **modally definable**. The same definition could be generalized to sequences, but this is of little benefit. We say that a logic is $\mathfrak{X}$–**elementary** if all of its axioms are elementary in $\mathfrak{X}$. In addition to this definition of elementarity, which for distinction will be called **local**, there is also a *global elementarity*. Namely, we say that φ is **globally elementary** in $\mathfrak{X}$ if there exists an $\mathcal{L}^f$–sentence ζ such that for all $\mathfrak{F} \in \mathfrak{X}$ we have $\mathfrak{F} \vDash \varphi$ iff $\mathfrak{F} \vDash \zeta$. Global elementarity is weaker as we have seen earlier, and it will be of little importance henceforth. Nevertheless, the following theorems can be stated for global rather than local elementarity.

Proposition 5.4.2. *Let $\mathfrak{X}$ be closed under the map $\langle \mathfrak{f}, \mathbb{F}\rangle \mapsto \mathfrak{f}$. Then if a logic is globally $\mathfrak{X}$–elementary, it is complete with respect to the Kripke–frames of $\mathfrak{X}$.*

Examples of such classes are the class of differentiated frames, of refined frames, the class of canonical frames together with the class of Kripke–frames. With respect to a class $\mathfrak{X}$ we say that a logic Λ is **persistent** if for all $\mathfrak{F} \in \mathfrak{X}$ we can infer $\mathfrak{F}_\sharp \vDash \Lambda$ from $\mathfrak{F} \vDash \Lambda$. This is the general scheme. Moreover, we have *d–persistence*, which is persistence with respect to $\mathfrak{D}$, *r–persistence*, which is persistence with respect to $\mathfrak{R}$ etc.

Proposition 5.4.3. *If Λ is $\mathfrak{X}$–persistent and $\mathfrak{X}$–complete, Λ is $\mathfrak{X}_\sharp$–complete.*

Proof. Let $\varphi \notin \Lambda$. Then since Λ is $\mathfrak{X}$–complete there is a frame $\mathfrak{F} \in \mathfrak{X}$ such that $\mathfrak{F} \vDash \Lambda$ but $\mathfrak{F} \nvDash \varphi$. Since Λ is $\mathfrak{X}$–persistent, $\mathfrak{F}_\sharp \vDash \Lambda$. But $\mathfrak{F}_\sharp \nvDash \varphi$ as well, showing Λ to be $\mathfrak{X}_\sharp$–complete. ◻

Proposition 5.4.4. *If Λ is globally $\mathfrak{X} \cup \mathfrak{X}_\sharp$–elementary it is $\mathfrak{X}$–persistent.*

Proof. Let $\Lambda = \mathbf{K}_\kappa \oplus \Delta$. Each $\varphi \in \Delta$ is $\mathfrak{X} \cup \mathfrak{X}_\sharp$–elementary, whence there is an elementary sentence ζ_φ such that $\mathfrak{F} \vDash \varphi$ iff $\mathfrak{F} \vDash \zeta_\varphi$ for all $\mathfrak{F} \in \mathfrak{X} \cup \mathfrak{X}_\sharp$. Now assume $\mathfrak{F} \vDash \Lambda$. Then $\mathfrak{F} \vDash \Delta$. By $\mathfrak{X}$–elementarity of Δ, $\mathfrak{F} \vDash \zeta_\varphi$ for all $\varphi \in \Delta$, hence $\mathfrak{F}_\sharp \vDash \zeta_\varphi$, for all $\varphi \in \Delta$ since each ζ_φ is an $\mathcal{L}^f$–sentence. So $\mathfrak{F}_\sharp \vDash \Delta$, by $\mathfrak{X}_\sharp$–elementarity of Δ. Consequently, $\mathfrak{F}_\sharp \vDash \Lambda$. ◻

Let us now turn to the problem of determining which statements of the form '$\zeta \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi}$' are valid. In the sequel we will be concerned with five basic choices for $\mathfrak{X}$, namely the class of all frames, the class of differentiated frames, tight frames, refined frames and of descriptive frames. Of course if $\zeta \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi}$ and $\mathfrak{X} \supseteq \mathfrak{Y}$ then also $\zeta \leftrightsquigarrow_{\mathfrak{Y}} \vec{\varphi}$ as well, so not all work has to be done separately. In this section we will introduce a calculus for deriving correspondence statements in a quasi–logical style, with axioms and rules. This calculus, called **Seq**, will be correct for $\mathfrak{G}$ and hence for all classes. As starting sequences one can in all cases take a constant formula. Recall namely, that if φ is constant, so is the standard translation $ST(\varphi, x)$. Consequently, the standard translation is first–order. Now we always have

$$\mathfrak{F} \vDash (\forall \vec{p})(\forall x)(x \in \varphi. \equiv .ST(\varphi, x))$$

In this special case, the propositional quantifier is superfluous since there are no variables to be bound. This shows that we always have

(axiom.) $\qquad ST(\varphi, x_0)[x_0] \leftrightsquigarrow_{\mathfrak{X}} \varphi$ if $var(\varphi) = \varnothing$

Of course, this can be stated for sequences as well. So we do have some nontrivial correspondences to start with. The next set of rules is rather obvious. We may add, for example, an inessential variable. If $\zeta[x_0 \cdot \ldots \cdot x_{n-1}]$ is a condition on the n–tuple $\langle x_0, \ldots, x_{n-1}\rangle$, we can nevertheless view it as a condition on $n+1$–tuples $\langle x_0, \ldots, x_n\rangle$ (which we abbreviate by $\vec{x} \cdot x_n$). There is then no condition on x_n, and so on the modal side this corresponds to adding $\top$ at the end of the sequence. Furthermore, we can permute sequences. Given a permutation $\pi : n \to n$, we can write $\pi(\vec{x})$ for the sequence $\langle x_{\pi(0)}, \ldots, x_{\pi(n-1)}\rangle$, and similarly for the modal side. Then the following rules are valid.

$$(\textbf{exp}.) \quad \frac{\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}}{\zeta[\vec{x} \cdot x_n] \leftrightsquigarrow \vec{\varphi} \cdot \top} \qquad\qquad (\textbf{per}.) \quad \frac{\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}}{\zeta[\pi(\vec{x})] \leftrightsquigarrow \pi(\vec{\varphi})}$$

In both rules we used a double line separating the top row from the bottom row. This means that the rules can be applied top–to–bottom or bottom–to–top, which in the case of (**exp**.) amounts to killing an unnecessary variable. Next consider the operation of renaming the variables in $\vec{\varphi}$ by a substitution σ. Obviously, if p^s is another variable and $p \mapsto p^\sigma$ is injective, that is, no two variables are identified, then this is just a harmless operation, with no bearing on the property described by the sequence.

$$(\textbf{ren}.) \quad \frac{\zeta \leftrightsquigarrow \vec{\varphi}}{\zeta \leftrightsquigarrow \vec{\varphi}^\sigma} \qquad \sigma : p\text{–}var \rightarrowtail p\text{–}var$$

Similarly, consider swapping $\neg p$ and p (denoted by $\varphi[\neg p \rightleftharpoons p]$), or alternatively, replacing p by $\neg p$ and killing double negation. Again, this does not change the elementary property described.

$$(\textbf{swap}.) \quad \frac{\zeta \leftrightsquigarrow \vec{\varphi}}{\zeta \leftrightsquigarrow \vec{\varphi}[\neg p \rightleftharpoons p]}$$

Next consider the operation of replacing in $\zeta(\vec{x})$ a variable, say x_{n-1}, by another, say x_{n-2}. This amounts on the modal part to replacing φ_{n-2} by $\varphi_{n-2} \wedge \varphi_{n-1}$. Also consider iterating a condition on another variable.

$$(\textbf{cnt}.)\qquad \frac{\zeta(\vec{x},y,z)[\vec{x}\cdot y\cdot z] \leftrightsquigarrow \vec{\varphi}\cdot\varphi_{n-2}\cdot\varphi_{n-1}}{\zeta(\vec{x},y,y)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\varphi_{n-2}\wedge\varphi_{n-1}}$$

$$(\textbf{iter}.)\qquad \frac{\zeta(\vec{x},y)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\chi}{(\zeta(\vec{x},y)\ \wedge\ \zeta(\vec{x},z))[\vec{x}\cdot y\cdot z] \leftrightsquigarrow \vec{\varphi}\cdot\chi\cdot\chi}$$

Suppose, namely, that $\zeta[\vec{x}\cdot y\cdot z] \leftrightsquigarrow \vec{\varphi}\cdot\varphi_{n-2}\cdot\varphi_{n-1}$. Then for all n–tuples $\vec{w}\subseteq f$ of worlds in $\mathfrak{F}\in\mathfrak{X}$, we have $\zeta[\vec{w}]$ iff for some valuation β, $w_i\in\overline{\beta}(\varphi_i)$ for all $i<n$. Now choose an $n-1$–tuple $\vec{w}\subseteq f$. By assumption, $\zeta[\vec{w}\cdot w_{n-2}]$ iff for some valuation $w_i\in\overline{\beta}(\varphi_i)$ for all $i<n-2$ and $w_{n-2}\in\overline{\beta}(\varphi_{n-2})$ as well as $w_{n-2}\in\overline{\beta}(\varphi_{n-1})$, so that $w_{n-2}\in\overline{\beta}(\varphi_{n-2}\wedge\varphi_{n-1})$, and conversely. Thus the rule is correct. Notice that we must reduce the number of arguments here. We cannot conclude, for example, that $\zeta\wedge x_{n-2}\doteq x_{n-1}$ is describable, for this would require simultaneous fixing of x_{n-2} and x_{n-1} to the same value. The rule (**iter**.) is likewise straightforward. For the statement of the following rule let $\vec{\varphi}^1$ and $\vec{\varphi}^2$ be two sequences of length n. Then $\vec{\varphi}^1\wedge\vec{\varphi}^2 := \langle\varphi^1_i\wedge\varphi^2_i : i<n\rangle$.

$$(\wedge\textbf{–I}.)\qquad \frac{\zeta^1[\vec{x}] \leftrightsquigarrow \vec{\varphi}^1 \qquad \zeta^2[\vec{x}] \leftrightsquigarrow \vec{\varphi}^2}{(\zeta^1\ \wedge\ \zeta^2)[\vec{x}] \leftrightsquigarrow \vec{\varphi}^1\wedge\vec{\varphi}^2}\qquad \text{if } var(\vec{\varphi}^1)\cap var(\vec{\varphi}^2)=\varnothing$$

For a proof assume $(\zeta^1\ \wedge\ \zeta^2)[\vec{w}]$. By assumption we have a valuation β^1 such that $w_i\in\overline{\beta}^1(\varphi^1_i)$ for all $i<n$, and a valuation β^2 such that $w_i\in\overline{\beta}^2(\varphi^2_i)$ for all $i<n$. Define β as follows. $\beta(p):=\beta^1(p)$ if $p\in var(\vec{\varphi})$, $\beta(p):=\beta^2(p)$ otherwise. This is well defined by the assumption that $\vec{\varphi}^1$ and $\vec{\varphi}^2$ are disjoint in variables. Then $w_i\in\overline{\beta}(\varphi^1_i)$ as well as $w_i\in\overline{\beta}(\varphi^2_i)$ for all $i<n$, and so $w_i\in\overline{\beta}(\varphi^1_i\wedge\varphi^2_i)$, as required. Conversely, if $w_i\in\overline{\beta}(\varphi^1_i\wedge\varphi^2_i)$ for all $i<n$ then $w_i\in\overline{\beta}(\varphi^1_i)$ as well as $w_i\in\overline{\beta}(\varphi^2_i)$ for all $i<n$ and so by assumption $\zeta^1[\vec{w}]$ as well as $\zeta^2[\vec{w}]$.

$$(\vee\textbf{–I}.)\qquad \frac{\zeta^1[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\chi \qquad \zeta^2[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\psi}{(\zeta^1\vee\zeta^2)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\chi\vee\psi}$$

$$(\Diamond\textbf{–I}.)\qquad \frac{\zeta(\vec{x},y)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\chi}{(\exists z \rhd_i y)\zeta(\vec{x},z)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\varphi}\cdot\Diamond_i\chi}$$

To see the correctness of the first rule, assume the premisses hold, and that $\zeta^1\vee\zeta^2[\vec{w}]$ for an $n+1$–tuple $\vec{w}$. Then either $\zeta^1[\vec{w}]$ or $\zeta^2[\vec{w}]$. Assume without loss of generality the first. Then there is a valuation β such that $w_i\in\overline{\beta}(\varphi_i)$ all $i<n$ and $w_n\in\overline{\beta}(\chi)$. Then also $w_n\in\overline{\beta}(\chi\vee\psi)$. Conversely, assume that $w_i\in\overline{\beta}(\varphi_i)$ for all $i<n$ and $w_n\in\overline{\beta}(\chi\vee\psi)$. Then either $w_n\in\overline{\beta}(\chi)$ or $w_n\in\overline{\beta}(\psi)$. Assume without loss of generality the first. Then, by the left hand premiss of the rule $\zeta^1[\vec{w}]$, whence $(\zeta^1\vee\zeta^2)[\vec{w}]$. Next the rule ($\Diamond$**–I**.). Assume $((\exists z\rhd_i y)\zeta)(\vec{x},z)$. Then there is a sequence $\vec{w}\cdot v$ such that $(\exists z\rhd_i y)\zeta[\vec{w}\cdot v]$, that is, there is a u such that $v\lhd_i u$ and $\zeta[\vec{w}\cdot u]$. By assumption, we can find a valuation β such that $w_i\in\overline{\beta}(\varphi_i)$ for all i and $u\in\overline{\beta}(\chi)$. Then $v\in\overline{\beta}(\Diamond_i\chi)$,

which had to be shown. Conversely, assume $w_i \in \overline{\beta}(\varphi_i)$ for all i and $v \in \overline{\beta}(\Diamond_i \chi)$. Then there is a u such that $v \lhd_i u$ and $u \in \overline{\beta}(\chi)$. Thus, by the premiss of the rule, $\zeta(\vec{w} \cdot u)$. Then, however, $((\exists z \rhd_i v)\zeta(\vec{v}, z))[\vec{w} \cdot v]$, as required.

Now let **Seq**, consist of the axioms **(axiom.)** and of all the rules (**exp**.), (**per**.), (**ren**.), (**swap**.), (**cnt**.), (**iter**.), ($\wedge$–**I**.), ($\vee$–**I**.) and ($\Diamond$–**I**.). We call **Seq** the **base calculus**. Let $\mathcal{C}$ be a calculus of internal descriptions, consisting of axioms and rules. A statement '$\zeta \leftrightsquigarrow \vec{\varphi}$' is **derivable** in $\mathcal{C}$ if it can be produced from the axioms with the help of the rules in finitely many steps. We say that $\vec{\varphi}$ is **derivable** in $\mathcal{C}$ if there exists a $\zeta \in \mathcal{L}^f$ and a sequence $\vec{x}$ such that '$\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}$' is derivable in **Seq**; and that $\zeta[\vec{x}]$ is **derivable** in $\mathcal{C}$ if there exists a sequence $\vec{\varphi}$ such that '$\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}$' is derivable in $\mathcal{C}$. A calculus $\mathcal{C}$ of correspondence statements is **sound** for a class $\mathfrak{X}$ if '$\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}$' is derivable in $\mathcal{C}$ only if $\vec{\varphi}$ internally describes $\zeta[\vec{x}]$ in $\mathfrak{X}$. $\mathcal{C}$ is called **complete** for $\mathfrak{X}$ if whenever $\vec{\varphi}$ internally describes $\zeta[\vec{x}]$ in $\mathfrak{X}$, '$\zeta[\vec{x}] \leftrightsquigarrow \vec{\varphi}$' is derivable in $\mathcal{C}$.

Theorem 5.4.5. ***Seq*** *is sound for all classes of frames.*

An important consequence is the following.

Theorem 5.4.6. *Let $\mathfrak{X}$ be any class of frames. If $\zeta(x_0)$ is obtained from formulae internally describable in $\mathfrak{X}$ with the help of conjunction, disjunction or restricted existential quantification, then $\zeta(x_0)$ is internally describable in $\mathfrak{X}$.*

Proof. From Lemma 5.4.7 we conclude that the set of internally describable $\zeta(\vec{x})$ is closed under $\wedge$. It is clearly also closed under restricted $\exists$. Now let $\zeta(x_0)$ be composed from internally describable formulae with conjunction, disjunction and restricted existential quantification. Then, by some straightforward manipulations, $\zeta(x_0)$ is equivalent in predicate logic to a disjunction of formulae $\eta_i(x_0)$, $i < n$, each of which is made from describable formulae using only $\wedge$ and restricted $\exists$. Then, for all $i < n$, $\eta_i(x_0)$ is describable in $\mathfrak{X}$, and by Lemma 5.4.8, $\zeta(x_0)$ is describable in $\mathfrak{X}$. ◻

Lemma 5.4.7. *Let $\mathfrak{X}$ be a class, and let $\zeta(\vec{x})$ and $\eta(\vec{x})$ be describable in $\mathfrak{X}$. Then $(\zeta \wedge \eta)(\vec{x})$ is describable in $\mathfrak{X}$.*

Proof. By assumption, there exists sequence $\vec{\varphi}$ such that $\zeta(\vec{x}) \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi}$ and a sequence $\vec{\psi}$ such that $\eta(\vec{x}) \leftrightsquigarrow_{\mathfrak{X}} \vec{\psi}$. Now let $\vec{\chi}$ result from renaming variables of ψ in such a way that they become disjoint to the variables of $\vec{\varphi}$. Then, by (**ren**.), $\eta(\vec{x}) \leftrightsquigarrow_{\mathfrak{X}} \vec{\psi}$. Finally, by ($\wedge$–**I**.), $\vec{\varphi} \wedge \vec{\chi}$ describes $(\zeta \wedge \eta)(\vec{x})$ in $\mathfrak{X}$. ◻

Lemma 5.4.8. *Let $\mathfrak{X}$ be a class, and let $\zeta(x_0)$ and $\eta(x_0)$ be describable in $\mathfrak{X}$. Then $(\zeta \vee \eta)(x_0)$ is describable in $\mathfrak{X}$.*

The proof of this theorem is similar. Notice that in proving the correctness of the rules we have always shown how to construct a valuation for a given sequence of worlds. Now we consider two rules.

($\neq$–**I**.) $\quad x_0 \neq x_1 \leftrightsquigarrow p \cdot \neg p \qquad$ ($\not\lhd$–**I**.) $\quad x_0 \not\lhd_i x_1 \leftrightsquigarrow \Box_i p \cdot \neg p$

Theorem 5.4.9. ***Seq*** *+ ($\not\equiv$–**I**.) is sound for* $\mathfrak{Df}$*, the class of differentiated frames.*

The proof is an exercise.

Theorem 5.4.10. ***Seq*** *+ ($\not\lhd$–**I**.) is sound for* $\mathfrak{Ti}$*, the class of tight frames.*

Again the proof is an exercise. We now give some examples of the calculus.

Example 1. **K.T** is ti–persistent. For a proof consider the following derivation.

$$\frac{x_0 \not\lhd x_1 \leftrightsquigarrow \Box p \cdot \neg p}{x_0 \not\lhd x_0 \leftrightsquigarrow \Box p \wedge \neg p}$$

The first line is true in $\mathfrak{Ti}$. Hence $\Box p \to p$ locally defines $x_0 \lhd x_0$ in the class of tight frames.

Example 2. **K**.t is ti–persistent.

$$\begin{array}{c} x_0 \not\lhd_0 x_1 \leftrightsquigarrow \Box_0 p \cdot \neg p \\ \hline (\exists u \rhd_1 x_0)(u \not\lhd_0 x_1) \leftrightsquigarrow \Diamond_1 \Box_0 p \cdot \neg p \\ \hline (\exists u \rhd_1 x_0)(u \not\lhd_0 x_0) \leftrightsquigarrow \Diamond_1 \Box_0 p \wedge \neg p \end{array}$$

$$\begin{array}{c} x_0 \not\lhd_1 x_1 \leftrightsquigarrow \Box_1 p \cdot \neg p \\ \hline (\exists u \rhd_0 x_0)(u \not\lhd_1 x_1) \leftrightsquigarrow \Diamond_0 \Box_1 p \cdot \neg p \\ \hline (\exists u \rhd_0 x_0)(u \not\lhd_1 x_0) \leftrightsquigarrow \Diamond_0 \Box_1 p \wedge \neg p \end{array}$$

These two derivations show that in the class of tight bimodal frames the formulae $\Diamond_0 \Box_1 p \to p$ and $\Diamond_1 \Box_0 p \to p$ locally correspond to $(\forall u \rhd_1 x_0)(u \lhd_0 x_0)$ and $(\forall u \rhd_0 x_0)(u \lhd_1 x_0)$, respectively.

Example 3. **K.alt**$_1$ is df–persistent. The following derivation is a proof of this fact.

$$\begin{array}{c} x_0 \not\equiv x_1 \leftrightsquigarrow p \cdot \neg p \\ \hline (\exists x_0 \rhd u)(u \not\equiv x_1) \leftrightsquigarrow \Diamond p \cdot \neg p \\ \hline (\exists x_1 \rhd v)(\exists u \rhd x_0)(u \not\equiv v) \leftrightsquigarrow \Diamond p \cdot \Diamond \neg p \\ \hline (\exists x_0 \rhd v)(\exists u \rhd x_0)(u \not\equiv v) \leftrightsquigarrow \Diamond p \wedge \Diamond \neg p \end{array}$$

Notice that the axiom $\Diamond p \wedge \Diamond q. \to .\Diamond(p \wedge q)$ is not derivable with the help of the calculus for differentiated frames.

Theorem 5.4.11. *Let $\zeta(x_0) \in \mathcal{R}^f$ be universal, restricted and positive. Then $\zeta(x_0)$ is internally definable in $\mathfrak{R}$, the class of refined frames.*

Proof. $\neg\zeta(x_0)$ is negative, existential and restricted. Thus, it is composed from formulae of the form $x_i \not\equiv x_j$, $x_i \not\lhd_k x_j$ with the help of conjunction, disjunction and restricted existential quantification. Whence it is derivable in **Seq** + ($\not\equiv$ **–I**.) + ($\not\lhd$ **–I**.). ∎

Let us end this section by considering the question of the completeness of the basic calculus. We will show in Section 9.5 that for a modal formula φ it is undecidable whether it axiomatizes the inconsistent logic. Hence the set of sequences '$\varphi \leftrightsquigarrow \mathrm{f}$' is undecidable. The idea behind setting up modal equivalence not as a matter of theorems of second–order logic but as a deductive calculus is that while the former is undecidable, the latter may be decidable, however at the price of being incomplete. On the other hand, we will see later (Theorem 5.8.6) that **Seq** is complete in the following weaker sense. If '$\varphi \leftrightsquigarrow \zeta(x)$' holds in $\mathfrak{G}$ then there exist ψ and η such that $\mathbf{K}_\kappa \oplus \psi = \mathbf{K}_\kappa \oplus \varphi$ and $(\forall x)\zeta(x) \equiv (\forall x)\eta(x)$ (in predicate logic) and '$\psi \leftrightsquigarrow \eta(x)$' is derivable in **Seq**. It is not clear whether this generalizes to arbitrary sequences, but that seems to be the case. So, rather than generating all facts we generate at least a representative class of them. We could in principle add rules that would make the calculus complete (by closing under equivalence), but that would make it undecidable and useless for practical purposes.

Exercise 176. Show that (**iter.**) is sound for all classes.

Exercise 177. Show that inequality is internally describable in $\mathfrak{X}$ iff $\mathfrak{X}$ is a class of differentiated frames. This proves Theorem 5.4.9.

Exercise 178. Show that j–inaccessibility (i. e. $\ntriangleleft_j$) is internally describable in $\mathfrak{X}$ for all $j < \kappa$ iff $\mathfrak{X}$ is a class of tight frames. This proves Theorem 5.4.10.

Exercise 179. Show that the set of $\mathfrak{X}$–elementary logics is closed under finite joins and finite meets in the lattice of all modal logics.

*__Exercise 180.__ Show that $\Box\Diamond\Box\Box p \to \Diamond\Diamond\Box\Diamond p$ is globally $\mathfrak{Krp}$–elementary and corresponds to $(\forall x)(\exists y)(x \lhd y)$. However, this axiom is not locally elementary. (The first half is not so difficult, only the failure of local elementarity. For those who want to see a proof, it can be found in van Benthem [**10**], page 82.)

5.5. Sahlqvist's Theorem

Two classes of modal formulae will play a fundamental role, *monotone* and $\wedge$–*distributive* formulae. A formula $\varphi(p, \vec{q})$ is **monotone in** p if $\vdash \varphi(p \wedge r, \vec{q}) \to \varphi(p, \vec{q})$, where $\vdash \varphi$ abbreviates $\varphi \in \mathbf{K}_\kappa$, and **$\wedge$–distributive in** p if $\vdash \varphi(p \wedge r, \vec{q}). \leftrightarrow .\varphi(p, \vec{q}) \wedge \varphi(r, \vec{q})$. φ is called **monotone** (**$\wedge$–distributive**) if it is monotone ($\wedge$–distributive) in all occurring variables. The notions **antitone** and **$\vee$–distributive** are dual notions. That is to say, φ is **antitone** ($\vee$–distributive) in p iff $\neg\varphi$ is monotone ($\wedge$–distributive) in p. All of these notions can be characterized syntactically. We will give sufficient criteria here. Call a formula $\varphi(p, \vec{q})$ **positive** in p if all occurrences of p are in the scope of an even number of negations. Call a formula **strongly positive in** p if p does not occur in the scope of $\neg$ (or any $\Diamond_j$ for $j < \kappa$ if $\Diamond_j$ is a primitive symbol).

φ is **positive** (**strongly positive**) if it is positive (strongly positive) in all occurring variables. A formula is **negative** (**strongly negative**) if it can be obtained from a positive (strongly positive) formula by replacing each occurrence of a variable p by $\neg p$. Notice that we can characterize a positive formula also as follows.

Proposition 5.5.1. *$\varphi(p, \vec{q})$ is positive in p iff there exists a formula ψ which is built from the letter p and formulae not containing p with the help of $\wedge$, $\vee$, $\Box_i$, $\Diamond_i$, $i < \kappa$, such that $\varphi(p) \leftrightarrow \psi \in \boldsymbol{K}_\kappa$.*

The proof is an exercise. Notice that the occurring constant subformulae can be arbitrarily complex. The formula $\Box_0((\Diamond_1(\top \wedge \neg\Box_0\bot) \vee p) \wedge \Diamond_0 p)$ is positive. Likewise, $\Box_0(\Diamond_1\top \vee \Box_0 p)$ is strongly positive.

Proposition 5.5.2. *The following holds.*

(1) *If $\varphi(p, \vec{q})$ is positive in p it is monotone in p.*
(2) *If $\varphi(p, \vec{q})$ is strongly positive in p it is $\wedge$–distributive in p.*

Proof. (1.) By induction on $\varphi(p, \vec{q})$. By Proposition 5.5.1 we can assume that φ is built from the letter p and formulae not containing p with the help of $\wedge$, $\vee$, $\Box_j$ and $\Diamond_j$, $j < \kappa$. The formula p is monotone in p; likewise a formula not containing p is obviously monotone in p. Suppose $\varphi = \psi_1 \wedge \psi_2$. By induction hypothesis, $\psi_1(p \wedge r, \vec{q}) \vdash \psi_1(p, \vec{q})$ and $\psi_2(p \wedge r, \vec{q}) \vdash \psi_2(p, \vec{q})$. Hence $\varphi(p \wedge r, \vec{q}) \vdash \psi_1(p, \vec{q}) \wedge \psi_2(p, \vec{q})(= \varphi(p, \vec{q}))$. Similarly for $\varphi = \psi_1 \vee \psi_2$. Now suppose that $\varphi = \Box_i\psi_1$. Then by induction hypothesis $\psi_1(p \wedge r, \vec{q}) \vdash \psi_1(p, \vec{q})$. From this we get $\Box_i\psi(p \wedge r, \vec{q}) \vdash \Box_i\psi(p, \vec{q})$, as required. Similarly for $\varphi = \Diamond_i\psi_1$. (2.) Again by induction. A formula not containing p is $\wedge$–distributive in p by the fact that $\vdash \psi \leftrightarrow \psi \wedge \psi$. Now let $\varphi = \psi_1 \wedge \psi_2$, ψ_1 and ψ_2 strongly positive. By induction hypothesis both are $\wedge$–distributive in p. Then we have $\vdash \psi_i(p \wedge r, \vec{q}) \leftrightarrow \psi_i(p, \vec{q}) \wedge \psi_i(r, \vec{q})$ and so

$$\begin{aligned}
\varphi(p \wedge r, \vec{q}) \quad & \dashv\vdash \psi_1(p \wedge r, \vec{q}) \wedge \psi_2(p \wedge r, \vec{q}) \\
& \dashv\vdash \psi_1(p, \vec{q}) \wedge \psi_1(r, \vec{q}) \wedge \psi_2(p, \vec{q}) \wedge \psi_2(r, \vec{q}) \\
& \dashv\vdash \varphi(p, \vec{q}) \wedge \varphi(r, \vec{q}) \ .
\end{aligned}$$

Similarly for $\varphi = \Box_i\psi$. ∎

A sequence $\vec{\varphi}$ of formulae is called a **spone** if each of its members is either strongly positive in all variables or negative in all variables. (The name *spone* is an acronym from *strongly positive and negative*.) We will usually write a spone in the form $\vec{\pi} \cdot \vec{\nu}$, where $\vec{\pi}$ is the subsequence of the strongly positive formulae and $\vec{\nu}$ the subsequence of the negative formulae. We will show that in the class $\mathfrak{Krp} \cup \mathfrak{D}$ all spones are elementary, and this will prove Sahlqvist's Theorem. In the basic calculus **Seq** not all spones can be derived. However, in the class $\mathfrak{Krp} \cup \mathfrak{D}$ another rule is valid. The key to this rule is the following lemma. To state this lemma properly recall from Section 2.9 the concept of an *upward directed family of sets*. Let a frame $\mathfrak{F}$ be given. Consider a set I of indices which is partially ordered by $\leq$, and for each pair i, j there is a k such that $i \leq k$ and $j \leq k$. A family over $\langle I, \leq\rangle$ is simply a collection of sets c_i, $i \in I$ (or a function from I into $\mathbb{F}$). This family is **upward directed** if $i \leq j$ implies $c_i \subseteq c_j$. Given such a family $\mathcal{D} = \langle d_i : i \in I\rangle$ we write $\Box_j\mathcal{D}$ for the family $\langle \Box_j d_i : i \in I\rangle$. An upward directed family has a **limit**, $\lim \mathcal{D}$, which is just the union $\bigcup_{i\in I} d_i$. The notion of a downward directed family is dual, that is, we require $i \leq j \Rightarrow d_i \supseteq d_j$ instead. The following theorem is due to Leo Esakia in [**58**].

Lemma 5.5.3 (Esakia). *Let $\mathfrak{F}$ be a tight and compact frame and $\mathcal{D} = \langle d_i : i \in I\rangle$ be an upward directed family of sets in $\mathbb{F}$. Then*

$$\blacksquare_i \lim \mathcal{D} = \lim \blacksquare_i \mathcal{D}.$$

Proof. $\mathcal{D}$ is upward directed, and so $-\mathcal{D} = \langle -d_i : i \in I\rangle$ is downward directed. It will be sufficient to show that for a downward directed family $\mathcal{E}$, $\blacksquare_i \lim -\mathcal{E} = \lim \blacksquare_i - \mathcal{E}$. This is the same as $-\blacksquare_i \lim -\mathcal{E} = -\lim \blacksquare_i - \mathcal{E}$ since lim commutes with $-$. This is finally equivalent to

$$\blacklozenge_i \lim \mathcal{E} = \lim \blacklozenge_i \mathcal{E}.$$

(For an upward directed family this is clear, but now $\mathcal{E}$ is downward directed.) ($\subseteq$.) Let $x \in \blacklozenge_i \lim \mathcal{E}$. Then there is a $y \in \lim \mathcal{E} = \bigcap \mathcal{E}$ such that $x \lhd_i y$. Thus for all $j \in I$, $x \in \blacklozenge_i d_j$ and so $x \in \lim \blacklozenge_i \mathcal{E}$. ($\supseteq$.) Suppose $x \notin \blacklozenge_j \lim \mathcal{E}$. Pick a $y \in \lim \mathcal{E}$. Then $x \ntriangleleft_j y$. By tightness of $\mathfrak{F}$ there is an internal set a_y such that $y \in a_y$, but $x \in \blacksquare_j - a_y$. Since the union of the a_y contains $\lim \mathcal{E}$, there is a finite subset $Y \subseteq \lim \mathcal{E}$ such that $\lim \mathcal{E} \subseteq \bigcup_{y\in Y} a_y$. (This follows from the compactness of the frame.) Let $b := \bigcup_{y\in Y} a_y$. Then $\lim \mathcal{E} \subseteq b$ and $x \in \blacksquare_j - b$. Moreover, by compactness of $\mathfrak{F}$ again, there is an $e \in \mathcal{E}$ such that $e \subseteq b$. Then $\blacklozenge_j e \subseteq \blacklozenge_j b$. Since $x \notin \blacklozenge_j b$, $x \notin \blacklozenge_j e$. Therefore $x \notin \lim \blacklozenge_j \mathcal{E}$. ∎

Theorem 5.5.4. *The rule* ($\Box$–*I*.) *is sound for* $\mathfrak{Krp} \cup \mathfrak{D}$.

$$(\Box\text{–I.}) \qquad \frac{\zeta[\vec{x}\cdot y] \leftrightsquigarrow \vec{\rho}\cdot\mu}{(\forall z \rhd_j y)\zeta(\vec{x}, z)[\vec{x}\cdot y] \leftrightsquigarrow \vec{\rho}\cdot\Box_j\mu} \qquad (\vec{\rho}\ \textit{a spone}, \mu\ \textit{negative})$$

Proof. Assume that $\zeta[\vec{x}\cdot y] \leftrightsquigarrow \vec{\rho}\cdot\mu$ in the class $\mathfrak{Krp}\cup\mathfrak{D}$. Let $\vec{x}$ be of length n. Take a frame $\mathfrak{F}$ from $\mathfrak{Krp}\cup\mathfrak{D}$ and a valuation β such that $w_i \in \overline{\beta}(\rho_i)$ for all $i < n$ and $v \in \overline{\beta}(\Box_j\mu)$. Then for all u such that $v \lhd_j u$ we have $u \in \overline{\beta}(\mu)$, and so by assumption $\zeta[\vec{w}\cdot u]$. Hence $((\forall y \rhd_j v)\zeta(\vec{x}, y))[\vec{w}\cdot v]$, as required. For the converse direction choose points $w_0, \dots, w_{n-1}, v$ such that for all u with $v \lhd_j u$ we have $\zeta[\vec{w}\cdot u]$. Let $A := \{u : v \lhd_j u\}$. By assumption, for each $u \in A$ there is a valuation β_u such that $w_i \in \overline{\beta}_u(\rho_i)$ and $u \in \overline{\beta}_u(\mu)$.

Case 1. $\mathfrak{F}$ is a Kripke–frame. Then put $\beta(p) := \bigcap_{u\in A}\beta_u(p)$. Then $w_i \in \overline{\beta}(\rho_i)$ for all i. For either ρ_i is negative (and the claim easily follows), or ρ_i is strongly positive, in which case $\overline{\beta}(\rho_i) = \bigcap_{u\in A}\overline{\beta_u}(\rho_i)$ and so $w_i \in \overline{\beta}(\varphi_i)$. Furthermore, $u \in \overline{\beta}(\mu)$ for all $u \in A$, since μ is negative. Thus $v \in \overline{\beta}(\Box_j\mu)$, and so everything is shown.

Case 2. $\mathfrak{F}$ is a descriptive frame. Let I be the set of finite subsets of A. For $S \in I$ let $\beta_S(p) := \bigcap_{u\in S}\beta_u(p)$. It is not hard to verify that $w_i \in \overline{\beta}_S(\rho_i)$ for all $i < n$. Our aim is to find an $S \in I$ such that $v \in \overline{\beta}_S(\Box_j\mu)$. With respect to $\subseteq$, I is ordered and $\langle \beta_S(p) : S \in I\rangle$ is a downward directed family of sets. Now, $u \in \overline{\beta}_S(\mu)$ if $u \in S$. Hence $A \subseteq \lim \overline{\beta}_S(\mu)$ and so $v \in \blacksquare \lim \overline{\beta}_S(\mu)$. By Esakia's Lemma $v \in \blacksquare_j \lim \overline{\beta}_S(\mu) = \lim \blacksquare_j \overline{\beta}_S(\mu) = \lim \overline{\beta}_S(\Box_j\mu)$. By compactness of $\mathfrak{F}$ there is an $S \in I$ such that $v \in \overline{\beta}_S(\Box_j\mu)$; this had to be shown. ∎

We define a new calculus $\mathbf{Seq}^+$, which is $\mathbf{Seq}$ enriched by ($\neq$–**I**.) and ($\Box$–**I**.) and a new rule for conjunction, ($\wedge$–$\mathbf{I}_2$.). To define it, let $\vec{\varphi}^1$ and $\vec{\varphi}^2$ be two n–long sequences. Recall that $\vec{\varphi}^1 \wedge \vec{\varphi}^2 := \langle \varphi^1_i \wedge \varphi^2_i : i < n\rangle$.

$$(\wedge\text{–}\mathbf{I}_2.)\qquad \frac{\zeta_1[\vec{x}] \leftrightsquigarrow \vec{\pi}_1 \cdot \vec{v}_1 \qquad \zeta_2[\vec{x}] \leftrightsquigarrow \vec{\pi}_2 \cdot \vec{v}_2}{\zeta_1 \wedge \zeta_2[\vec{x}] \leftrightsquigarrow (\vec{\pi}_1 \wedge \vec{\pi}_2)\cdot(\vec{v}_1 \wedge \vec{v}_2)} \quad \text{for } \vec{\pi}_i \cdot \vec{v}_i \text{ spones}$$

It is left as an exercise to show the soundness of this rule. Moreover, this rule is sound in all classes, so it can actually be added to **Seq**. It is only for ease of exposition that we have chosen to ignore this rule previously.

Theorem 5.5.5 (Sahlqvist). *Let χ be a modal formula of the form*

$$\boxplus(\varphi \to \psi)$$

where $\boxplus$ is a compound modality. Suppose that

(1) *ψ is positive.*
(2) *φ is composed from strongly positive formulae using only $\wedge$, $\vee$ and $\Diamond_j$, $j < \kappa$.*

Then $\boldsymbol{K}_\kappa \oplus \chi$ is locally d–persistent and locally elementary in $\mathfrak{Krp} \cup \mathfrak{D}$.

Proof. It is enough if we can show that there is a ζ such that $\zeta \leftrightsquigarrow \neg\chi$. Now $\neg\chi = \neg \boxplus \neg(\varphi \wedge \neg\psi)$. By repeated use of ($\Diamond$–**I**.) and ($\vee$–**I**.) backwards this can be reduced to showing that $\varphi \wedge \neg\psi$ is derivable. By **(cnt.)** it is enough that $\varphi \cdot \neg\psi$ is derivable. $\neg\psi$ is negative, and φ is composed from strongly positive formulae using $\wedge$, $\vee$ and $\Diamond_j$. By appealing to the rules (**cnt**.), ($\vee$–**I**.), ($\Diamond$–**I**.) this can be reduced to the problem of showing that sequences of the form $\vec{\pi} \cdot v$ are derivable, where $\vec{\pi}$ is a sequence of strongly positive formulae and v negative. The theorem below establishes this. ∎

Lemma 5.5.6. *All spones are derivable in $\boldsymbol{Seq}^+$.*

Proof. Let us concentrate on the modal formulae in $\mathbf{Seq}^+$. We proceed in several steps.
Step 1. All spones $\vec{\pi} \cdot \vec{v}$ are derivable where for some p, $\pi_i = p$ and $v_j = \neg p$, $\top$ or $\bot$. Simply start with $p \cdot \neg p$ and use (**iter**.) and (**per**.).
Step 2. All spones are derivable in which the v_i are constant or a negated variable. Namely, by (**swap**.) it is enough to show that if all π_i are variables or constants and v_j are strongly negative, the spone is derivable. So there are three possibilities, (i) v_j is constant, (ii) $v_j = \mu_1 \vee \mu_2$, (iii) $v_j = \Diamond_j\mu$. We deal with these cases in turn,

assuming without loss of generality that $j = n-1$, the last entry of the list. So we have the spone $\vec{\rho} \cdot \nu_{n-1}$. **Case (i).** If $\vec{\rho}$ is derivable, so is $\vec{\rho} \cdot \top$. Moreover, ν_{n-1} is derivable, and by (**exp**.) also $\vec{\top} \cdot \nu_{n-1}$, and so by ($\wedge$–**I**.) $\vec{\rho} \cdot \nu_{n-1}$. **Case (ii).** If $\vec{\rho} \cdot \mu_1$ and $\vec{\rho} \cdot \mu_2$ are derivable, so is $\vec{\rho} \cdot \mu_1 \vee \mu_2$, by ($\vee$–**I**.). **Case (iii).** If $\vec{\rho} \cdot \mu$ is derivable, so is $\vec{\rho} \cdot \Diamond_j \mu$, by ($\Diamond$–**I**.).

Step 3. All spones are derivable. We may start from spones in which the strongly positive part can be complex. The reduction is similar to that in Step 2, with two more cases to be considered, namely (iv) $\nu = \mu_1 \wedge \mu_2$, (v) $\nu = \Box_j \mu$. Case (iv) is dealt with by using the rule ($\wedge$–**I**$_2$.), and Case (v) is dealt with by using ($\Box$–**I**.). $\Box$

To master the syntactic description of the theorem requires some routine. We note that for example the Geach formula $\Diamond\Box p \to \Box\Diamond p$ satisfies the conditions of the theorem while the McKinsey formula $\Box\Diamond p \to \Diamond\Box p$ does not.

Example 1. The Geach formula is elementary in $\mathfrak{Krp} \cup \mathfrak{D}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{x_0 \not\doteq x_1 \leftrightsquigarrow p \cdot \neg p}{(\forall u \rhd x_0)(u \not\doteq x_1) \leftrightsquigarrow \Box p \cdot \neg p}}{(\forall v \rhd x_1)(\forall u \rhd x_0)(u \not\doteq v) \leftrightsquigarrow \Box p \cdot \Box\neg p}}{(\exists u' \rhd x_0)(\forall v \rhd x_1)(\forall u \rhd u')(u \not\doteq v) \leftrightsquigarrow \Diamond\Box p \cdot \Box\neg p}}{(\exists v' \rhd x_1)(\exists u' \rhd x_0)(\forall v \rhd v')(\forall u \rhd u')(u \not\doteq v) \leftrightsquigarrow \Diamond\Box p \cdot \Diamond\Box\neg p}}{(\exists v' \rhd x_0)(\exists u' \rhd x_0)(\forall v \rhd v')(\forall u \rhd u')(u \not\doteq v) \leftrightsquigarrow \Diamond\Box p \wedge \Diamond\Box\neg p}$$

Example 2. The axiom $\Diamond p \wedge \Diamond q \to \Diamond(p \wedge q)$ defines a df–persistent logic. Nevertheless, we have seen that this cannot be shown inside the calculus for differentiated frames. However, in the extended calculus it is derivable. The derivation is somewhat contrived. First, from $x_0 \not\doteq x_2 \leftrightsquigarrow p \cdot \top \cdot \neg p$ and $\mathsf{t} \leftrightsquigarrow \top \cdot \top \cdot q$ we get $x_0 \not\doteq x_2 \leftrightsquigarrow p \cdot q \cdot \neg p$. Likewise, $x_1 \not\doteq x_2 \leftrightsquigarrow p \cdot q \cdot \neg q$ is derived. From this we get $x_0 \not\doteq x_2 \vee x_1 \not\doteq x_2 \leftrightsquigarrow p \cdot q \cdot \neg p \vee \neg q$. Now we get

$$\dfrac{\dfrac{(\forall y \rhd x_2)(x_0 \not\doteq y \vee x_1 \not\doteq y) \leftrightsquigarrow p \cdot q \cdot \Box(\neg p \vee \neg q)}{(\exists u \rhd x_0)(\exists v \rhd x_1)(\forall y \rhd x_2)(u \not\doteq y \vee v \not\doteq y) \leftrightsquigarrow \Diamond p \cdot \Diamond q \cdot \Box(\neg p \vee \neg q)}}{(\exists u \rhd x_0)(\exists v \rhd x_0)(\forall y \rhd x_0)(v \not\doteq y \vee u \not\doteq y) \leftrightsquigarrow \Diamond p \wedge \Diamond q \wedge \neg\Diamond(p \wedge q)}$$

The formula $(\exists u \rhd x_0)(\exists v \rhd x_0)(\forall y \rhd x_0)(u \not\doteq y \vee v \not\doteq y)$ is equivalent in predicate logic to $(\exists u \rhd x_0)(\exists v \rhd x_0)(u \not\doteq v)$. Hence $\Diamond p \wedge \Diamond q \to \Diamond(p \wedge q)$ corresponds to $(\forall u \rhd x_0)(\forall v \rhd x_0)(u \doteq v)$.

A formula satisfying the conditions of the theorem will be called a (**modal**) **Sahlqvist formula**. A logic axiomatizable by Sahlqvist formulae is called a **Sahlqvist logic**. We note that formulae of the form $\varphi \to \psi$, where both φ and ψ are positive and free of $\Box$, are all Sahlqvist formulae. There are some stronger versions of this theorem. For example the following characterization due to Johan van Benthem [**10**], which uses the notions of positive and negative *occurrences* of a variable. These are defined as follows. Let φ be a formula with n occurrences of p. For ease of reference, we consider the different occurrences as being numbered from 0 to $n-1$. Replace each occurrence by a distinct variable p_i, and call the resulting formula φ^x. p **occurs positively at** j in φ if φ^x is positive in p_j. p **occurs negatively at** j if it does not occur positively at j. For example, in $\varphi = p \wedge \Diamond\neg p$, then the first occurrence of p is positive and the second occurrence negative. We have $\varphi^x = p_1 \wedge \Diamond\neg p_2$, which is positive in p_1 and negative in p_2.

Definition 5.5.7. *Call a formula φ a **Sahlqvist–van Benthem formula** if for all occurring variables p either (i) no positive occurrence of p in φ is in a subformula of the form $\psi \wedge \chi$ or $\Box_j\psi$ if that subformula is in the scope of a $\Diamond_k$ or (ii) no negative occurrence of p in φ is in a subformula of the form $\psi \wedge \chi$ or $\Box_j\psi$ if that subformula is in the scope of some $\Diamond_k$.*

Not every Sahlqvist formula is a Sahlqvist–van Bethem formula. For example the formula $\Diamond(p \wedge \Box\Diamond\neg p) \rightarrow (\Diamond\Box p \vee \Box\Box\neg p)$ is Sahlqvist–van Benthem, but not Sahlqvist. It was shown in [**10**] that logics axiomatized by a Sahlqvist–van Benthem formula are canonical. In the next section we will establish that the class of logics axiomatizable by Sahlqvist–van Benthem formulae is actually not larger than the class of Sahlqvist logics, so that this result actually immediately follows.

Here we will show a theorem to the effect that the class of Sahlqvist logics can be axiomatized by simpler axioms than originally described by Sahlqvist. Namely, one can dispense with the operator prefix. This is quite suitable for certain applications. We should perhaps note that our definition of a Sahlqvist formula is not exactly Sahlqvist's own. In fact, he allows only an operator prefix of the form $\Box^k$. (His proof is only for monomodal logics, but is easily extended to polymodal logics.) We have allowed ourselves to define a slightly larger class, which is also somewhat easier to define in a polymodal setting. In view of the next theorem, these differences are rather marginal. The class of Sahlqvist logics remains the same, no matter what definition we choose.

Theorem 5.5.8. *Let $\Box^s$ be a compound modality. Let $\chi = \Box^s(\varphi \rightarrow \psi)$ be a Sahlqvist formula and $q \notin var(\chi)$. Put $\chi^\circ := \Diamond^s(q \wedge \varphi) \rightarrow \Box^s(q \vee \psi)$. χ° is Sahlqvist and $\boldsymbol{K}_\kappa \oplus \chi = \boldsymbol{K}_\kappa \oplus \chi^\circ$.*

Proof. It is clear that χ° is Sahlqvist. Therefore, let us show the second claim. $\varphi \rightarrow \psi$ is Sahlqvist. Let $\varphi \cdot \neg\psi$ describe $\zeta(x, y)[x \cdot y]$. Then $\varphi \wedge \neg\psi$ describes $\zeta(x, x)$, by (**cnt**.), and $\Diamond^s(\varphi \wedge \neg\psi)$ describes $(\exists y \rhd^s x)\zeta(y, y)$. Hence the elementary condition of χ is $(\forall y \rhd^s x)\neg\zeta(y, y)$. Now, $q \cdot \neg q$ describes $(x \not\doteq y)[x \cdot y]$. Hence, by ($\wedge$–**I**.), $q \wedge \varphi \cdot \neg q \wedge \neg\psi$ describes the formula $(\zeta(x, y) \wedge x \not\doteq y)[x \cdot y]$. Therefore, $\Diamond^s(q \wedge \varphi) \cdot \Diamond^s(\neg q \wedge \neg\psi)$ describes

$$(\exists x' \rhd^s x)(\exists y' \rhd^s y)(\zeta(x', y') \wedge x' \not\doteq y')[x \cdot y] \; .$$

$\Diamond^s(q \wedge \varphi) \wedge \Diamond^s(\neg q \wedge \neg\psi)$ describes $(\exists x' \rhd^s x)(\exists y' \rhd x)(\zeta(x', y') \wedge x' \not\doteq y')$, by (**cnt**.). Hence χ°, which is the negation, defines

$$(\forall x' \rhd^s x)(\forall y' \rhd^s)(x' \doteq y' \rightarrow \neg\zeta(x', y'))$$

This is the same as $(\forall x' \rhd^s x)(\neg\zeta(x', x'))$. ∎

Exercise 181. Show Proposition 5.5.1.

Exercise 182. Name formulae which are monotone but not positive, and formulae which are $\wedge$–distributive without being strongly positive.

Exercise 183. Show the soundness of the rule ($\wedge$**–I**$_2$.).

Exercise 184. Suppose φ contains only positive or only negative occurrences of p. Show that either $\mathbf{K}_\kappa \oplus \varphi = \mathbf{K}_\kappa \oplus \varphi[\top/p]$ or $\mathbf{K}_\kappa \oplus \varphi = \mathbf{K}_\kappa \oplus \varphi[\bot/p]$. Thus, to be essential a variable must occur at least once positively and once negatively.

Exercise 185. Show that if φ is Sahlqvist there exists a Sahlvist formula ψ such that $\mathbf{K}_\kappa \oplus \varphi = \mathbf{K}_\kappa \oplus \psi$, and every variable of ψ occurs exactly once positively and once negatively.

Exercise 186. Generally, modal algebras are not closed under infinitary intersections and unions. This can be remedied as follows. Given a modal algebra $\mathfrak{A}$, the completion of $\mathfrak{A}$ is defined by $\mathfrak{Em}\,\mathfrak{A} := ((\mathfrak{A}^+)_\sharp{}^\sharp)_+$. (See Section 4.8.) It consists of all sets which can be generated as arbitrary intersections and unions of sets in $\mathfrak{A}$. Show that if φ is a Sahlqvist–formula then $\mathfrak{A} \vDash \varphi$ implies $\mathfrak{Em}(\mathfrak{A}) \vDash \varphi$.

5.6. Elementary Sahlqvist Conditions

In this section we will characterize those elementary conditions which are determined by axioms of the form considered in Sahlqvist's Theorem. It is clear that all derivable sequents '$\zeta \leftrightsquigarrow \vec{\varphi}$' state a local correspondence and $\zeta \in \mathcal{R}^f$. We will for simplicity always assume that the situation never arises that a world–variable v occurs both free and bound in a subformula. Such a formula is called **clean**. Not clean is $x \lhd_1 y \wedge (\exists y \rhd_2 x)(y \lhd_1 x)$. Every unclean formula is equivalent to a clean formula in the predicate calculus. In a clean formula ζ a variable y is **inherently existential** if (i) all occurrences of y are free or (ii) ζ has a subformula $\eta = (\exists y \rhd_k x)\theta$ which is not in the scope of a universal quantifier. Likewise y is **inherently universal** if either all occurrences of y are free in ζ or ζ contains a subformula $\eta = (\forall y \rhd_k x)\theta$ which is not in the scope of an existential quantifier.

Now recall the notation $x \lhd^s y$ for sets s of sequences of indices $< \kappa$. It means that x can see y through one of the sequences in s. The formula $x \doteq y$ can be represented by $x \lhd^{\{\epsilon\}} y$. We will change our language for frames. Recall from Section 5.1 that the language $\mathcal{S}^f$ is obtained from $\mathcal{R}^f$ by adding $x \lhd^s y$ as primitive expressions. This is only a technical move to simplify the terminology somewhat. The theorem we will prove is the following.

Theorem 5.6.1. *Suppose $\zeta(x)$ is a positive $\mathcal{S}^f$–formula such that every subformula $x \lhd^s y$ contains at least one inherently universal variable. Then there exists a Sahlqvist formula φ which locally corresponds to $\zeta(x)$ in $\mathfrak{Krp} \cup \mathfrak{D}$. Conversely, any Sahlqvist formula corresponds in $\mathfrak{Krp} \cup \mathfrak{D}$ to an elementary formula of this kind.*

A first–order $\mathcal{S}^f$–formula in $\mathcal{L}^f$ satisfying the conditions of the theorem will henceforth be called an **elementary** or **first–order Sahlqvist formula**. Generalizing this somewhat, a formula $\zeta(\vec{x})$ of $\mathcal{S}^f$ is called **Sahlqvist** if it is positive and every atomic subformula contains at least one inherently universal variable. If $\zeta(\vec{x})$ is Sahlqvist, $\neg\zeta(\vec{x})$ will be called **negative Sahlqvist**. On the way to prove Theorem 5.6.1 we will derive some useful facts.

Lemma 5.6.2. *Let $\zeta(\vec{x}, y)$ be a $\mathcal{R}^f$–formula such that every **atomic** subformula contains at most one variable $\notin \vec{x}$. Then there exists a clean $\eta(\vec{x}, y)$ such that $(\forall y)(\forall \vec{x})(\zeta(\vec{x}, y) \equiv \eta(\vec{x}, y))$ in predicate logic and **every** subformula of η has at most one free variable outside of $\vec{x}$.*

Proof. By induction on ζ. We can assume that negation occurs only in front of the atomic formulae. Moreover, we can assume that ζ is clean and that a subformula η is in the scope of a quantifier only if the quantifier binds a free variable of η. The claim holds by virtue of the assumptions for positive and negative atomic subformulae. Now let $\zeta(\vec{x}, y) = \eta_1(\vec{x}, y) \wedge \eta_2(\vec{x}, y)$. By hypothesis, every atomic subformula of ζ has at most one free variable $\notin \vec{x}$. This holds also of the η_i. By induction hypothesis there exist $\delta_1(\vec{x}, y)$ and $\delta_2(\vec{x}, y)$ such that $(\forall y)(\forall \vec{x})(\eta_1(\vec{x}, y) \equiv \delta_1(\vec{x}, y))$ and $(\forall y)(\forall \vec{x})(\eta_2(\vec{x}, y) \equiv \delta_2(\vec{x}, y))$ in predicate logic and every subformula of δ_1 and δ_2 contains at most one free variable $\notin \vec{x}$. Then put $\theta(\vec{x}, y) := \delta_1(\vec{x}, y) \wedge \delta_2(\vec{x}, y)$. This satisfies the claim. θ is clean if δ_1 and δ_2 are. For a subformula θ' of θ either $\theta' = \theta$ or θ' is a subformula of δ_1 or δ_2. In the first case only y is a free variable $\notin \vec{x}$. In the second case we know that every subformula of θ' contains at most one free variable $\notin \vec{x}$. Similarly the case $\zeta(\vec{x}, y) = \eta_1(\vec{x}, y) \vee \eta_2(\vec{x}, y)$ is treated. Next assume $\zeta(\vec{x}, y) = (\forall w \rhd_k v)\theta(\vec{x}, w, y)$. Then $v \in \vec{x}$ or $v = y$. Assume $\theta(\vec{x}, w, y) = \delta_1(\vec{x}, w, y) \wedge \delta_2(\vec{x}, w, y)$. Then distribute the quantifier over the conjunction. The formula $(\forall w \rhd_k v)(\delta_1(\vec{x}, w, y)) \wedge (\forall w \rhd_k v)(\delta_2(\vec{x}, w, y))$ is equivalent to ζ in predicate logic; it is clean and every atomic subformula contains at most one variable $\notin \vec{x}$. Now back to the case of $\zeta = \eta_1 \wedge \eta_2$. Assume $\theta(\vec{x}, w, y) = \delta_1(\vec{x}, w, y) \vee \delta_2(\vec{x}, w, y)$. We may assume by induction hypothesis that every subformula of δ_1 and δ_2 contains at most one free variable outside of $\vec{x}$; this variable is either w or y. Hence several cases may arise.

Case 1. $w \notin \mathit{fvar}(\delta_1)$, $w \notin \mathit{fvar}(\delta_2)$, where $\mathit{fvar}(\zeta)$ denotes the set of variables occurring free in ζ. Then put $\eta := \theta_1(\vec{x}, y) \vee \theta_2(\vec{x}, y)$; η fulfills the requirements.

Case 2. $w \in \mathit{fvar}(\theta_1)$ and $w \in \mathit{fvar}(\theta_2)$. Then consider an atomic subformula containing the variable y. By assumption on ζ it must be of the form $x_i \lhd^s y$ or $y \lhd^s x_i$ for some i and s. By assumption on ζ this cannot happen.

Case 3. $w \in fvar(\theta_1)$ and $w \notin fvar(\theta_2)$. Then

$$\eta(\vec{x}, y) := (\forall w \rhd_k v)\theta_1(\vec{x}, w). \vee .\theta_2(\vec{x}, y) .$$

Then since $y = v$ or $y \in \vec{x}$, η fulfills the claim.
Case 4. $w \notin fvar(\theta_1)$ and $w \in fvar(\theta_2)$. Then

$$\eta(\vec{x}, y) := \theta_1(\vec{x}, y). \vee .(\forall w \rhd_k v)\theta_2(\vec{x}, w) .$$

This fufills the requirements. Proceed dually in the case of an existential quantifier. ∎

The previous theorem shows that if the atomic subformulae are of the form $x \lhd^s y$ with $x \in fvar(\zeta)$, then ζ can be written in an essentially 'modal' way. If we do not require ζ to be clean we can actually arrange that ζ contains very few variables by reusing bound variables every time they are no longer needed. Two more variables than occur free in ζ are therefore needed; in particular, for $\zeta = (\forall x)\eta(x)$ with no free variables, then if ζ is of this form, it has exactly three variables. This follows from the next theorem. It has been shown by Dov Gabbay ([**72**]) (see the exercises).

Proposition 5.6.3 (Gabbay). *Let $\zeta(\vec{x}) \in \mathcal{R}^f$ be Sahlqvist, $\vec{x}$ of length n. Then $\zeta \equiv \eta$ for an η which contains at most $n + 2$ variables.*

Lemma 5.6.4. *Every $\zeta(\vec{x}, y) \in \mathcal{S}^f$ which is negative and in which every subformula has exactly one free variable outside of $\vec{x}$ is derivable in* **Seq**$^+$ *and corresponds to a spone $\vec{\pi} \cdot v$.*

Proof. By induction on $\zeta(\vec{x}, y)$, using the rules ($\wedge$–**I**$_2$.), ($\vee$–**I**.), ($\Diamond$–**I**.) and ($\Box$–**I**.). The starting sequences are $x_0 \not\lhd^s y_0 \leftrightsquigarrow \Box^s p \cdot \neg p$, which in turn are derivable in **Seq**$^+$. ∎

The proof of Theorem 5.6.1 is now quite short. Suppose that $\zeta(x_0)$ is a negative Sahlqvist formula. According to Theorem 5.4.6 it is enough to prove the claim for negative Sahlqvist formulae of the form $(\forall y' \rhd_k x_i)\eta(\vec{x}, y')$. The latter formula results from $\delta := (\forall y' \rhd_k y_0)\eta(\vec{x}, y')$ by applying (**cnt**.). Hence it is enough to show the claim for the latter formula. By Lemma 5.6.2, we can assume that every subformula of δ contains at most one variable $\notin \vec{x}$. Those subformulae with free variables completely in $\vec{x}$ can be moved outside the scope of any quantifier using laws of predicate logic. So the problem is reduced to the case where every subformula of ζ contains exactly one extra free variable. By Lemma 5.6.4, those are derivable in **Seq**$^+$. The proof of Theorem 5.6.1 is now complete.

Let us discuss the theorem to get a better insight into the classes of formulae that are at issue here. Clearly, a somewhat more satisfying result would be one in which we had only the restriction that $\zeta \in \mathcal{R}^f$ and that ζ is positive. Better than that we can never do. (See the next sections.) The really hairy part is the conditions on variables. Moreover, two questions arise. (1.) Why have we chosen $\mathcal{R}^f$ rather than $\mathcal{L}^f$ to begin with? (2.) Why use $\mathcal{S}^f$ rather than $\mathcal{R}^f$ in **Seq**$^+$? The answer to both questions is: this is a matter of utility. For example, it is hard to state what we mean by **positive in** $\mathcal{R}^f$ inside $\mathcal{L}^f$ without actually talking about restricted quantifiers in some hidden form. The second question is somewhat harder to answer, but it will become clear by discussing some cases. Suppose first that $\zeta \in \mathcal{R}^f$ and that ζ contains no existential quantifiers; then the restriction on variables is vacuous. By Theorem 5.4.11 we know already that all elementary conditions are derivable. Next consider the formulae of the form $\forall\exists$ which are not $\forall$. These formulae have existential quantifiers inside universal quantifiers, but not conversely. In this case the condition says that in a subformula $x \lhd^s y$ not both x and y may be existentially bound variables. If we replace the clause $x \lhd^s y$ by $(\exists z \rhd^s x)(z \doteq y)$, then both z and y may be existentially bound so the condition on the variables cannot be stated in the same way. Finally, consider the case where ζ is of the form $\forall\exists\forall$. Here, if we rewrite the clauses $x \lhd^s y$ the quantifier alternations increase; the resulting formula is of the form $\forall\exists\forall\exists$. Moreover, the newly introduced existentials are innermost, that is, closest to the atomic subformulae. Let us now assume that we have eliminated the expressions $x \lhd^s y$. Then both x and y may be existentially bound. But there is a difference between variables that are introduced from rewriting the complex accessibility clauses and the original variables. The former must be bound by an innermost existential which in turn must have a restrictor which is inherently universal. No such restriction applies to the other variables. Although the restriction can be restated in this way, it is clear that this characterization is much less straightforward.

As an application, the following result will be proved.

Theorem 5.6.5. *Let φ be a Sahlqvist–van Benthem formula. Then the logic $\boldsymbol{K}_\kappa \oplus \varphi$ is locally d–persistent and locally elementary in $\mathfrak{Krp} \cup \mathfrak{D}$. Moreover, there exists a Sahlqvist formula ψ such that $\boldsymbol{K}_\kappa \oplus \varphi = \boldsymbol{K}_\kappa \oplus \psi$.*

Proof. We shall show that $\neg\varphi$ corresponds to an elementary negative Sahlqvist–formula. First we rewrite $\neg\varphi$ so that it contains only variables, negated variables, $\Diamond_j$, $\Box_j$, $\wedge$ and $\vee$. Recall the definition of a Sahlqvist–van Benthem formula. The negation of such a formula satisfies the dual of that condition. This is the following condition: for every variable p, either (i) no positive occurrence of p is a subformula of the form $\psi \vee \chi$ or $\Diamond_j\psi$ if that subformula is in the scope of a $\Box_k$ or (ii) no positive occurrence of p is a subformula of the form $\psi \vee \chi$ or $\Diamond_j\psi$ if that subformula is in the scope of a $\Box_k$. By substituting $\neg p$ for p, we can arrange it that for every variable p only (i) obtains. Call such a formula **good**. A formula is good if every subformula $\Box_k\chi$ is either negative or strongly positive. Notice that the set of good formulae is closed under subformulae. We will now show by induction on the constitution of the formulae that each sequence $\vec{\chi} = \chi_0 \cdot \chi_1 \cdot \ldots \cdot \chi_{n-1}$ of good formulae corresponds to a negative Sahlqvist formula. To start, assume every χ_i is either positive or negative. In that case, because the χ_i are good, the positive χ_i are actually strongly positive. Hence $\vec{\chi}$ is a spone, and it corresponds to a negative Sahlqvist–formula. If this does not obtain, $\vec{\chi}$ contains a formula, say χ_0, which is neither positive nor negative. Then one of the following cases obtains. **Case 1.** $\chi_0 = \Diamond_j\tau$. Then $\tau \cdot \chi_1 \cdot \ldots \cdot \chi_{n-1}$ corresponds by induction hypothesis to an elementary negative Sahlqvist formula $\zeta(\vec{x})$. Then $\vec{\chi}$

corresponds to $((\exists y \rhd_j x_0)\zeta[y/x_0])[\vec{x}]$, which is negative, elementary Sahlqvist. **Case 2.** $\chi_0 = \tau_1 \vee \tau_2$. By induction hypothesis and closure of elementary negative Sahlqvist formulae under disjunction. **Case 3.** $\chi_0 = \tau_1 \wedge \tau_2$. Then by induction hypothesis $\tau_1 \cdot \tau_2 \cdot \chi_1 \cdot \ldots \cdot \chi_{n-1}$ corresponds to some $\zeta(\vec{x})$, which is elementary negative Sahlqvist. Then $\vec{\chi}$ corresponds to $\zeta[x_1/x_0]$, which is also elementary negative Sahlqvist. ◻

For future reference we will introduce the **Sahlqvist Hierarchy** to measure the complexity of descriptions given by Sahlqvist formulae in $\mathcal{R}^f$. This will be measured roughly by the number of quantifier alternations occurring in the formula. This can be defined as follows. Let η be an occurrence of a subformula of ζ. Then by replacing t by $x_i \doteq x_i$ and by suitably renaming the occurrences of the variables in ζ we can achieve it that each subformula occurs only once. Let $\zeta^\dagger$ denote a formula resulting from ζ by this operation. $\zeta^\dagger$ need not be unique. Under this condition, the following is well–defined (and does not depend on a particular choice of $\zeta^\dagger$).

$$\begin{array}{lcl}
\textit{sq-rank}(\zeta^\dagger, \zeta^\dagger) & = & 0 \\
\textit{sq-rank}(\zeta^\dagger, \neg\eta) & = & \textit{sq-rank}(\zeta^\dagger, \eta) \\
\textit{sq-rank}(\zeta^\dagger, \eta_1 \wedge \eta_2) & = & \textit{sq-rank}(\zeta^\dagger, \eta_1) \\
\textit{sq-rank}(\zeta^\dagger, \eta_1 \wedge \eta_2) & = & \textit{sq-rank}(\zeta^\dagger, \eta_2) \\
\textit{sq-rank}(\zeta^\dagger, (\exists y \rhd_j x)\eta) & = & \begin{cases} \textit{sq-rank}(\zeta^\dagger, \eta) \\ \qquad \text{if } \textit{sq-rank}(\zeta^\dagger, (\exists y \rhd_j x)\eta) \text{ is odd} \\ \textit{sq-rank}(\zeta^\dagger, \eta) + 1 \\ \qquad \text{if } \textit{sq-rank}(\zeta^\dagger, (\exists y \rhd_j x)\eta) \text{ is even} \end{cases} \\
\textit{sq-rank}(\zeta^\dagger, (\forall y \rhd_j x)\eta) & = & \begin{cases} \textit{sq-rank}(\zeta^\dagger, \eta) \\ \qquad \text{if } \textit{sq-rank}(\zeta^\dagger, (\forall y \rhd_j x)\eta) \text{ is even} \\ \textit{sq-rank}(\zeta^\dagger, \eta) + 1 \\ \qquad \text{if } \textit{sq-rank}(\zeta^\dagger, (\forall y \rhd_j x)\eta) \text{ is odd} \end{cases}
\end{array}$$

Call a formula **constant** if all atomic subformulae are of the form f or $x_i \doteq x_i$. Finally,

$$\textit{sq-rank}(\zeta) := \textit{max}\{\textit{sq-rank}(\zeta^\dagger, \eta) : \eta \in \textit{sf}(\zeta^\dagger), \eta \text{ not constant}\}$$

Here, the maximum over an empty set is defined to be 0. This defines the **rank of an elementary formula**. For example, a universal restricted formula comes out with rank 0, while an existential formula has rank 1. This is desired even though the rank counts quantifier *alternations*. Let us note, namely, that a $\mathcal{R}^f$ formula must contain at least one free variable since all quantifiers are restricted and the free variables are assumed to be quantified universally (though by an unrestricted quantifier). Thus, all formulae invariably start with a universal quantifier, and this causes the asymmetry. Notice further that the rank does not increase in case of a constant formula (even though this has not been noted explicitly in the informal definition, but should be clear from the first clause). This is another deviation from classical quantifier complexity. It is essential for many reasons that constant formulae have zero complexity. If for some reason we want to count the constant formulae as well, we speak of the **pure rank**. It is defined by

$$\mathit{sq\text{-}rank}(\zeta) := \mathit{max}\{\mathit{sq\text{-}rank}(\zeta^{\dagger}, \eta) : \eta \in \mathit{sf}(\zeta^{\dagger})\}$$

We denote by $\mathbf{Sq}_n$ the class of Sahlqvist formulae of rank n and the logics axiomatizable by such formulae by $\mathfrak{Sq}_n$. So, $\Diamond\top \in \mathbf{Sq}_0$, but it has pure rank 1. Likewise, $\mathbf{alt}_1 \in \mathbf{Sq}_0$, $\mathbf{.4} \in \mathbf{Sq}_0$. Finally, it is helpful to distinguish the rank we obtain as above from a rank in which $x \lhd_j y$ is not a primitive formula, but equal to $(\exists z \rhd_j x)(z \doteq y)$. This we call the **special rank**. Notice that the special rank of ζ is equal to the Sahlqvist–rank of ζ rank if the latter is odd or if the atomic formulae are of the form $x \doteq y$ (i. e. not using $\lhd_j$), and $= \mathit{sq\text{-}rank}(\zeta) + 1$ otherwise. This is so, because for the special rank we only have to eliminate the formulae $x \lhd_j y$, introducing an existential quantifier.

Notes on this section. Let FO^k be the set of expressions of predicate logic in which at most k distinct variables occur. It is known that FO^3 is generally undecidable. However, M. Mortimer has shown in [**158**] that FO^2 has the finite model property and is therefore decidable. Given φ, the size of model of a minimal model for φ is exponential in the length of φ. Hence polymodal **K** is decidable. This does not extend to polymodal logics in general. For there exist finitely axiomatizable Sahlqvist–logics without the finite model property; the first system of this kind is the one of David Makinson in [**145**]. Makinson only shows that his logic is complete but does not possess the finite model property. His paper predates that of Sahlqvist and does therefore not discuss the fact that this logic is Sahlqvist. See also [**106**] for a discussion. Examples of Sahlqvist logics without the finite model property can be found in this book in Section 9.4. These logics are elementary, but the corresponding first–order property is in FO^3. It cannot be in FO^2, by Mortimer's result.

Exercise 187. (Gabbay [**72**].) Show Proposition 5.6.3.

Exercise 188. Show that $\mathfrak{Sq}_n$ is closed under arbitrary unions and finite intersections.

Exercise 189. Show that a Sahlqvist logic can be axiomatized by formulae of the form $\boxplus(\varphi \to \psi)$ where both φ and ψ are positive and each occurring variable occurs exactly once in φ and exactly once in ψ. Moreover, if that axiom corresponds to ζ, the number of variables is at most the number of atomic subformulae in ζ.

Exercise 190. A naive approach to correspondence is to take a formula and regard the variables as referring to worlds rather than sets of worlds. For example, in $p \to \Diamond p$ or $p \to \Box\Diamond p$ we can get the desired first–order correspondent by thinking of p as denoting a single world. This is not in general a correct approach (for example, it fails with the Geach Axiom). Show however, that if $\varphi \to \psi$ is Sahlqvist, and φ contains no box–modalities and each variable in the antecedent at most once, then pretending p to denote single worlds yields a correct first–order correspondent. *Hint.* Consider valuations of the form $\beta(p) = \{w\}$ and show that they can detect a failure of the axiom.

Exercise 191. Show that the Sahlqvist–van Benthem formulae are Seq^{+}–derivable.

5.7. Preservation Classes

The next two sections will require some techniques from model theory, though quite basic ones. In contrast to the previous sections, which characterized those statements '$\zeta \leftrightsquigarrow_{\mathfrak{X}} \vec{\varphi}$' which are valid, we will now derive facts about the correspondence statements that are *not* valid for any ζ. We will prove in this section that if a logic is complete and closed under elementary equivalence it is canonical, which is the same as being d–persistent, by Theorem 4.8.6. Whether the converse holds is unknown and has resisted any attempt to solve it. Moreover, in the next section we will elucidate the connection between the syntactic form of elementary conditions and persistence with respect to a class. Both questions receive only partial answers; for the most interesting class, $\mathfrak{Krp} \cup \mathfrak{D}$, they are in effect still unsolved. The first result is that when $\mathfrak{X}$ contains all Kripke–frames and is $\mathcal{L}^{e}$–definable, any $\mathfrak{X}$–persistent logic is elementary. Furthermore, we will show that if $\mathfrak{X}$ includes the class of Kripke–frames and if ζ is internally describable then ζ is equivalent to a positive and restricted formula. So all that is needed in order to show that Sahlqvist's Theorem is optimal is to derive the condition on variables occurring in atomic subformulas. No one has solved that yet.

Recall now from model theory the construction of an *ultraproduct*. In connection with algebra this construction has been introduced in Section 4.1. An ultraproduct can be defined for Kripke–frames and for generalized frames as well. However, these are two different constructions. One is the ultraproduct of Kripke–frames defined in the usual way. The other is the ultraproduct of Kripke–frames viewed as general frames, i. e. as full frames. Take an index set I, an ultrafilter U on I and a family $\langle \mathfrak{f}_i : i \in I\rangle$ of Kripke–frames. The **ultraproduct** of the $\mathfrak{f}_i$ modulo U, denoted by $\prod_U \mathfrak{f}_i$ is defined as follows. The worlds are equivalence classes of sequences $\vec{w} = \langle w_i : i \in I\rangle$ modulo $\approx$, which is defined by $\vec{v} \approx \vec{w}$ iff $\{i : v_i = w_i\} \in U$. We write $\vec{w}_U$ for $\{\vec{v} : \vec{v} \approx \vec{w}\}$ and $\prod_U f_i$ for the set $\{\vec{v}_U : \vec{v} \in \mathsf{X}_{i\in I} f_i\}$. (Mostly, we will write $\vec{w}$ rather than $\vec{w}_U$ if no confusion arises.) We put $\vec{v}_U \lhd_j \vec{w}_U$ iff $\{i : v_i \lhd_j w_i\} \in U$. This definition does not depend on the choice of representatives, as is easily checked. Finally, $\prod_U \mathfrak{f}_i := \langle \prod_U f_i, \langle \lhd_j : j < \kappa\rangle\rangle$. Now to the ultraproduct of generalized frames. Let $\mathfrak{F}_i$, $i \in I$, be a family of frames. The ultraproduct $\prod_U \mathfrak{F}_i$ of the frames is defined as follows. The underlying Kripke–frame of $\prod_U \mathfrak{F}_i$ is the frame $\prod_U \mathfrak{f}_i$, where $\mathfrak{f}_i := (\mathfrak{F}_i)_\sharp$. The internal sets are as equivalence classes of sequences $\vec{a} = \langle a_i : i \in I\rangle$ modulo $\approx$, where $\vec{a} \approx \vec{b}$ iff $\{i : a_i = b_i\} \in U$. We write $\vec{a}_U$ (or mostly simply $\vec{a}$) for the set $\{\vec{b} : \vec{b} \approx \vec{a}\}$. Furthermore, $\vec{w}_U \in \vec{a}_U$ iff $\{i : w_i \in a_i\} \in U$. It is straightforward to verify that this definition does not depend on the choice of representatives of the classes. This is the only nontrivial fact here. Notice that elementhood really has to be defined, so we had actually better use a different symbol here. In the same way as in model theory we can conclude that $\prod_U \mathfrak{F}_i$ is a frame. Namely, by Proposition 5.7.1 below the so defined set of internal sets is closed under the operations.

Proposition 5.7.1. *Let $\mathfrak{F}_i$, $i \in I$, be a family of frames, and U an ultrafilter over I. Let $\prod_U \mathfrak{F}_i$ be the ultraproduct of the $\mathfrak{F}_i$ with respect to U. Then (i) $\vec{w} \in -\vec{a}$ iff $\vec{w} \notin \vec{a}$, (ii) $\vec{w} \in \vec{a} \cap \vec{b}$ iff $\vec{w} \in \vec{a}$ and $\vec{w} \in \vec{b}$, (iii) $\vec{w} \in \blacklozenge_j \vec{a}$ iff there is a $\vec{v}$ such that $\vec{w} \lhd_j \vec{v}$ and $\vec{v} \in \vec{a}$.*

The proof of this theorem is left as an exercise. Moreover, the identity map is an isomorphism between $(\prod_U \mathfrak{F}_i)_+$ and $\prod_U (\mathfrak{F}_i)_+$. This is not hard to see. A valuation on an ultraproduct can be seen as the equivalence class of a sequence $\vec{\gamma}$ $(\vec{\iota})$ of valuations on the individual factors.

Theorem 5.7.2. *Let $\mathfrak{F}_i$ be a family of Kripke–frames and $\zeta \in \mathcal{L}^e$. Then $\langle \prod_U \mathfrak{F}_i, \vec{\gamma}, \vec{\iota}\rangle \vDash \zeta$ iff $\{i : \langle \mathfrak{F}_i, \gamma_i, \iota_i\rangle \vDash \zeta\} \in U$.*

Proof. Analogous to the elementary case. For example, let $\zeta = (\exists p)\eta$. Then $\langle \prod_U \mathfrak{F}_i, \vec{\gamma}, \vec{\iota}\rangle \vDash \zeta$ iff there is a valuation $\vec{\gamma'} = \langle \gamma'_i : i \in I\rangle$ different from γ at most in p such that $\langle \prod_U \mathfrak{F}_i, \vec{\gamma'}, \vec{\iota}\rangle \vDash \eta$ iff there is a valuation $\vec{\gamma'}$ different from γ at most in p such that $\{i : \langle \mathfrak{F}_i, \gamma'_i, \iota_i\rangle \vDash \eta\} \in U$ iff $\{i : \langle \mathfrak{F}_i, \gamma_i, \iota_i\rangle \vDash (\exists p)\eta\} \in U$. ∎

Lemma 5.7.3 (Goldblatt). *The ultraproduct $\prod_U \mathfrak{f}_i$ is a generated subframe of the ultrapower $\prod_U(\bigoplus_{i \in I} \mathfrak{f}_i)$.*

Proof. Take $h : \vec{w} \mapsto \{\langle i, w_i\rangle : i \in I\}$. Then $\vec{w} \lhd_j \vec{v}$ iff $\{i : w_i \lhd_j v_i\} \in U$. However, $\{i : w_i \lhd_j v_i\} = \{i : \langle i, w_i\rangle \lhd_j \langle i, v_i\rangle\}$, and so $\vec{w} \lhd_j \vec{v}$ iff $h(\vec{w}) \lhd_j h(\vec{v})$. Similarly it is shown that $\vec{v} = \vec{w}$ iff $h(\vec{v}) = h(\vec{w})$, so h is injective. And similarly for the other properties. It is easy to see that the map defines a p–morphism, for if $h(\vec{w}) \lhd_j \vec{v}$ then for almost all i, $\langle i, w_i\rangle \lhd_j \langle i, v_i\rangle$. In that case, define $\vec{v}'$ by $\langle i, v'_i\rangle := \langle i, v_i\rangle$ if $w_i \lhd_j v_i$ and $\langle i, v'_i\rangle := \langle i, w_i\rangle$ otherwise. We have $\vec{v}' \approx \vec{v}$, so they are equal in the ultraproduct, and for $\vec{u} = \langle v'_i : i \in I\rangle$ we have $h(\vec{u}) = \vec{v}'$. ∎

Corollary 5.7.4. *Any class of Kripke–frames closed under generated subframes, disjoint unions, isomorphic copies and ultrapowers is closed under ultraproducts.*

A class of first–order structures is called **elementary** if it is characterized by a single sentence, and Δ–**elementary** if it is characterized by a set of sentences. It is called Σ–**elementary** if it is a union of elementary classes, and ΣΔ–**elementary** if it is a union of Δ–elementary classes. It can be shown that a class is closed under elementary equivalence iff it is ΣΔ–elementary. Moreover, Theorem 6.1.15 of [**45**] states that two structures are elementarily equivalent iff some ultrapowers of the structures are isomorphic. Hence a class is elementary iff it and its complement are closed under ultraproducts and isomorphisms. In the present case, the hierarchy collapses.

Theorem 5.7.5 (van Benthem). *Let $\mathfrak{X}$ be a $\Sigma\Delta$–elementary class of Kripke–frames closed under generated subframes and disjoint unions. Then $\mathfrak{X}$ is Δ–elementary. Under the same conditions $\mathfrak{X}$ is elementary if it is Σ–elementary.*

Proof. $\mathfrak{X}$ is closed under elementary equivalence and hence under ultrapowers. By Lemma 5.7.3 it is also closed under ultraproducts. By a standard model theoretic result $\mathfrak{X}$ is Δ–elementary. (For example, see 4.1.12(i) in [**45**]. There a class is called elementary if it is Δ–elementary in our sense.) Now assume that $\mathfrak{X}$ is Σ–elementary. Then its complement is Δ–elementary hence closed under ultraproducts as well. So, $\mathfrak{X}$ is elementary. ◻

Theorem 5.7.6. *A modal formula globally corresponds in $\mathfrak{Krp}$ to an $\mathfrak{L}^f$–sentence iff it is preserved in $\mathfrak{Krp}$ under $\mathfrak{L}^f$–elementary equivalence iff it is preserved under ultrapowers of Kripke–frames.*

Proof. If φ corresponds to $\zeta \in \mathfrak{L}^f$, ζ a sentence, then it is closed under elementary equivalence. And if it is preserved under elementary equivalence then it must be closed under ultrapowers. By the previous corollary we have that the class of Kripke–frames for φ, $\mathsf{Krp}(\varphi)$, in addition to being closed under isomorphic copies and generated subframes is also closed under ultraproducts. Now, consider the complement of $\mathsf{Krp}(\varphi)$. It is definable by $(\exists x_0)(\exists \vec{p})(x_0 \in \varphi)$. By 4.1.14 in [**45**] we have that the class of models of this formula is closed under ultraproducts, too. Both classes are therefore closed under ultraproducts and isomorphic images. Therefore $\mathsf{Krp}(\varphi)$ is elementary. Hence φ corresponds to an $\mathfrak{L}^f$–sentence. ◻

We can immediately boost this up. Let $\mathfrak{X}$ be class of general frames which can be defined by a set Φ of $\mathfrak{L}^e$–sentences. Examples are the classes $\mathfrak{G}$, $\mathfrak{Df}$, $\mathfrak{Ti}$, $\mathfrak{R}$. We can define a modal logic Λ to be **Φ–persistent** if for all frames $\mathfrak{F}$ such that $\mathfrak{F} \vDash \Phi$ we can infer $\mathfrak{f} \vDash \Lambda$ from $\mathfrak{F} \vDash \Lambda$.

Theorem 5.7.7 (Goldblatt). *Let Φ a set of $\mathfrak{L}^e$–sentences true in all Kripke–frames. Then if a finitely axiomatizable logic Λ is Φ–persistent, it is globally elementary in $\mathfrak{Krp}$.*

Proof. We want to proceed as before and show that the class of frames for Λ and its complement are closed under ultraproducts. For the complement there is no problem, we appeal again to 4.1.14 of [**45**]. For the class itself notice that if we take the ultraproduct of the Kripke–frames as full frames then we may from $\mathfrak{f}_i^\sharp \vDash \Lambda$ still conclude $\prod_U \mathfrak{f}_i^\sharp \vDash \Lambda$. The latter is not in general a Kripke–frame. But we can use the fact that the underlying frame is in fact $\prod_U \mathfrak{f}_i$, the desired ultraproduct, and that it is in the class defined by Φ, by assumption. Furthermore, we have Φ–persistence, so $\prod_U \mathfrak{f}_i \vDash \Lambda$. ◻

This shows that g–persistent, df–persistent and ti–persistent logics are Δ–elementary and also the

Corollary 5.7.8 (Fine). *Let Λ be r–persistent. Then Λ is Δ–elementary.*

We half–complete the circle by showing the following famous theorem of Kit Fine in [**65**] below. To prove it, let us recall some notions from model theory. Let x_i, $i < n$, be a set of variables, $\mathfrak{M}$ a first–order structure for a first–order language $\mathcal{L}$. Let $\Gamma = \{\gamma_i(\vec{x}) : i \in I\}$ be a set of formulae in the variables x_i, $i < n$. An n–tuple $\vec{u} := \langle u_i : i < n\rangle$ **realizes** Γ if $\mathfrak{M} \vDash \gamma[\vec{u}]$ for all $\gamma \in \Gamma$. Γ is **finitely realizable** in $\mathfrak{M}$ if every finite subset of Γ is realized by some $\vec{u}$. $\mathfrak{M}$ is n–**saturated** if every finitely realizable set in n variables can be realized. $\mathfrak{M}$ is $\aleph_0$–**saturated** if it is n–saturated for every $n < \aleph_0$. It is a well–known fact of model theory that for every $\mathcal{L}$–structure $\mathfrak{M}$ there exists an elementary extension $\mathfrak{G}$ which is $\aleph_0$–saturated.

Definition 5.7.9. *A frame $\mathfrak{F}$ is called* ***modally 1–saturated*** *if for every set $U \subseteq \mathbb{F}$ with the finite intersection property, $\bigcap U \neq \varnothing$. $\mathfrak{F}$ is called* ***modally 2–saturated*** *if for every $j < \kappa$ and every set $U \subseteq \mathbb{F}$ such that $x \in \blacklozenge_j U$ there is a $y \rhd_j x$ such that $y \in \bigcap U$. $\mathfrak{F}$ is* ***modally saturated*** *if it is modally 1– and 2–saturated.*

Clearly, a frame is 1–saturated iff it is compact. Recall from Section 4.6 the notion of the refinement map. We show that on modally saturated frames the refinement map is a p–morphism, and the image is a descriptive frame.

Lemma 5.7.10. *Let $\mathfrak{F}$ be modally saturated. Put $U_x := \{a \in \mathbb{F} : x \in a\}$. Define $\sim\, \subseteq f \times f$ by $x \sim y$ iff $U_x = U_y$. Then $\sim$ is a net on $\mathfrak{F}$, and $\mathfrak{F}/{\sim}$ is descriptive. The algebra of sets of $\mathfrak{F}$ is isomorphic to the algebra of sets of $\mathfrak{F}/{\sim}$.*

Proof. Let $v \sim v'$ and $v \lhd_j w$. Then $v \in \bigcap \blacklozenge_j U_w$. Then, by definition of $\sim$, $v' \in \bigcap \blacklozenge_j U_w$. By 2–saturation, there exists a $w' \in \bigcap U_w$ such that $v' \lhd_j w'$. Again by definition of $\sim$ and of U_w, $w' \sim w$. The algebra of sets over $\mathfrak{F}/{\sim}$ is defined as the set of sets $[c] := \{[x] : x \in c\}$. Given two internal sets $b, c \in \mathbb{F}$ there exists an x such that $x \in b$, but $x \notin c$. Then $[x] \in [b]$ but $[x] \notin [c]$. Hence $b \mapsto [b]$ is bijective. This shows the last of the claims. In $\mathfrak{F}/{\sim}$ we have $U_x = U_y$ iff $x = y$, and so it is refined. Moreover, it is compact, since $\mathfrak{F}$ is. Finally, let $[v] \ntriangleleft [w]$. Then for no $w' \sim w$, $v \lhd_j w'$. By 2–saturatedness of $\mathfrak{F}$, therefore, $v \notin \bigcap \blacklozenge_j U_w$. So there exists a $c \in U_w$ such that $v \notin \blacklozenge_j c$. ◻

Theorem 5.7.11 (Fine). *If Λ is $\mathfrak{Krp}$–complete and $\mathfrak{Krp}$–ΣΔ–elementary then Λ is $\aleph_1$–canonical.*

Proof. If φ is Λ–consistent there is a Λ–model $\langle \mathfrak{f}_\varphi, \beta_\varphi, x_\varphi\rangle \vDash \varphi$. Let $\mathfrak{f} := \bigoplus_\varphi \mathfrak{f}_\varphi$ and $\beta := \bigoplus_\varphi \beta_\varphi$. Now let $\mathbb{F}$ consist of all the sets of the form $\overline{\beta}(\psi)$ for some ψ. Then $\langle \mathfrak{f}, \mathbb{F}\rangle$ is a general frame, since $\mathbb{F}$ is closed under the usual operations. Also, $\mathfrak{F}_+ \cong \mathfrak{Fr}_\Lambda(\aleph_0)$. For consider the map $\varphi \mapsto \overline{\beta}(\varphi)$; we show that it is an isomorphism. If $\mathfrak{Fr}_\Lambda(\aleph_0) \vDash \varphi \approx \psi$ then $\Lambda \vdash \varphi \leftrightarrow \psi$ and so $\mathfrak{F} \vDash \varphi \leftrightarrow \psi$, thus $\overline{\beta}(\varphi) = \overline{\beta}(\psi)$.

Further, if $\mathfrak{Fr}_\Lambda(\aleph_0) \nvDash \varphi \approx \psi$ then $\Lambda \nvdash \varphi \leftrightarrow \psi$, then $\chi := \neg(\varphi \leftrightarrow \psi)$ has a model $\langle \mathfrak{f}_\chi, \beta_\chi, x_\chi \rangle$. Hence the world $\langle \chi, x_\chi \rangle \in \mathfrak{f}$ is a member of $\overline{\beta}(\chi)$, that is, $\langle \mathfrak{F}, \beta \rangle \nvDash \neg\chi$, or equivalently, $\langle \mathfrak{F}, \beta \rangle \nvDash \varphi \leftrightarrow \psi$. Now adjoin for each $c \in \mathfrak{F}$ a unary predicate $\underline{c}$ to $\mathcal{L}^f$. This defines the language $\mathcal{L}^f(\mathfrak{F})$. Expand $\mathfrak{f}$ to an $\mathcal{L}^f(\mathfrak{F})$–structure $\mathfrak{f}^\circ$ by interpreting $\underline{c}$ as the set c itself. (This allows to forget the set structure on $\mathfrak{F}$ for a while.) There exists an elementary extension of $\mathfrak{f}^\circ$ in the language $\mathcal{L}^f(\mathfrak{F})$, denoted by $\mathfrak{g}^\circ$, which is $\aleph_0$–saturated. Put $[\underline{c}] := \{w \in g : \underline{c}(w)\}$. The $[\underline{c}]$ are closed under intersection, complement and $\blacksquare_j$. In particular, we have $-[\underline{c}] = [\underline{-c}]$, $[\underline{c}] \cap [\underline{d}] = [\underline{c \cap d}]$ and $\blacksquare_j[\underline{c}] = [\underline{\blacksquare_j c}]$. For these are elementary statements which hold in $\mathfrak{f}^\circ$, thus they hold in $\mathfrak{g}^\circ$. The map $k : c \mapsto [\underline{c}]$ is therefore an isomorphism of the algebras. Let $\mathfrak{g}$ be the $\mathcal{L}^f$–reduct of $\mathfrak{g}^\circ$. Put $\mathbb{G} := \{[\underline{c}] : c \in \mathbb{F}\}$. We have managed to get a structure $\mathfrak{G} = \langle \mathfrak{g}, \mathbb{G} \rangle$ such that $\mathfrak{G}_+ \cong \mathfrak{Fr}_\Lambda(\aleph_0)$ with $\mathfrak{g}^+$ $\aleph_0$–saturated. We show that $\mathfrak{G}$ is modally saturated. Namely, if $A \subseteq \mathbb{G}$ has the finite intersection property, then $\bigcap A \neq \varnothing$. For under the given assumption, $\{\underline{c}(x_0) : [\underline{c}] \in A\}$ is finitely satisfiable; by saturatedness of $\mathfrak{g}^\circ$ there exists a w such that $\underline{c}(w)$ for all $[\underline{c}] \in \mathbb{G}$. Hence $w \in \bigcap A$. So, $\mathfrak{G}$ is 1–saturated. Second, if A is a set such that for every finite subset A_0, $v \in \blacklozenge_j \bigcap A_0$, then there exists a w such that $v \lhd_j w$ and $w \in c$ for all $c \in A$. For by assumption on A, the set $\{v \lhd_j y \wedge \underline{c}(y) : [\underline{c}] \in A\}$ is finitely satisfiable. Hence by saturatedness there exists a w such that $v \lhd_j w$ and $\underline{c}(w)$ for all $[\underline{c}] \in A$. So, $\mathfrak{G}$ is modally 2–saturated. By Lemma 5.7.10 the refinement map is a p–morphism from $\mathfrak{G}$ onto a descriptive frame whose set algebra is isomorphic to $\mathfrak{Fr}_\Lambda(\aleph_0)$. Hence there is a p–morphism $\mathfrak{G} \twoheadrightarrow \mathfrak{Fr}_\Lambda(\aleph_0)$. Now, for all φ we had $\mathfrak{f}_\varphi \vDash \Lambda$ and so $\mathfrak{f} \vDash \Lambda$. $\mathfrak{g}$, being an elementary extension, also satisfies Λ, by assumption that $\mathsf{Krp}\,\Lambda$ is closed under elementary equivalence. Finally, $\mathfrak{Fr}_\Lambda(\aleph_0)_\sharp \vDash \Lambda$ by closure under p–morphisms. ◻

The proof works analogously for any cardinal $\alpha \geq \aleph_1$. Using this theorem we can obtain a partial converse of Theorem 5.7.11. Namely, if we have a logic which is $\aleph_1$–canonical and elementary, then we can get $\mathfrak{Fr}_\Lambda(\alpha)$ in a similar process from finite models. Moreover, the following holds as well.

Theorem 5.7.12 (Fine). *Suppose that Λ is $\mathfrak{Krp}$–complete and $\mathfrak{Krp}$–$\Sigma\Delta$–elementary. Then Λ is canonical.*

The connection between canonicity and elementarity is a very delicate one as Fine shows in [**65**]. Consider the logic $\Theta := \mathbf{K} \oplus \varphi$ where

$$\varphi := \Diamond\Box p \to \Diamond\Box(p \wedge q) \vee \Diamond\Box(p \wedge \neg q)$$

We claim the following: (a) Θ is canonical, (b) $\mathsf{Krp}(\Theta)$ is not $\Sigma\Delta$–elementary and (c) Θ is complete with respect to some elementary class of Kripke–frames. So, canonicity does in general not imply elementarity. It is believed until today that it does imply completeness with respect to some (Δ–)elementary class of frames. Θ is a case in point. Let us prove the stated properties of Θ. Consider the formula

$$\varepsilon(x) := (\forall y \rhd x)(\exists z \rhd x)(\forall u, v \rhd z)(u \doteq v \wedge y \lhd z)$$

Lemma 5.7.13. *Let $\mathfrak{F}$ be a frame. If $\mathfrak{F} \vDash (\forall x)\varepsilon(x)$ then $\mathfrak{F} \vDash \varphi$.*

Proof. Assume $\mathfrak{F} \vDash (\forall x)\varepsilon(x)$ and let β and x be such that $\langle \mathfrak{F}, \beta, x\rangle \vDash \Diamond\Box p$. Then for some $y \rhd x$ we have $\langle \mathfrak{F}, \beta, y\rangle \vDash \Box p$. Now, by our assumptions, there is a $z \rhd x$ such that z is either a dead end or has exactly one successor. If z is a dead end, $\Box(p \wedge q)$ is true at z and so $\Diamond\Box(p \wedge q)$ is true at x. So, let us assume that z actually has a successor, u. Then $y \lhd u$. Since $\Box p$ holds at y, p is true at u. Now either q is true at u or not. In the first case, $\langle \mathfrak{F}, \beta, z\rangle \vDash \Box(p \wedge q)$ and so $\langle \mathfrak{F}, \beta, x\rangle \vDash \Diamond\Box(p \wedge q)$. In the second case $\langle \mathfrak{F}, \beta, x\rangle \vDash \Diamond\Box(p \wedge \neg q)$, as required. ∎

Lemma 5.7.14. *Let $\mathfrak{F}$ be a canonical Θ–frame. Then $\mathfrak{F}$ satisfies $(\forall x)\varepsilon(x)$.*

Proof. Since $\mathfrak{F}$ is canonical, worlds are actually maximally consistent sets of formulae. Let X be a point of $\mathfrak{F}$. In case that X has no successor or sees a dead end (that is, in case $\Box\bot \in X$ or $\Diamond\Box\bot \in X$) ε holds trivially of X. So, let us consider the case where this is not so. Then pick a successor Y. We have to show that there is a Z with exactly one successor, call it U, such that $Y \lhd U$. Now choose an enumeration χ_n of the language. Inductively we define formulae α_n as follows. $\alpha_0 := \top$. $\alpha_{n+1} := \alpha_n \wedge \chi_n$ if $\Diamond\Box(\alpha_n \wedge \chi_n) \in X$ and $\Box(\alpha_n \wedge \chi_n) \in Y$, and $\alpha_{n+1} := \alpha_n \wedge \neg\chi_n$ if $\Diamond\Box(\alpha_n \wedge \neg\chi_n) \in X$. Finally, we let U be the MP–closure of the set $\{\alpha_n : n \in \omega\}$. By force of the axiom φ, this is actually well–defined and if $\Diamond\Box\alpha_n \in X$ then also $\Diamond\Box\alpha_{n+1} \in X$. Moreover, $\Diamond\Box\alpha_0 \in X$, and so $\Diamond\Box\alpha_n \in X$ for all $n \in \omega$. It follows that U is consistent. For if not, for some n, α_n is inconsistent, that is, $\alpha_n \vdash \bot$. But since $\Diamond\Box\alpha_n \in X$, also $\Diamond\Box\bot \in X$, which we have excluded. Furthermore, by construction, for every n either $\chi_n \in U$ or $\neg\chi_n \in U$. So, U is a maximally consistent set, and so a world of $\mathfrak{F}$. Also, for every $\Box\chi \in Y$, $\chi \in U$, again by construction. It follows that $Y \lhd U$. Finally, put $A := \{\Box\chi : \chi \in U\}$ and $D := \{\chi : \Box\chi \in X\}$. We claim that $A \cup D$ is consistent. If not, there is $\delta \in D$ and $\Box\alpha \in A$ such that $\delta; \Box\alpha \vdash \bot$. ($A$ and D are closed under conjunction.) So, $\Box\alpha \vdash \neg\delta$ and hence $\Diamond\Box\alpha \vdash \Diamond\neg\delta$. Since $\Diamond\Box\alpha \in X$ we also have $\neg\Box\delta \in X$, against the definition of D. So, $A \cup D$ is consistent and is contained in a maximally consistent set Z. Then if $\Box\chi \in Z$, $\chi \in U$ and so $Z \lhd U$. Furthermore, if $Z \lhd V$ then $A \subseteq V$, from which $U \subseteq V$. Since U is maximally consistent, $U = V$. Finally, assume that $\Box\chi \in X$. Then $\chi \in D$ and so $\chi \in Z$, by construction. This shows that Z has the desired properties. ∎

Corollary 5.7.15. *Θ is canonical and complete with respect to an elementary class of frames.*

Proof. Let $\mathfrak{F}$ be some canonical frame for Θ. Then by Lemma 5.7.14 $\mathfrak{F} \vDash (\forall x)\varepsilon(x)$, whence also $\mathfrak{F}_\sharp \vDash (\forall x)\varepsilon(x)$. By Lemma 5.7.13, $\mathfrak{F}_\sharp \vDash \varphi$. Hence Θ is canonical. Clearly, this shows that Θ is complete with respect to the class of Kripke–frames satisfying $(\forall x)\varepsilon(x)$. ∎

Lemma 5.7.16. *$\mathsf{Krp}(\Theta)$ is not $\Sigma\Delta$–elementary.*

Proof. Let F be a nonprincipal ultrafilter on ω. Let O be the union of ω, F and $\{F\}$. It is easy to see that this union is disjoint. Finally, put $\lhd := \{\langle x, y\rangle \in O^2 : y \in x\}$ and $\Omega := \langle O, \lhd\rangle$. It is not hard to show that $\Omega \vDash \varphi$. O is not countable. By the downward Löwenheim–Skolem theorem, Ω has an elementary countable submodel $\mathfrak{p} = \langle P, \lhd_P\rangle$, where $P \subseteq O$ and $\lhd_P = \lhd \cap P^2$. We will show that $\mathfrak{p} \nvDash \varphi$. This establishes the claim. The following properties of $\mathfrak{p}$ are easy to establish. (1) $\mathfrak{p}$ contains the root of Ω. (2) $P \cap F$ is infinite. (3) For any two $M, M' \in P \cap F$ the set $succ(M) \cap succ(M')$ is infinite. Now let $P \cap F = \{M_i : i \in \omega\}$ be an enumeration of $P \cap F$. We can assume by (3) that there are sequences $\langle a_i : i \in \omega\rangle$ and $\langle b_i : i \in \omega\rangle$ of natural numbers which are all distinct such that all $a_i, b_i \in M_0 \cap P$, and $M_n \lhd a_n, b_n$. Put $\beta(p_0) := M_0 \cap P$ and $\beta(p_1) := \{a_i : i \in \omega\}$. Then $\langle \mathfrak{p}, \beta, F\rangle \vDash \Diamond\Box p_0$ since $\langle \mathfrak{p}, \beta, M_0\rangle \vDash \Box p_0$. Now take an M such that $F \lhd M$. Then for some n, $M = M_n$. Then $M_n \vDash \Diamond(p_0 \wedge p_1)$ since $M_n \lhd a_n$ and $a_n \vDash p_0; p_1$. On the other hand $b_n \vDash p_0; \neg p_1$, and since $M_n \lhd b_n$ we also have $M_n \vDash \Diamond(p_0 \wedge \neg p_1)$. It follows that $\langle \mathfrak{p}, \beta, F\rangle \vDash \neg\Diamond\Box(p_0 \wedge p_1); \neg\Diamond\Box(p_0 \wedge \neg p_1)$. So, $\mathfrak{p} \nvDash \varphi$. $\Box$

Let us close with a characterization of classes of Kripke–frames which are both modally definable and elementarily definable. Those classes must be closed under the standard operations of generated subframes, disjoint unions and p–morphic images. However, notice that we also have in the case of generalized frames the closure under biduals for both the class and its complement. The bidual of a Kripke–frame is not necessarily a Kripke–frame; however, let us define for a Kripke–frame $\mathfrak{f}$ the **ultrafilter extension** $\mathsf{ue}(\mathfrak{f})$ to be the Kripke–frame underlying the bidual of $\mathfrak{f}$, that is, $\mathsf{ue}(\mathfrak{f}) = (\mathfrak{f}^{\sharp}{}_{+})^{+}{}_{\sharp}$. The following has been proved in [**83**].

Theorem 5.7.17 (Goldblatt & Thomason). *A class of Kripke–frames is both modally and elementarily definable iff it is closed under generated subframes, disjoint unions, p–morphic images and ultrafilter extensions, while its complement is also closed under ultrafilter extensions.*

If we analyse the proof of Theorem 5.7.11 we see that it proves that the ultrafilter extension of a frame is a p–morphic image of some ultrapower. This is a rather useful fact.

Theorem 5.7.18. *The ultrafilter extension of a Kripke–frame $\mathfrak{g}$ is a contractum of an ultrapower of $\mathfrak{g}$.*

Notes on this section. The theorem by Kit Fine was proved again by Bjarni Jónsson in [**112**] using methods from universal algebra. Later, Yde Venema has generalized the results as follows (see [**222**]). Call an elementary condition α a *pseudo–correspondent* of χ if α holds in the Kripke–frames underlying each canonical frame for $\mathbf{K} \oplus \chi$, and every frame satisfying α also satisfies χ. Obviously, the formula $(\forall x)\varepsilon(x)$ above is a pseudo–correspondent of φ. Now let $\pi(p)$ be a positive formula in one variable. Then the formula $\pi(p \vee q) \leftrightarrow \pi(p) \vee \pi(q)$ has a first–order pseudo–correspondent. Hence it is canonical, complete with respect to an elementary class of Kripke–frame, but its class of Kripke–frames is in general not $\Sigma\Delta$–elementary. The formula of Kit Fine falls into this class of formulae, as is easily demonstrated.

Exercise 192. Show Proposition 5.7.1.

Exercise 193. Show Theorem 5.7.17.

5.8. Some Results from Model Theory

In this section we will use some techniques from model theory which enable us to derive characterizations of modally definable elementary conditions. For an extensive exposition the reader is referred to van Benthem [**10**] and also for a survey to [**11**]. Here we will basically prove two of the results, which we consider to be the most central.

Definition 5.8.1. *A first–order formula $\alpha(\vec{x})$ on frames is* ***preserved*** *under generated subframes if for each model $\langle \mathfrak{g}, \iota\rangle \vDash \alpha(\vec{x})$ and each $\mathfrak{f} \rightarrowtail \mathfrak{g}$ such that $\iota(\vec{x}) \subseteq f$ also $\langle \mathfrak{f}, \iota\rangle \vDash \alpha(\vec{x})$. $\alpha(\vec{x})$ is* ***reflected*** *under generated subframes if $(\neg\alpha)(\vec{x})$ is preserved under generated subframes. $\alpha(\vec{x})$ is* ***invariant*** *under generated subframes if it is both preserved and reflected under generated subframes.*

The following theorem is an analogue of a theorem by Goldblatt (modeled after Feferman [**59**]).

Theorem 5.8.2. *A first–order formula $\alpha(\vec{x})$ with at least one free variable is invariant under generated subframes iff it is equivalent to a $\beta(\vec{x}) \in \mathcal{R}^f$.*

Proof. Surely, if α is equivalent to a restricted β in the same variables, then α is invariant under generated subframes. The converse needs to be established. Thus assume that $\alpha(\vec{x})$ is invariant under generated subframes, and let α be consistent. Moreover, α contains at least one free variable, say x_0. Let $R(\alpha) := \{\beta(\vec{x}) : \beta \in \mathcal{R}^f \text{ and } \alpha \vDash \beta\}$. Thus if we can show that $R(\alpha) \vDash \alpha$ we are done. Namely, by compactness there is a finite set $\Delta \subseteq R(\alpha)$ such that $\Delta \vDash \alpha$ and so, taking β to be the conjunction of Δ, we have found our desired formula. Thus assume $R(\alpha)$ has a model $\mathfrak{F}_0 = \langle \mathfrak{f}_0, \iota\rangle$. (From now on we will suppress the explicit mentioning of the valuation; thus when we speak of a model $\mathfrak{F}_0$ we mean the frame endowed with a valuation.) The language we are using is called $\mathcal{L}_0$, for future reference. Now form $\mathcal{L}_1$ by adjoining a constant c_i for each variable x_i occurring free in α. Expand the model $\mathfrak{F}_0$ to a model $\mathfrak{F}_1$ by interpreting the constants c_i by $\iota(x_i)$. We claim that the set

$$\Sigma := \{\delta \in \mathcal{R}_1 : \mathfrak{F}_1 \vDash \delta\} \cup \{\alpha[\vec{c}/\vec{x}]\}$$

is consistent. Otherwise there is a finite set — or indeed, a single formula $\delta(\vec{x})$, by closure of the set under conjunction — such that $\alpha[\vec{c}/\vec{x}] \vDash (\neg\delta)[\vec{c}/\vec{x}]$, whence $\neg\delta \in R(\alpha)$. But this contradicts the definition of $R(\alpha)$, by consistency of α. So, $\Sigma \subseteq \mathcal{L}_1$ has a model, $\mathfrak{G}_1$. Moreover, every restricted $\mathcal{L}_1$–sentence is true in $\mathfrak{F}_1$ iff it is is true in $\mathfrak{G}_1$. (We say briefly that $\mathfrak{F}_1$ and $\mathfrak{G}_1$ are r–equivalent.) For, if δ is restricted and holds in $\mathfrak{F}_1$ then it is in Σ and holds in $\mathfrak{G}_1$, too. And if δ fails in $\mathfrak{F}_1$, then $\neg\delta$ holds in $\mathfrak{F}_1$, so $\neg\delta$ holds in $\mathfrak{G}_1$ as well. This is now the starting base. We have two models $\mathfrak{F}_1$ and $\mathfrak{G}_1$ over a common language $\mathcal{L}_1$, such that $\mathfrak{G}_1$ is a model for $\alpha[\vec{c}/\vec{x}]$ and both are r–equivalent. We will now construct sequences $\mathcal{L}_i$ of languages, and $\mathfrak{F}_i$ and $\mathfrak{G}_i$ such that

(1) $\mathfrak{F}_i$, $\mathfrak{G}_i$ are $\mathcal{L}_i$–models.
(2) $\mathfrak{G}_i \vDash \alpha[\vec{c}/\vec{x}]$.
(3) The same $\mathcal{L}_i$–sentences hold in $\mathfrak{G}_i$ and $\mathfrak{F}_i$. ($\mathfrak{F}_i$ and $\mathfrak{G}_i$ are r–equivalent.)
(4) $\mathfrak{F}_{i-1}$ is an $\mathcal{L}_{i-1}$–elementary substructure of $\mathfrak{F}_i$ and $\mathfrak{G}_{i-1}$ is an $\mathcal{L}_{i-1}$–elementary substructure of $\mathfrak{G}_i$.

We have started the construction with $i = 1$. Now let the construction proceed as follows.

Case 1. i is odd. First we define $\mathcal{L}_{i+1}$. Assume that we have a constant c in $\mathcal{L}_i$ and that $\mathfrak{G}_i \vDash c \lhd_j x$ for some $x \in \mathfrak{G}_i$. Then adjoin a constant $\underline{x}$ for x. Do this for all x of this kind. This defines $\mathcal{L}_{i+1}$. $\mathfrak{G}_{i+1}$ is defined as the expansion of $\mathfrak{G}_i$ in which $\underline{x}$ is interpreted as x. Then $\mathfrak{G}_i$ is an $\mathcal{L}_i$–elementary substructure of $\mathfrak{G}_{i+1}$, which we abbreviate by $\mathfrak{G}_i \prec_i \mathfrak{G}_{i+1}$. Now form the set

$$\Sigma := \{\delta \in \mathcal{R}_{i+1} : \mathfrak{G}_{i+1} \vDash \delta\}$$

This set is finitely satisfiable in $\mathfrak{F}_i$. To show this it is enough to see that every formula $\delta \in \Sigma$ is satisfiable in $\mathfrak{F}_i$. Let $\delta \in \Sigma$ be given. Now retract the new constants in $\mathcal{L}_{i+1}$ as follows. Let $\underline{x} \in \mathcal{L}_{i+1} - \mathcal{L}_i$ occur in δ. Then there is a c such that $\mathfrak{G}_i \vDash c \lhd_j x$. Put $\delta_1 := (\exists y \rhd_j c)\delta[y/\underline{x}]$. Continuing this process until there are no free variables outside $\mathcal{L}_i$ left, we get a formula $\delta_\star \in \mathcal{R}_i$. Now since $\mathfrak{G}_i \vDash \delta_\star$ and $\mathfrak{F}_i$ is r–equivalent to $\mathfrak{G}_i$, we get $\mathfrak{F}_i \vDash \delta_\star$. Hence δ is consistent, $\delta_\star$ being the existential closure of δ. Therefore, Σ is consistent. So there is a model $\mathfrak{F}_{i+1}$ in the language $\mathcal{L}_{i+1}$ such that $\mathfrak{F}_i \prec_i \mathfrak{F}_{i+1}$ and $\mathfrak{G}_{i+1}$ and $\mathfrak{F}_{i+1}$ are r–equivalent in $\mathcal{L}_{i+1}$.

Case 2. This step is dual. Now adjoin constants for worlds x such that $\mathfrak{F}_i \vDash c \lhd_j x$, $c \in \mathcal{L}_i$. Interchange the roles of $\mathfrak{F}_i$ and $\mathfrak{G}_i$ in the above construction.

$$\begin{tikzcd}
\mathfrak{F}_1 \arrow[d, equal] & \prec_1 & \mathfrak{F}_2 \arrow[d, equal] & \prec_2 & \mathfrak{F}_3 \arrow[d, equal] & \prec_3 & \mathfrak{F}_4 \arrow[d, equal] & \cdots & \mathfrak{F}^\star \arrow[d, equal] & \mathfrak{F}^\circ \arrow[l] \arrow[d, "f"] \\
\mathfrak{G}_1 & \prec_1 & \mathfrak{G}_2 & \prec_2 & \mathfrak{G}_3 & \prec_3 & \mathfrak{G}_4 & \cdots & \mathfrak{G}^\star & \mathfrak{G}^\circ \arrow[l] \\
\mathcal{L}_1 & & \mathcal{L}_2 & & \mathcal{L}_3 & & \mathcal{L}_4 & \cdots & \mathcal{L}_\omega &
\end{tikzcd}$$

Let $\mathcal{L}_\omega := \bigcup_i \mathcal{L}_i$. Now consider the structure $\mathfrak{F}^\star := \bigcup_i \mathfrak{F}_i$. Since $\mathfrak{F}_i$ is an $\mathcal{L}_i$–elementary substructure and also contained in $\mathfrak{F}_{i+1}$, we can invoke Tarski's Lemma on elementary chains to conclude that $\mathfrak{F}_i$ is $\mathcal{L}_i$–elementary in $\mathfrak{F}^\star$. Likewise for $\mathfrak{G}^\star$. Consider the substructures induced by the constants. That is, let $\mathfrak{F}^\circ$ be the subframe of $\mathfrak{F}^\star$ based on all interpretations of constants $c \in \mathcal{L}_\omega$ and likewise $\mathfrak{G}^\circ$ the subframe of all interpretations for constants in $\mathfrak{G}^\star$. We claim that $\mathfrak{F}^\circ \rightarrowtail \mathfrak{F}^\star$ and $\mathfrak{G}^\circ \rightarrowtail \mathfrak{G}^\star$. Namely, if $x \in \mathfrak{F}^\star$ is the value of a constant c in $\mathcal{L}_\omega$, there is an i such that $c \in \mathcal{L}_i$, and so c is interpreted in $\mathfrak{F}_i$. Let $x \lhd_j y$ for some $y \in f_i$. If i is even then there is a $\underline{y} \in \mathcal{L}_{i+1}$ such that $\underline{y}$ is interpreted in $\mathfrak{F}_{i+1}$, and its interpretation is y by construction. If i is odd, then a $\underline{y}$ will be introduced in $\mathcal{L}_{i+2}$. Similarly for $\mathfrak{G}^\circ$. Now, define a map $f : \mathfrak{F}^\circ \to \mathfrak{G}^\circ$ as follows. We put $f(x) := y$ if there is a constant c such that c is interpreted as x in $\mathfrak{F}^\circ$ but as y in $\mathfrak{G}^\circ$. We claim that this function is an isomorphism. (i.) It it is defined on $\mathfrak{F}^\circ$. (ii.) If c and d are two constants such that their interpretations coincide, then they coincide in a $\mathfrak{F}_i$ for some i, and so $\mathfrak{G}_i \vDash c \doteq d$, from which follows that $\mathfrak{G}^\star \vDash c \doteq d$. So the definition of f is sound. (iii.) f is onto, by definition of $\mathfrak{G}^\circ$. (iv.) f is injective. For if c and d are interpreted by different worlds in $\mathfrak{F}^\star$ then $\mathfrak{F}^\star \vDash \neg(c \doteq d)$, whence $\mathfrak{F}_i \vDash \neg(c \doteq d)$, for all i such that $c, d \in \mathcal{L}_i$. Then $\mathfrak{G}_i \vDash \neg(c \doteq d)$, by r–equivalence. And (v.) $\mathfrak{F}^\circ \vDash c \lhd_j d$ iff $\mathfrak{G}^\circ \vDash c \lhd_j d$, by the fact that the two are generated subframes and the formula is restricted.

Thus, since $\mathfrak{G}_1 \vDash \alpha[\vec{c}/\vec{x}]$, we also have $\mathfrak{G}^\star \vDash \alpha[\vec{c}/\vec{x}]$ and therefore $\mathfrak{G}^\circ \vDash \alpha[\vec{c}/\vec{x}]$, since α was assumed to be invariant under generated subframes. Now also $\mathfrak{F}^\circ \vDash \alpha[\vec{c}/\vec{x}]$, by the fact that the two models are isomorphic. Again by invariance under generated subframes, we get $\mathfrak{F}^\star \vDash \alpha[\vec{c}/\vec{x}]$ and finally $\mathfrak{F}_1 \vDash \alpha[\vec{c}/\vec{x}]$, which is to say $\langle \mathfrak{f}_0, \iota\rangle \vDash \alpha$. This had to be shown. ◻

Definition 5.8.3. *A first–order condition $\alpha(\vec{x})$ is said to be* ***preserved*** *under contractions if whenever $\langle \mathfrak{f}, \iota\rangle \vDash \alpha(\vec{x})$ and $p : \mathfrak{f} \twoheadrightarrow \mathfrak{g}$ is a p–morphism then for $\iota'(x_i) := p(\iota(x_i))$ we have $\langle \mathfrak{g}, \iota'\rangle \vDash \alpha(\vec{x})$. $\alpha(\vec{x})$ is* ***reflected*** *under contractions if $(\neg\alpha)(\vec{x})$ is preserved under contractions, and it is called* ***invariant*** *under contractions if it is both preserved and reflected under contractions.*

Below we will prove that a formula with at least one free variable is invariant under generated subframes and preserved under p–morphisms iff it is equivalent to a positive $\mathcal{R}^f$–formula. The proof is taken from van Benthem [**10**]. This shows that if α defines a modal class in $\mathfrak{Krp} \cup \mathfrak{D}$ then it is equivalent to a restricted positive formula. To show that Sahlqvist's Theorem is the best possible result for local correspondence, two things need to be established. (1.) Every locally $\mathfrak{D}$–persistent logic is elementary, (2.) every restricted positive α defining a modal class in $\mathfrak{Krp} \cup \mathfrak{D}$ is equivalent to a Sahlqvist formula. Both are still open questions. Notice that for (1.) and (2.) the choice of the class seems essential. We have seen (Corollary 5.7.8) that (1.) holds if the class $\mathfrak{R}$ is chosen instead. If it holds for the class $\mathfrak{D}$ we would have the converse of Theorem 5.7.11. It is unknown whether or not this holds. Also, with respect to (2.) it is also not known whether it holds. In the exercises we will ask the reader to show that if $\alpha \in \mathcal{L}^f$ is at most $\forall\exists$ and modally definable, then α is equivalent to some restricted formula.

Theorem 5.8.4 (van Benthem). *A first–order formula $\alpha(\vec{x})$ with at least one free variable is invariant under generated subframes and preserved under contractions iff it is equivalent to a positive restricted formula in the same free variables.*

Proof. We proceed as in the previous proof. The construction is somewhat more complicated, though. It is clear that a positive $\mathcal{R}^f$–formula is invariant under generated subframes and preserved under p–morphisms. The converse is the difficult part of the proof. Let $\alpha = \alpha(\vec{x})$ be a formula in x_i, $i < n$, with $n > 0$. Put

$$RP(\alpha) := \{\beta(\vec{x}) : \beta \in \mathcal{R}^f, \beta \text{ positive and } \alpha \vDash \beta\} \ .$$

If we can show that $RP(\alpha) \vDash \alpha$ we are done. For then, by compactness of first–order logic, there is a finite subset Δ of $RP(\alpha)$ such that $\Delta \vDash \alpha$. The conjunction δ of the members of Δ is also in $RP(\alpha)$. It follows that $\alpha \vDash \delta \vDash \alpha$, as desired. Let $\mathfrak{F}$ be a model for $RP(\alpha)$. Let x_i be interpreted by w_i. The language $\mathcal{L}_1$ is obtained from our initial language by adjoining a constant $\underline{w}_i$, $i < n$. $\mathfrak{F}_1$ is the expansion of $\mathfrak{F}$ in which $\underline{w}_i$ is interpreted by w_i. Consider now the set

$$\Sigma := \{\alpha[\vec{w}/\vec{x}]\} \cup \{\neg\beta : \beta \text{ a restricted positive } \mathcal{L}_1\text{–sentence}, \mathfrak{F}_1 \vDash \neg\beta\} \ .$$

Σ is finitely satisfiable. For if not, there is a finite subset $\Sigma_0 := \{\alpha[\vec{w}/\vec{x}]\} \cup \{\neg\beta_i : i < p\}$ such that $\Sigma_0 \vDash \mathsf{f}$. This however implies that $\alpha[\vec{w}/\vec{x}] \vDash \bigvee_{i<p}\beta_i$. So, $\alpha \vDash (\bigvee_{i<p}\beta_i)[\vec{x}/\underline{\vec{w}}](=: \beta)$. (Here, bound variables of the form x_i are suitably renamed.) Thus $\beta \in RP(\alpha)$. But then, by definition of $\mathfrak{F}_1$, $\mathfrak{F}_1 \vDash \beta[\vec{w}/\vec{x}]$, in contradiction to the definition of Σ, because $\mathfrak{F}_1 \vDash \neg\beta_i$ for all $i < n$. Hence, Σ_0 is consistent. So, Σ is finitely consistent and therefore has a model, $\mathfrak{G}_1$ (for example, an $\aleph_0$–saturated extension). The following holds of $\mathfrak{G}_1$.

(i) $\mathfrak{G}_1 \vDash \alpha[\underline{\vec{w}}/\vec{x}]$.
(ii) Every positive restricted $\mathcal{L}_1$–sentence which holds in $\mathfrak{G}_1$ also holds in $\mathfrak{F}_1$.

We will now construct a series of languages $\mathcal{L}_i$, $i < n$, such that $\mathcal{L}_i \subseteq \mathcal{L}_{i+1}$, $\mathcal{L}_i$ an expansion of $\mathcal{L}_0$ by constants, and two series of models, $\mathfrak{F}_i$ and $\mathfrak{G}_i$, such that for $i > 1$

1. $\mathfrak{F}_i$, $\mathfrak{G}_i$ are $\mathcal{L}_i$–models.
2. $\mathfrak{G}_i \vDash \alpha[\underline{\vec{w}}/\vec{x}]$.
3. Every positive restricted $\mathcal{L}_i$–sentence which holds in $\mathfrak{G}_i$ also holds in $\mathfrak{F}_i$.
4. $\mathfrak{F}_{i-1}$ is an $\mathcal{L}_{i-1}$–elementary substructure of $\mathfrak{F}_i$ and $\mathfrak{G}_{i-1}$ is an $\mathcal{L}_{i-1}$–elementary substructure of $\mathfrak{G}_i$.

We repeat the picture from the previous proof in a modified form.

$$\begin{tikzcd}
\mathfrak{F}_1 \arrow[r, phantom, "\prec_1"] \arrow[d, equal] & \mathfrak{F}_2 \arrow[r, phantom, "\prec_2"] \arrow[d, equal] & \mathfrak{F}_3 \arrow[r, phantom, "\prec_3"] \arrow[d, equal] & \mathfrak{F}_4 \arrow[r, phantom, "\cdots"] \arrow[d, equal] & \mathfrak{F}^\star \arrow[d, equal] & \mathfrak{F}^\circ \arrow[l] \\
\mathfrak{G}_1 \arrow[r, phantom, "\prec_1"] & \mathfrak{G}_2 \arrow[r, phantom, "\prec_2"] & \mathfrak{G}_3 \arrow[r, phantom, "\prec_3"] & \mathfrak{G}_4 \arrow[r, phantom, "\cdots"] & \mathfrak{G}^\star & \mathfrak{G}^\circ \arrow[l] \arrow[u, "\pi"'] \\
\mathcal{L}_1 & \mathcal{L}_2 & \mathcal{L}_3 & \mathcal{L}_4 \arrow[r, phantom, "\cdots"] & \mathcal{L}_\omega &
\end{tikzcd}$$

Assume that we have $\mathfrak{L}_n$ and $\mathfrak{F}_n$, $\mathfrak{G}_n$ satisfying (0.) – (3.). We show how to produce $\mathfrak{L}_{n+1}$, $\mathfrak{F}_{n+1}$ and $\mathfrak{G}_{n+1}$ satisfying (0.) – (3.). Let c be a constant of $\mathfrak{L}_n$ and $w \in \mathfrak{G}_n$ such that $\mathfrak{G}_n \vDash (c \lhd_j x)[w/x]$. Then add a new constant $\underline{w}$ to $\mathfrak{L}_n$. $\mathfrak{L}^1_n$ shall denote the language obtained by adding all such constants. $\mathfrak{G}_n$ is expanded to an $\mathfrak{L}^1_1$–structure by interpreting $\underline{w}$ by w. Let

$$\Delta := \{\beta : \beta \text{ a restricted positive } \mathfrak{L}^1_n\text{–sentence such that } \mathfrak{G}^1_n \vDash \beta\} .$$

We show that each finite subset of Δ can be satisfied in a model which is an expansion of $\mathfrak{F}_n$. Let namely $\Delta_0 := \{\beta_i : i < p\}$ be such a set, and denote its conjunction by δ. Assume that $\underline{w}_j$, $j < k$, are the constants of $\mathfrak{L}^1_n - \mathfrak{L}_n$ occurring in δ. By construction of $\mathfrak{L}^1_n$ there exist constants c_j, indices $j(i) < \kappa$ and variables x_i for all $i < k$ such that

$$\mathfrak{G}_n \vDash \{(c_i \lhd_{j(i)} x_i)[w_i/x_i] : i < k\} \cup \{\delta[\vec{w}/\underline{\vec{w}}]\}$$

$$\mathfrak{G}_n \vDash (\exists x_0 \rhd_{j(0)} c_0)(\exists x_1 \rhd_{j(1)} c_1) \dots (\exists x_{k-1} \rhd_{j(k-1)} c_{k-1})(\delta[\vec{x}/\underline{\vec{w}}])$$

This formula is an $\mathfrak{L}_n$–sentence. By (2.) it holds in $\mathfrak{F}_n$. Hence we find corresponding values for the constants $\underline{w}_j$, $j < k$. There exists an $\mathfrak{F}^1_n$ such that

(a) $\mathfrak{F}^1_n$ is an $\mathfrak{L}_n$–structure.
(b) $\mathfrak{F}_n$ is an $\mathfrak{L}_n$–elementary substructure of $\mathfrak{F}^1_n$.
(c) Every positive restricted $\mathfrak{L}^1_n$–sentence which holds in $\mathfrak{G}^1_n$ also holds in $\mathfrak{F}^1_n$.

Now we turn to the dual step. For each constant c of $\mathfrak{L}^1_n$ and each w in $\mathfrak{F}^1_n$ such that $\mathfrak{F}^1_n \vDash (c \lhd_j x)[w/x]$ add a new constant $\gamma(c, j, w)$. Let $\mathfrak{L}_{n+1}$ be the language obtained by adding all such constants. Expand $\mathfrak{F}^1_n$ to an $\mathfrak{F}_{n+1}$–structure by interpreting $\gamma(c, j, w)$ by w. Let

$$\begin{aligned}\Xi := \quad & \{\neg\beta : \beta \text{ a restricted positive } \mathfrak{L}_{n+1} \text{ sentence such that } \mathfrak{F}_{n+1} \vDash \neg\beta\} \\ & \cup \{c \lhd_j \gamma(c, j, w) : \gamma(c, j, w) \in \mathfrak{L}_{n+1} - \mathfrak{L}^1_n\}\end{aligned}$$

We show that each finite subset of Ξ is satisfiable in an expansion of $\mathfrak{F}^1_n$. For suppose that Ξ_0 is a finite subset of Ξ. It consists of a set of $\neg\beta_i$, $i < s$, from the first set, and a set $C := \{c_i \lhd_{j_i} \gamma(c_i, j_i, x_i) : i < t\}$. The conjunction of the $\neg\beta_i$ is denoted by ξ. ξ is also in Ξ, so we may take ξ in place of the first subset of Ξ_0. We may additionally assume that C already contains all $\gamma(c, j, w) \in \mathfrak{L}_{n+1} - \mathfrak{L}^1_n$ that appear in ξ. Now suppose that $C \cup \{\xi\}$ is not satisfiable in any expansion of $\mathfrak{G}^1_n$. Then

$$\mathfrak{G}^1_n \vDash (\forall x_0 \rhd_{j_0} c_0)(\forall x_1 \rhd_{j_1} c_1) \dots (\forall x_{t-1} \rhd_{j_{t-1}} c_{t-1})(\neg\xi[x_i/\gamma(c_i, j_i, w_i) : i < t])$$

This is a positive restricted $\mathfrak{L}^1_n$–sentence, so it holds in $\mathfrak{F}^1_n$, by (c.). But

$$\mathfrak{F}^1_n \vDash \{c_i \lhd_{j_i} x_i : i < t\}; \xi[w_i/\gamma(c_i, j_i, w_i) : i < t]$$

This is a contradiction. Hence, every finite subset of Ξ is satisfiable in an expansion of $\mathfrak{G}^1_n$, and so there exists a $\mathfrak{G}_{n+1}$ such that

(a) $\mathfrak{G}_{n+1}$ is an $\mathfrak{L}_{n+1}$–structure.
(b) $\mathfrak{G}^1_{n+1}$ is an $\mathfrak{L}^1_n$–elementary substructure of $\mathfrak{G}_{n+1}$.

(c) Every positive restricted $\mathfrak{L}_{n+1}$–sentence which holds in $\mathfrak{G}_{n+1}$ also holds in $\mathfrak{F}_{n+1}$.

The claims (1.) – (4.) now easily follow for $\mathfrak{L}_{n+1}$, $\mathfrak{F}_{n+1}$ and $\mathfrak{G}_{n+1}$ using (a.), (b.) and (c.). We have that $\mathfrak{F}_n$ is an $\mathfrak{L}_n$–elementary substructure of $\mathfrak{F}_{n+1}$. By induction it follows that $\mathfrak{F}_n$ is an $\mathfrak{L}_n$–elementary substructure of $\mathfrak{F}_{n+m}$ for every m. Similarly for $\mathfrak{G}_n$. Let $\mathfrak{L}_\omega := \bigcup_{i\in\omega} \mathfrak{L}_i$, and let $\mathfrak{F}^\star$ be the union of the $\mathfrak{F}_i$, $\mathfrak{G}^\star$ the union of the $\mathfrak{G}_i$. Then $\mathfrak{F}^\star$ and $\mathfrak{G}^\star$ are $\mathfrak{L}_\omega$–structures, of which $\mathfrak{F}_n$ and $\mathfrak{G}_n$ are respective $\mathfrak{L}_n$–elementary substructures. Furthermore, let $\mathfrak{F}^\circ$ be the transit of the w_i in $\mathfrak{F}^\star$ and $\mathfrak{G}^\circ$ the transit of the interpretation of the $\underline{w}_i$ in $\mathfrak{G}^\star$. By construction, the constants of $\mathfrak{L}^\omega$ all take values in $\mathfrak{F}^\circ$ ($\mathfrak{G}^\circ$), and every element of $\mathfrak{F}^\circ$ ($\mathfrak{G}^\circ$) is the interpretation of a constant from $\mathfrak{L}_\omega$. Define $\pi : \mathfrak{G}^\circ \to \mathfrak{F}^\circ$ as follows. If $x \in \mathfrak{G}^\circ$, let c be a constant with interpretation x. Then $\pi(x) := y$, where y is the interpretation of c in $\mathfrak{F}^\circ$. (1.) π is well–defined. Suppose that c and d are interpreted by x. Then there is a $n \in \omega$ such that $c \in \mathfrak{L}_n$ and $d \in \mathfrak{L}_n$. Then $\mathfrak{G}_n \vDash c \doteq d$, and so $\mathfrak{G}^\star \vDash c \doteq d$, being an $\mathfrak{L}_n$–elementary superstructure. Hence $\mathfrak{F}^\star \vDash c \doteq d$, and this shows that the interpretation of c is equal to the interpretation of d in $\mathfrak{F}^\star$ (and it is in $\mathfrak{F}^\circ$). (2.) π is onto. For since every element of $\mathfrak{F}^\circ$ is in the interpretation of some c. Take $x \in \mathfrak{G}^\circ$ such that x is the interpretation of c in $\mathfrak{G}^\circ$. Then $\pi(x) = y$. (3.) π is a p–morphism. (a) Suppose $x \lhd_j y$. Let x be the interpretation of c in $\mathfrak{G}^\circ$ and y the interpretation of d. Then $\mathfrak{G}^\star \vDash c \lhd_j d$. There is an n such that $c \in \mathfrak{L}_n$ and $d \in \mathfrak{L}_n$. For this n, $\mathfrak{G}_n \vDash c \lhd_j d$. Hence $\mathfrak{F}_n \vDash c \lhd_j d$, and so $\mathfrak{F}^\star \vDash c \lhd_j d$. Therefore $\pi(x) \lhd_j \pi(y)$. (b) Let $\pi(x) \lhd_j u$. There exists a constant c such that c is interpreted by $\pi(x)$ in $\mathfrak{F}^\star$. Then consider $\gamma(c, j, u)$. By construction, $\mathfrak{G}^\star \vDash c \lhd_j \gamma(c, j, u)$. Let y be the interpretation of $\gamma(c, j, u)$ in $\mathfrak{G}^\star$. The interpretation of $\gamma(c, j, u)$ in $\mathfrak{F}^\star$ is just u. Therefore, $\pi(y) = u$, and $x \lhd_j y$, as required.

Finally, we know that $\mathfrak{G}_1 \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. Hence, $\mathfrak{G}^\star \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. Since α is invariant under generated subframes, $\mathfrak{G}^\circ \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. α is also preserved under contractions, and so $\mathfrak{F}^\circ \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. Again by invariance under generated subframes, $\mathfrak{F}^\star \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. Finally, since $\mathfrak{F}_1$ is a $\mathfrak{L}_1$–elementary substructure, $\mathfrak{F}_1 \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. By construction, this is the same as $\mathfrak{F} \vDash \alpha[\vec{w}/\vec{x}]$. This is the desired conclusion. □

The proof in Marcus Kracht [**124**], supposed to be a construction of the equivalent formula, is actually incorrect. There is a plethora of similar results concerning the interplay between syntactic form (up to equivalence) and invariance properties with respect to class operators. A particular theorem is the following.

Theorem 5.8.5. *Call a formula $\zeta(\vec{x})$* ***constant*** *if no prime subformula containing a variable is in the scope of a quantifier. Let $\alpha(\vec{x})$ be a formula with at least one free variable. Assume that α is invariant under generated subframes and contractions. Then α is equivalent to a constant formula $\chi(\vec{x})$.*

Proof. The direction from right to left is easy. So, assume that α is invariant under contractions and generated subframes. Let $\alpha(\vec{x})$ be given. The initial language is $\mathfrak{R}^f$. Put

$$C(\alpha) := \{\beta(\vec{x}) : \beta \text{ constant}, \alpha \vDash \beta\}$$

Clearly, $\alpha \vDash C(\alpha)$. We show that $C(\alpha) \vDash \alpha$. To do that, let $\mathfrak{F} \vDash C(\alpha)$, where x_i is mapped to w_i. Adjoin new constants $\underline{w}_i$. This yields the language $\mathcal{L}_1$. Make $\mathfrak{F}$ into a $\mathcal{L}_1$–structure $\mathfrak{F}_1$ by interpreting $\underline{w}_i$ by w_i. Now let

$$\Xi := \{\alpha[\underline{\vec{w}}/\vec{x}]\} \cup \{\delta : \delta \text{ a constant } \mathcal{L}_1\text{–sentence and } \mathfrak{F}_1 \vDash \delta\}$$

Ξ is finitely satisfiable, and so there is a model $\mathfrak{G}_1$. Let $\mathfrak{F}^\star$ be a $\aleph_0$–saturated expansion of $\mathfrak{F}_1$ and $\mathfrak{G}^\star$ a $\aleph_0$–saturated expansion of $\mathfrak{G}_1$. Further, let $\mathfrak{F}^\circ$ be the subframe of $\mathfrak{F}^\star$ generated by w_i, $i < n$, and $\mathfrak{G}^\circ$ the subframe of $\mathfrak{G}^\star$ generated by the interpretation of $\underline{w}_i$. Now, define a binary relation $\sim$ on $\mathfrak{F}^\circ$ by $u \sim u'$ iff u and u' satisfy the same constant $\mathcal{R}^f$–formulae. It is not hard to show that this is a net. For if $u \sim u'$ and $u \lhd_j v$, let $D(y)$ be the set of constant $\mathcal{R}^f$–formulae in y satisfied by v. Then let $\Diamond_j D(y) := \{(\exists y \rhd_j x)\delta(y) : \delta \in D(y)\}$. u satisfies $\Diamond_j D(y)$. Hence u' satisfies $\Diamond_j D(y)$. By saturatedness, there is a successor v' of u' satisfying $D(y)$. Similarly, define $\approx$ on $\mathfrak{G}^\circ$ by $u \approx u'$ iff they satisfy the same constant $\mathcal{R}^f$–sentences. Now, let t_i interpret $\underline{w}_i$ in $\mathfrak{G}_1$. Then t_i satisfies the same constant formulae as does w_i. Therefore, it can be shown by induction on the depth that $\mathfrak{F}^\circ/{\sim}$ is isomorphic to $\mathfrak{G}^\circ/{\approx}$. Now the conclusion is easily established. $\mathfrak{G}_1 \vDash \alpha[\underline{\vec{w}}/\vec{x}]$. By elementary embedding, $\mathfrak{G}^\star \vDash \alpha[\underline{\vec{w}}/\vec{x}]$ and so $\mathfrak{G}^\circ \vDash \alpha[\underline{\vec{w}}/\vec{x}]$, $\mathfrak{G}^\circ/{\approx} \vDash \alpha[\underline{w}/\vec{x}]$ by preservation under generated subframes and contractions. Then $\mathfrak{F}^\circ/{\sim} \vDash \alpha[\underline{w}/\vec{x}]$, since it is isomorphic to the latter structure. It follows that $\mathfrak{F}_1 \vDash \alpha[\underline{\vec{w}}/\vec{x}]$, and finally $\mathfrak{F} \vDash \alpha[\vec{w}/\vec{x}]$. ◻

Theorem 5.8.6 (van Benthem). *A logic $\boldsymbol{K} \oplus \chi$ is g–persistent iff there exists a constant formula $\gamma(x_0)$ such that the class of frames for that logic is defined by $\gamma(x_0)$.*

These theorems establish the weak form of completeness of the calculus **Seq** for $\mathfrak{G}$ discussed at the end of Section 5.4.

Exercise 194. ($\kappa < \aleph_0$.) Let $\alpha(x) \in \mathcal{L}^f$ be modally definable in $\mathfrak{Krp}$. We can assume that α is in prenex normal form; furthermore, assume that α contains only $\forall$–quantifiers. Show that α is equivalent to a formula β in which every universal quantifier $\forall y$ is replaced by a restricted quantifier $(\forall y \rhd^s x)$ for some x and some finite set s of finite sequences over κ. *Hint.* Let $\delta_i(x)$ be obtained by replacing each quantifier $\forall y$ by $(\forall y \rhd^{s_n} x)$, where s_n consists of all sequences of length $\leq n$. Then $\delta_n(x) \vdash \delta_{n+1}(x)$ in predicate logic, and $\delta_n(x) \vdash \alpha(x)$. It remains to prove that some n exists such that $\alpha(x) \vdash \delta_n(x)$. Suppose that no such n exists; then the set $\{\neg\delta_n(x) : n \in \omega\} \cup \{\alpha(x)\}$ is consistent. It has a model $\langle \mathfrak{f}, \iota \rangle$. Now use the fact that $\alpha(x)$ is closed under generated subframes.

Exercise 195. As above, but for α of the complexity $\forall\exists$. (Assume $\kappa < \aleph_0$; see [**128**] for a proof.)

Exercise 196. The same as the previous exercise, but without the restriction $\kappa < \aleph_0$.

CHAPTER 6

Reducing Polymodal Logic to Monomodal Logic

6.1. Interpretations and Simulations

The main body of technical results in modal logic is within monomodal logic, for example extensions of **K4**. The theory of one operator is comparatively speaking well–understood. Many applications of modal logic, be they in philosophy, computer science, linguistics or mathematics proper, require several operators, sometimes even infinitely many. Moreover, in the theory of modal logic many counterexamples to conjectures can be found easily if one uses logics with several operators. So there is a real need for a theory of polymodal logic. On the other hand, if such a theory is needed and we have developed a theory of a single operator, it is most desirable if we could so to speak transfer the results from the one operator setting to several operators. This, however, is not straightforward. It has often been deemed a plausible thing to do but turned out to be notoriously difficult. Only fairly recently methods have been developed that allow to transfer results of reasonable generality. They go both ways. It is possible to interpret a monomodal logic as a polymodal logic, which involves axioms for one of the operators only. Let us call these *one–operator logics*. They were introduced by S. K. Thomason [**211**] and systematically studied in Kit Fine and Gerhard Schurz [**67**] and also Marcus Kracht and Frank Wolter [**132**]. If we fix an operator, we have a natural embedding of monomodal logic into polymodal logic. We can also study unions of one–operator logics, where the distinguished operators differ. Such logics are called *independently axiomatizable*. For independently axiomatizable logics there exist a number of strong transfer results. Basically, all common properties of the one operator logics transfer to the union.

This direction is unsurprising, perhaps. Moreover, polymodal logics contain monomodal logics. The converse, however, is prima facie unplausible for it suggests that we can model several operators with the help of a single one. Yet, exactly this is the case. S. K. Thomason has proved a lot of negative results for monomodal logic by reducing polymodal logic and other sorts of logics to monomodal logic. However, it has gone unnoticed that not only negative properties of logics such as incompleteness, undecidability and so on are transferred, but also positive ones. Once this is noticed, we derive a plethora of strong results concerning monomodal logic. In this way we can gain insight not only into polymodal logic but also into the theory of a single operator.

Before we enter the discussion of polymodal versus monomodal logic, we need to make our ideas precise concerning *reduction*. We have mentioned two cases of reduction, one from polymodal logic to monomodal logic, and another from monomodal logic to polymodal logic. More precisely, these reductions consist of a *translation* of one language into another. Moreover, this translation reduces a logic in the first language to a logic in the second language if it is faithful with respect to the deducibilities. This is made precise as follows. Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two propositional languages with variables drawn from *var*. An **interpretation** of $\mathcal{L}_1$ in $\mathcal{L}_2$ is a map which assigns to the variables uniformly an expression of $\mathcal{L}_2$ and to each connective of $\mathcal{L}_1$ a possibly complex functional expression of $\mathcal{L}_2$. This means that for $I : \mathcal{L}_1 \to \mathcal{L}_2$ to be an interpretation it must satisfy

$$(f(\varphi_0, \ldots, \varphi_{k-1}))^I = (f(p_0, \ldots, p_{k-1}))^I[\varphi_0^I/p_0, \ldots, \varphi_{k-1}^I/p_{k-1}]$$

for all connectives f in $\mathcal{L}_1$, formulae $\varphi_0, \ldots, \varphi_{k-1} \in \mathcal{L}_1$, and variables p_i, $i < k$; and for all variables p, q

$$q^I = p^I[q/p]$$

In brief, an interpretation is fixed by the term it assigns to a simple expression; interpretations are not to be confused with homomorphisms. First of all, an interpretation is a map between languages with possibly different signatures. Moreover, even when the signatures are not different, the concept itself may still differ. For example, the duality map is an interpretation, though strictly speaking not a homomorphism, because conjunction and disjunction may not be interchanged by a homomorphism. Furthermore, an interpretation is free to assign a complex term to a simple term. For example, we might choose to interpret $\Box$ as $\blacksquare\Diamond$ (for example, in interpreting non–classical monomodal logics as bimodal logics, see [**133**]).

The definitions have some noteworthy consequences. First of all, a variable p is translated into an expression p^I which contains at most the variable p, that is, $var(p^I) \subseteq \{p\}$. For if $q \neq p$ we have $p^I = q^I[p/q]$, thus $q \notin var(p^I)$. Likewise, for any expression φ we have $var(\varphi^I) \subseteq var(\varphi)$.

Now consider two logics $\langle \mathcal{L}_1, \vdash_1 \rangle$ and $\langle \mathcal{L}_2, \vdash_2 \rangle$ and an interpretation I. Then $\vdash_2$ **simulates** $\vdash_1$ with respect to I if for all $\Gamma \subseteq \mathcal{L}_1$ and $\varphi \in \mathcal{L}_1$

$$\Gamma \vdash_1 \varphi \qquad \text{iff} \qquad \Gamma^I \vdash_2 \varphi^I .$$

Denote by $\mathcal{S}_I(\vdash_1)$ the set of all consequence relations $\vdash$ over $\mathcal{L}_2$ which simulate $\vdash_1$ with respect to I. It is readily checked that $\mathcal{S}_I(\vdash_1)$ contains a minimal element. The following is a fundamental property of simulations.

Proposition 6.1.1. *Suppose that $\vdash_2$ simulates $\vdash_1$ with respect to some interpretation I. Then if $\vdash_2$ is decidable, so is $\vdash_1$.*

For a proof just observe that by definition the problem $\Gamma \vdash_1 \varphi$ is equivalent to $\Gamma^I \vdash_2 \varphi^I$. A priori, a connective can be translated by an arbitary expression. However, under mild conditions the interpretation of a boolean connective $\odot$ must be an expression equivalent to $\odot$. In the case of modal logics this means that under these conditions only the modal operators receive a nontrivial interpretation. Call an interpretation I **atomic** if $p^I = p$ for all propositional variables p. In this case $f^I(\varphi_0, \ldots, \varphi_{k-1})$ will be used instead of $f(p_0, \ldots, p_{k-1})^I[\varphi_0/p_0, \ldots, \varphi_{k-1}/p_{k-1}]$. The following then holds, as has been oberserved in [**133**].

Proposition 6.1.2 (Wolter). *Write $\varphi \equiv_2 \chi$ for $\varphi \vdash_2 \chi$ and $\chi \vdash_2 \varphi$. Suppose that $\wedge$ and $\neg$ are both symbols of $\mathcal{L}_1$ and $\mathcal{L}_2$, that the restrictions of $\vdash_1$ and $\vdash_2$ to expressions containing $\neg$ and $\wedge$ equals the propositional calculus over $\neg$ and $\wedge$. And suppose that we have the* **replacement rule** *for $\vdash_2$; that is, if $\varphi_1 \equiv_2 \varphi_2$ then $\psi[\varphi_1/p] \equiv_2 \psi[\varphi_2/p]$. Finally, suppose that I is an atomic interpretation. Then if $\vdash_2$ simulates $\vdash_1$ with respect to I the following holds:*

(1) $p \wedge q \equiv_2 p \wedge^I q$.
(2) $\neg p \equiv_2 \neg^I p$.

Proof. (1.) We have $p \wedge q \vdash_2 \{p, q\} \vdash_2 p \wedge^I q \vdash_2 \{p, q\} \vdash_2 p \wedge q$. (2.) It is readily checked that φ is $\vdash_1$–inconsistent iff φ^I is $\vdash_2$–inconsistent. Hence, $p; \neg^I p$ is $\vdash_2$–inconsistent, since $p; \neg p$ is $\vdash_1$–inconsistent. Hence $\neg^I p \vdash_2 \neg p$. It remains to show that $\neg\neg^I p; \neg p$ is $\vdash_2$–inconsistent. But this follows with $q \vdash_2 ((p \wedge q) \vee (\neg p \wedge q))^I$ and (i) by

$$\begin{aligned}
&\neg\neg^I p \wedge \neg p \\
&\quad \vdash_2 (p \wedge^I (\neg\neg^I p \wedge \neg p)) \vee^I (\neg^I p \wedge^I (\neg\neg^I p \wedge \neg p)) \\
&\quad \vdash_2 (p \wedge^I \neg p) \vee^I (\neg^I p \wedge^I \neg\neg^I p) \\
&\quad \vdash_2 ((p \wedge \neg p) \vee (p \wedge \neg p))^I
\end{aligned}$$

and the $\vdash_2$–inconsistency of this last formula. $\Box$

It is worthwile to reflect on the notion of a simulation. We will use it also to show undecidability of certain logics. The rationale will be to use well–known undecidable problems, in this case facts about word–problems in semigroups, and simulate these problems in polymodal logics. Furthermore, as polymodal logics can themselves be simulated by monomodal logics, this yields undecidable problems for monomodal logic. We shall indicate here that the notion of simulation is quite similar to a notion that is defined in Ryszard Wójcicki [**231**].

6.2. Some Preliminary Results

In the next sections we are dealing with the following standard situation. We have a bimodal language $\mathcal{L}_2$, denoted here by $\mathcal{L}_{\Box\blacksquare}$, and two monomodal fragments, $\mathcal{L}_\Box$ and $\mathcal{L}_\blacksquare$. Naturally arising objects such as formulae and consequence relations are subject to the same notation, which we assume to be clear without explanation. There are two possible interpretations of a single operator — denoted here by $\boxminus$ — in bimodal logic over $\Box$ and $\blacksquare$. We may read it as $\Box$ or as $\blacksquare$. Notice that these symbols are used in place of $\Box_\lambda$. Although from a technical viewpoint, if, say $\Box = \Box_0$ and $\blacksquare = \Box_1$, then $\Box$ and $\boxminus$ are the same, we wish to make notation independent from an accidental choice of interpretation for $\Box$ and $\blacksquare$. Seen this way, we are now dealing with three independent operators, $\boxminus$, $\Box$ and $\blacksquare$.

Define two translations, $\tau_\Box$ and $\tau_\blacksquare$ in the following way.

$$\begin{array}{lcl@{\qquad\qquad}lcl}
\tau_\Box(p) & := & p & \tau_\blacksquare(p) & := & p \\
\tau_\Box(\top) & := & \top & \tau_\blacksquare(\top) & := & \top \\
\tau_\Box(\neg\varphi) & := & \neg\tau_\Box(\varphi) & \tau_\blacksquare(\neg\varphi) & := & \neg\tau_\blacksquare(\varphi) \\
\tau_\Box(\varphi \wedge \chi) & := & \tau_\Box(\varphi) \wedge \tau_\Box(\chi) & \tau_\blacksquare(\varphi \wedge \chi) & := & \tau_\blacksquare(\varphi) \wedge \tau_\blacksquare(\chi) \\
\tau_\Box(\boxminus\varphi) & := & \Box\tau_\Box(\varphi) & \tau_\blacksquare(\boxminus\varphi) & := & \blacksquare\tau_\blacksquare(\chi)
\end{array}$$

We speak of the translation $\tau_\blacksquare$ as the *dual* of $\tau_\Box$, by which we want to imply that the roles of $\blacksquare$ and $\Box$ are interchanged. (So, $\tau_\Box$ likewise is the dual of $\tau_\blacksquare$.) It is in this sense that we want to be understood when we talk about *duality* in this chapter. This will frequently arise in proofs, where we will perform the argument with one operator, and omit the case of the other operator. Given two modal logics, Λ and Θ, the **fusion** is defined as in Section 2.5 by

$$\Lambda \otimes \Theta := \mathbf{K}_2 \oplus \tau_\Box[\Lambda] \oplus \tau_\blacksquare[\Theta]$$

This defines an operation $-\otimes- : (\mathcal{E}\mathbf{K}_1)^2 \to \mathcal{E}\mathbf{K}_2$. $\otimes$ is a $\bigsqcup$-homomorphism in both arguments. Moreover, it is easy to see that

$$(\mathbf{K}_1 \oplus X) \otimes (\mathbf{K}_1 \oplus Y) = \mathbf{K}_2 \oplus \tau_\Box[X] \oplus \tau_\blacksquare[Y]$$

(Namely, observe that $\tau_\Box$ and $\tau_\blacksquare$ translate valid derivations in $\mathbf{K}_1$ into valid derivations in $\mathbf{K}_2$. So, if Λ derivable from X by (mp.), (mn.) and substitution in $\mathbf{K}_1$, all formulae of $\tau_\Box[\Lambda]$ are derivable from $\tau_\Box[X]$ by means of (mp.), (mn.) and substitution in $\mathbf{K}_2$.) We call a bimodal logic Ξ **independently axiomatizable** if there exist Λ and Θ such that $\Xi = \Lambda \otimes \Theta$. The following theorem will be made frequent use of.

Lemma 6.2.1. $\langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle \vDash \Lambda \otimes \Theta$ *iff* $\langle g, \lhd, \mathbb{G}\rangle \vDash \Lambda$ *and* $\langle g, \blacktriangleleft, \mathbb{G}\rangle \vDash \Theta$.

The easy proof is left to the reader. Moreover, it should be clear that if $\langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$ is a bimodal frame, $\langle g, \lhd, \mathbb{G}\rangle$ and $\langle g, \blacktriangleleft, \mathbb{G}\rangle$ are monomodal frames. Given a bimodal logic Λ define

$$\begin{array}{lcl}
\Lambda_\Box & := & \tau_\Box^{-1}[\Lambda] \\
\Lambda_\blacksquare & := & \tau_\blacksquare^{-1}[\Lambda]
\end{array}$$

There are certain easy properties of these maps which are noteworthy. Fixing the second argument we can study the map $-\otimes\Theta : \mathcal{E}\mathbf{K}_1 \to \mathcal{E}\mathbf{K}_2$. This is a $\bigsqcup$–homomorphism. The map $-_\Box : \mathcal{E}\mathbf{K}_2 \to \mathcal{E}\mathbf{K}_1 : \Lambda \mapsto \Lambda_\Box$ will be shown to be almost the inverse of $-\otimes\Theta$.

Lemma 6.2.2. *(1.) Let Λ be a normal modal logic. Then $(\Lambda \otimes \Theta)_\Box \supseteq \Lambda$. (2.) Let Ξ be a normal bimodal logic Ξ. Then $\Xi_\Box \otimes \Xi_\blacksquare \subseteq \Xi$. Moreover, Ξ is independently axiomatizable iff $\Xi = (\Xi_\Box) \otimes (\Xi_\blacksquare)$.*

Proof. $\langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle \vDash \Xi$ implies $\langle g, \lhd, \mathbb{G}\rangle \vDash \Xi_\Box$ and $\langle g, \blacktriangleleft, \mathbb{G}\rangle \vDash \Xi_\blacksquare$. This implies in turn that $\langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle \vDash \Xi_\Box \otimes \Xi_\blacksquare$. Consequently, if Ξ is independently axiomatizable then $\langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$ is a general Ξ–frame iff $\langle g, \lhd, \mathbb{G}\rangle$ is a general $\Xi_\Box$–frame and $\langle g, \blacktriangleleft, \mathbb{G}\rangle$ is a general $\Xi_\blacksquare$–frame. ◻

Given a monomodal frame $\mathfrak{G} := \langle g, \leqslant, \mathbb{G}\rangle$, put $\mathfrak{G}^\bullet := \langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$, where $\lhd := \leqslant$, and $\blacktriangleleft := \varnothing$, and put $\mathfrak{G}^\circ := \langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$ where we have $\lhd := \leqslant$ and $\blacktriangleleft := \{\langle x, x\rangle : x \in g\}$. It is easy to check that both $\mathfrak{G}^\bullet$ and $\mathfrak{G}^\circ$ are bimodal frames. (To see that, one only has to verify that for $b \in \mathbb{G}$ also $\blacksquare b \in \mathbb{G}$; but this is straightforward.) The following was shown in [**211**].

Theorem 6.2.3 (Thomason). $(\Lambda \otimes \Theta)_\Box = \Lambda$ *iff* $\bot \notin \Theta$ *or* $\bot \in \Lambda$.

Proof. ($\Rightarrow$) Suppose $\bot \in \Theta$ and $\bot \notin \Lambda$. Then $\bot \in \Lambda \otimes \Theta$ and hence $\bot \in (\Lambda \otimes \Theta)_\Box$, so hat $\Lambda \neq (\Lambda \otimes \Theta)_\Box$.

($\Leftarrow$) Suppose $\bot \in \Lambda$. Then $\bot \in \Lambda \otimes \Theta$ and so $\bot \in (\Lambda \otimes \Theta)_\Box$ from which $\Lambda = (\Lambda \otimes \Theta)_\Box$. Now suppose $\bot \notin \Lambda$. Then $\bot \notin \Theta$ and hence either $\boxed{\bullet} \vDash \Theta$ or $\boxed{\circ} \vDash \Theta$. Let $\mathfrak{G} = \langle g, \lhd, \mathbb{G}\rangle$ be a Λ–frame. Then put $\mathfrak{G}^\bullet := \langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$ as above with $\blacktriangleleft := \varnothing$ and $\mathfrak{G}^\circ = \langle g, \lhd, \blacktriangleleft, \mathbb{G}\rangle$ with $\blacktriangleleft := \{\langle x, x\rangle : x \in g\}$. If $\boxed{\bullet} \vDash \Theta$ then $\mathfrak{G}^\bullet$ is a $\Lambda \otimes \Theta$–frame and if $\boxed{\circ} \vDash \Theta$ then $\mathfrak{G}^\circ$ is a $\Lambda \otimes \Theta$–frame. For $\varphi \in \Lambda_\Box$ we have $\mathfrak{G} \vDash \varphi \Leftrightarrow \mathfrak{G}^\bullet \vDash \varphi \Leftrightarrow \mathfrak{G}^\circ \vDash \varphi$. Thus $(\Lambda \otimes \Theta)_\Box \subseteq \Lambda$ and therefore $(\Lambda \otimes \Theta)_\Box = \Lambda$. ◻

The theorem states that if $\bot \in \Lambda$ or $\bot \notin \Theta$ then $\Lambda \otimes \Theta$ is a *conservative* extension of Λ. Thus given two logics Λ, Θ we have both $\Lambda = (\Lambda \otimes \Theta)_\Box$ and $\Theta = (\Lambda \otimes \Theta)_\blacksquare$ iff $\bot \in \Lambda \Leftrightarrow \bot \in \Theta$. In all the theorems that will follow the case that $\bot \in \Lambda$ or $\bot \in \Theta$ will be excluded. These cases are trivial anyway, so nothing is lost. The way in which Makinson's theorem has been used to build a minimal extension of a mono–modal frame to a bimodal frame is worth remembering. It will occur quite often later on. Although Makinson's theorem has no analogue for bimodal logics as there are infinitely many maximal consistent bimodal logics, at least for independently axiomatizable logics the following holds.

Corollary 6.2.4. *Suppose that Λ is a consistent, independently axiomatizable bimodal logic. Then there is a Λ–frame based on one point.*

Theorem 6.2.5. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is finitely axiomatizable (recursively axiomatizable) iff both Λ and Θ are.*

Proof. If Λ and Θ are recursively axiomatizable, so is clearly their fusion. And if the fusion is, then the theorems are recursively enumerable, and hence also $(\Lambda \otimes \Theta)_\Box$ and $(\Lambda \otimes \Theta)_\blacksquare$. Thus Λ and Θ are recursively axiomatizable. Now for finite axiomatizability. Only the direction from left to right is not straightforward. Assume therefore that $\Lambda \otimes \Theta$ is finitely axiomatizable, say $\Lambda \otimes \Theta = \mathbf{K}_2(Z)$. Let X and Y be such that $\Lambda = \mathbf{K}_1(X)$, $\Theta = \mathbf{K}_1(Y)$. Then $Z \subseteq \mathbf{K}_2(\tau_\Box[X] \cup \tau_\blacksquare[Y])$. By the Compactness Theorem we have finite sets $X_0 \subseteq X$, $Y_0 \subseteq Y$ such that $Z \subseteq \mathbf{K}_2(\tau_\Box[X_0] \cup \tau_\blacksquare[Y_0])$. But then $\Lambda \otimes \Theta = \mathbf{K}_2(\tau_\Box[X_0] \cup \tau_\blacksquare[Y_0]) = \mathbf{K}_1(X_0) \otimes \mathbf{K}_1(Y_0)$ and hence $\Lambda = \mathbf{K}_1(X_0)$ and $\Theta = \mathbf{K}_1(Y_0)$. ◻

Theorem 6.2.6. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is r–persistent iff both Λ and Θ are.*

Proof. ($\Leftarrow$) Suppose that both Λ and Θ are r–persistent. Further assume that $\langle g, \triangleleft, \blacktriangleleft, \mathbb{G}\rangle \vDash \Lambda \otimes \Theta$. Then $\langle g, \triangleleft, \mathbb{G}\rangle \vDash \Lambda$ and $\langle g, \blacktriangleleft, \mathbb{G}\rangle \vDash \Theta$. By assumption, $\langle g, \triangleleft\rangle \vDash \Lambda$ and $\langle g, \blacktriangleleft\rangle \vDash \Theta$ and so $\langle g, \triangleleft, \blacktriangleleft\rangle \vDash \Lambda \otimes \Theta$. ($\Rightarrow$) Suppose that $\bot \notin \Lambda$ and that Λ is not r–persistent. We have to show that $\Lambda \otimes \Theta$ is also not r–persistent. We know that there is a Λ–frame $\mathfrak{G} = \langle g, \triangleleft, \mathbb{G}\rangle$ such that $\langle g, \triangleleft\rangle \nvDash \Lambda$. On the condition that $\mathfrak{G}^\bullet$ and $\mathfrak{G}^\circ$ are both refined the theorem is proved. For either $\mathfrak{G}^\bullet \vDash \Lambda \otimes \Theta$ or $\mathfrak{G}^\circ \vDash \Lambda \otimes \Theta$, but $\langle g, \triangleleft, \blacktriangleleft\rangle \nvDash \Lambda \otimes \Theta$ since $\langle g, \triangleleft\rangle \nvDash \Lambda$.

Both $\mathfrak{G}^\bullet$ and $\mathfrak{G}^\circ$ are differentiated and tight. That $\mathfrak{G}^\circ$ satisfies tightness for $\blacksquare$ is seen as follows. If $x = y$ then for all $c \in \mathbb{G}$, $x \in \blacksquare c$ implies $x \in c$ since $\blacksquare c = c$. But if $x \neq y$ there is a $c \in \mathbb{G}$ such that $x \in c$, $y \notin c$. Then $x \in \blacksquare c$, $y \notin c$, as required. Similarly, $\mathfrak{G}^\bullet$ satisfies tightness for $\blacksquare$ since for arbitrary x, y there is $c \in \mathbb{G}$ with $y \notin c$. Moreover, $x \in \blacksquare c$, since $\blacksquare c = 1$. ◻

Theorem 6.2.7. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is d–persistent iff both Λ and Θ are.*

Proof. As in the previous theorem. One only has to check that if $\mathfrak{G}$ is descriptive, so are $\mathfrak{G}^\bullet$ and $\mathfrak{G}^\circ$. We have seen that if $\mathfrak{G}^\bullet$ and $\mathfrak{G}^\circ$ are both refined, so we only have to check compactness. So let U be an ultrafilter $\mathfrak{G}^\bullet$ ($\mathfrak{G}^\circ$). Then U is also an ultrafilter of $\mathfrak{G}$, since the underlying boolean algebras are identical. Since $\mathfrak{G}$ is compact we have $\bigcap U \neq \varnothing$, as required. ◻

Corollary 6.2.8. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is canonical iff both Λ and Θ are.*

Proof. By Theorem 4.8.6. ◻

Exercise 197. Let Θ be a consistent logic. Define $e : \mathcal{E}\,\mathbf{K}_1 \to \mathcal{E}\,\mathbf{K}_2$ by $\Lambda \mapsto \Lambda \otimes \Theta$ and let $r : \mathcal{E}\,\mathbf{K}_2 \to \mathcal{E}\,\mathbf{K}_1$ be defined by $\Gamma \mapsto \Gamma_\Box$. Both e and r are not necessarily lattice homomorphisms. However, both are isotonic (prove this). Show that e is left adjoined to r. That is, for monomodal logics Λ and bimodal logics

$$e(\Lambda) \subseteq \Gamma \qquad \Leftrightarrow \qquad \Lambda \subseteq r(\Gamma)$$

It follows that $ere(\Lambda) = e(\Lambda)$ as well as $rer(\Gamma) = r(\Gamma)$. Moreover, show that e preserves infinite joins, while r preserves infinite meets. In both cases, give a direct proof and a proof using the adjoinedness of the maps.

Exercise 198. Show that d is a left adjoint of r, where $d : \mathcal{E}\,\mathbf{K}_1 \to \mathcal{E}\,\mathbf{K}_2$ is defined by $d : \Lambda \mapsto \Lambda \otimes \Lambda$.

Exercise 199. Let t, the *twist map*, be defined as follows.

$$\begin{array}{lcl} p^t & := & p \\ (\neg\varphi)^t & := & \neg\varphi^t \\ (\varphi \wedge \psi)^t & := & \varphi^t \wedge \psi^t \\ (\Box\varphi)^t & := & \blacksquare\varphi^t \\ (\blacksquare\varphi)^t & := & \Box\varphi^t \end{array}$$

For a logic let $\Lambda^t := \{\varphi^t : \varphi \in \Lambda\}$. Show that Λ^t is again a logic, that $\Lambda^{tt} = \Lambda$ and that $(-)^t : \mathcal{E}\,\mathbf{K}_2 \to \mathcal{E}\,\mathbf{K}_2$ is an automorphism of the lattice of bimodal logics.

Exercise 200. (Continuing the previous exercise.) Show that $\mathbf{K}_2(X)^t = \mathbf{K}_2(X^t)$. Hence show that for all bimodal logics Γ, $\Delta := \Gamma \sqcup \Gamma^t$ is the smallest logic containing Γ such that $\Delta^t = \Delta$. Show also that if $\Gamma = \Lambda \otimes \Theta$ then $\Gamma \sqcup \Gamma^t = (\Lambda \sqcup \Theta) \otimes (\Lambda \sqcup \Theta)$.

Exercise 201. This exercise is given to motivate the notation of fusion as $\otimes$. Show namely that $\otimes$ distributes over arbitrary joins, thus behaves like a meet operator. Show also that it distributes over arbitary meets!

Exercise 202. Show that the lattice $\mathcal{E}\,\mathbf{K}_2 \oplus (\Box p \leftrightarrow \blacksquare p)$ is isomorphic to $\mathcal{E}\,\mathbf{K}_1$.

6.3. The Fundamental Construction

In this section we will prove Theorem 6.3.6. It says that a consistent bimodal logic $\Lambda \otimes \Theta$ is complete with respect to atomic frames iff both Λ and Θ are complete with respect to atomic frames. The proof is a successive construction of a model, and it allows similar results concerning completeness and finite model property. It allows to reduce the decision procedure in the bimodal logic $\Lambda \otimes \Theta$ to a decision procedure in the logic Λ (or Θ) given an $\Lambda \otimes \Theta$–oracle for formulae of smaller complexity. However, it is conditional on the completeness of Λ and Θ with respect to atomic frames. A proof of this fact without this assumption will be proved in the next section.

For a proper understanding of the method some terminology needs to be introduced. For each formula $\Box\psi, \blacksquare\psi \in \mathcal{L}_{\Box\blacksquare}$ we reserve a variable $q_{\Box\psi}$ and $q_{\blacksquare\psi}$ respectively, which we call the **surrogate** of $\Box\psi$ ($\blacksquare\psi$). $q_{\Box\psi}$ is called a **$\Box$–surrogate** and $q_{\blacksquare\psi}$ a **$\blacksquare$–surrogate**. We assume that the set of surrogate variables is distinct from our original set of variables. Any variable which is not a surrogate is called a **p–variable** and every formula composed exclusively from p–variables a **p–formula**. A p–variable is denoted by $p, p_1, \dots, p_i, \dots$ and an arbitrary variable by q. Finally, if φ is a formula, then $var^p(\varphi)$ denotes the set of p–variables of φ, and likewise the $var^\Box(\varphi)$, $var^\blacksquare(\varphi)$ denote the set of $\Box$–surrogates of φ and the set of $\blacksquare$–surrogates. The set of p–variables in $\mathcal{L}_{\Box\blacksquare}$ is assumed to be countably infinite.

Definition 6.3.1. *For a p–formula* φ *we define the* $\Box$*–**ersatz*** $\varphi^{\Box} \in \mathcal{L}_{\Box}$ *as follows.*

$$\begin{array}{rcl} q^{\Box} & := & q \\ (\varphi_1 \wedge \varphi_2)^{\Box} & := & \varphi_1^{\Box} \wedge \varphi_2^{\Box} \\ (\neg\varphi)^{\Box} & := & \neg\varphi^{\Box} \\ (\Box\varphi)^{\Box} & := & \Box\varphi^{\Box} \\ (\blacksquare\varphi)^{\Box} & := & q_{\blacksquare\varphi} \end{array}$$

Notice that the ersatz of φ is computed outside in and not inside out, which is typical for an inductive definition. For a set Γ of p–formulae call $\Gamma^{\Box} := \{\varphi^{\Box} : \varphi \in \Gamma\}$ the $\Box$–ersatz of Γ. Dually for $\blacksquare$.

Now let φ be composed either without $\blacksquare$–surrogates or without $\Box$–surrogates. Then we define the **reconstruction** of φ, $\uparrow\varphi$, as follows.

$$\begin{array}{rcl} \uparrow\varphi & := & \psi(\blacksquare\varphi_0^{\blacksquare}/q_{\blacksquare\varphi_0}, \ldots, \blacksquare\varphi_{\ell-1}^{\blacksquare}/q_{\blacksquare\varphi_{\ell-1}}, p_0, \ldots, p_{m-1}) \\ \uparrow\varphi & := & \varphi(\Box\varphi_0^{\Box}/q_{\Box\varphi_0}, \ldots, \Box\varphi_{\ell-1}^{\Box}/q_{\Box\varphi_{\ell-1}}, p_0, \ldots, p_{m-1}) \end{array}$$

Note that if $\uparrow$ is defined on φ it is also defined on $\uparrow\psi$; for if φ was free of $\blacksquare$–surrogates, $\uparrow\varphi$ is free of $\Box$–surrogates and vice versa. Now if φ is a p–formula then $\varphi^{\Box}$ is free of $\Box$–surrogates and therefore the reconstruction operator is defined on φ. Also, if $\uparrow$ is defined on φ then for some $n \in \omega$, $\uparrow^{n+1}\varphi = \uparrow^{n}\varphi$ (where $\uparrow^{n}$ denotes the nth iteration of $\uparrow$) which is the case exactly if $\uparrow^{n}\varphi$ is a p–formula. We then call $\uparrow^{n}\varphi$ the **total reconstruction** of φ and denote it by $\varphi^{\uparrow}$. $\varphi^{\uparrow}$ results from φ by replacing each occurrence of a surrogate q_{χ} in φ by χ. Now let φ be a p–formula. Then we put $\varphi_n := \uparrow^{n}(\varphi^{\Box})$. It is clear that $(\varphi^{\Box})^{\uparrow} = \varphi$. The $\Box$–**alternation–depth** of φ — $adp^{\Box}(\varphi)$ — is defined by $adp^{\Box}(\varphi) = min\{n : \varphi_n = \varphi\}$. For $m > adp^{\Box}(\varphi)$, $\varphi_m = \varphi_{m-1}$. The $\blacksquare$–**alternation depth**, $adp^{\blacksquare}(\varphi)$ is defined dually. Finally

$$adp(\varphi) := (adp^{\Box}(\varphi) + adp^{\blacksquare}(\varphi))/2\ .$$

It is easy to show that $|adp^{\Box}(\varphi) - adp^{\blacksquare}(\varphi)| \leq 1$. For example, if $\varphi \in \mathcal{L}_{\Box}$ then $adp^{\Box}(\varphi) = 0$ and $adp^{\blacksquare}(\varphi) = 1$ and so $adp(\varphi) = 1/2$. Conversely, $adp(\varphi) = 1/2$ implies $\varphi \in \mathcal{L}_{\Box} \cup \mathcal{L}_{\blacksquare}$. Finally, let us define the $\Box$**–depth** of a p–formula φ, $dp^{\Box}(\varphi)$, as follows.

$$\begin{array}{lcl} dp^{\Box}(p) & := & 0 \\ dp^{\Box}(\top) & := & 0 \\ dp^{\Box}(\neg\varphi) & := & dp^{\Box}(\varphi) \\ dp^{\Box}(\varphi \wedge \psi) & := & max\{dp^{\Box}(\varphi), dp^{\Box}(\psi)\} \\ dp^{\Box}(\Box\varphi) & := & 1 + dp^{\Box}(\varphi) \\ dp^{\Box}(\blacksquare\varphi) & := & dp(\varphi) \end{array}$$

Dually the $\blacksquare$**–depth** is defined.

Definition 6.3.2. *Let Λ be a (bimodal) logic and $\Delta \subseteq \mathcal{L}_{\Box\blacksquare}$ be a finite set. For a set $A \subseteq \Delta$ let*

$$\psi_A := \bigwedge\langle \chi : \chi \in A\rangle \wedge \bigwedge\langle \neg\chi : \chi \notin A\rangle \ .$$

The ***consistency set*** *of Δ, $C(\varphi)$, is defined by*

$$C(\Delta) := \{\psi_A : A \subseteq \Delta, \psi_A \text{ is } \Lambda\text{–consistent}\} \ .$$

The ***consistency formula*** *$\Sigma(\Delta)$ of Δ (with respect to Λ) is defined by*

$$\Sigma(\Delta) := \bigvee\langle \chi : \chi \in C(\Delta)\rangle \ .$$

If Δ is an infinite set then we define

$$\begin{array}{rcl} C(\Delta) & := & \bigcup\langle C(\Delta') : \Delta' \subseteq \Delta, \Delta' \textit{ finite}\rangle \\ \Sigma(\Delta) & := & \{\Sigma(\Delta') : \Delta' \subseteq \Delta, \Delta' \textit{ finite}\} \end{array}$$

Note that the consistency formulae are Λ–theorems. They depend of course on Λ, but we write $\Sigma(\Delta)$ rather than $\Sigma_\Lambda(\Delta)$. We abbreviate the consistency formula for the set $sf\{\psi : q_\psi \in var(\varphi^\Box)\} \cup var^p(\varphi)$ by $\Sigma_\Box(\varphi)$.

Theorem 6.3.3. *Let Λ and Θ be complete with respect to atomic frames. Then $\Lambda \otimes \Theta$ is also complete with respect to atomic frames.*

Proof. In the proof of Theorem 6.3.3 we construct not ordinary models but *partial models*. If $\mathfrak{G}$ is a frame and V a set of variables then $\beta : V \to \{0, 1, *\}^g$ is called a **partial valuation** if $\beta^{-1}(0)$, $\beta^{-1}(1)$ and $\beta^{-1}(*)$ are internal. Here, $0, 1$ are called the **standard** truth values and $*$ is the **undefined** or — to avoid confusion — the **nonstandard** truth value. We define the value of a formula according to the three–valued logic of 'inherent undefinedness' (or Weak Kleene Logic). It has the following truth tables

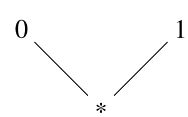
	$\neg$
0	1
1	0
$*$	$*$

$\wedge$	0	1	$*$
0	0	0	$*$
1	0	1	$*$
$*$	$*$	$*$	$*$

We put

$$\begin{array}{rcl} \overline{\beta}(\neg\varphi, x) & := & \neg\overline{\beta}(\varphi, x) \\ \overline{\beta}(\varphi \wedge \psi, x) & := & \overline{\beta}(\varphi, x) \wedge \overline{\beta}(\psi, x) \\ \overline{\beta}(\Box\varphi, x) & := & \bigwedge\langle\overline{\beta}(\varphi, y) : x \lhd y\rangle \end{array}$$

Note that by definition $\Box\varphi$ and $\Diamond\varphi$ receive a standard truth value iff *every* successor receives a standard truth value. We define the following order on the truth values



In the sequel we will assume that all valuations are defined on the entire set of variables. In contrast to what is normally considered a partial valuation, namely a partial function from the set of variables, the source of partiality or undefinedness is twofold. It may be *local*, when a variable or formula fails to be standard at a single world, or *global*, when a variable or formula is nonstandard throughout a frame. Our proof relies crucially on the ability to allow for local partiality. The **domain** of a valuation $\beta : V \to \{0, 1, *\}^g$ is the set of variables on which β is not globally partial i. e. $dom(\beta) := \{q : (\exists x \in g)\beta(q, x) \neq *\}$. If $\beta, \gamma : V \to \{0, 1, *\}^g$ we define $\beta \leq \gamma$ if $\beta(p, x) \leq \gamma(p, x)$ for all $p \in V$ and all $x \in g$. It is easy to see that if $\beta \leq \gamma$ then for all $x \in g$ and all φ with $var(\varphi) \subseteq V$: $\overline{\beta}(\varphi, x) \leq \overline{\gamma}(\varphi, x)$. Hence if β and γ are comparable then they assign equal *standard* truth values to formulae to which they both assign a standard truth value. In the proof we will only have the situation where a partial valuation β is nonstandard either on all $\Box$–surrogates or on all $\blacksquare$–surrogates. In the latter case we define for a point $x \in g$ and a set Δ of formulae

$$\begin{array}{rl} X_{\Box}^{\beta,\Delta}(x) := & \{\psi : \psi \in \Delta, \overline{\beta}(\psi^{\Box}, x) = 1\} \\ \cup & \{\neg\psi : \psi \in \Delta, \overline{\beta}(\psi^{\Box}, x) = 0\} \end{array}$$

and call $X_{\Box}^{\beta,\Delta}(x)$ the **characteristic set** of x in $\langle g, \beta\rangle$. If $X_{\Box}^{\beta,\Delta}(x)$ is finite (for example, if Δ is finite), then $\chi_{\Box}^{\beta,\Delta}(x) := \bigwedge X_{\Box}^{\beta,\Delta}(x)$ is the **characteristic formula** of x. And dually $X_{\blacksquare}^{\beta,\Delta}(x)$ and $\chi_{\blacksquare}^{\beta,\Delta}(y)$ are defined. We call a set Δ **sf–founded** if for all $\chi \in \Delta$ and $\tau \in sf(\chi)$ then either $\tau \in \Delta$ or $\neg\tau \in \Delta$.

Now we begin the actual proof of the theorem. Assume $\nvdash_{\Box\blacksquare} \neg\varphi$ and $adp^{\Box}(\varphi) = n$. Let

$$\mathbb{S}_i := sf\{\psi : q_\psi \in var(\varphi_i)\} \cup var^p(\varphi)$$

For $i = 0$ this is exactly the set of formulae on which the consistency formula for φ is defined. We will use an inductive construction to get a $\Lambda \otimes \Theta$–frame for φ. We will build a sequence $\langle\langle\mathfrak{G}_i, \beta_i, w_0\rangle : i \in \omega\rangle$ of models, which will be stationary for $i \geq adp^{\Box}(\varphi)$. The construction of the models shall satisfy the following conditions, which we spell out for $i = 2k$; for odd indices the conditions are dual.

[a]$_{2k}$ $\langle \mathfrak{G}_{2k}, \beta_{2k}, w_0\rangle \vDash \varphi_{2k}$

[b]$_{2k}$ $dom(\beta_{2k}) = var(\mathbb{S}^{\square}_{2k})$

[c]$_{2k}$ $\langle g_{2k}, \lhd_{2k}, \mathbb{G}_{2k}\rangle = \langle g_{2k-2}, \lhd_{2k-2}, \mathfrak{G}_{2k-2}\rangle \oplus \mathfrak{H}$ for some $\mathfrak{H}$, and
$\blacktriangleleft_{2k} = \blacktriangleleft_{2k-1}$

[d]$_{2k}$ $\mathfrak{G}_{2k}$ is an atomic frame and $\mathfrak{G}_{2k} \vDash \Lambda$

[e]$_{2k}$ For $x \in g_{2k-1}$:

(1) $\beta_{2k}(p, x) = \beta_{2k-1}(p, x), \quad p \in var(\varphi)$

(2) $\beta_{2k}(q_{\blacksquare\psi}, x) \leq \beta_{2k-1}(\blacksquare\psi^{\blacksquare}, x), \quad q_{\blacksquare\psi} \in var(\mathbb{S}^{\blacksquare}_{2k})$

(3) $\beta_{2k-1}(q_{\square\psi}, x) \leq \beta_{2k}(\square\psi^{\square}, x), \quad q_{\square\psi} \in var(\mathbb{S}^{\square}_{2k-1})$

[f]$_{2k}$ $X^{2k}(x) := X^{\beta_{2k},\mathbb{S}_{2k}}_{\blacksquare}(x)$ is $\Lambda \otimes \Theta$–consistent and sf–founded for
$x \in g_{2k} - g_{2k-1}$

We begin the construction as follows. Let φ be given; put $\delta := dp^{\square}(\varphi)$. Since φ is $\Lambda \otimes \Theta$–consistent, so is $\square^{\leq\delta}\Sigma_{\square}(\varphi)^{\square} \wedge \varphi^{\square}$. For $\Sigma_{\square}(\varphi)$ is a theorem of $\Lambda \otimes \Theta$. A fortiori, $\square^{\leq\delta}\Sigma_{\square}(\varphi)^{\square} \wedge \varphi^{\square}$ is Λ–consistent and has a model

$$\langle\langle g_0, \lhd_0, \mathbb{G}_0\rangle, \gamma_0, w_0\rangle \vDash \square^{\leq\delta}\Sigma_{\square}(\varphi)^{\square}; \varphi^{\square}$$

such that $\langle g_0, \lhd_0, \mathbb{G}_0\rangle$ is atomic and $dom(\gamma_0) = var(\mathbb{S}^{\square}_0)$. We put $\blacktriangleleft_0 := \varnothing$; this way, $\mathfrak{G}_0 := \langle g_0, \lhd_0, \blacktriangleleft_0, \mathfrak{G}_0\rangle$ is a bimodal frame. We may assume that the frame is rooted at w_0. Furthermore, we may assume that the valuation is nowhere partial for $q \in dom(\gamma_0)$, that is, $\gamma_0(q, x) \neq *$ for all x and all q. Let now I_0 be the set of points for which $X^{\gamma_0}_{\square}(x)$ is $\Lambda \otimes \Theta$–inconsistent; and let $I_k := \Diamond^k I_0 - \Diamond^{\leq k-1} I_0$ for all $k < \delta$. Finally, put $I_\delta := g_0 - \bigcup_{j<\delta} I_j$. All sets I_k are internal. (To see that, notice that the set of points with a given consistency set is internal; I_0 is a finite union of such sets.) We define a partial valuation $\beta_0 \leq \gamma_0$ as follows. $\beta_0(q_\psi, x) := *$ iff for some k, $x \in I_k$ and $dp^{\square}(\psi) > k$. Since $w_0 \in I_\delta$, $\varphi^{\square}$ is defined at w_0 and since $\beta_0 \leq \gamma_0$

$$\langle \mathfrak{G}_0, \beta_0, w_0\rangle \vDash \varphi^{\square} \ (= \varphi_0) \ .$$

Therefore, $[a]_0$, $[b]_0$ and $[d]_0$ hold. $[c]_0$ and $[e]_0$ are void, there is nothing to show. For $[f]_0$ note that $X^0(x) \subseteq X^{\gamma_0}(x)$; and the latter is consistent. To show that $X^0(x)$ is sf–founded notice that: (i) $\mathbb{S}_0$ is sf–founded; (ii) $\chi \in X^0(x)$ iff $\beta(\chi, x) \neq *$, by definition. And so it is enough to show that

$$(\ddagger) \qquad \text{If } \beta(\chi, x) \neq * \text{ and } \tau \in sf(\chi) \text{ then also } \beta_0(\tau, x) \neq * \ .$$

This is, however, immediate; for if $\beta_0(\tau, x) = *$ then there exists a k such that $x \in I_k$ and $dp^{\square}(\tau) > k$. From this follows $dp^{\square}(\chi) > k$, and so $\beta_0(\chi, x) = *$ as well.

The inductive step is done only for the case $i = 2k > 0$. For odd i the construction is dual. Assume $[a]_{2k}$–$[f]_{2k}$. For every point $y \in g_{2k} - g_{2k-1}$ we build a model

$$\langle\langle h_y, \blacktriangleleft_y, \mathbb{H}_y\rangle, \gamma_y, y\rangle \vDash \blacksquare^{dp^{\blacksquare}(\chi^{2k}(t))}\Sigma_{\blacksquare}(\chi^{2k}(t))^{\blacksquare}; \chi^{2k}(t)^{\blacksquare}$$

based on an atomic Θ–frame $\mathfrak{H}_y := \langle h_y, \blacktriangleleft_y, \mathbb{H}_y\rangle$, where $\chi^{2k}(y) := \chi^{\beta_{2k},\mathbb{S}_{2k}}_{\blacksquare}(y)$. This is possible since all the characteristic formulae are $\Lambda \otimes \Theta$–consistent and so their

$\blacksquare$–ersatz is Θ–consistent. We assume that $h_y \cap h_{y'} = \varnothing$ for $y \neq y'$, $h_y \cap g_{2k} = \{y\}$ and $h_y = Tr^{\blacksquare}(y, h_y)$. Now call x and y **C–equivalent** if $\chi^{2k}(x) = \chi^{2k}(y)$. We will assume that if x and y are C–equivalent there is an isomorphism ι_{xy} from the model $\langle \mathfrak{H}_x, \gamma_x, x\rangle$ to the model $\langle \mathfrak{H}_y, \gamma_y, y\rangle$. (This means that ι_{xy} is an isomorphism of the frames, that $\gamma_x(q, u) = \gamma_y(q, \iota_{xy}(u))$ for all $u \in h_x$ and that $\iota_{xy}(x) = y$.) Finally, the following composition laws shall hold for all $x, y, z \in g_{2k} - g_{2k-1}$: $\iota_{xx} = id$, $\iota_{yx} = \iota_{xy}^{-1}$, $\iota_{xz} = \iota_{xy} \circ \iota_{yz}$. This is always possible. For let Y be the set of all points C–equivalent to a given point y_0. For every $y \in Y$ let κ_y be an isomorphism from $\langle \mathfrak{H}_{y_0}, \gamma_{y_0}, y_0\rangle$ onto $\langle \mathfrak{H}_y, \gamma_y, y\rangle$. Then put $\iota_{xy} := \kappa_y \circ \kappa_x^{-1}$.

In case that $dp^{\blacksquare}(\chi^{2k}(y)) = 0$ we set in particular

$$\begin{aligned} h_y &:= \{y\} \\ \blacktriangleleft_y &:= \begin{cases} \{\langle y, y\rangle\} & \text{if } \boxed{\circ} \vDash \Theta \\ \varnothing & \text{otherwise} \end{cases} \end{aligned}$$

Clearly then $\beta_{2k}(q, y) = \gamma_y(q, y)$ for $q \in var(\mathbb{S}^{\blacksquare}_{2k})$. As before, I_0 is the set of z such that $\chi^{\gamma_y}_{\Box}(z)$ is inconsistent. Let $\delta := dp^{\blacksquare}(\chi^{\gamma_y}_{\Box}(z))$ and $k < \delta$. Then $I_k := \blacklozenge^k I_0 - \blacklozenge^{\leq k-1} I_0$, $I_\delta := h_y - \bigcup_{j<\delta} I_j$. We put $\beta_y(q, x) := *$ for $q \notin var(X^{2k}(y)^{\blacksquare})$ and $\beta_y(q_{\Box\psi}, x) = *$ if there is a k such that $x \in I_k$ and $dp^{\blacksquare}(\psi) > k$. In all other cases $\beta_y(q, x) := \gamma_y(q, x)$. Clearly, $\beta_y \leq \gamma_y$. Now observe that $var(\mathbb{S}^{\blacksquare}_{2k}) = var(\mathbb{S}^{\blacksquare}_{2k+1})$ and therefore $var(\chi^{2k}(y)^{\blacksquare}) \subseteq var(\mathbb{S}^{\blacksquare}_{2k+1})$. We can conclude that (1) $\chi^{2k}(y)^{\blacksquare}$ is defined at y in $\langle \mathfrak{H}_y, \beta_y\rangle$ and therefore $\langle\langle h_y, \blacktriangleleft_y, \mathbb{H}_y\rangle, \beta_y, y\rangle \vDash \chi^{2k}(y)^{\blacksquare}$ and that (2) $X^{\beta_y}(x)$ is consistent and sf–founded (using (‡)). Now let

$$\begin{aligned} g_{2k+1} &:= g_{2k} \cup \bigcup \langle h_y : y \in g_{2k} - g_{2k-1}\rangle \\ \lhd_{2k+1} &:= \lhd_{2k} \\ \blacktriangleleft_{2k+1} &:= \blacktriangleleft_{2k} \cup \bigcup \langle \blacktriangleleft_y : y \in g_{2k} - g_{2k-1}\rangle \end{aligned}$$

For the internal sets, some care is needed. Put $h := \bigcup\langle h_y : y \in g_{2k} - g_{2k-1}\rangle$. A subset $T \subseteq h$ is called **homogeneous** with **support** a if (1.) $a \subseteq g_{2k} - g_{2k-1}$, $a \in \mathbb{G}_{2k}$, (2.) $T = \bigcup_{x \in a} T \cap h_x$, (3.) all $x, y \in a$ are C–equivalent and $\iota_{xy}(T \cap h_x) = T \cap h_y$. A set T is called **homogeneous** if there exists an a such that T is homogeneous with support a. T is called **semihomogeneous** if it is a finite union of homogeneous sets (not necessarily with identical support).

$$\begin{aligned} \mathbb{G}_{2k+1} &:= \{S \cup T : S \in \mathbb{G}_{2k-1}, T \text{ semihomogeneous}\} \\ \mathfrak{G}_{2k+1} &:= \langle g_{2k+1}, \lhd_{2k+1}, \blacktriangleleft_{2k+1}, \mathbb{G}_{2k+1}\rangle \end{aligned}$$

We have to show that $\mathfrak{G}_{2k+1}$ is a (bimodal) frame. $\mathbb{G}_{2k+1}$ is clearly closed under finite unions. To see that it is closed under complements it is enough to show that the relative complement of a homogeneous set T in h, $h - T$, is semihomogeneous. $h - T$ is the union of the sets $u_y := h_y - T$, $y \in g_{2k} - g_{2k-1}$. Let a be the support of T, $U := \bigcup_{y \in a} u_y$ and $V := \bigcup_{y \in -a} u_y$. Then U is homogeneous with support a, and V is homogeneous with support $-a$. Moreover, $h - T = U \cup V$, and so the complement of T is semihomogeneous. Furthermore, we have to show that the internal sets are closed under $\Diamond$ and $\blacklozenge$. First $\Diamond$. Let $S \in \mathbb{G}_{2k+1}$. Then $\Diamond S = \Diamond(S \cap g_{2k})$. g_{2k} is the union of g_{2k-1} and $g_{2k} - g_{2k-1}$. g_{2k-1} is internal. $g_{2k} - g_{2k-1}$ is semihomogeneous with support $g_{2k} - g_{2k-1}$, for by construction, all frames $\mathfrak{H}_y$ are atomic, and so in particular the set $\{y\}$ is internal in $\mathfrak{H}_y$. Hence $\Diamond(S \cap g_{2k})$ is an internal set of $\mathfrak{G}_{2k}$ and so by the same argument an internal set of $\mathfrak{G}_{2k+1}$. Next, closure under $\blacklozenge$ must be shown. Take a set $S \cup T$, where $S \in \mathbb{G}_{2k-1}$ and T semihomogeneous. Then $\blacklozenge(S \cup T) = (\blacklozenge S) \cup (\blacklozenge T)$. By construction, $\blacklozenge S \in \mathbb{G}_{2k-1}$. Moreover, it is easy to see that if T' is homogeneous with support a, so is $\blacklozenge T'$. Thus, $\blacklozenge T$ is semihomogeneous, since T is. Finally, $\mathfrak{G}_{2k+1}$ is atomic. For if $x \in g_{2k-1}$ then $\{x\}$ is internal by $[\mathrm{d}]_{2k-1}$. If $x \notin g_{2k-1}$ then there exists a y such that $x \in h_y$. Then $\{x\}$ is an internal set of $\mathfrak{H}_y$, and since $\{y\}$ is internal in $\mathfrak{G}_{2k}$, $\{x\}$ is homogeneous with support $\{y\}$. (Here we use the fact that $\iota_{yy}(x) = x$.)

Define β_{2k+1} as follows. Put $\beta_{2k+1}(q, x) := \beta_y(q, x)$ for $x \in h_y$ and $\beta_{2k+1}(q, x) := \beta_{2k-1}(q, x)$ for $x \in g_{2k-1}, q \in var(\mathbb{S}^{\blacksquare}_{2k+1})$; in all other cases $\beta_{2k+1}(q, x) := *$. By construction, $[\mathrm{b}]_{2k+1}$ holds. $[\mathrm{c}]_{2k+1}$ holds because

$$\langle g_{2k+1}, \blacktriangleleft_{2k+1}, \mathbb{G}_{2k+1}\rangle = \langle g_{2k-1}, \blacktriangleleft_{2k-1}, \mathbb{G}_{2k-1}\rangle \oplus \mathfrak{H} ;$$

moreover, $\mathfrak{H}$ is based on the disjoint union of the $\langle h_y, \blacktriangleleft_y\rangle$, taking semihomogeneous sets as internal sets. And $\lhd_{2k+1} = \lhd_{2k}$. $[\mathrm{d}]_{2k+1}$ is immediate from the observation above that the frame is atomic, and from $[\mathrm{c}]_{2k+1}$, $[\mathrm{d}]_{2k-1}$ and $\mathfrak{H}_y \vDash \Theta$. Now we show $[\mathrm{e}]_{2k+1}$.

Ad (1). Let $x \in g_{2k-1}$. Then by $[\mathrm{e}]_{2k}$, $\beta_{2k}(p, x) = \beta_{2k-1}(p, x) = \beta_{2k+1}(p, x)$. If $x \in g_{2k} - g_{2k-1}$ then

$$\begin{array}{ll} & \beta_{2k+1}(p, x) = 1 \\ \Leftrightarrow & \beta_x(p, x) = 1 \\ \Leftrightarrow^{\dagger} & p \in X^{2k}(x) \\ \Leftrightarrow & \beta_{2k}(p, x) = 1 \end{array}$$

Here, $\Leftrightarrow^{\dagger}$ is true since $X^{2k}(x)$ is sf–founded and $dom(\beta_x) = var(X^{2k}(x))$. Similarly, $\beta_{2k+1}(p, x) = 0 \Leftrightarrow \beta_{2k}(p, x) = 0$ is shown.

Ad (2). If $x \in g_{2k-1}$ we have $\beta_{2k+1}(q_{\Box\psi}, x) = \beta_{2k-1}(q_{\Box\psi}, x) \leq \beta_{2k}(\Box\psi^{\Box}, x)$, by $[\mathrm{e}]_{2k}$. If $x \in g_{2k} - g_{2k-1}$ then

$$\begin{array}{ll} & \beta_{2k+1}(q_{\Box\psi}, x) = 1 \\ \Leftrightarrow & \beta_x(q_{\Box\psi}, x) = 1 \\ \Leftrightarrow & \Box\psi \in X^{2k}(x) \\ \Leftrightarrow & \beta_{2k}(\Box\psi^{\Box}, x) = 1 \end{array}$$

The argument continues as in (1).

Ad (3). If $x \in g_{2k-1}$ the claim follows by $[\mathrm{e}]_{2k}$. Now let $x \in g_{2k} - g_{2k-1}$. If $\beta_{2k}(q_{\blacksquare\psi}, x) = *$ then there is nothing to show. However, if $\beta_{2k}(q_{\blacksquare\psi}, x) \neq *$ then

$\blacksquare\psi \in X^{2k}(x)$ or $\neg\blacksquare\psi \in X^{2k}(x)$ and thus

$$\begin{array}{rl} & \beta_{2k+1}(\blacksquare\psi^{\blacksquare}, x) = 1 \\ \Leftrightarrow & \beta_x(\blacksquare\psi^{\blacksquare}, x) = 1 \\ \Leftrightarrow & \blacksquare\psi \in X^{2k}(x) \\ \Leftrightarrow & \beta_{2k}(q_{\blacksquare\psi}, x) = 1 \end{array}$$

$[f]_{2k+1}$ holds because of $[c]_{2k+1}$ and by the definition of β_{2k+1} and finally because of (2) of $[e]_{2k+1}$. $[a]_{2k+1}$ follows directly from $[e]_{2k+1}$ (1) and (3).

If $n = adp^{\Box}(\varphi)$ we have $g_{n+1} = g_n$ and $dp^{\blacksquare}(\chi^n(y)) = dp^{\Box}(\chi^n(y)) = 0$ for all y since $\mathbb{S}_n = var(\varphi)$ and therefore $dom(\beta_n) = var(\varphi)$, by $[b]_n$. By construction, the $\mathfrak{H}_y$ are based on a single point and thus g_{n+1} is based on the same points as g_n. Moreover, by $[d]_n$ and $[d]_{n+1}$, $\mathfrak{G}_{n+1} \vDash \Lambda \otimes \Theta$ and by $[a]_{n+1}$, $\langle \mathfrak{G}_{n+1}, \beta_{n+1}, w_0\rangle \vDash \varphi_{n+1}$ $(= \varphi)$. Take any valuation $\gamma \geq \beta_{n+1}$ which is standard for the p–variables. Then $\langle \mathfrak{G}_{n+1}, \gamma, w_0\rangle \vDash \varphi$. □

Corollary 6.3.4. *Suppose that $\Lambda \otimes \Theta$ are complete with respect to atomic frames. Let φ be a bimodal formula, $m \geq dp^{\Box}(\varphi)$, and $n \geq dp^{\blacksquare}(\varphi)$. Then the following are equivalent.*

(1) $\vdash_{\Box\blacksquare} \varphi$.
(2) $\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box}$.
(3) $\vdash_{\blacksquare} \blacksquare^{\leq n}\Sigma_{\blacksquare}(\varphi)^{\blacksquare} \to \varphi^{\blacksquare}$.

The proof of the latter theorem is evident from the construction used in the previous proof.

Theorem 6.3.5. *Suppose that Λ and Θ are complete with respect to atomic frames. Then $\Lambda \otimes \Theta$ is decidable iff both Λ and Θ are decidable.*

Proof. By induction on $n := adp(\varphi)$. If $n = 0$, φ is boolean and since $\bot \notin \Lambda \otimes \Theta$ $\vdash_{\Box\blacksquare} \varphi$ iff φ is a boolean tautology. Since the propositional calculus is decidable, this case is settled. Now suppose that for all ψ with $adp(\psi) < n$ we have shown the decidability of $\vdash_{\Box\blacksquare} \psi$. We know by Theorem 6.3.4 that for $m \geq max\{dp^{\Box}(\varphi), dp^{\blacksquare}(\varphi)\}$

$$\begin{array}{rcl} \vdash_{\Box\blacksquare} \varphi & \Leftrightarrow & \vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box} \\ \vdash_{\Box\blacksquare} \varphi & \Leftrightarrow & \vdash_{\blacksquare} \blacksquare^{\leq m}\Sigma_{\blacksquare}(\varphi)^{\blacksquare} \to \varphi^{\blacksquare} \end{array}$$

Therefore we can decide $\vdash_{\Box\blacksquare} \varphi$ on the condition that either $\Sigma_{\Box}(\varphi)$ or $\Sigma_{\blacksquare}(\varphi)$ can be constructed. But now either $adp(\Sigma_{\Box}(\varphi)) < n$ or $adp(\Sigma_{\blacksquare}(\varphi)) < n$. This is seen as follows. Suppose that $adp^{\Box}(\varphi) \leq adp^{\blacksquare}(\varphi)$. Then there is a maximal chain of nested alternating modalities starting with $\Box$. Then any maximal chain of nested alternating modalities in $\Sigma_{\Box}(\varphi)$ starts with $\blacksquare$ (!) and is a subchain of of such a chain in φ. Consequently, we have $adp^{\blacksquare}(\Sigma_{\Box}(\varphi)) < adp^{\blacksquare}(\varphi)$ and with $adp^{\Box}(\Sigma_{\Box}(\varphi)) \leq adp^{\Box}(\varphi)$ the claim follows. Now assume that $adp(\Sigma_{\Box}(\varphi)) < adp(\varphi)$ is the case. Then

$$\Sigma_{\Box}(\varphi) = \bigvee\langle \psi_c : c \subseteq C, \nvdash_{\Box\blacksquare} \neg\psi_c\rangle$$

Consequently, $\Sigma_{\Box}(\varphi)$ can be constructed if only $\vdash_{\Box\blacksquare} \neg\psi_c$ is decidable for all c. But this is so because $adp(\neg\psi_c) < n$. ☐

Note that for $m > 1$ we have $adp^{\Box}(\Box^{\leq m}\Sigma_{\Box}(\varphi)) \leq adp^{\Box}(\varphi)$. However,

$$adp^{\blacksquare}(\Box^{\leq m}\Sigma_{\Box}(\varphi)) \leq adp^{\blacksquare}(\varphi) + 1 \ .$$

A case where the inequalities are sharp is given by $\varphi = \blacksquare p$. But in all these cases $adp^{\Box}(\varphi) > adp^{\blacksquare}(\varphi)$ in which case we also have

$$adp^{\Box}(\blacksquare^{\leq m}\Sigma_{\blacksquare}(\varphi)) \leq adp^{\Box}(\varphi)$$

and

$$adp^{\blacksquare}(\Sigma_{\blacksquare}(\varphi)) < adp^{\Box}(\blacksquare^{\leq m}\Sigma_{\blacksquare}(\varphi)) \leq adp^{\Box}(\varphi)$$

and therefore $adp(\blacksquare^{\leq m}\Sigma_{\blacksquare}(\varphi)) \leq adp(\varphi)$.

Corollary 6.3.6. *Suppose* $\bot \notin \Lambda, \Theta$. *Then* $\Lambda \otimes \Theta$ *is complete iff both* Λ *and* Θ *are complete.*

Corollary 6.3.7. *Suppose that* $\bot \notin \Lambda, \Theta$. *Then* $\Lambda \otimes \Theta$ *has the finite model property iff both* Λ *and* Θ *have the finite model property.*

Notes on this section. Edith Spaan [**202**] has investigated the complexity of the fusion of modal logics. If Λ and Θ are both $\mathcal{C}$–hard, so is $\Lambda \otimes \Theta$. If Λ and Θ are both in PSPACE, then so is $\Lambda \otimes \Theta$. The fusion may however become PSPACE–hard even if the individual factors are in NP. For example, Joseph Halpern and Y. Moses [**95**] have shown that $\mathbf{S5} \otimes \mathbf{S5}$ is PSPACE hard. However, **S5** is itself in NP by a result of R. Ladner [**137**]. (See also the exercises of Section 3.1.)

Exercise 203. Show with a specific example that there are monomodal logics Λ and Θ which are consistent and tabular such that $\Lambda \otimes \Theta$ is not tabular. Show furthermore that $\Lambda \otimes \Theta$ has the finite model property and that it is tabular iff one of Λ, Θ is of codimension 1.

***Exercise 204.** (Wolter [**237**].) As the previous exercise has shown, there are logics whose lattice of extension is finite, yet the lattice of extensions of their fusion is infinite. Now show that the lattice $\mathcal{E}\,\mathbf{K.T}$ is isomorphic to the lattice $\mathcal{E}(\mathbf{K.alt}_1 \otimes \mathbf{S5})$. As we will see, each of the lattices $\mathcal{E}\,\mathbf{K.alt}_1$ and $\mathcal{E}\,\mathbf{S5}$ is countable, but $\mathcal{E}(\mathbf{K} \oplus p \to \Diamond p)$ is uncountable.

***Exercise 205.** Let $\Lambda = \mathbf{K.alt}_1.\Box^3\bot$. Show that $\mathcal{E}(\Lambda \otimes \Lambda)$ has $2^{\aleph_0}$ elements.

***Exercise 206.** The following three exercises will establish that the requirement of completeness for atomic frames can be lifted in the Consistency Reduction Theorem when we consider only extensions of **K4**. The next section will establish this for all logics but with a different technique. Call a **K4**–frame $\mathfrak{F}$ **separable** if for all $x \in f$ there exists a $a \in \mathbb{F}$ such that $x \lhd y$ and $y \in a$ implies $y \lhd x$. Show that every extension of **K4** is complete with respect to separable frames. *Hint.* Take the canonical frame $\mathfrak{Can}_\Lambda(n)$. Call W *eliminable* if for all formulae φ such that $W \in \widehat{\varphi}$ there exists a $V \in \widehat{\varphi}$ such that $W \lhd V \ntriangleleft W$. Show that if W is eliminable and $\varphi \in W$ then there exists a noneliminable V such that $W \lhd V$ and $\varphi \in V$. Now take the set N of noneliminable points and let $\mathfrak{N}$ be the frame based on N. (Since N is not internal, this is not a subframe.) Show that $\mathsf{Th}\,\mathfrak{N} = \mathsf{Th}\,\mathfrak{Can}_\Lambda(n)$. (The term *separable* is taken from Rautenberg [**169**]. The completeness result is from Fine [**63**].)

Exercise 207. (Continuing the previous exercise.) Call a frame $\mathfrak{G}$ **hooked** if it is rooted, and that for the root w_0 the set $\{w_0\}$ is internal. Show that the results of this section can be generalized to logics which are complete with respect to hooked frames.

Exercise 208. (Continuing the previous exercise.) Show that all extensions of **K4** are complete with respect to hooked frames.

6.4. A General Theorem for Consistency Reduction

The results and proofs of this section are from Frank Wolter [**243**]. They make use of another type of completeness for logics, namely with respect to algebras which are atomless. For simplicity we assume that $\kappa \leq \aleph_0$ throughout this section. Recall that an element x of a boolean algebra is an atom if for all y such that $0 < y \leq x$ we have $y = x$. A boolean algebra is called **atomless** if it has no atoms. A modal algebra is called atomless, if its boolean reduct is atomless. An atomless algebra is either isomorphic to **1** or infinite. The following is a well–known fact about boolean algebras (see [**117**]).

Proposition 6.4.1. *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two countably infinite atomless boolean algebras. Then $\mathfrak{A} \cong \mathfrak{B}$.*

Moreover, let $\mathfrak{A}$ be atomless and countable and $a > 0$. Then let $\mathfrak{A}_a$ be the algebra based on all sets $b \leq a$, with the operations being intersection and relative complement. Then $\mathfrak{A}_a$ is also countable and atomless, and $\mathfrak{A}_a$ is not isomorphic to **1**. Hence, it is countably infinite and so by the previous theorem isomorphic to $\mathfrak{A}$.

Definition 6.4.2. *Let $\mathfrak{A}$ be a boolean algebra. A family $\{a_i : i \in n\}$ is a **partition** of $\mathfrak{A}$ if (1.) $a_i \neq 0$ for all $i < n$, (2.) $a_i \cap a_j = 0$ if $i \neq j$ and (3.) $\bigcup\langle a_i : i < n\rangle = 1$.*

Lemma 6.4.3. *Let $\mathfrak{A}$ be a countably infinite atomless boolean algebra and $\{a_i : i < n\}$ be a partition of $\mathfrak{A}$. Then $\sigma : \mathfrak{A} \to \prod_{i<n} \mathfrak{A}_{a_i}$ defined by $\sigma(x) := \langle x \cap a_i : i < n\rangle$ is an isomorphism. Moreover, if $\{a_i : i < n\}$ is a partition of $\mathfrak{A}$ and $\{b_i : i < n\}$ a partition of $\mathfrak{B}$, then there exists an isomorphism $\tau : \mathfrak{A} \to \mathfrak{B}$ such that $\tau(a_i) = b_i$.*

Proof. The proof of the first claim is as follows. The map $x \mapsto x \cap a$ is a homomorphism from $\mathfrak{A}$ onto $\mathfrak{A}_a$. Hence, σ defined above is a homomorphism. Now suppose that $\sigma(x) = \sigma(y)$. Then $x \cap a_i = y \cap a_i$ for all $i < n$. So we have $x \cap \bigcup_{i<n} a_i =$

$y \cap \bigcup_{i<n} a_i$. Since $\bigcup_{i<n} a_i = 1$, $x = y$. Hence σ is an isomorphism. Now let $\{a_i : i < n\}$ be a partition of $\mathfrak{A}$, and $\{b_i : i < n\}$ a partition of $\mathfrak{B}$. Let $\sigma : x \mapsto \langle x \cap a_i : i < n\rangle$, and $\widehat{\sigma} : x \mapsto \langle x \cap b_i : i < n\rangle$. These are isomorphisms. Furthermore, there exist isomomorphisms $h_i : \mathfrak{A}_{a_i} \to \mathfrak{B}_{b_i}$. Put $\upsilon := \prod_{i<n} h_i$. Finally, $\tau := \widehat{\sigma}^{-1} \circ \upsilon \circ \sigma$. This is an isomorphism, and

$$\begin{aligned} \tau(a_i) &= \widehat{\sigma}^{-1} \circ \upsilon(\langle 0, \ldots, 0, 1, 0, \ldots, 0\rangle) \\ &= \widehat{\sigma}^{-1}(\langle 0, \ldots, 0, 1, 0, \ldots, 0\rangle) \\ &= b_i \end{aligned}$$

$\square$

Proposition 6.4.4. ($\kappa < \aleph_1$.) *Let Λ be a κ–modal logic and λ an infinite cardinal. Then $\mathfrak{Fr}_\Lambda(\lambda)$ is atomless. Moreover, $\mathfrak{Fr}_\Lambda(\aleph_0)$ is countable.*

If Λ is inconsistent then $\mathfrak{Fr}_\Lambda(\lambda)$ is finite and isomorphic to $\mathbf{1}$. In all other cases it is infinite.

Definition 6.4.5. *Let Λ a κ–modal logic. By $\mathsf{Atg}\,\Lambda$ we denote the set of countably infinite atomless algebras $\mathfrak{A}$ of $\mathsf{Alg}\,\Lambda$.*

The following is now immediate.

Proposition 6.4.6. ($\kappa \leq \aleph_0$.) *Let Λ be a consistent κ–modal logic. Then $\Lambda = \mathsf{Th}\,\mathsf{Atg}\,\Lambda$.*

Theorem 6.4.7 (Wolter, Global Consistency Reduction). ($\kappa < \aleph_1$.) *Let Λ and Θ be consistent monomodal logics. Then the following are equivalent*

(1) $\varphi \nVdash_{\blacksquare\Box} \psi$
(2) *There exists $\Delta \subseteq C_\Box(\varphi; \psi)$ such that*
  (a) $\varphi^\Box; (\bigvee \Delta)^\Box) \nVdash_\Box \psi^\Box$
  (b) *for all $\chi \in \Delta$,* $\varphi^\Box; (\bigvee \Delta)^\Box \nVdash_\Box \neg\chi^\Box$ *and* $(\bigvee \Delta)^\blacksquare \nVdash_\blacksquare \neg\chi^\blacksquare$
(3) *There exists $\Delta \subseteq C_\blacksquare(\varphi; \psi)$ such that*
  (a) $\varphi^\blacksquare; (\bigvee \Delta)^\blacksquare) \nVdash_\blacksquare \psi^\blacksquare$
  (b) *for all $\chi \in \Delta$,* $\varphi^\blacksquare; (\bigvee \Delta)^\blacksquare \nVdash_\blacksquare \neg\chi^\blacksquare$ *and* $(\bigvee \Delta)^\Box \nVdash_\Box \neg\chi^\Box$

If Δ satisfies 2., then $\varphi; \bigvee \Delta \nVdash_{\Box\blacksquare} \psi$ and $\bigvee \Delta \nVdash_{\Box\blacksquare} \neg\chi$ for all $\chi \in \Delta$.

Proof. Obviously, it is enough to prove the equivalence of (1.) and (2.). Moreover, (1.) implies (2.). For suppose that $\varphi \nVdash_{\Box\blacksquare} \psi$. Then there exists a bimodal algebra $\mathfrak{B}$ and a valuation β such that $\overline{\beta}(\varphi) = 1$ but $\overline{\beta}(\psi) \neq 1$. Let $\Delta := \{\chi \in sf^\Box(\varphi; \psi) : \overline{\beta}(\chi) > 0\}$. Then Δ is as required for (2.). So, in the remainder of the proof we will show that (2.) implies (1.). Assume therefore that $\Delta \subseteq C_\Box(\varphi; \psi)$ exists satisfying the requirements under (2.). Then for each $\chi \in \Delta \cup \{\neg\psi\}$ there exists a $\mathfrak{B}_\chi \in \mathsf{Atg}\,\Lambda$ and a valuation β_χ such that

$$\overline{\beta}_\chi((\bigvee \Delta)^\Box \wedge \varphi^\Box) = 1 \text{ and } \overline{\beta}_\chi(\chi^\Box) > 0\ .$$

Let

$$\mathfrak{B} := \prod \langle \mathfrak{B}_\chi : \chi \in \Delta \cup \{\neg\psi\}\rangle$$

and

$$\beta^\Box := \prod \langle \beta_\chi : \chi \in \Delta \cup \{\neg\psi\}\rangle$$

Then $\mathfrak{B} \in \mathsf{Atg}\,\Lambda$, $\overline{\beta}^\Box(\varphi^\Box) = 1$, $\overline{\beta}^\Box(\psi^\Box) \neq 1$ and the set

$$\{\overline{\beta}^\Box(\chi^\Box) : \chi \in \Delta\}$$

is a partition of $\mathfrak{B}$. Since $(\bigvee \Delta)^\blacksquare \nVdash_\blacksquare \neg\chi^\blacksquare$ for all $\chi \in \Delta$ we get in a similar way a $\mathfrak{C} \in \mathsf{Atg}\,\Theta$ and a valuation $\beta^\blacksquare$ such that the set

$$\{\overline{\beta}^\blacksquare(\chi^\blacksquare) : \chi \in \Delta\}$$

is a partition of $\mathfrak{C}$. By Proposition 6.4.3 there is an isomorphism σ from the boolean reduct of $\mathfrak{C}$ onto the boolean reduct of $\mathfrak{B}$ such that for all $\chi \in \Delta$

$$\sigma(\overline{\beta}^\blacksquare(\chi^\blacksquare)) = \overline{\beta}^\Box(\chi^\Box)\ .$$

Via this isomorphism we define the following algebra

$$\mathfrak{A} := \langle B, 1, -, \cap, \Box', \blacksquare'\rangle$$

where $\Box' a := \Box a$ and $\blacksquare' a := \sigma^{-1}(\blacksquare\sigma(a))$. Clearly, $\mathfrak{A} \in \mathsf{Atg}(\Lambda \otimes \Theta)$. It holds for all $\eta \in sf^\Box(\varphi;\psi)$ that

$$\begin{aligned}\overline{\beta}^\Box(\eta^\Box) &= \bigvee\langle \overline{\beta}^\Box(\chi^\Box) : \chi \in \Delta, \eta \text{ a conjunct of } \chi\rangle\\ &= \bigvee\langle \overline{\beta}^\blacksquare(\chi^\blacksquare) : \chi \in \Delta, \eta \text{ a conjunct of } \chi\rangle\\ &= \overline{\beta}^\blacksquare(\eta^\blacksquare)\end{aligned}$$

We now define a valuation γ on the p–variables of φ and ψ by

$$\gamma(p) := \beta^\Box(p) \quad (= \beta^\blacksquare(p))\ .$$

We claim that for all $\eta \in sf^\Box(\varphi;\psi)$

$$\gamma(\eta) = \overline{\beta}^\Box(\eta^\Box) \quad (= \overline{\beta}^\blacksquare(\eta^\blacksquare))\ .$$

The proof is by induction on the complexity of η. The claim holds by construction for the p–variables. Moreover, the steps for $\neg$ and $\wedge$ are straightforward. Now let $\eta = \Box\theta$. Then

$$\begin{aligned}\overline{\gamma}(\Box\theta) &= \Box'\overline{\gamma}(\theta)\\ &= \Box'\overline{\gamma}^\Box(\theta^\Box)\\ &= \Box'\overline{\beta}^\Box(\theta)\\ &= \overline{\beta}^\Box(\Box\theta)\end{aligned}$$

Analogously, the case $\eta = \blacksquare\theta$ is dealt with. With the claim being proved, we have $\overline{\gamma}(\varphi) = \overline{\beta}^\Box(\varphi^\Box) = 1$ and $\overline{\gamma}(\psi) = \overline{\beta}^\Box(\psi^\Box) \neq 1$. And that had to be shown. ◻

THEOREM 6.4.8 (Wolter, Local Consistency Reduction). ($\kappa < \aleph_1$.) *Let* Λ *and* Θ *be consistent normal monomodal logics. Let* φ *be a bimodal formula, and* $m \geq dp^{\Box}(\varphi)$, $n \geq dp^{\blacksquare}(\varphi)$. *Then the following are equivalent*

(1) $\vdash_{\blacksquare\Box} \varphi$.
(2) $\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box}$.
(3) $\vdash_{\blacksquare} \blacksquare^{\leq n}\Sigma_{\blacksquare}(\varphi)^{\blacksquare} \to \varphi^{\blacksquare}$.

LEMMA 6.4.9. *Suppose that* $\Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box} \notin \Lambda$. *Then there exists* $\mathfrak{A} \in \mathsf{Atg}\,\Lambda$, *a valuation* $\beta^{\Box}$ *and a sequence* $\langle a_i : 0 \leq i \leq m\rangle$ *such that the following holds:*

(a1) $a_{i+1} \leq a_i \cap \Box a_i$ *for all* $i < m$.
(a2) $a_i \leq \overline{\beta}^{\Box}(\Box^{\leq i}\Sigma_{\Box}(\varphi)^{\Box})$, *for all* $0 \leq i \leq m$.
(a3) $a_m \cap \overline{\beta}^{\Box}(\neg\varphi^{\Box}) \neq 0$.
(a4) *The set* $\{\overline{\beta}^{\Box}(\chi^{\Box}) \cap a_m : \chi \in \Sigma_{\Box}(\varphi)\}$ *is a partition of* $\mathfrak{A}_{a_m}$.
(a5) *The sets* $\{\overline{\beta}^{\Box}(\chi^{\Box}) \cap (a_i - a_{i+1}) : \chi \in \Sigma_{\Box}(\varphi)\}$
are partitions of $\mathfrak{A}_{(a_i - a_{i+1})}$, *for all* $i < m$.

PROOF. There exists a $\mathfrak{B} \in \mathsf{Atg}\,\Lambda$ and a valuation γ such that

$$\overline{\gamma}(\neg\varphi^{\Box} \wedge \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box}) > 0 \ .$$

We put for $0 \leq i \leq m$,

$$b_i := \overline{\gamma}(\Box^{\leq i}\Sigma_{\Box}(\varphi)^{\Box}) \ .$$

Take for each $i \leq m$ an algebra $\mathfrak{C}_i \in \mathsf{Atg}\,\Lambda$ and valuations δ_i such that

$$\{\overline{\delta}_i(\chi^{\Box}) : \chi \in \Sigma_{\Box}(\varphi)\}$$

is a partition of $\mathfrak{C}_i$. (For example, $\mathfrak{C}_i := \mathfrak{Fr}_{\Lambda}(X)$, where $X := var(\Sigma_{\Box}(\varphi)^{\Box})$, and $\delta_i : p \mapsto \widehat{p}$ the natural valuation, as defined in Section 2.8.) Put

$$\mathfrak{A} := \mathfrak{B} \times \prod \langle \mathfrak{C}_i : i \leq m\rangle, \qquad \beta^{\Box} := \gamma \times \prod \langle \delta_i : i \leq m\rangle$$

and define the a_i by

$$\begin{array}{lcl} a_0 & := & \langle b_0, 1, 1, \dots, 1\rangle \\ a_1 & := & \langle b_1, 0, 1, \dots, 1\rangle \\ & \dots & \dots \\ a_m & := & \langle b_m, 0, 0, \dots, 0, 1\rangle \ . \end{array}$$

Then the elements $\langle a_i : 0 \leq i \leq m\rangle$ and the valuation $\beta^{\Box}$ satisfy (a1) – (a5).

Ad (a1). $\Box a_i = \langle b_{i+1}, \Box 0, \Box 0, \dots, \Box 0, 1, \dots, 1\rangle$. Therefore we have $a_i \cap \Box a_i = \langle b_{i+1}, 0, \dots, 0, 1, \dots, 1\rangle$ (i times 0). Thus $a_{i+1} \leq a_i \cap \Box a_i$.

Ad (a2). By definition of δ_i, $\overline{\delta}_i(\Sigma_{\Box}(\varphi)^{\Box}) = 1$. Therefore $\overline{\beta}^{\Box}(\Box^{\leq i}\Sigma_{\Box}(\varphi)^{\Box}) = \langle b_i, 1, \dots, 1\rangle = a_i$.

Ad (a3). $a_m \cap \overline{\beta}^{\Box}(\neg\varphi^{\Box}) = \langle \overline{\gamma}(\neg\varphi^{\Box}), 0, \dots, 0, \overline{\delta}_i(\neg\varphi^{\Box})\rangle > 0$.

Ad (a4). Clearly, $\overline{\beta}^{\Box}(\chi_1^{\Box}) \cap a_m$ and $\overline{\beta}^{\Box}(\chi_2^{\Box}) \cap a_m$ are disjoint for $\chi_1 \neq \chi_2$, and the sum of all of these sets is $= a_m$. Moreover, by choice of a_m and the definition of $\Sigma_{\Box}(\varphi)$,

these elements are all distinct from 0.
Ad (a5). Again, the members $\{\overline{\delta}_i(\chi^\Box) : \chi \in \Sigma_\Box(\varphi)\}$ are pairwise distinct and their sum is $a_i - a_{i+1}$. Also, by choice of the a_i, none of the elements is $= 0$. $\Box$

Lemma 6.4.10. *Suppose that $\Box^{\leq m}\Sigma_\Box(\varphi)^\Box \to \varphi^\Box \notin \Lambda$. Then there exists a $\mathfrak{A} \in \mathsf{Atg}\,\Lambda \otimes \Theta$ such that there are valuations $\beta^\Box$ and $\beta^\blacksquare$ and a sequence $\langle a_i : 0 \leq i \leq m\rangle$ satisfying (a1) – (a5) and*

(a6) $\overline{\beta}^\Box(\chi^\Box) \cap a_0 = \overline{\beta}^\blacksquare(\chi^\blacksquare) \cap a_0$*, for all* $\chi \in C(\varphi)$*.*

(a7) *For all* $0 \leq i \leq m$ *and all* $b \in A$*,* $\quad a_i \cap \blacksquare b = \blacksquare(a_i \cap b)$*.*

Proof. For each $i \leq m$ take a $\mathfrak{B}_i \in \mathsf{Atg}\,\Theta$ and a valuation γ_i such that

$$\{\overline{\gamma}_i(\chi^\blacksquare) : \chi \in C(\varphi)\}$$

is a partition of $\mathfrak{B}_i$. In addition, take an arbitrary $\mathfrak{B}_{-1}$ and an arbitrary valuation γ_{-1} on $\mathfrak{B}_{-1}$. Put

$$\begin{array}{lcl} \mathfrak{C} & := & \prod\langle \mathfrak{B}_i : -1 \leq i \leq m\rangle \\ \gamma^\blacksquare & := & \prod\langle \gamma_i : -1 \leq i \leq m\rangle \end{array}$$

Now take $\mathfrak{D} \in \mathsf{Atg}\,\Lambda$ and a valuation $\gamma^\Box$ such that the conditions (a1) – (a5) are satisfied. Put $a_{-1} := 1^{\mathfrak{D}} - a_0$. Without loss of generality we may assume that $a_{-1} \neq 0$. There exist boolean isomorphisms $\sigma_m : \mathfrak{B}_m \to \mathfrak{D}_{a_i}$ and $\sigma_i : \mathfrak{B}_i \to \mathfrak{D}_{a_i - a_{i+1}}$ such that for all $i < m$

$$\sigma_m(\overline{\gamma}^\blacksquare(\chi^\blacksquare)) = \overline{\gamma}^\Box(\chi^\Box) \cap a_m \quad \text{and} \quad \sigma_i(\overline{\gamma}^\blacksquare(\chi^\blacksquare)) = \overline{\gamma}^\Box(\chi^\Box) \cap (a_i - a_{i+1})\ .$$

Fix an arbitrary boolean isomorphism $\sigma_{-1} : \mathfrak{B}_{-1} \to \mathfrak{D}_{a_{-1}}$. Now the map σ defined by $\sigma := \prod\langle \sigma_i : -1 \leq i \leq m\rangle$ is an isomorphism from $\mathfrak{C}$ to $\mathfrak{D}$. Using this isomorphism we define $\mathfrak{A}$ as follows.

$$\begin{array}{lcl} \mathfrak{A} & := & \langle D, 1, -, \cap, \Box', \blacksquare'\rangle \\ \Box' a & := & \Box a \\ \blacksquare' a & := & \sigma^{-1}(\blacksquare\sigma(a)) \end{array}$$

Put $\beta^\blacksquare(p) := \sigma(\gamma^\blacksquare(p))$ and $\beta^\Box(p) := \gamma^\blacksquare(p)$. Then **(a6)** holds by construction. **(a7)** holds by virtue of the fact that $\mathfrak{A}$ is isomorphic to a product of $\blacksquare$–algebras based on the relative boolean algebras $\mathfrak{A}_{a_i}$ and $\mathfrak{A}_{1-a_i}$, for all $0 \leq i \leq m$. $\Box$

Proof of Theorem 6.4.8. The equivalence between (1.) and (3.) is proved in the same way as the equivalence between (1.) and (2.). So we perform only the latter. Moreover, (2.) implies (1.) by definition of the consistency formulae. So, let us assume that (2.) fails. We will show that (1.) fails as well. Suppose therefore that $\Box^{\leq m}\Sigma_\Box(\varphi)^\Box \to \varphi^\Box \notin \Lambda$, where $m \geq dp^\Box(\varphi)$. Take $\mathfrak{A} \in \mathsf{Atg}\,\Lambda \otimes \Theta$ and valuations $\beta^\Box$ and $\beta^\blacksquare$ satisfying (a1) – (a7). Using (a4), (a5) and (a6) and the definition of $\Sigma_\Box(\varphi)$ it is shown that for all $\eta \in sf^{\Box}(\varphi)$

$$\begin{array}{rcl} a_0 \cap \overline{\beta}^{\Box}(\eta^{\Box}) & = & a_0 \cap \bigcup\langle \overline{\beta}^{\Box}(\chi^{\Box}) : \chi \in \Sigma_{\Box}(\varphi), \eta \text{ a conjunct of } \chi\rangle \\ & = & a_0 \cap \bigcup\langle \overline{\beta}^{\blacksquare}(\chi^{\blacksquare}) : \chi \in \Sigma_{\Box}(\varphi), \eta \text{ a conjunct of } \chi\rangle \\ & = & a_0 \cap \overline{\beta}^{\blacksquare}(\eta^{\blacksquare}) \end{array}$$

We define a valuation β on $\mathfrak{A}$ by putting for all p–variables p of φ

$$\beta(p) := \beta^{\Box}(p) .$$

We claim that for all $0 \leq k \leq m$ and for all $\eta \in sf^{\Box}(\varphi)$ such that $dp^{\Box}(\eta) \leq k$:

$$a_k \cap \overline{\beta}(\eta) = a_k \cap \overline{\beta}^{\Box}(\eta^{\Box}) .$$

The proof is by induction on η. If η is a p–variable, the claim is obviously correct. The inductive steps for $\neg$ and $\wedge$ are straightforward. So, we are left with the cases $\eta = \Box\theta$ and $\eta = \blacksquare\theta$.

Assume that $\eta = \Box\theta$. By induction hypothesis, $a_k \cap \overline{\beta}(\theta) = a_k \cap \overline{\beta}^{\Box}(\theta^{\Box})$. So,

$$\begin{array}{rcl} a_k \cap \Box a_k \cap \overline{\beta}(\eta) & = & a_k \cap \Box a_k \cap \Box\overline{\beta}(\theta) \\ & = & a_k \cap \Box a_k \cap \Box\overline{\beta}^{\Box}(\theta^{\Box}) \\ & = & a_k \cap \Box a_k \cap \overline{\beta}^{\Box}(\eta^{\Box}) \end{array}$$

(This makes use of the following. If $a \cap c = a \cap b$, then $\Box a \cap c \cap \Box c = \Box b \cap c \cap \Box c$.) Since $a_{k+1} \leq a_k \cap \Box a_k$ (by (a1)), the claim now holds for η. Next assume that $\eta = \blacksquare\theta$. We know by induction hypothesis that $a_k \cap \overline{\beta}(\theta) = a_k \cap \overline{\beta}^{\blacksquare}(\theta^{\blacksquare})$. Therefore

$$\blacksquare(a_k \cap \overline{\beta}(\theta)) = \blacksquare(a_k \cap \overline{\beta}^{\blacksquare}(\theta^{\blacksquare})) .$$

Using (a7) we conclude

$$a_k \cap \blacksquare\overline{\beta}(\theta) = a_k \cap \blacksquare\overline{\beta}^{\blacksquare}(\theta^{\blacksquare}) .$$

This gives $a_k \cap \overline{\beta}(\eta) = a_k \cap \overline{\beta}^{\blacksquare}(\theta^{\blacksquare})$. It follows that $a_k \cap \overline{\beta}(\theta) = a_k \cap \overline{\beta}^{\Box}(\theta^{\Box})$ since $a_k \cap \overline{\beta}^{\Box}(\eta^{\Box}) = a_k \cap \overline{\beta}^{\blacksquare}(\eta^{\blacksquare})$. Finally, $a_m \cap \overline{\beta}(\varphi) = a_m \cap \overline{\beta}^{\Box}(\varphi^{\Box}) > 0$, by (a3). Thus $\varphi \notin \Lambda \otimes \Theta$. This concludes the proof.

6.5. More Preservation Results

Theorem 6.5.1. *Let $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is compact iff both Λ and Θ are compact.*

Proof. Proceed as in the proof of Theorem 6.3.3. The only difference is that we work with sets of formulae rather than a single formula. Call the starting set Δ. We get a sequence of models $\langle \mathfrak{g}_k, \beta_k, w_0\rangle$ satisfying $[a]_k$ – $[f]_k$ for $\uparrow^k\Delta^{\Box}$. Now put

$$\begin{array}{rcl} g_\omega & := & \bigcup_{k\in\omega} g_k \\ \lhd_k & := & \bigcup_{k\in\omega} \lhd_k \\ \blacktriangleleft_k & := & \bigcup_{k\in\omega} \blacktriangleleft_k \end{array}$$

Furthermore, for a p–variable p put $\beta_\omega(p,x) := 1$ if $\beta_k(p,x) = 1$ for almost all $k \in \omega$, $\beta_\omega(p,x) := 0$ if $\beta_k(p,x) = 0$ for almost all k. It is checked that if $\beta_k(p,x) = 1$ then also $\beta_{k+1}(p,x) = 1$, and if $\beta_k(p,x) = 0$ then $\beta_{k+1}(p,x) = 0$. Finally, let γ be any standard valuation such that $\beta_\omega \leq \gamma$. It is not hard to show, using the properties [a] – [f], that $\langle \mathfrak{g}_\omega, \gamma, w_0 \rangle \vDash \Delta$. ◻

Theorem 6.5.2. *Suppose $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is globally complete (globally compact) iff both Λ and Θ are globally complete (globally compact).*

For weak compactness this method fails to yield a transfer theorem. Indeed, a counterexample can be contructed as follows. Take a monomodal logic Λ which is weakly compact but not compact, for example **Grz.3**. As is shown in [**66**], **Grz.3** is weakly compact. To show this is an exercise, see below. Earlier, in Section 3.2, we have shown that **Grz.3** is not $\aleph_1$–compact. Now we show that **Grz.3** $\otimes$ **K** is not even 1–compact. For there exists a set Δ which is **Grz.3**–consistent but lacks a Kripke–model. Then Δ is based on infinitely many variables, say $var[\Delta] = \{p_i : i \in \omega\}$; now let $\widetilde{\Delta}$ result from Δ by replacing the variable p_i by the formula $\blacksquare^{i+1}\bot \wedge \neg\blacksquare^{i}\bot$ for each $i \in \omega$. Then $var[\widetilde{\Delta}] = \varnothing$ and so $\widetilde{\Delta}$ is based on no variables. Clearly, $\widetilde{\Delta}$ is consistent; but if $\widetilde{\Delta}$ has a model based on a Kripke–frame then this allows a direct construction of a Kripke–model for Δ. Thus **Grz.3** $\otimes$ **K** is indeed not 1–compact. It is to be noted that this argument uses only the fact that **K** has infinitely many constant formulae. Any other weakly compact logic with these properties will do.

Now we will turn to interpolation and Halldén–completeness. The proof in both cases consists in a close analysis of the consistency formulae $\Sigma_\Box(\varphi \vee \psi)$ and $\Sigma_\Box(\varphi \to \psi)$. Since both are identical, it suffices to concentrate on the latter. We can write $\Sigma_\Box(\varphi) = \bigvee\langle\widetilde{\varphi}_c : c \in C\rangle$ and $\Sigma_\Box(\psi) = \bigvee\langle\widetilde{\psi}_d : d \in D\rangle$. Then obviously $\Sigma_\Box(\varphi \to \psi)$ is (up to boolean equivalence) a suitable disjunction of $\widetilde{\varphi}_c \wedge \widetilde{\psi}_d$; namely, this disjunction is taken over the set E of all pairs $\langle c, d\rangle$ such that $\widetilde{\varphi}_c \wedge \widetilde{\psi}_d$ is consistent. Equivalently, we can write

$$\Sigma_\Box(\varphi \to \psi) = \Sigma_\Box(\varphi) \wedge \Sigma_\Box(\psi) \wedge \bigwedge\langle\widetilde{\varphi}_c \to \neg\widetilde{\psi}_d : \langle c, d\rangle \notin E\rangle$$

We abbreviate the third conjunct by $\nabla(\varphi;\psi)$ (or, to be more precise we would again have to write $\nabla_\Box(\varphi;\psi)$). Obviously, $\nabla(\varphi;\psi)$ serves to cut out the unwanted disjuncts. In some sense $\nabla(\varphi;\varphi)$ measures the extent to which φ and ψ are interdependent. So if $\nabla(\varphi;\psi) = \top$ both are independent. It is vital to observe that all reformulations are classical equivalences.

Theorem 6.5.3. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ is Halldén–complete iff both Λ and Θ are.*

Proof. ($\Rightarrow$) Suppose $\varphi \vee \psi \in \Lambda$ and $var(\varphi) \cap var(\psi) = \varnothing$. Then $\varphi \vee \psi \in \Lambda \otimes \Theta$ and so either $\varphi \in \Lambda \otimes \Theta$ or $\psi \in \Lambda \otimes \Theta$ and thus either $\varphi \in \Lambda$ or $\psi \in \Lambda$, since $\Lambda \otimes \Theta$ is a conservative extension of Λ. ($\Leftarrow$) By induction on $n := adp(\varphi \vee \psi)$. For $n = 0$ this follows from Theorem 1.7.14. Now assume that $n > 0$ and that the theorem is proved for all formulae of alternation depth $< n$. Take $\varphi \vee \psi$ such that $var(\varphi) \cap var(\psi) = \varnothing$

and $adp(\varphi \vee \psi) = n$. Assume $adp^{\Box}(\Sigma_{\Box}(\varphi \vee \psi)) < adp^{\Box}(\varphi \vee \psi)$. (By the calculations following Theorem 6.3.5 we may assume that $adp^{\Box}(\Sigma_{\Box}(\varphi \vee \psi)) < adp^{\Box}(\varphi \vee \psi)$ or that $adp^{\blacksquare}(\Sigma_{\blacksquare}(\varphi \vee \psi)) < adp^{\blacksquare}(\varphi \vee \psi)$. We only show how to deal with the first case; the other case is dual.) Then by Theorem 6.4.8,

$$\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi \vee \psi)^{\Box} \to \varphi^{\Box} \vee \psi^{\Box}$$

for large m, by which

$$\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \wedge \Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \wedge \Box^{\leq m}\nabla(\varphi; \psi)^{\Box} \to \varphi^{\Box} \vee \psi^{\Box}$$

The crucial fact now is that $\nabla(\varphi; \psi) = \top$. For if $\widetilde{\varphi}_c$ and $\widetilde{\psi}_d$ are both $\Lambda \otimes \Theta$–consistent, then since $var(\widetilde{\varphi}_c) \cap var(\widetilde{\psi}_d) \subseteq var(\varphi) \cap var(\psi) = \varnothing$ and $adp(\widetilde{\varphi}_c), adp(\widetilde{\psi}_d) < n$, $\widetilde{\varphi}_c \wedge \widetilde{\psi}_d$ is $\Lambda \otimes \Theta$–consistent by induction hypothesis. Thus $\nabla(\varphi; \psi)$ is an empty conjunction. Consequently, we can rewrite the above to

$$\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \wedge \Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \to \varphi^{\Box} \vee \psi^{\Box}$$
$$\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box}. \vee .\Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \to \psi^{\Box}$$

Now since Λ is Halldén–complete, we have $\Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \to \varphi^{\Box} \in \Lambda$ or $\Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \to \psi^{\Box} \in \Lambda$ from which by Theorem 6.4.8 $\varphi \in \Lambda \otimes \Theta$ or $\psi \in \Lambda \otimes \Theta$. ◻

Theorem 6.5.4. *Suppose that $\bot \notin \Lambda, \Theta$. Then $\Lambda \otimes \Theta$ has interpolation iff both Λ and Θ have interpolation. Moreover, if $\varphi \to \psi \in \Lambda \otimes \Theta$ then an interpolant χ can be found such that $adp^{\Box}(\chi) \leq min\{adp^{\Box}(\varphi), adp^{\Box}(\psi)\}$ and $adp^{\blacksquare}(\chi) \leq min\{adp^{\blacksquare}(\varphi), adp^{\blacksquare}(\psi)\}$.*

Proof. ($\Rightarrow$) Let $\varphi \to \psi \in \Lambda$. Then by hypothesis there is a χ such that $\varphi \to \chi, \chi \to \psi \in \Lambda \otimes \Theta$ based on the common variables of φ and ψ. Now, by Makinson's Theorem, either $\Theta(p \leftrightarrow \blacksquare p)$ or $\Theta(\blacksquare p)$ is consistent. Let the former be the case. Then let χ° result from χ by successively replacing a subformula $\blacksquare\psi$ by ψ. Then $\chi^{\circ} \in \mathcal{L}_{\Box}$ and $\chi \leftrightarrow \chi^{\circ} \in \Theta(p \leftrightarrow \blacksquare p)$. Hence, as $\varphi \to \chi \in \Lambda \otimes \Theta$, then also $\varphi \to \chi^{\circ} \in \Lambda \otimes \Theta(p \leftrightarrow \blacksquare p)$. But $\Lambda \otimes \Theta(p \leftrightarrow \blacksquare p)$ is a conservative extension of Λ and therefore $\varphi \to \chi^{\circ} \in \Lambda$. In the case where $\Theta(\blacksquare p)$ is consistent, define χ° to be the result of replacing subformulas of type $\blacksquare\psi$ by $\top$. Then use the same argument as before.

($\Leftarrow$) By induction on $n := adp(\varphi \to \psi)$. The case $n = 0$ is covered by Theorem 1.7.14. Now suppose that $n > 0$ and that the theorem has been proved for all formulae of alternation depth $< n$. Let $\varphi \to \psi \in \Lambda \otimes \Theta$. We may assume that $adp^{\Box}(\varphi \to \psi) \leq adp^{\blacksquare}(\varphi \to \psi)$ and thus

$$adp^{\blacksquare}(\Sigma_{\Box}(\varphi \to \psi)) < adp^{\blacksquare}(\varphi \to \psi)$$

(see the calculations following Theorem 6.3.5). Then $adp(\Sigma_{\Box}(\varphi \to \psi)) < adp(\varphi \to \psi)$. By Theorem 6.4.8, for sufficiently large m,

$$\vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi \to \psi) \to \varphi^{\Box} \to \psi^{\Box}$$
$$(\dagger) \quad \vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \wedge \varphi^{\Box} \wedge \Box^{\leq m}\nabla(\varphi; \psi)^{\Box} \to \Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \to \psi^{\Box}$$

Let $\widetilde{\varphi}_c \to \neg\widetilde{\psi}_d$ be a conjunct of $\nabla(\varphi;\psi)$. By induction hypothesis and the fact that $adp(\widetilde{\varphi}_c), adp(\widetilde{\psi}_d) < adp(\varphi \to \psi)$ (since we have $adp^{\blacksquare}(\widetilde{\varphi}_c), adp^{\blacksquare}(\neg\widetilde{\psi}_d) < adp^{\blacksquare}(\varphi \to \psi)$ and $adp^{\Box}(\widetilde{\varphi}_c), adp^{\Box}(\neg\widetilde{\psi}_d) \leq adp^{\Box}(\varphi \to \psi)$) there is an interpolant $Q_{c,d}$ for $\widetilde{\varphi}_c$ and $\widetilde{\psi}_d$. Note that $var(Q_{c,d}) = var^p(Q_{c,d}) \subseteq var(\varphi) \cap var(\psi)$ and that

$$adp^{\Box}(Q_{c,d}) \leq min\{adp^{\Box}(\widetilde{\varphi}_c), adp^{\Box}(\widetilde{\psi}_d)\} \leq min\{adp^{\Box}(\varphi), adp^{\Box}(\psi)\} \qquad .$$

Likewise for $\blacksquare$. Again by Theorem 6.4.8 we get

$$\begin{aligned} \vdash_{\Box} \Box^{\leq m}\Sigma_{\Box}(\widetilde{\varphi}_c \to Q_{c,d})^{\Box} \wedge \Box^{\leq m}\Sigma_{\Box}(Q_{c,d} \to \neg\widetilde{\psi}_d)^{\Box}. \\ \to .(\widetilde{\varphi}_c \to Q_{c,d})^{\Box} \wedge (Q_{c,d} \to \neg\widetilde{\varphi}_d)^{\Box} \end{aligned}$$

and therefore with $F := C \times D - E$ (recall the definition of ∇)

$$\begin{aligned} & \textstyle\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(\widetilde{\varphi}_c \to Q_{c,d})^{\Box} \wedge \bigwedge_F \Box^{\leq m}\Sigma_{\Box}(Q_{c,d} \to \neg\widetilde{\psi}_d)^{\Box} \\ & \quad \vdash_{\Box} \textstyle\bigwedge_F(\widetilde{\varphi}_c \to Q_{c,d})^{\Box} \wedge \bigwedge_F(Q_{c,d} \to \neg\widetilde{\psi}_d)^{\Box} \\ & \quad \vdash_{\Box} \nabla(\varphi;\psi)^{\Box} \end{aligned}$$

Thus (†) can be rewritten modulo boolean equivalence into

$$\begin{aligned} & \Box^{\leq m}\Sigma_{\Box}(\varphi)^{\Box} \wedge \varphi^{\Box} \wedge \textstyle\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(\widetilde{\varphi}_c \to Q_{c,d})^{\Box} \\ & \quad \vdash_{\Box} \textstyle\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(Q_{c,d} \to \neg\widetilde{\psi}_d)^{\Box} \wedge \Box^{\leq m}\Sigma_{\Box}(\psi)^{\Box} \to \psi^{\Box} \end{aligned}$$

Abbreviate the formula to the left by η_ℓ and the one to the right by η_r. Then

$$\begin{aligned} adp^{\Box}(\uparrow \eta_\ell) = \quad & max\{adp^{\Box}(\Box^{\leq m}\Sigma_{\Box}(\varphi)), adp^{\Box}(\varphi), \\ & \qquad adp^{\Box}(\textstyle\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(\widetilde{\varphi}_c \to Q_{c,d}))\} \\ = \quad & adp^{\Box}(\varphi)\ , \end{aligned}$$

since we have that $adp^{\Box}(\Box^{\leq m}\Sigma_{\Box}(\varphi)) \leq adp^{\Box}(\varphi)$ by an earlier observation and

$$adp^{\Box}(\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(\widetilde{\varphi}_c \to Q_{c,d})) \leq adp^{\Box}(\Box^{\leq m}\Sigma_{\Box}(\varphi)).$$

By a similar argument

$$\begin{aligned} adp^{\Box}(\uparrow \eta_r) = \quad & max\{adp^{\Box}(\Box^{\leq m}\Sigma_{\Box}(\psi)), adp^{\Box}(\psi), \\ & \qquad adp^{\Box}(\textstyle\bigwedge_F \Box^{\leq m}\Sigma_{\Box}(Q_{c,d} \to \neg\widetilde{\psi}_d))\} \\ = \quad & adp^{\Box}(\psi)\ . \end{aligned}$$

Likewise we argue with $adp^{\blacksquare}$. By assumption on Λ, there is an interpolant χ for η_ℓ and η_r. By definition, χ is based on the same surrogate variables as η_ℓ and η_r. Therefore for the total reconstruction $\chi^{\uparrow}$ of χ

$$adp^{\Box}(\chi^{\uparrow}) \leq min\{adp^{\Box}(\uparrow \eta_\ell), adp^{\Box}(\eta_r^{\uparrow})\} = min\{adp^{\Box}(\varphi), adp^{\Box}(\psi)\}$$

and similarly for $adp^{\blacksquare}$. It is easily verified that $var^p(\chi^{\uparrow}) \subseteq var^p(\varphi) \cap var^p(\psi)$. Moreover, from $\eta_\ell = \eta_\ell^{\uparrow\Box} \vdash_{\Box} \uparrow \chi^{\Box}$ with Theorem 6.4.8 and the fact that the consistency formulae are $\Lambda\otimes\Theta$–theorems we conclude that $\varphi \vdash_{\Box\blacksquare} \chi^{\uparrow}$ and likewise that $\chi^{\uparrow} \vdash_{\Box\blacksquare} \psi$. $\Box$

Theorem 6.5.4 implies a stronger interpolation property for $\Lambda \otimes \Theta$. Namely, if $\varphi \to \psi \in \Lambda \otimes \Theta$ then an interpolant exists which is not only based on the common variables but also contains only the modalities which occur in both φ and ψ.

Exercise 209. The logic **Here** defined by the axiom $p \to \Box p$ is Sahlqvist. It corresponds to $(\forall y \rhd x)(x = y) \vee (\forall y \rhd x)\mathsf{f}$. Define for a modal logic Λ the logic $\Lambda(\mathsf{c})$, which is the result of adding a new propositional constant c to the language. Now define a translation from $\mathcal{P}_\kappa(\mathsf{c})$ to $\mathcal{P}_{\kappa+1}$ by translating c by $\Diamond_\kappa \top$. (Further, the translation commutes with $\neg$, $\wedge$ and $\Diamond_\lambda$, $\lambda < \kappa$.) Show that the translation induces an isomorphism from $\mathcal{E}\,\Lambda(\mathsf{c})$ onto $\mathcal{E}\,\Lambda \otimes$ **Here**.

Exercise 210. (Fine [**63**] and [**66**].) Show that **Grz.3** is weakly compact. *Hint.* First of all, the canonical frame based on n generators is linear. Now show that there is no infinite properly ascending chain of points by showing that the underlying Kripke–frame is isomorphic to $\langle \alpha, \geq \rangle$, where α is an ordinal number. To do that, let $\mathfrak{F}$ be a **Grz.3**–frame such that $\mathfrak{F}_+$ is generated by the elements a_i, $i < n$. Define the **character** of a point x to be the set of all subsets C of n such that

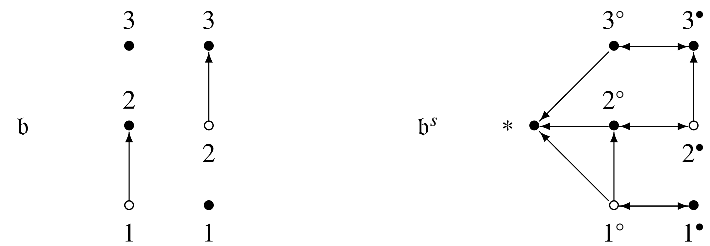
$$x \in \bigcap_{i \in C} a_i \cap \bigcap_{i \notin C} -a_i$$

A point is of **minimal character** in $M \subseteq \alpha$ iff for all $y \rhd x$, and $y \in M$, y has the same character as x. Show that if x is of minimal character in M and $x \lhd y$ for $y \in M$ then $x = y$. Let x^0 be the point of minimal character in α. In general, let x^β be the point of minimal character in the set $M^\beta := \{y : (\forall \gamma < \beta)(y \lhd x^\gamma \text{ and } y \neq x^\gamma)\}$. Whenever M^β is not empty, there exists a point of minimal character. Now define as the new frame $\mathfrak{G}$ the subframe induced on the set g of all x^β. Show that $\mathfrak{G}_+$ is isomorphic to $\mathfrak{F}_+$.

6.6. Thomason Simulations

Now that we have seen how to embed modal logics into polymodal logics, we will turn to the question of simulating polymodal logics by monomodal logics. The results can be found also in [**133**], though sometimes with different proofs. Again, it is useful to restrict the discussion to the case of bimodal logics. It is a priori not clear that we can use a single operator to simulate two operators but it will turn out that the situation is as good as possible. Not only can we perform such a simulation, we can also map the whole lattice of extensions of $\mathbf{K}_2$ isomorphically onto the interval $[\mathbf{Sim}, \mathsf{Th}\boxed{\bullet}]$, where **Sim** is some finitely axiomatizable (monomodal) logic. This will have numerous consequences for the theory of modal logic, as we will see. The construction works only for normal logics, and in the exercises we give some indication as to why it fails for quasi–normal logics. The basic idea of simulations is due to S. K. Thomason, who used it in [**209**], [**208**] and [**210**] as a tool to derive certain negative facts about monomodal logics. We explain this simulation first by using frames and then go over to algebras.

Figure 6.1. A Simulation



Take a bimodal frame $\mathfrak{B} = \langle b, \lhd, \blacktriangleleft, \mathbb{B}\rangle$. The *simulation* of $\mathfrak{B}$ is a frame $\mathfrak{B}^s = \langle b^s, \leqslant, \mathbb{B}^s\rangle$ for the logic with the operator $\boxminus$. It is defined as follows. We let $\nabla := \{\circ, \bullet, *\}$. Put $b^s := \{*\} \cup b \times \{\circ, \bullet\}$. We will denote $\langle x, \circ\rangle$ by x° and $\langle x, \bullet\rangle$ by $x^\bullet$. Moreover, if $x \in b$ then x^* will be another name for $*$. The notation is extended to sets of points. For a set $A \subseteq b$, $A^\circ := A \times \{\circ\}$, $A^\bullet := A \times \{\bullet\}$ and $A^* := \{*\}$ if A is nonempty, and $\varnothing^* := \varnothing$. We use the symbols $\flat$ and $\natural$ as variables ranging over $\nabla(= \{\circ, \bullet, *\})$. The relation $\leqslant$ is defined as follows.

$$x^\flat \leqslant y^\natural \qquad \text{if} \qquad \begin{cases} & \flat = \circ, \natural = *, \\ \textit{or} & \flat = \circ, \natural = \bullet, \quad x = y \\ \textit{or} & \flat = \bullet, \natural = \circ, \quad x = y \\ \textit{or} & \flat = \circ, \natural = \circ, \quad x \lhd y \\ \textit{or} & \flat = \bullet, \natural = \bullet, \quad x \blacktriangleleft y \end{cases}$$

$$\mathbb{B}^s := \{c^* \cup d^\circ \cup e^\bullet : c, d, e \in \mathbb{B}\}$$

This defines $\mathfrak{B}^s$. Proposition 6.6.1 asserts that this is indeed a frame. Given a Kripke–frame $\mathfrak{b}$, $\mathfrak{b}^s := \langle b^s, \leqslant\rangle$. For example, let $\mathfrak{b}$ consist of three points, 1, 2 and 3. Let $1 \lhd 1$, $1 \lhd 2$ and $2 \blacktriangleleft 2$ and $2 \blacktriangleleft 3$. Then $\mathfrak{b}^s$ is shown in Figure 6.1. If $\mathfrak{B}$ is the empty frame then $\mathfrak{B}^s = \boxed{\bullet}$. This is a useful fact. It accounts for the fact that the simulation of the inconsistent bimodal logic is not the inconsistent monomodal logic, but rather the logic of $\boxed{\bullet}$.

Proposition 6.6.1. *Let $\mathfrak{B}$ be a bimodal frame, and let $\mathfrak{B}^s$ be defined as above. Then $\mathfrak{B}^s$ is a frame.*

Proof. Closure under complement and negation is clear. Now we show closure under $\Diamond\!\!\!\cdot\;$. We can reduce this to a discussion of three cases, namely $c = d^*$, $c = d^\circ$ and $c = d^\bullet$. Let $c = d^*$. If $d = \varnothing$, $\Diamond\!\!\!\cdot\; d^* = \varnothing = d^*$. If $d \neq \varnothing$, then $\Diamond\!\!\!\cdot\; d^* = b^\circ$. Next, let $c = d^\circ$. Then $\Diamond\!\!\!\cdot\; c = d^\bullet \cup (\Diamond d)^\circ$. Finally, let $c = d^\bullet$. Then $\Diamond\!\!\!\cdot\; c = d^\circ \cup (\blacklozenge d)^\bullet$. Since both $\Diamond d$ and $\blacklozenge d$ are in $\mathbb{B}$, we are done. ∎

Let us note that the sets $\{*\}$, b° and $b^\bullet$ are definable by *constant* formulae. This will be of extreme importance. Namely, put

$$\begin{aligned} \omega &:= \boxminus\bot \\ \alpha &:= \Diamond \boxminus \bot \\ \beta &:= \neg\Diamond \boxminus \bot \wedge \neg \boxminus \bot \end{aligned}$$

We will usually also denote by α, β and ω the sets of points defined by α, β and ω in a given frame. It is easy to verify that $\omega = \{*\}$, $\alpha = b^\circ$ and $\beta = b^\bullet$. We can now conclude that if we have $\mathbb{B}$ then $\mathbb{B}^s$ is generated by the sets $c^\circ \cup c^\bullet$ using the operations $-$, $\cap$ and $\boxminus$. Now let $\mathfrak{B}$ and $\mathfrak{C}$ be bimodal frames and $\pi : \mathfrak{B} \to \mathfrak{C}$ a p–morphism. Put

$$\pi^s(x^\flat) := \pi(x)^\flat$$

We claim that $\pi^s : \mathfrak{B}^s \to \mathfrak{C}^s$. So, let us check the first condition on p–morphisms. Assume $x^\flat \leqslant y^\natural$. (1.) $\flat = \circ$, $\natural = *$. Then $\pi^s(x^\flat) = \pi(x)^\circ \leqslant \pi(y)^* = \pi^s(y^\natural)$. (2.) $\flat = \circ$, $\natural = \bullet$ and $x = y$. Then $\pi^s(x^\flat) = \pi(x)^\circ \leqslant \pi(x)^\bullet = \pi^s(y^\natural)$. (3.) $\flat = \bullet$, $\natural = \circ$ and $x = y$. As in (2.). (4.) $\flat = \natural = \circ$. Then $x \lhd y$ and so $\pi(x) \lhd \pi(y)$ by assumption on π. Therefore, $\pi^s(x^\flat) = \pi(x)^\circ \leqslant \pi(y)^\circ = \pi^s(y^\natural)$. (5.) $\flat = \natural = \bullet$. Analogous to (4.). This exhausts all cases. Now let us check the second condition. Assume that $\pi(x)^\flat \leqslant u^\natural$ for some $u \in c$, c the underlying set of $\mathfrak{C}$. By definition of π^s, $\pi^s(x^\flat) \leqslant u^\natural$. Hence $\flat \neq *$. If $\natural = *$, then $\flat = \circ$ and we may take $u := x$. Then $u^\natural = *$, and $\pi^s(x^\flat) = \pi(x)^\circ \leqslant \pi(x)^* = u^\natural$. So assume from now on that $\natural \neq *$ and $\flat \neq *$. Assume first that $\flat = \circ$. If also $\natural = \circ$ then we must have $\pi(x) \lhd u$, by construction of $\mathfrak{C}^s$. Since π is a p–morphism there is a y such that $\pi(y) = u$ and $x \lhd y$. Then $x^\flat \leqslant y^\natural$ and $\pi^s(y^\natural) = \pi(y)^\natural = u^\natural$, as required. If $\natural = \bullet$ then $u = \pi(x)$. In that case, we have $x^\flat \leqslant x^\natural$ and $\pi^s(x^\natural) = \pi(x)^\natural = u^\natural$, again as desired. Assume next that $\flat = \bullet$. The argumentation is parallel to the first case. Finally, we must show that for every internal set d of $\mathfrak{C}^s$ the set $(\pi^s)^{-1}[d]$ is internal in $\mathfrak{B}^s$. So let $d = d_w^\circ \cup d_b^\bullet \cup d_o^*$ for some sets $d_w, d_b, d_o \in \mathbb{C}$. Then

$$(\pi^s)^{-1}[d] = \pi^{-1}[d_w]^\circ \cup \pi^{-1}[d_b]^\bullet \cup \pi^{-1}[d_o]^*$$

This is an internal set of $\mathfrak{B}^s$.

Theorem 6.6.2. *The map $(-)^s$ is a functor from the category of bimodal frames into the category of monomodal frames. Moreover, for a bimodal frame $\mathfrak{B}$, $(\mathfrak{B}_\sharp)^s = (\mathfrak{B}^s)_\sharp$ and for a bimodal Kripke–frame $\mathfrak{b}$, $(\mathfrak{b}^\sharp)^s = (\mathfrak{b}^s)^\sharp$.*

Proposition 6.6.3. *Let $\mathfrak{B}$ be a bimodal frame.*

(1) *$\mathfrak{B}^s$ is differentiated iff $\mathfrak{B}$ is.*
(2) *$\mathfrak{B}^s$ is refined iff $\mathfrak{B}$ is.*
(3) *$\mathfrak{B}^s$ is compact iff $\mathfrak{B}$ is.*

Proof. (1.) If $\mathfrak{B}^s$ is differentiated, then surely $\mathfrak{B}$ is differentiated. Now assume that $\mathfrak{B}$ is differentiated. Since b°, $b^\bullet$ and $\{*\}$ are internal sets, $x^\flat$ and $y^\natural$ can be discriminated by an internal set if $\flat \neq \natural$. So, assume that $\flat = \natural$. In case $\flat = *$

we have $x = y$ by construction. So, finally, $\flat = \natural \in \{\circ, \bullet\}$. In that case $x \neq y$ and there exists a set $c \in \mathbb{B}$ such that $x \in c$ but $y \notin c$. Then $x^\flat \in c^\flat$ but $y^\natural \notin c^\flat$.
(2.) Suppose that $\mathfrak{B}^s$ is tight. Then $\mathfrak{B}$ is tight as well. For let $x \not\lhd y$. Then $x^\circ \not\lessdot y^\circ$ and so there is a set a such that $x^\circ \in \boxminus a$ but $y^\circ \notin a$. Let $a = a_1^\circ \cup a_2^\bullet \cup a_3^*$ for some internal sets a_1, a_2, a_3 of $\mathfrak{B}$. Then $x \in \Box a_1$ but $y \notin a_1$. Now let $x \blacktriangleleft y$ not be the case. Then $x^\bullet \not\lessdot y^\bullet$. So there is a set a such that $x^\bullet \in \boxminus a$ but $y^\bullet \notin a$. Let $a = a_1^\bullet \cup a_2^\circ \cup a_3^*$ for some internal sets a_1, a_2 and a_3 of $\mathfrak{B}$. Then $y \notin a_1$, but

$$x^\circ \in \alpha \cap \boxminus(\beta \to \boxminus(\beta \to \boxminus(\alpha \to a_1^\circ))) \ .$$

From this follows $x \in \blacksquare a_1$, as desired. Now it follows by (1.) that if $\mathfrak{B}^s$ is refined, $\mathfrak{B}$ is refined as well. Now assume that $\mathfrak{B}$ is refined. Then $\mathfrak{B}^s$ is differentiated, by (1.). We have to show that it is tight. Let $x \not\lessdot y$. We have to find a such that $x \in \boxminus a$ but $y \notin a$. **Case 1.** $x = *$ or $y = *$. If $x = *$ put $a := \varnothing$. If $y = *$ then $x \in \beta$. So put $a := -\beta$. **Case 2.** $x = u^\circ$ and $y = v^\bullet$ (or $x = u^\bullet$ and $y = v^\circ$). Then $u \neq v$ and so there is a set c such that $u \in c$ but $v \notin c$. Then $x \in \boxminus(-\beta \cup c^\bullet)$, since every successor which does not satisfy β must be of the form $u^\bullet \in c^\bullet$. Moreover, $y \notin -\beta \cup c^\bullet$. So, $a := -\beta \cup c^\bullet$. **Case 3.** $x^\circ \not\lessdot y^\circ$ or $x^\bullet \not\lessdot y^\bullet$. Let the first be the case. (The other case is dual.) Here, we can use the assumption of tightness for $\mathfrak{B}$ to get a set c such that $x \in \Box c$ but $y \notin c$. Then $x^\circ \in \boxminus(c^\circ \cup -\alpha)$ but $y^\circ \notin -\alpha \cup c^\circ$. Put $a := -\alpha \cup c^\circ$.
(3.) Assume $\mathfrak{B}^s$ is compact, and let U be an ultrafilter on $\mathfrak{B}$. Put $U^\alpha := \{a^\circ : a \in U\}$. U^α is an ultrafilter on α in $\mathfrak{B}^s$. U^α can be expanded to an ultrafilter U^+ on $\mathfrak{B}^s$, taking all sets c such that $c \cap \alpha \in U^\alpha$. We have $\alpha \in U^+$. U^+ has nonempty intersection by assumption on $\mathfrak{B}^s$. Let $x \in \bigcap U^+$. Then $x \in \alpha$, so $x = u^\circ$ for some u. Thus $u \in \bigcap U$, showing the compactness of $\mathfrak{B}$. Now assume that $\mathfrak{B}$ is compact and let $U \subseteq \mathbb{B}^s$ be an ultrafilter on $\mathfrak{B}^s$. We have $\omega \cup \alpha \cup \beta \in U$. Thus three cases arise. **Case 1.** $\omega \in U$. Then $\{*\} \in U$ and so $* \in \bigcap U$. **Case 2.** $\alpha \in U$. Then the trace $U_\alpha = \{a \cap \alpha : a \in U\}$ is an ultrafilter on α. Hence it corresponds to the ultrafilter $V := \{v : v^\circ \in U_\alpha\}$ on $\mathfrak{B}$. By assumption, there is a y such that $y \in \bigcap V$. Then $y^\circ \in \bigcap U_\alpha = \bigcap U$. This completes the second case. **Case 3.** $\beta \in U$. Then the trace $U_\beta := \{a \cap \beta : a \in U\}$ is an ultrafilter on β. But then $U' = \{\alpha \cap \Diamond_\beta a : a \in U_\beta\}$ is an ultrafilter on α, since we replace sets of the form $a^\bullet$ by sets of the form a°. We then get a $y^\circ \in \bigcap U'$ and so $y^\bullet \in \bigcap U_\beta = \bigcap U$, as required. ∎

The construction can be defined on algebras as well. Namely, assume that $\mathfrak{B} = \langle B, 1, -, \cap, \Box, \blacksquare\rangle$ is a bimodal algebra. Then let $B^s := B \times B \times 2$ and

$$\boxminus\langle a, b, c\rangle := \begin{cases} \langle b \cap \Box a, a \cap \blacksquare b, 1\rangle & \text{if } c = 1 \\ \langle 0, a \cap \blacksquare b, 1\rangle & \text{if } c = 0 \end{cases}$$

Then $\mathfrak{B}^s := \langle B^s, 1, -, \cap, \boxminus\rangle$ is a monomodal algebra, where $1 := \langle 1, 1, 1\rangle$. We use α, β and ω for the value of α, β and ω, respectively, in $\mathfrak{A}$. Given a homomorphism $h : \mathfrak{B} \to \mathfrak{C}$ of bimodal algebras, the map $h \times h \times id(2)$ can be shown to be a homomorphism of monomodal algebras.

Theorem 6.6.4. *$(-)^s$ is a functor from the category of bimodal algebras to the category of monomodal algebras.*

Theorem 6.6.5. *Let $\mathfrak{B}$ be a bimodal algebra. Then $(\mathfrak{B}^+)^s \cong (\mathfrak{B}^s)^+$. Let $\mathfrak{C}$ be a bimodal frame. Then $(\mathfrak{C}_+)^s \cong (\mathfrak{C}^s)_+$.*

Proof. The second claim is rather straightforward. Let $\mathfrak{C}$ be given. Then the maps $d \mapsto d \cap \alpha$, $d \mapsto d \cap \beta$ and $d \mapsto d \cap \omega$ are projections of the boolean algebra of internal sets of $\mathfrak{C}^s$ onto the boolean algebra of internal sets of $\mathfrak{C}$, $\mathfrak{C}$ and onto 2, respectively. They induce an isomorphism $\mathbb{C}^s \cong \mathbb{C} \times \mathbb{C} \times 2$. The operation $\boxminus$ on $\mathbb{C}^s$ is checked to be exactly as in the definition of $(\mathfrak{C}_+)^s$. This shows the second claim. Now to the first claim. We define a bijection ι from the set of points of $(\mathfrak{B}^s)^+$ onto the set of points of $(\mathfrak{B}^+)^s$. Take an ultrafilter U of $\mathfrak{B}^s$. This is an ultrafilter in the boolean algebra $\mathbf{B} \times \mathbf{B} \times \mathbf{2}$, where $\mathbf{B} = \langle B, 1, -, \cap\rangle$. Now, $\langle 1, 1, 1\rangle \in U$, and $\langle 1, 1, 1\rangle = \langle 1, 0, 0\rangle \cup \langle 0, 1, 0\rangle \cup \langle 0, 0, 1\rangle$ and so either $\langle 1, 0, 0\rangle \in U$, or $\langle 0, 1, 0\rangle \in U$ or $\langle 0, 0, 1\rangle \in U$. In the first case, there exists an ultrafilter V on $\mathbf{B}$ such that $U = \{\langle a, b, c\rangle : a \in V, b \in B, c \in 2\}$. We then put $\iota(U) := V^\circ$. In the second case there exists an ultrafilter V on $\mathbf{B}$ such that $U = \{\langle a, b, c\rangle : a \in B, b \in V, c \in 2\}$. Then we put $\iota(U) := V^\bullet$. In the third case $U = \{\langle a, b, 1\rangle : a, b \in B\}$. Then we put $\iota(U) = *$ or $\iota(U) := V^*$, where V is any ultrafilter of $\mathbf{B}$ (by our naming convention this is the same). By careful checking of cases it is shown that ι is an isomorphism. ◻

Given a class $\mathcal{K}$ of frames we write $\mathcal{K}^s$ for the class $\{\mathfrak{B}^s : \mathfrak{B} \in \mathcal{K}\}$ and similarly for classes of algebras.

Definition 6.6.6. *Let Θ be a bimodal normal logic. Then we define $\Theta^s := \mathsf{Th}\,(\mathsf{Frm}\,\Theta)^s$. This is called the* **(monomodal) simulation** *of Θ.*

Our aim is to show that the map $\Theta \mapsto \Theta^s$ actually is an isomorphism from the lattice $\mathcal{E}\,\mathbf{K}_2$ onto an interval in the lattice $\mathcal{E}\,\mathbf{K}_1$. To show this we first show how to axiomatize Θ^s. Before we can give such an axiomatization, we state a simple lemma.

Lemma 6.6.7. *Let $\mathfrak{M}$ be a monomodal frame. $\mathfrak{M} \cong \mathfrak{B}^s$ for some bimodal frame iff*

(A) *every point satisfies either ω, α or β,*
(B) *there exists exactly one point in ω,*
(C) *every point in α sees one point in ω,*
(D) *no point of β sees a point in ω,*
(E) *every point in α has exactly one successor in β,*
(F) *every point in β has exactly one successor in α,*
(G) *for every point x in α and every point y in β, $x \leqslant y$ iff $y \leqslant x$.*

It is clear that the properties (A), (C) – (G) are modally characterizable and also df–persistent. The problem lies in condition (B). However, notice the following.

Lemma 6.6.8. *Let $\mathfrak{M}$ be a rooted monomodal frame. Let $\mathfrak{M}$ satisfy the properties (A), (C) – (G) of Lemma 6.6.7. Then $\mathfrak{M}$ satisfies (B) iff it satisfies (H).*

(H) for all x and all $y, y' \in \omega$ such that $x \leqslant^{\leq 3} y; y'$ we have $y = y'$.

Proof. It is clear that (H) rules out that a point has more than one (direct) ω–successor. On the other hand, each point sees an ω–point in at most two steps. If that point is always the same for each point reachable from the root, we have only one ω–point. So, if ω contains more than one point, there must be a point $x \notin \omega$ seeing in at most two steps an ω–point u, and a $\leqslant$–successor y of x such that y sees in at most 2 steps an ω–point u' different from u. Then x sees in at most three steps two different ω–points. This is ruled out by (H), however. ◻

Definition 6.6.9. *The logic $\boldsymbol{Sim}$ is the extension of $\boldsymbol{K}_1$ by the axioms (a) – (h).*

$$\begin{array}{ll}
(a) & \omega \vee \alpha \vee \beta \\
(b) & \alpha. \to .\Diamond(\omega \wedge p) \to \boxminus(\omega \to p) \\
(c) & \alpha. \to .\Diamond\beta \wedge \Diamond\omega \\
(d) & \alpha. \to .\Diamond(\beta \wedge p) \to \boxminus(\beta \to p) \\
(e) & \beta. \to .\Diamond\alpha \wedge \neg\Diamond\omega \\
(f) & \beta. \to .\Diamond(\alpha \wedge p) \to \boxminus(\alpha \to p) \\
(g) & \alpha. \to .p \to \boxminus(\beta \to \boxminus(\alpha \to p)) \\
(h) & \Diamond^{\leq 3}(\omega \wedge p) \to \boxminus^{\leq 3}(\omega \to p)
\end{array}$$

Some of the axioms are satisfied by choice of α, β and ω, but this is unimportant. The following is easy to verify.

Proposition 6.6.10. *$\boldsymbol{Sim}$ is df–persistent. It is Sahlqvist and of special rank 0.*

(For a proof, notice that the axioms state properties that assert (i.) the existence of successors or (ii.) the uniqueness of successors. These are of special rank 0.) Before we prove that **Sim** is correctly defined let us introduce some notation. Put $\Diamond_\alpha\chi := \Diamond(\alpha \wedge \chi)$, $\boxminus_\alpha\chi := \boxminus(\alpha \to \chi)$ and likewise for β and ω. Then

$$\begin{array}{lllll}
\Diamond\chi & = & \bigvee_{\flat\in\nabla} \Diamond_\flat & = & \Diamond_\alpha\chi \vee \Diamond_\beta\chi \vee \Diamond_\omega\chi \\
\boxminus\chi & = & \bigwedge_{\flat\in\nabla} \boxminus_\flat\chi & = & \boxminus_\alpha\chi \wedge \boxminus_\beta\chi \wedge \boxminus_\omega\chi
\end{array}$$

Clearly, $\Diamond_\flat\chi$ is satisfied at a point x in a frame if there exists a successor y which is in $\flat$ and satisfies χ. $\boxminus_\flat\chi$ is satisfied at a point x if all successors y of x which are in $\flat$ satisfy χ. This allows to rewrite some of the postulates.

$$\begin{array}{ll}
(b') & \alpha. \to .\Diamond_\omega p \to \boxminus_\omega p \\
(d') & \alpha. \to .\Diamond_\beta p \to \boxminus_\beta p \\
(f') & \beta. \to .\Diamond_\alpha p \to \boxminus_\beta p \\
(g') & \alpha. \to .p \to \boxminus_\beta \boxminus_\alpha p
\end{array}$$

Moreover, we have the following theorems, which follow from Lemma 6.6.7 and 6.6.8, respectively.

Corollary 6.6.11. *Let* $\mathfrak{M}$ *be a differentiated monomodal frame whose theory contains (a) to (g). Let* ω *have a single point as its extension in* $\mathfrak{M}$*. Then for some bimodal frame* $\mathfrak{B}$ *we have* $\mathfrak{M} \cong \mathfrak{B}^s$.

Corollary 6.6.12. *Let* $\mathfrak{M}$ *be a rooted, differentiated frame for* $\boldsymbol{Sim}$*. Then* $\mathfrak{M} \cong \mathfrak{B}^s$ *for some (possibly empty) differentiated bimodal frame* $\mathfrak{B}$.

The previous theorem shows that **Sim** axiomatizes the right kind of frames, assuming that we deal with differentiated frames. To obtain an axiomatization of Θ^s on the basis of an axiomatization for Θ we must also find a syntactic correlate of the frame simulation. The simulation of a bimodal formula φ is defined as follows.

$$\begin{array}{lcl} p^s & := & p \\ (\neg\varphi)^s & := & \alpha \wedge \neg(\varphi^s) \\ (\varphi \wedge \psi)^s & := & \varphi^s \wedge \psi^s \\ (\Diamond\varphi)^s & := & \alpha \wedge \Diamond\varphi^s \\ (\blacklozenge\varphi)^s & := & \alpha \wedge \Diamond(\beta \wedge \Diamond(\beta \wedge \Diamond(\alpha \wedge \varphi^s))) \end{array}$$

Proposition 6.6.13. *If* φ *is a Sahlqvist formula, so is* φ^s.

Proposition 6.6.14. *Let* $\mathfrak{B}$ *be a bimodal frame and* γ *a valuation on* $\mathfrak{B}^s$ *such that* $\gamma(p) \cap b^\circ = (\beta(p))^\circ$*. Then*

$$\langle \mathfrak{B}, \beta, x \rangle \vDash \varphi \quad \textit{iff} \quad \langle \mathfrak{B}^s, \gamma, x^\circ \rangle \vDash \varphi^s$$

Proof. By induction on φ. ☐

Proposition 6.6.15. *Let* $\langle \mathfrak{B}, \beta, x \rangle$ *be a bimodal model. Put* $\beta^s(p) := \beta(p)^\circ$*. Then the following holds.*

$$\begin{array}{rcr} \langle \mathfrak{B}, \beta, x \rangle \vDash \varphi & \Leftrightarrow & \langle \mathfrak{B}^s, \beta^s, x^\circ \rangle \vDash \varphi^s \\ \langle \mathfrak{B}, \beta \rangle \vDash \varphi & \Leftrightarrow & \langle \mathfrak{B}^s, \beta^s \rangle \vDash \alpha \to \varphi^s \\ \mathfrak{B} \vDash \varphi & \Leftrightarrow & \mathfrak{B}^s \vDash \alpha \to \varphi^s \end{array}$$

Proof. The first is an immediate consequence of the previous lemma. The second also follows, since $x^\flat$ satisfies α in $\mathfrak{B}^s$ iff $\flat = \circ$. The third is proved thus. From right to left is a consequence of the previous line. From left to right is not entirely obvious. Pick a valuation γ on $\mathfrak{B}^s$. Then let $\beta(p) := \{x : x^\circ \in \gamma(p)\}$. Then $\beta(p) \cap b^\circ = (\gamma(p))^\circ$ and so by the previous lemma $\langle \mathfrak{B}, \beta, x \rangle \vDash \varphi$ iff $\langle \mathfrak{B}^s, \gamma, x^\circ \rangle \vDash \alpha \to \varphi^s$. Moreover, $\langle \mathfrak{B}^s, \gamma, x^\bullet \rangle \vDash \alpha \to \varphi^s$ as well as $\langle \mathfrak{B}^s, \gamma, x^* \rangle \vDash \alpha \to \varphi^s$. So, $\langle \mathfrak{B}, \beta \rangle \vDash \varphi$ iff $\langle \mathfrak{B}^s, \gamma \rangle \vDash \alpha \to \varphi^s$. Since γ was arbitrary and $\mathfrak{B} \vDash \varphi$, we conclude that $\mathfrak{B}^s \vDash \alpha \to \varphi^s$. ☐

Theorem 6.6.16. *Let* Θ *be a bimodal logic and* $\Theta = \boldsymbol{K}_2 \oplus X$*. Then* $\Theta^s = \boldsymbol{Sim} \oplus \{\alpha \to \varphi^s : \varphi \in X\}$.

Proof. Let $\Lambda := \mathbf{Sim} \oplus \{\alpha \to \varphi^s : \varphi \in X\}$. A differentiated rooted frame for Λ is of the form $\mathfrak{B}^s$ for some bimodal frame $\mathfrak{B}$ by Corollary 6.6.12. $\mathfrak{B}^s \vDash \alpha \to \varphi^s$ for all $\varphi \in X$. Hence $\mathfrak{B} \vDash \varphi$ for all $\varphi \in X$. So, $\mathfrak{B}$ is a Θ–frame. The converse also holds. Hence, $\Lambda = \Theta^s$. ◻

Now that we have defined the simulation of a bimodal formula let us see whether there is a possibility of recovering from a bimodal frame the monomodal frame of which it is a simulation; this we call *unsimulating*. We will deal with unsimulations of frames that are frames for **Sim** and differentiated. It is possible to generalize this somewhat (see the exercises) but there is no benefit for the aims that we are pursuing here. First, the previous theorems guarantee that for a rooted differentiated **Sim**–frame $\mathfrak{M}$ there is exactly one $\mathfrak{B}$ such that $\mathfrak{B}^s \cong \mathfrak{M}$. We can make $\mathfrak{B}$ unique by the following construction.

Definition 6.6.17. *A* ***standard simulation frame*** *is a differentiated monomodal* ***Sim****–frame such that ω has a single point as its extension. The category of standard simulation frames and p–morphisms is denoted by* StSim. *Let $\mathfrak{M}$ be a standard simulation frame. Then $\mathfrak{M}_s$ is defined as follows. $m_s := m \cap \alpha$, $x \lhd y$ iff $x \leqslant y$ and $x \blacktriangleleft y$ iff there exist $x', y' \in \beta$ such that $x \leqslant x' \leqslant y' \leqslant y$. Finally, $\mathbb{M}_s := \{a \cap \alpha : a \in \mathbb{M}\} = \{a \in \mathbb{M} : a \leq \alpha\}$. Then $\mathfrak{M}_s := \langle m_s, \lhd, \blacktriangleleft, \mathbb{M}\rangle$. $\mathfrak{M}_s$ is called the* ***unsimulation*** *of $\mathfrak{M}$.*

Let $\mathfrak{M}$ and $\mathfrak{N}$ be standard simulation frames and $\pi : \mathfrak{M} \to \mathfrak{N}$ a p–morphism. Then put $\pi_s(x) := \pi(x)$, $x \in \alpha$. It is not hard to verify that this is a p–morphism from $\mathfrak{M}_s$ to $\mathfrak{N}_s$. Moreover, each p–morphism $\rho : \mathfrak{M}_s \to \mathfrak{N}_s$ is actually of the form π^s for some $\pi : \mathfrak{M} \to \mathfrak{N}$.

Theorem 6.6.18. *Let* Df_2 *denote the category of differentiated bimodal frames.* $(-)_s$ *is a functor from* StSim *to* Df_2. *Moreover, the two categories are naturally equivalent; there is a natural transformation from the identity functor on* StSim *to the functor* $((-)^s)_s$ *and a natural transformation from the identity functor on* Df_2 *to* $((-)_s)^s$. *In particular, for a bimodal frame $\mathfrak{B}$ and a monomodal frame $\mathfrak{M}$ we have $(\mathfrak{B}^s)_s \cong \mathfrak{B}$, $(\mathfrak{M}_s)^s \cong \mathfrak{M}$.*

Proof. Let $\mathfrak{B}$ be a bimodal frame. Then $\eta(\mathfrak{B}) : x \mapsto x^\circ$ is an isomorphism from $\mathfrak{B}$ to $(\mathfrak{B}^s)_s$. η is a natural transformation from the identity on Df_2 to $((-)^s)_s$ as is straightforward to check. Now let $\mathfrak{M}$ be a standard simulation frame. Then $\epsilon(\mathfrak{M})$ is defined as follows. $\epsilon(\mathfrak{M})(x) := x^\circ$ if $x \in \alpha$, $\epsilon(\mathfrak{M})(*) = *$. If $x \in \beta$, $\epsilon(\mathfrak{M})(x) := y^\bullet$, where $y \in \alpha$ and $x \leqslant y$. Again, it is straightforward to verify that $\epsilon(\mathfrak{M})$ is an isomorphism from $\mathfrak{M}$ onto $(\mathfrak{M}_s)^s$ and ϵ a natural transformation from the identity functor on StSim to $((-)_s)^s$. ◻

Notice the following particular consequence. The category of bimodal frames has coproducts, namely the disjoint union. So, by equivalence, the category of standard simulation frames must have coproducts, too. However, it is not closed under disjoint unions. So, the coproduct of frames exists but is not the disjoint union. Since we have an equivalence, the frame operation $\widehat{\bigoplus}$ defined below is a coproduct.

$$\widehat{\bigoplus_{i \in I}} \mathfrak{M}^i := (\bigoplus_{i \in I} \mathfrak{M}^i_s)^s \ .$$

There is an alternative way of defining $\widehat{\bigoplus}$. Take a differentiated **Sim**–frame $\mathfrak{M}$. Let $\sim$ be defined by $x \sim y$ iff $x, y \in \omega$ or $x = y$. Then $\sim$ is a net and the natural factorization $\mathfrak{M}/\sim$ is denoted by $\mathfrak{M}_\star$. It is easy to verify that $\mathfrak{M}_\star$ is a standard simulation frame. Moreover, for every point x, $\mathsf{Th}\,\langle \mathfrak{M}, x\rangle = \mathsf{Th}\,\langle \mathfrak{M}_\star, [x]\rangle$. (This follows from the fact that the transit of x in $\mathfrak{M}$ has a unique ω–point in it, and so is isomorphic to the transit of $[x]$ in $\mathfrak{M}_\star$.) It is easy to see that

$$\widehat{\bigoplus_{i \in I}} \mathfrak{M}^i \cong (\bigoplus_{i \in I} \mathfrak{M}^i)_\star \ .$$

Similar definitions can be made with respect to the coproduct of descriptive frames.

$$\widehat{\coprod_{i \in I}} \mathfrak{M}^i := (\coprod_{i \in I} \mathfrak{M}^i)_\star$$

Call this construction the **reduced coproduct**. Now recall from Theorem 4.7.5 that a class of descriptive frames is modally definable iff it is closed under coproducts, p–morphic images and generated subframes. It follows that if a class of bimodal descriptive frames is modally definable, its simulation image is closed under reduced coproducts, p–morphic images and generated subframes. The converse also holds. A class of descriptive **Sim**–frames is modally definable iff its intersection with the class of standard **Sim**–frames is closed under reduced coproducts, generated subframes and p–morphic images. Putting this together we obtain the following result.

Theorem 6.6.19. *The map $\Theta \mapsto \Theta^s$ is an isomorphism from the lattice $\mathcal{E}\,\boldsymbol{K}_2$ onto the interval $[\boldsymbol{Sim}, \mathsf{Th}\,\boxed{\bullet}\,]$ in $\boldsymbol{K}_1$.*

Definition 6.6.20. *Let Λ be an extension of $\boldsymbol{Sim}$ contained in $\mathsf{Th}\,\boxed{\bullet}$. Then let Λ_s be the unique bimodal logic Θ such that $\Theta^s = \Lambda$.*

It is clear that the map $\Lambda \mapsto \Lambda_s$ is a lattice isomorphism from the interval $[\mathbf{Sim}, \mathsf{Th}\,\boxed{\bullet}\,]$.

Exercise 211. Let $\mathfrak{B}$ be a bimodal algebra. Show that $Sub(\mathfrak{B}^s) \cong Sub(\mathfrak{B})$ and $Con(\mathfrak{B}^s) \cong Con(\mathfrak{B}) + 1$. (Here, $\mathfrak{L} + 1$ denotes the addition of a new element at the top of the lattice $\mathfrak{L}$.)

Exercise 212. Call a monomodal algebra **local** if it has a unique maximal congruence $\neq \nabla$. Show that a **Sim**–algebra is local iff its dual is a standard simulation frame. Moreover, show that every **Sim**–algebra has a largest local subalgbra.

Exercise 213. Call a monomodal frame $\mathfrak{M}$ **local** if (1.) $\mathfrak{F}$ is differentiated, (2.) the subframe $\mathfrak{G}$ of points x such that $\langle \mathfrak{F}, x\rangle \vDash \mathbf{Sim}$ is a standard simulation frame. Define $(-)_t$ on local frames as follows. Let $\mathfrak{G}$ be the subframe of points x such that $\langle \mathfrak{F}, x\rangle \vDash \mathbf{Sim}$. Then $\mathfrak{F}_t := \mathfrak{G}_s$. Show that $(-)_t$ is a functor. Now define $(-)^t$, a functor from the category of bimodal frames into the category of local frames simply by $\mathfrak{B}^t := \mathfrak{B}^s$. Show that $(-)_t$ is right adjoint to $(-)^t$.

Exercise 214. Let Θ be a quasinormal bimodal logic. Then define

$$\Theta^q := \bigcap_{\langle \mathfrak{B}, x\rangle \vDash \Theta} \mathsf{Th}\,\langle \mathfrak{B}^s, x^\circ\rangle \cap \mathsf{Th}\,\langle \mathfrak{B}^s, x^\bullet\rangle \cap \mathsf{Th}\,\langle \mathfrak{B}^s, x^*\rangle$$

Show that the map $\Theta \mapsto \Theta^q$ maps $\mathcal{Q}\,\mathbf{K}_2$ isomorphically into $\mathcal{Q}\,\mathbf{Sim}$. (This map does not need to be onto!) Moreover, show that if Θ is normal, $\Theta^q = \Theta^s$.

Exercise 215. Show that there exists no lattice isomorphism $(-)_q$ from $\mathcal{Q}\,\mathbf{Sim}$ onto $\mathcal{Q}\,\mathbf{K}_2$ such that for a normal logic Λ, $\Lambda_q = \Lambda_s$. Show furthermore that there is no homomorphism from $\mathcal{Q}\,\mathbf{Sim}$ into $\mathcal{Q}\,\mathbf{K}_2$ such that for normal bimodal logics Θ, $\Lambda_q = \Theta$ iff $\Lambda = \Theta^s$. *Hint.* Consider the bimodal frame $\mathfrak{f} := \langle \{0, 1\}, \lhd, \blacktriangleleft\rangle$ where $\blacktriangleleft = \varnothing$ and $\lhd = \{\langle 0, 1\rangle\}$. Put $\Theta := \mathsf{Th}\,\mathfrak{f}$. Show that there is no lattice isomorphism from $\mathcal{Q}\,\Theta^s$ onto $\mathcal{Q}\,\Theta$ with the required properties.

6.7. Properties of the Simulation

In this section we will investigate what properties of logics are preserved under simulation and unsimulation. As on similar occasions, we say that a property $\mathfrak{P}$ is **preserved** under simulation if for a given bimodal logic Λ, Λ^s has $\mathfrak{P}$ if Λ has $\mathfrak{P}$. $\mathfrak{P}$ is reflected under simulation if Λ has $\mathfrak{P}$ whenever Λ^s has $\mathfrak{P}$; and $\mathfrak{P}$ is invariant under simulation if it is both preserved and reflected under simulation. As a first step we investigate the problem of axiomatizing Λ_s on the basis of an axiomatization of Λ.

Axiomatization. In analogy to the construction of φ^s from a bimodal formula φ we want to construct for a given monomodal formula φ a bimodal formula ψ such that φ is equivalent to ψ^s. This is not possible; just take the formula $\alpha \wedge \Diamond(\beta \wedge p)$. The problem is that the bimodal frame is internally reconstructed as the area defined by α. So the values under a valuation in the β and ω region cannot be distinguished. We can, however, mimick the behaviour of a monomodal valuation as follows. For each variable p we introduce three new variables, p^*, p° and $p^\bullet$. Given our original set V or variables, we obtain three new sets, $V^\flat := \{p^\flat : p \in V\}$, $\flat \in \{\bullet, \circ, *\}$. All four sets are assumed to be pairwise disjoint.

In this section we are working in the language $\boxminus_\alpha$, $\boxminus_\beta$ and $\boxminus_\omega$ together with their duals $\Diamond_\alpha$, $\Diamond_\beta$ and $\Diamond_\omega$. Let φ be a formula and χ a subformula of φ. Fix an occurrence of χ in φ. A **modal cover** of that occurrence of χ is a minimal subformula ψ of modal degree greater than χ containing that occurrence of χ. We also say that this particular occurrence of ψ **modally covers** χ. If χ has a modal cover, it is unique and a formula beginning with a modal operator. (We will often speak of *formulae* rather than *occurrences of formulae*, whenever the context allows this.) Now let φ be a formula of the language with operators $\boxminus_\flat$, $\diamondminus_\flat$, $\flat \in \{\alpha, \beta, \omega\}$. Let us agree to say that an occurrence of a formula χ in φ is $\flat$**–covered** if it modally covered by a formula of the form $\boxminus_\flat\psi$ or $\diamondminus_\flat\psi$. Call φ and χ **white–equivalent** if $\alpha \vdash_{\mathbf{Sim}} \varphi \leftrightarrow \chi$ and **black–equivalent** if $\beta \vdash_{\mathbf{Sim}} \varphi \leftrightarrow \chi$. Given a formula, we say that a subformula occurs **white** if it is not in the scope of a modal operator or else is α–covered. A subformula occurs **black** if it is β–covered. If φ occurs white (black) in ψ, and φ is white–equivalent (black–equivalent) to χ, then that occurrence of φ may be replaced in ψ by χ preserving white–equivalence. By axioms (c) and (d) of **Sim**, $\boxminus_\beta\tau$ is white–equivalent to $\diamondminus_\beta\tau$ and by (e) and (f), $\boxminus_\alpha\tau$ is black–equivalent to $\diamondminus_\alpha\tau$.

Lemma 6.7.1. *Let φ be a formula in the language with $\boxminus_\flat$ and $\diamondminus_\flat$, $\flat \in \{\alpha, \beta, \omega\}$. There exists a finite number n and formulae χ_i and ψ_i, $i < n$, such that χ_i is nonmodal for all $i < n$, and ψ_i is in the language with $\boxminus_\alpha, \boxminus_\beta$, $\diamondminus_\alpha$ and $\diamondminus_\beta$ for all $i < n$, and φ is white–equivalent to the formula*

$$\bigvee_{i<n} (\diamondminus_\omega\chi_i) \wedge \psi_i$$

Proof. By a straightforward semantic argument one shows for every n

$$\alpha; \boxminus_\omega\chi \Vdash_{\mathbf{Sim}} \boxminus_\omega\chi$$

Take $S \subseteq var(\varphi)$. Put

$$\chi(S) := \bigwedge_{p \in S} p \wedge \bigwedge_{p \in var(\varphi) - S} \neg p$$

Then

$$\alpha \vdash_{\mathbf{sim}} \varphi \quad \leftrightarrow \quad \bigvee_{S \subseteq var(\varphi)} \varphi \wedge \diamondminus_\omega\chi(S)$$

Now by the above consideration, any occurrence of $\diamondminus_\omega\chi(S)$ as a subformula of φ can be replaced by $\top$ in the formula $\varphi \wedge \diamondminus_\omega\chi(S)$. It remains to be shown that any subformula $\boxminus_\omega\chi$ is of this type. Clearly, by some **Sim**–equivalences, any subformula $\boxminus_\omega\chi$ can be transformed into a subformula $\boxminus_\omega\chi'$ where χ' is nonmodal. Moreover, we are allowed to replace χ' by $\chi(S) \wedge \chi'$. The latter reduces via some equivalences to either $\chi(S)$ or to $\bot$. Therefore, the subformulae of the form $\boxminus_\omega\bot$ need to be treated. Several cases need to be distinguished. (a) $\boxminus_\omega\bot$ is ω–covered. Then it can be replaced by $\top$. (b) $\boxminus_\omega\bot$ is α–covered. Then it can be replaced by $\bot$. (c) $\boxminus_\omega\top$ is β–covered. Then it can be replaced by $\top$. (d) $\boxminus_\omega\bot$ is not in the scope of an operator. Then it can be replaced by $\bot$. All these replacements are **Sim**–equivalences. This shows the lemma. ◻

Definition 6.7.2. *A monomodal formula φ is called* **simulation transparent** *if it is of the form p, $\neg p$, $\diamondminus_\beta p$, $\neg\diamondminus_\beta p$, $\diamondminus_\omega p$, $\neg\diamondminus_\omega p$, p a variable, or of the form $\psi \wedge \chi$, $\psi \vee \chi$, $\diamondminus_\alpha\psi$, $\boxminus_\alpha\psi$, $\diamondminus_\beta\diamondminus_\beta\diamondminus_\alpha\psi$ or $\boxminus_\beta\boxminus_\beta\boxminus_\alpha\psi$ where ψ and χ are simulation transparent.*

Definition 6.7.3. *Call a formula φ **white based** if there do not exist occurrences of subformulae χ_0, χ_1, χ_2 and χ_3 such that χ_0 β–covers χ_1, χ_1 β–covers χ_2, and χ_2 β–covers χ_3.*

Lemma 6.7.4. *For every formula φ there exists a formula χ which is white–based and white–equivalent to φ.*

Proof. Suppose that there is a quadruple $\langle\chi_0, \chi_1, \chi_2, \chi_3\rangle$ of occurrences of subformulae such that χ_0 β–covers χ_1, χ_1 β–covers χ_2, and χ_2 β–covers χ_3. Then there exists such a quadruple in which χ_0 occurs white. Now replace the occurrence of χ_3 by $\boxminus_\alpha \boxminus_\beta \chi_3$. Since χ_3 is black equivalent with $\boxminus_\alpha \boxminus_\beta \chi_3$, this replacement yields a formula φ' which is white equivalent to φ'. Now repeat this procedure with φ'. It is not hard to see that this process terminates with a white based formula. (For example, count the number of occurrences of quadruples $\langle\chi_0, \chi_1, \chi_2, \chi_3\rangle$ such that χ_0 β–covers χ_1, χ_1 β–covers χ_2, and χ_2 β–covers χ_3. It decreases by at least one in passing from φ to φ'. If it is zero, the formula is white based.) ◻

Lemma 6.7.5. *Let φ be a monomodal formula. Then there exists a simulation transparent formula χ such that*

$$\alpha \vdash_{\mathbf{Sim}} \varphi \leftrightarrow \chi$$

Proof. First we simplify the problem somewhat. Namely, by some standard manipulations we can achieve it that no operator occurs in the scope of negation. We call a formula in such a form **basic**. So, let us assume φ to be basic. By Lemma 6.7.1 we can assume φ to be a disjunction of formulae of the form $\upsilon \wedge \psi$, where $\upsilon = \Diamond_\omega \chi$, for a nonmodal χ, and τ contains only $\boxminus_\alpha$, $\Diamond_\alpha$, $\boxminus_\beta$ and $\Diamond_\beta$. In general, if the claim holds for φ_1 and φ_2, then it also holds for $\varphi_1 \vee \varphi_2$ and $\varphi_1 \wedge \varphi_2$. Therefore, we have two cases to consider: (i) φ contains no occurrences of $\Diamond_\alpha$, $\boxminus_\alpha$, $\Diamond_\beta$ or $\boxminus_\beta$, or (ii) φ contains no occurrences of $\Diamond_\omega$ and $\boxminus_\omega$. In case (i), we know that $\Diamond_\omega$ distributes over $\vee$ and $\wedge$, so that we can reduce φ (modulo white–equivalence) to the form $\Diamond_\omega p$, and $\Diamond_\omega \neg p$. Now we have

$$\alpha \vdash_{\mathbf{Sim}} \Diamond_\omega \neg p \leftrightarrow \neg \Diamond_\omega p$$

So in Case (i) φ is white–equivalent to a simulation transparent formula. From now on we can assume to be in Case (ii). Furthermore, by Lemma 6.7.4, we can assume that φ is white based, and (inspecting the proof of that lemma) that φ is built from variables and negated variables, using $\wedge$, $\vee$, and the modal operators $\Diamond_\alpha$, $\Diamond_\beta$, $\boxminus_\alpha$ and $\boxminus_\beta$.

Let $\mu_\beta(\varphi)$ denote the maximum number of nestings of black operators ($\Diamond_\beta$, $\boxminus_\beta$) in φ. Call φ **thinner** than χ if either $\mu_\beta(\varphi) < \mu_\beta(\chi)$ or φ is a subformula of χ. We will show that for given white based, basic φ there exists a simulation transparent formula χ which is white–equivalent to φ on the condition that this holds already for all white based basic formulae φ' thinner than φ.

If $\varphi = p$ we are done; for φ is simulation transparent. Likewise, if $\varphi = \neg p$. Suppose $\varphi = \varphi_1 \wedge \varphi_2$. φ_1 and φ_2 are thinner than φ. Therefore there exist simulation transparent formulae χ_1 and χ_2 such that χ_i is white–equivalent to φ_i, $i \in \{1, 2\}$. Then $\chi_1 \wedge \chi_2$ is white–equivalent to $\varphi_1 \wedge \varphi_2$. Similarly for $\varphi = \varphi_1 \vee \varphi_2$. If $\varphi = \diamondminus_\alpha \varphi_1$ there exists a simulation transparent χ which is white–equivalent to φ_1. So $\alpha \to \varphi_1 \dashv\vdash_{\mathbf{Sim}} \alpha \to \chi$, and therefore $\diamondminus_\alpha \varphi_1 \dashv\vdash_{\mathbf{Sim}} \diamondminus_\alpha \chi$; it follows that $\diamondminus_\alpha \chi$ is white–equivalent to φ. Similarly for $\varphi = \boxminus_\alpha \varphi_1$. We are left with the case that φ is either $\diamondminus_\beta \tau$ or $\boxminus_\beta \tau$. By inductive hypothesis, for every basic white based χ such that $\mu_\beta(\chi) < \mu_\beta(\varphi)$ there is a simulation transparent ω such that ω is white–equivalent to χ. As φ occurs white, we can distribute $\boxminus_\beta$ and $\diamondminus_\beta$ over $\wedge$ and $\vee$, and so reduce τ to the form p, $\neg p$ or $\boxminus_\flat \rho$ or $\diamondminus_\flat \rho$, with $\flat \in \{\alpha, \beta\}$ and ρ basic. This reduction does not alter $\mu_\beta(\varphi)$.
Case 1. $\tau = p$. Then φ is simulation transparent.
Case 2. $\tau = \neg p$. Observe that $\diamondminus_\beta \neg p$ is white equivalent to $\neg \diamondminus_\beta p$. So we are done.
Case 3. $\tau = \boxminus_\alpha \rho$ or $\tau = \diamondminus_\alpha \rho$. Then φ is white–equivalent to ρ. The claim follows by induction hypothesis for ρ.
Case 4. $\tau = \boxminus_\beta \rho$ or $\tau = \diamondminus_\beta \rho$, ρ basic. Now let us look at ρ. ρ occurs black. Furthermore, ρ is the result of applying a lattice polynomial to formulae of the form p, $\neg p$, $\boxminus_\beta \mu$, $\diamondminus_\beta \nu$, $\boxminus_\alpha \zeta$, $\diamondminus_\alpha \eta$. (Here, a lattice polynomial is an expression formed from variables and constants using only $\wedge$ and $\vee$, but no other functions. It turns out that $\top$ and $\bot$ can be eliminated from this polynomial as long as it contains at least one variable. It it does not contain a variable, it is equivalent to either $\top$ or $\bot$. Finally, $\top$ may be replaced by $p \vee \neg p$ and $\bot$ by $p \wedge \neg p$, so we may assume that the polynomial does not contain any occurrences of $\top$ and $\bot$.) However, as φ is white based, formulae of the form $\diamondminus_\beta \mu$ or $\boxminus_\beta \nu$ do not occur. Furthermore, if $\mu_\beta(\rho) > 0$, we replace the unmodalized occurrences of p by $\boxminus_\alpha \boxminus_\beta p$, and the unmodalized occurrences of $\neg p$ by $\boxminus_\alpha \boxminus_\beta \neg p$. This replacement does not change $\mu_\beta(\rho)$. (The case $\mu_\beta(\rho) = 0$ needs some attention. Here we replace ρ by $\boxminus_\alpha \boxminus_\beta \rho$. Then φ is white equivalent to either $\boxminus_\beta \boxminus_\beta \boxminus_\alpha \boxminus_\beta \rho$ or $\diamondminus_\beta \diamondminus_\beta \diamondminus_\alpha \boxminus_\beta \rho$, ρ nonmodal. Now we are down to the case of a formula of the form $\boxminus_\beta \rho$. $\boxminus_\beta$ commutes with $\wedge$ and $\vee$, which leaves the cases $\boxminus_\beta p$ and $\boxminus_\beta \neg p$ to consider. These are immediate.) Now, after this replacement, ρ is a lattice polynomial over formulae of the form $\boxminus_\alpha \zeta$, $\diamondminus_\alpha \eta$. The latter occur black, so $\boxminus_\alpha \zeta$ is intersubstitutable with $\diamondminus_\alpha \zeta$ and $\diamondminus_\alpha \eta$ is intersubstitutable with $\boxminus_\alpha \eta$. Finally, $\diamondminus_\alpha \top$ is intersubstitutable (modulo equivalence) with $\top$. So ρ is without loss of generality of the form $f(\langle \boxminus_\alpha \delta_i | i < n \rangle)$ for some lattice polynomial f. Thus, ρ can be substituted by the formula $\boxminus_w f(\langle \delta_i | i < n \rangle)$. Therefore, ρ can be reduced to the form $\boxminus_\alpha \delta$ for some basic δ. δ is white based. Therefore, by induction hypothesis, δ is white–equivalent to a simulation transparent formula θ. τ has the form $\boxminus_\beta \boxminus_\alpha \delta$ (**Case A**) or $\diamondminus_\beta \boxminus_\alpha \delta$ (**Case B**), so φ is white–equivalent to either $\chi_1 := \boxminus_\beta \boxminus_\beta \boxminus_\alpha \theta$ (Case A) or $\boxminus_\beta \diamondminus_\beta \boxminus_\alpha \theta$ (Case B). The latter is white–equivalent to $\chi_2 := \diamondminus_\beta \diamondminus_\beta \diamondminus_\alpha \theta$. Both χ_1 and χ_2 are simulation transparent, by assumption on θ. ◻

LEMMA 6.7.6. *Let φ be a monomodal formula in the variables $\{p_i | i < n\}$. There exists a bimodal formula χ in the variables $\{p_i^\circ, p_i^\bullet, p_i^* | i < n\}$ such that*

$$(\ddagger): \qquad \alpha \vdash_{\mathbf{Sim}} \varphi \leftrightarrow \chi^s[p_i/p_i^\circ, \blacklozenge_\beta p_i/p_i^\bullet, \blacklozenge_\omega p_i/p_i^* | i < n]$$

χ is called an ***unsimulation*** *of φ.*

PROOF. There exists a simulation transparent τ which is white–equivalent to φ. χ is obtained from τ by applying the following translation outside in.

$$\begin{array}{llllll}
(\chi_1 \wedge \chi_2)_\tau & := & (\chi_1)_\tau \wedge (\chi_2)_\tau & (\chi_1 \vee \chi_2)_\tau & := & (\chi_1)_\tau \vee (\chi_2)_\tau \\
(\blacklozenge_\alpha \chi)_\tau & := & \Diamond \chi_\tau & (\boxminus_\alpha \chi)_\tau & := & \Box \chi_\tau \\
(\blacklozenge_\beta \blacklozenge_\beta \blacklozenge_\alpha \chi)_\tau & := & \blacklozenge \chi_\tau & (\boxminus_\beta \boxminus_\beta \boxminus_\alpha \chi)_\tau & := & \blacksquare \chi_\tau \\
(\blacklozenge_\omega p)_\tau & := & p^* & (\neg \blacklozenge_\omega p)_\tau & := & \neg p^* \\
(\blacklozenge_\beta p)_\tau & := & p^\bullet & (\neg \blacklozenge_\beta p)_\tau & := & \neg p^\bullet \\
p_\tau & := & p^\circ & (\neg p)^\tau & := & \neg p^\circ
\end{array}$$

This concludes the proof. □

The formula χ is of course not uniquely determined by τ, but is unique up to equivalence. The proof of Lemma 6.7.5 is actually a construction of χ, and so let us denote by χ_s the particular formula that is obtained by performing that construction. Now take a set Δ of monomodal formulae; put $\Delta_s := \{\varphi_s : \varphi \in \Delta\}$. Assume that $\mathfrak{M}$ is a simulation frame and $\langle \mathfrak{M}, \beta, x \rangle \vDash \alpha; \Delta$. Then we have

$$\langle \mathfrak{M}, \beta, x \rangle \vDash \alpha; \sigma((\Delta_s)^s),$$

where $\sigma(p^\circ) := p$, $\sigma(p^\bullet) := \blacklozenge_\beta p$, $\sigma(p^*) := \blacklozenge_\omega p$. Now define a valuation $\beta^\diamond$ of the set $\{p^\circ, p^\bullet, p^* : p \in var[\Delta]\}$ by

$$\beta^\diamond(p^\circ) := \beta(p) \cap f^\circ, \quad \beta^\diamond(p^\bullet) := \blacklozenge_b \beta(p) \cap f^\circ, \quad \beta^\diamond(p^*) := \blacklozenge_t \beta(p) \cap f^\circ$$

By definition of $\beta^\diamond$,

$$\langle \mathfrak{M}, \beta, x \rangle \vDash \alpha; \sigma(\psi) \qquad \Leftrightarrow \qquad \langle \mathfrak{M}, \beta^\diamond, x \rangle \vDash \alpha; \psi$$

Thus we conclude

$$\langle \mathfrak{M}, \beta^\diamond, x \rangle \vDash \alpha; (\Delta_s)^s$$

Define a valuation γ on $\mathfrak{M}_s$ by $\gamma(q) := \beta^\diamond(q)$. x is of the form y° for some $y \in m_s$; in fact, by construction, $y^\circ = x$. By the previous results,

$$\langle \mathfrak{M}, \beta^\diamond, x \rangle \vDash \alpha; (\Delta_s)^s \qquad \Leftrightarrow \qquad \langle \mathfrak{M}_s, \gamma, x \rangle \vDash \Delta_s.$$

It therefore turns out that the satisfaction of Δ in a simulation frame at a white point is equivalent to the satisfaction of Δ_s in the unsimulation of the frame. The satisfaction of Δ at a black point is equivalent to the satisfaction of $\boxminus_\beta \Delta := \{\boxminus_\beta \varphi : \varphi \in \Delta\}$ at a white point. The satisfaction of Δ at f^* is likewise reducible to satisfaction of $\boxminus_\omega \Delta$ at a white point, which is defined analogously.

Now assume that φ is a theorem of Λ, Λ monomodal and consistent. Then $\omega \to \varphi$, $\alpha \to \varphi$ and $\beta \to \varphi$ are theorems of Λ as well. By consistency, $\omega \to \varphi$

can only be a theorem if it is a tautology. $\beta \to \varphi$ is globally equivalent in **Sim** to $\alpha \to \boxminus_\beta \varphi$. Thus, we can always assume that an axiom is of the form $\psi := \alpha \to \varphi$ for some φ. (This follows independently from the surjectivity of the simulation map and the fact that an axiomatization of this form for simulation logics has been given above.) Now φ is falsified in a model based on $\mathfrak{M}$ iff φ is rejected at a white point iff φ_s is rejected in a model based on $\mathfrak{M}_s$.

We summarize our findings as follows. Given a monomodal rooted **Sim**–frame $\mathfrak{G}$, a set Δ, a valuation γ, we define γ_s by $\gamma_s(q) := \gamma^\circ(q)$, where q is a variable of the form p°, $p^\bullet$ or p^*.

$$\begin{array}{ccc} \langle \mathfrak{G}, \gamma, x\rangle \vDash \alpha \wedge \Delta & \Leftrightarrow & \langle \mathfrak{G}_s, \gamma_s, x\rangle \vDash \Delta_s \\ \langle \mathfrak{G}, \gamma\rangle \vDash \Delta & \Leftrightarrow & \langle \mathfrak{G}_s, \gamma_s\rangle \vDash \Delta_s; (\boxminus_\beta \Delta)_s; (\boxminus_\omega \Delta)_s \\ \mathfrak{G} \vDash \Delta & \Leftrightarrow & \mathfrak{G}_s \vDash \Delta_s; (\boxminus_\beta \Delta)_s; (\boxminus_\omega \Delta)_s \end{array}$$

Two cases may arise. Suppose that $\mathfrak{G}$ contains only one point. Then $\mathfrak{G}_s$ is empty, and no formula is satisfiable in it. This case has to be dealt with separately. Else, let $\mathfrak{G}$ have more than point, and be rooted. Then satisfiability of Δ in $\mathfrak{G}$ is reducible to satisfiability of either $\alpha; \Delta_s$ or $\alpha; \boxminus_\omega \Delta_s$ or $\alpha; \boxminus_\beta \Delta_s$. All problems are reducible to satisfiability of a set Δ in $\mathfrak{G}_s$. Global satisfiability of Δ in $\langle \mathfrak{G}, \gamma\rangle$ is equivalent to the global satisfiability of $\{\alpha \to \varphi : \varphi \in \Delta^+\}$, where $\Delta^+ := \Delta; \boxminus_\beta \Delta; \boxminus_\omega \Delta$.

Theorem 6.7.7. *Let Λ be a monomodal logic in the interval* $[\boldsymbol{Sim}, \mathsf{Th}\,\boxed{\bullet}\,]$. *Then Λ is axiomatizable as $\Lambda := \boldsymbol{K}_1 \oplus \{\alpha \to \delta : \delta \in \Delta\}$ for some Δ. Furthermore, $\Lambda_s = \boldsymbol{K}_2 \oplus \Delta_s \oplus (\Box_\alpha \Delta)_s \oplus (\boxminus_\omega \Delta)_s$.*

The following properties are now shown to be invariant under simulation: *finite axiomatizability*, *(strongly) recursive axiomatizability*, *0–axiomatizability*.
Decidability and Completeness.

Proposition 6.7.8. *Let Δ be a set of bimodal formulae, φ a bimodal formula, and Θ a bimodal logic. Then*

$$\begin{array}{ccc} \Delta \vdash_\Theta \varphi & \Leftrightarrow & \alpha \to \Delta^s \vdash_{\Theta^s} \alpha \to \varphi^s \\ \Delta \Vdash_\Theta \varphi & \Leftrightarrow & \alpha \to \Delta^s \Vdash_{\Theta^s} \alpha \to \varphi^s \end{array}$$

Proof. Assume $\Delta \vdash_\Theta \varphi$. Let $\mathfrak{M}$ be a rooted and differentiated Θ^s–frame, δ a valuation and x a world such that $\langle \mathfrak{M}, \delta, x\rangle \vDash \alpha \to \Delta^s$. Then $\mathfrak{M} \cong \mathfrak{B}^s$ for some rooted Θ–frame $\mathfrak{B}$. Two cases arise.
Case 1. $\langle \mathfrak{M}, \delta, x\rangle \vDash \neg\alpha$. Then $\langle \mathfrak{B}, \delta, x\rangle \vDash \alpha \to \varphi^s$.
Case 2. $x \vDash \alpha$. Then $x = y^\circ$ for some $y \in b$. and a valuation γ on $\mathfrak{B}$ such that $\gamma(p)^\circ = \delta(p) \cap b^\circ$. Then $\langle \mathfrak{B}, \gamma, y\rangle \vDash \Delta$. By assumption, $\langle \mathfrak{B}, \gamma, y\rangle \vDash \varphi$ and so $\langle \mathfrak{M}, \delta, x\rangle \vDash \alpha \to \varphi^s$. Thus $\alpha \to \Delta^s \vdash_{\Theta^s} \alpha \to \varphi^s$. Now assume that the latter holds. Take a bimodal model $\langle \mathfrak{B}, \gamma, y\rangle \vDash \Delta$. Then $\langle \mathfrak{B}^s, \gamma^s, y^\circ\rangle \vDash \alpha \to \Delta^s$ and so by assumption $\langle \mathfrak{B}^s, \gamma^s, y^\circ\rangle \vDash \alpha \to \varphi^s$ and so $\langle \mathfrak{B}^s, \gamma^s, y^\circ\rangle \vDash \varphi^s$. From this we get $\langle \mathfrak{B}, \gamma, y\rangle \vDash \varphi$. Hence $\Delta \vdash_\Theta \varphi$. Now for the global deducibility assume $\Delta \Vdash_\Theta \varphi$. Let $\langle \mathfrak{M}, \gamma\rangle \vDash \alpha \to \Delta^s$. Then we can assume that $\mathfrak{M} = \mathfrak{B}^s$ for some $\mathfrak{B}$ and that $\gamma = \delta^s$. We then have

$\langle\mathfrak{B}, \delta\rangle \vDash \Delta$. By assumption, $\langle\mathfrak{B}, \delta\rangle \vDash \varphi$. From this we get $\langle\mathfrak{B}^s, \delta^s\rangle \vDash \alpha \to \varphi$. Hence $\alpha \to \Delta^s \Vdash_{\Theta^s} \alpha \to \varphi^s$. Assume the latter holds and let $\langle\mathfrak{B}, \delta\rangle \vDash \Delta$. Then $\langle\mathfrak{B}^s, \delta^s\rangle \vDash \alpha \to \Delta^s$ and so $\langle\mathfrak{B}^s, \delta^s\rangle \vDash \alpha \to \varphi^s$. From this we obtain $\langle\mathfrak{B}, \delta\rangle \vDash \varphi$. Hence $\Delta \Vdash_\Theta \varphi$. ◻

Proposition 6.7.9. *Let Λ be a monomodal logic, Δ a set of monomodal formulae and φ a monomodal formula. Then*

$$\begin{array}{lcrrl} \Delta \vdash_\Lambda \varphi & \Leftrightarrow & & \Delta_s & \vdash_{\Lambda_s} \varphi_s \\ & & \textit{and} & (\boxminus_\beta \Delta)_s & \vdash_{\Lambda_s} (\boxminus_\beta \varphi)_s \\ & & \textit{and} & (\boxminus_\omega \Delta)_s & \vdash_{\Lambda_s} (\boxminus_\omega \varphi)_s \\ \Delta \Vdash_\Lambda \varphi & \Leftrightarrow & & \Delta_s; (\boxminus_\beta \Delta)_s; (\boxminus_\omega \Delta)_s & \Vdash_{\Lambda_s} \varphi_s; (\boxminus_\beta \varphi)_s; (\boxminus_\omega \varphi)_s \end{array}$$

Proof. Let $\Delta \vdash_\Lambda \varphi$. Suppose that $\langle\mathfrak{B}, \delta, x\rangle \vDash \Delta_s$. Then $\langle\mathfrak{B}^s, \delta^s, x^\circ\rangle \vDash \alpha \wedge \Delta$. By assumption, $\langle\mathfrak{B}^s, \delta^s, x^\circ\rangle \vDash \alpha \wedge \varphi$. Hence $\langle\mathfrak{B}, \delta, x\rangle \vDash \varphi_s$. This shows $\Delta_s \vdash_{\Lambda_s} \varphi_s$. Similarly, $(\boxminus_\beta\Delta)_s \vdash_{\Lambda_s} (\boxminus_\beta\Delta)_s$ and $(\boxminus_\omega\Delta)_s \vdash_{\Lambda_s} (\boxminus_\omega\Delta)_s$ are proved. Now assume that all three obtain. Let $\langle\mathfrak{M}, \gamma, x\rangle \vDash \Delta$.
Case 1. $x = u^\circ$. Then $\langle\mathfrak{M}_s, \gamma_s, u\rangle \vDash \Delta_s$ and so $\langle\mathfrak{M}_s, \gamma_s u\rangle \vDash \varphi_s$. Thus $\langle\mathfrak{M}, \gamma, u^\circ\rangle \vDash \varphi$.
Case 2. $x = u^\bullet$. Then $\langle\mathfrak{M}, \gamma, u^\circ\rangle \vDash \boxminus_\beta\Delta$. As in Case 1. we get $\langle\mathfrak{M}, \gamma, u^\circ\rangle \vDash \boxminus_\beta\varphi$. Therefore $\langle\mathfrak{M}, \gamma, u^\bullet\rangle \vDash \varphi$. **Case 3.** $x = u^*$. Then $\langle\mathfrak{M}, \gamma, u^\circ\rangle \vDash \boxminus_\omega\Delta$. As in Case 1. we get $\langle\mathfrak{M}, \gamma, u^\circ\rangle \vDash \boxminus_\omega\varphi$. Hence $\langle\mathfrak{M}, \gamma, u^*\rangle \vDash \varphi$. Hence, $\Delta \vdash_\Lambda \varphi$. For the global deducibility let $\Delta \Vdash_\Lambda \varphi$. Assume $\langle\mathfrak{B}, \delta\rangle \vDash \Delta_s; (\boxminus_\beta\Delta)_s; (\boxminus_\omega\Delta)_s$. Then $\langle\mathfrak{B}^s, \delta^s\rangle \vDash \Delta$. By assumption, $\langle\mathfrak{B}^s, \delta^s\rangle \vDash \varphi$; this gives $\langle\mathfrak{B}, \delta\rangle \vDash \varphi_s; (\boxminus_\beta\varphi)_s; (\boxminus_\omega\varphi)_s$. Therefore

$$\Delta_s; (\boxminus_\beta\Delta)_s; (\boxminus_\omega\Delta)_s \Vdash_{\Lambda_s} \varphi_s; (\boxminus_\beta\varphi)_s; (\boxminus_\omega\varphi)_s \ .$$

Assume finally that the latter holds. Let $\langle\mathfrak{M}, \gamma\rangle \vDash \Delta$. Then $\langle\mathfrak{M}_s, \gamma_s\rangle \vDash \Delta; (\boxminus_\beta\Delta)_s; (\boxminus_\omega\Delta)_s$. By our assumption, $\langle\mathfrak{M}_s, \gamma_s\rangle \vDash \varphi_s; (\boxminus_\beta\varphi)_s; (\boxminus_\omega\varphi)_s$. This gives $\langle\mathfrak{M}, \gamma\rangle \vDash \varphi$. Hence $\Delta \Vdash_\Lambda \varphi$. ◻

As a consequence of these two theorems, the following properties are invariant under simulations: *local decidability*, *global decidability*, *local completeness*, *global completeness*, *local finite model property*, *global finite model property*, *weak compactness*, *strong compactness*.
Complexity. Let φ be a bimodal formula. Then φ^s can be computed in quadratic time. Moreover, its length is $O(|\varphi|)$. So, if Θ^s is $\mathcal{C}$–computable (where $\mathcal{C}$ is for example NP, PSPACE or EXPTIME, but any class invariant under linear time reductions will do) then so is Θ. Similarly, if Θ^s is globally $\mathcal{C}$–computable, so is Θ. Conversely, let χ be a monomodal formula. Going carefully through the construction one can show that it takes polynomial time to construct χ_s from χ. Moreover, $|\chi_s|$ is of size $O(|\chi|)$. It follows that if Θ is in $\mathcal{C}$, so is Θ^s. Therefore, Θ and Θ^s belong to the same complexity class, as satisfiability problems are linearly interreducible. Hence the following properties of logics are invariant under simulation: *local and global $\mathcal{C}$–computability*, *local and global $\mathcal{C}$–hardness* and *local and global $\mathcal{C}$–completeness*.

Persistence. Assume that Λ is r–persistent. Then let $\mathfrak{M}$ be a refined Λ^s–frame. We know that $\mathfrak{M}_s$ is a Λ–frame, and is refined by Proposition 6.6.3. So $(\mathfrak{M}_s)_\sharp$ is a Λ–frame, by r–persistence of Λ. But $(\mathfrak{M}_s)_\sharp = (\mathfrak{M}_\sharp)_s$, so $\mathfrak{M}_\sharp$ is a Λ^s–frame as well. The same reasoning establishes preservation of df–persistence, d–persistence and c–persistence. And analogously the reflection of these persistence properties is shown. Only with properties such as α–canonicity one has to be careful, since the unsimulation uses more variables. Finally, if φ is constant, so is φ^s. And if ψ is constant, so is ψ_s. This shows invariance of g–persistence as well. The following properties have been shown to be invarariant under simulation: *g–persistence*, *df–persistence*, *r–persistence*, *d–persistence*, *κ–canonicity* (κ infinite).
Elementarity. The next lemma shows that a bimodal logic is $\mathfrak{Krp}$–elementary iff its simulation is $\mathfrak{Krp}$–elementary.

Lemma 6.7.10. *Let $\mathfrak{K}$ be a class of bimodal Kripke–frames, $\mathfrak{L}$ a class of standard simulation frames. Then*

$$\begin{aligned} \mathsf{Up}\,\mathfrak{K}^s &= (\mathsf{Up}\,\mathfrak{K})^s \\ \mathsf{Up}\,\mathfrak{L}_s &= (\mathsf{Up}\,\mathfrak{L})_s \end{aligned}$$

It is easy to define a translation for elementary properties under simulation.

$$\begin{aligned} \omega(x) &:= (\forall y \geqslant x)\neg(y \doteq y) \\ \alpha(x) &:= (\exists y \geqslant x)\omega(x) \\ \beta(x) &:= \neg\omega(x) \wedge (\forall y \geqslant x)\neg\omega(x) \end{aligned}$$

These formulae define the sets f^t, f^w and f^t in a **Sim**–frame. Now put

$$\delta^e := \bigwedge_{x \in fvar(\delta)} \alpha(x). \rightarrow .\delta^f$$

$$\begin{aligned} (x \doteq y)^f &:= x \doteq y \\ (x \lhd y)^f &:= x \leqslant y \\ (x \blacktriangleleft y)^f &:= (\exists v \geqslant x)(\exists w \geqslant y)(\beta(v) \wedge \beta(w) \wedge v \leqslant w) \\ (\delta_1 \wedge \delta_2)^f &:= \delta_1^f \wedge \delta_2^f \\ (\neg\delta)^f &:= \neg(\delta^f) \\ ((\exists x)\delta)^f &:= (\exists x)(\alpha(x) \wedge \delta^f) \end{aligned}$$

It is not difficult to show that for a sentence α,

$$\mathfrak{f} \vDash \delta \quad \Leftrightarrow \quad \mathfrak{f}^s \vDash \delta^e$$

It is not hard to see that the class of **Sim**–Kripke frames is elementary. This also follows from the fact that axioms for **Sim** are Sahlqvist.

The case of unsimulating elementary properties is more complex than the simulating part. Take a formula φ in the first–order language for 1–modal frames. We may assume (to save some notation) that the formula does not contain $\forall$. Furthermore, we may assume that the formula is a sentence, that is, contains no free variables. Finally, we do not assume that structures are nonempty. We introduce new quantifiers $\exists^\alpha$, $\flat \in \{\alpha, \beta, \omega\}$, which are defined by

$$(\exists^\flat x)\varphi(x) := (\exists x)(\flat(x) \wedge \varphi(x))$$

Furthermore, for each variable x we introduce three new variables, $x^\flat$, where $\flat \in \{\alpha, \beta, \omega\}$. Now define a translation $(-)^\dagger$ as follows.

$$\begin{array}{lcl}
(\varphi \wedge \psi)^\dagger & := & \varphi^\dagger \wedge \psi^\dagger \\
(\varphi \vee \psi)^\dagger & := & \varphi^\dagger \vee \psi^\dagger \\
(\neg\varphi)^\dagger & := & \neg\varphi^\dagger \\
((\exists x)\varphi(x))^\dagger & := & (\exists^\alpha x^\alpha)\varphi(x^\alpha)^\dagger \vee (\exists^\beta x^\beta)\varphi(x^\beta)^\dagger \vee (\exists^\omega x^\omega)\varphi(x^\omega)^\dagger
\end{array}$$

It is clear that φ and $\varphi^\dagger$ are deductively equivalent. In a next step replace $(\exists^\beta x^\beta)\varphi(x^\beta)$ by

$$(\exists^\alpha x^\alpha)(\exists^\beta x^\beta \geqslant x^\alpha)\varphi(x^\beta)$$

and $(\exists^\omega x^\omega)\varphi(x^\omega)$ by

$$(\exists^\alpha x^\alpha)(\exists^\omega x^\omega \geqslant x^\alpha)\varphi(x^\omega)$$

Call ψ^δ the result of applying this replacement to ψ. It turns out that (in the first–order logic of simulation frames)

$$(\exists x)w(x) \vdash \psi^\delta \leftrightarrow \psi$$

That means that ψ^δ and ψ are equivalent on all frames $\mathfrak{G}^s$ where $\mathfrak{G}$ is not empty. In ψ^δ the variables x^β and x^ω are bound by a restricted quantifier with restrictor x^α; x^α in turn is bound by $\exists^\alpha$. To see whether such formulae are valid in a frame we may restrict ourselves to assignments h of the variables in which $x^\flat$ is in the $\flat$–region for each $\flat$ and each x, and furthermore $h(x^\alpha) \geqslant h(x^\beta)$. In a final step, translate as follows

$$\begin{array}{lcl}
(\varphi \wedge \psi)^\ddagger & := & \varphi^\ddagger \wedge \psi^\ddagger \\
(\varphi \vee \psi)^\ddagger & := & \varphi^\ddagger \vee \psi^\ddagger \\
(\neg\varphi)^\ddagger & := & \neg\varphi^\ddagger \\
((\exists^\alpha x^\alpha)\varphi(x^\alpha))^\ddagger & := & (\exists x)\varphi(x^\alpha)^\ddagger \\
((\exists^\beta x^\beta \geqslant x^\alpha)\varphi(x^\beta))^\ddagger & := & \varphi(x^\beta)^\ddagger \\
((\exists^\omega x^\omega \geqslant x^\alpha)\varphi(x^\omega))^\ddagger & := & \varphi(x^\omega)^\ddagger
\end{array}$$

For atomic formulae, $\varphi^{\ddagger}$ is computed as follows:

$x^{\flat} \leqslant y^{\natural}$		$\natural \to$ w	b	t
$\flat$	w	$x \lhd y$	$x \doteq y$	$\top$
$\downarrow$	b	$x \doteq y$	$x \blacktriangleleft y$	$\bot$
	t	$\bot$	$\bot$	$\bot$

$x^{\flat} \doteq y^{\natural}$		$\natural \to$ w	b	t
$\flat$	w	$x \doteq y$	$\bot$	$\bot$
$\downarrow$	b	$\bot$	$x \doteq y$	$\bot$
	t	$\bot$	$\bot$	$\top$

Let ψ be a subformula of $\varphi^{\dagger\delta}$ for some sentence φ. Let $\mathfrak{G}$ be a 1–modal simulation frame. Then $\mathfrak{G}$ is isomorphic to $\mathfrak{F}^s$ for some bimodal frame $\mathfrak{F}$ (for example, $\mathfrak{F} := \mathfrak{G}_s$) and therefore $\mathfrak{G}$ has exactly one point in the ω–region. Suppose $\mathfrak{G} \vDash \psi[h]$. Then we may assume that $h(x^{\flat})$ is in the α–region for each $\flat$, and that $h(x^{\beta}) \prec h(x^{\alpha})$. Now put $k(x) := h(x^{\alpha})$. It is verified by induction on ψ that $\mathfrak{G} \vDash \psi[h]$ iff $\mathfrak{G}_s \vDash \psi^{\ddagger}[k]$. On the other hand, if k is given, define h as follows: $h(x^{\alpha}) := k(x)$, $h(x^{\beta})$ is the unique world u in the β–region such that $k(x) \geqslant u$, and $h(x^{\omega})$ is the unique world in the ω–region. h is uniquely determined by k, and again it is verified that $\mathfrak{G} \vDash \psi[h]$ iff $\mathfrak{G}_s \vDash \psi^{\ddagger}[k]$. In particular, for $\psi = \varphi^{\dagger\delta}$ we get $\mathfrak{G} \vDash \psi$ iff $\mathfrak{G}_s \vDash \psi^{\ddagger}$. Now we return to φ. We have $\varphi \equiv \varphi \wedge (\forall x)\omega(x). \vee .\varphi \wedge (\exists x)\neg\omega(x)$. The first formula is either equivalent to $\bot$ (**Case 1**) or to $(\forall x)\omega(x)$ (**Case 2**). **Case 1.** Put $\varphi_e := (\varphi^{\dagger\delta})^{\ddagger}$. Then $\mathfrak{G} \vDash \varphi$ iff $\mathfrak{G}_s \vDash \varphi_e$. **Case 2.** Put $\psi_e := ((\forall x)\neg(x \doteq x)) \vee (\varphi^{\dagger\delta})^{\ddagger}$.

Proposition 6.7.11. *Let $\mathfrak{X}$ be a class of simulation frames. If $\mathfrak{X}$ is elementary (Δ–elementary) so is $\mathfrak{X}_s$.*

The following properties have been shown to be invariant under simulation: $\mathfrak{Krp}$*–(Δ)–elementarity.*

Sahlqvist Logics. Likewise, by Proposition 6.6.13, if Θ is Sahlqvist, so is Θ^s. For the other direction, we need to be a little bit more careful. Assume that Λ is a monomodal logic and Sahlqvist. Then by Theorem 5.5.8 it is axiomatizable by formulae of the form $\varphi \to \psi$ where ψ is positive and φ is composed from strongly positive formulae using $\wedge$, $\vee$, and $\Diamond_j$. Now $(\varphi \to \psi)_s$ is the same as $\varphi_s \to \psi_s$. It is not hard to see that the unsimulation of a positive formula is positive, and that the unsimulation of a strongly positive formula is strongly positive. Moreover, ψ_s is composed from strongly positive formulae using $\wedge$, $\vee$ and $\Diamond_j$. So, it is Sahlqvist.

The rank as well as the special rank do not increase under simulation. This can be shown by observing a few equivalences. Namely, note that the following holds.

$$\begin{array}{lcl}
((\forall y \rhd x)\delta)^s & \equiv & (\forall y \geqslant x)(\alpha(y) \to \delta(x)^s \\
((\forall y \blacktriangleright x)\delta)^s & \equiv & (\forall \widehat{x} \geqslant x)(\forall \widehat{y} \geqslant y)(\forall y \geqslant \widehat{y})(\beta(\widehat{x}) \wedge \beta(\widehat{y}) \wedge \alpha(y). \to .\delta^s) \\
((\exists y \rhd x)\delta)^s & \equiv & (\exists y \geqslant x)(\alpha(y) \wedge \delta(y)^s) \\
((\exists y \blacktriangleright x)\delta)^s & \equiv & (\exists \widehat{x} \geqslant x)(\exists \widehat{y} \geqslant y)(\exists y \geqslant \widehat{y})(\beta(\widehat{x}) \wedge \beta(\widehat{y}) \wedge \alpha(y). \to .\delta^s) \\
(x \blacktriangleleft y) & \equiv & (\forall \widehat{x} \geqslant x)(\forall \widehat{y} \geqslant y)(\beta(\widehat{x}) \wedge \beta(\widehat{y}). \to .\widehat{x} \leqslant \widehat{y}) \\
(x \blacktriangleleft y) & \equiv & (\exists \widehat{x} \geqslant x)(\exists \widehat{y} \geqslant y)(\beta(\widehat{x}) \wedge \beta(\widehat{y}) \wedge \widehat{x} \leqslant \widehat{y})
\end{array}$$

So, we translate a restricted $\forall$ by a sequence of restricted $\forall$ and a restricted $\exists$ by a sequence of restricted $\exists$. The atomic formulae $x \lhd y$ and $x \blacktriangleleft y$ may be translated by existential formulae or universal formulae, and so neither the rank nor the special rank are changed under simulation. Thus the special rank of φ^s is equal to the special rank of φ. Conversely, let ψ be given, a monomodal Sahlqvist–formula of special rank n. Then it is likewise checked that the unsimulation ψ_s is a Sahlqvist formula of rank at most that of ψ. Moreover, ψ is deductively equivalent to a formula $\widehat{\psi}$ whose unsimulation has rank equal to that of $\widehat{\psi}$. (The construction of the unsimulated formula actually proceeds by producing $\widehat{\psi}$ and unsimulating $\widehat{\psi}$ rather than ψ.) Now, **Sim** is $\mathfrak{Df}$–elementary and of special rank 0. Hence, if $\mathbf{K}_2 \oplus \Delta$ is a Sahlqvist logic of (special) rank n, then its simulation $\mathbf{Sim} \oplus \Delta^s$ is of (special) rank n, the rank being the maximum of the ranks of **Sim** and of Δ. (Notice that the property of being Sahlqvist of rank n are here properties of logics, not of the actual axiomatization; otherwise the claim would be false.) Therefore the properties *being Sahlqvist of rank n* and *being Sahlqvist of special rank n* are invariant under simulation.

Interpolation. Finally, we take a look at interpolation. It is easier to show that interpolation is transferred from monomodal to bimodal logic using the characterization via the amalgamation property than using the simulation of formulae. However, this has the disadvantage that an interpolant is not explicitly constructed. Take a bimodal logic Θ. Assume that Θ^s has global interpolation. Let $\iota_1 : \mathfrak{B}_0 \rightarrowtail \mathfrak{B}_1$ and $\iota_2 : \mathfrak{B}_0 \rightarrowtail \mathfrak{B}_2$ be embeddings of Θ–algebras. Then $\iota_1^s : \mathfrak{B}_0^s \rightarrowtail \mathfrak{B}_1^s$ and $\iota_2^s : \mathfrak{B}_0^s \rightarrowtail \mathfrak{B}_2^s$ are embeddings of the simulations. Now, by the assumption that Θ^s has global interpolation we get a $\mathfrak{A}_3$ and embeddings $\zeta_1 : \mathfrak{B}_1^s \rightarrowtail \mathfrak{A}_3$, $\zeta_2 : \mathfrak{B}_2^s \rightarrowtail \mathfrak{A}_3$ such that $\zeta_1 \circ \iota_1^s = \zeta_2 \circ \iota_2^s$. Let $\mathfrak{B}_3 := (\mathfrak{A}_3)_s$ and put $\epsilon_1 := (\zeta_1)_s$ and $\epsilon_2 := (\zeta_2)_s$. Then

$$\begin{array}{rcl}
\epsilon_1 \circ \iota_1 & = & (\zeta_1)_s \circ (\iota_1^s)_s \\
& = & (\zeta_1 \circ \iota_1^s)_s \\
& = & (\zeta_2 \circ \iota_2^s)_s \\
& = & (\zeta_2)_s \circ (\iota_2^s)_s \\
& = & \epsilon_2 \circ \iota_2
\end{array}$$

Therefore, the variety of Θ–algebras has amalgamation. Suppose that the variety of Θ^s–algebras has superamalgamation. Then we can show that the variety of Θ–algebras has superamalgamation, too. For assume that $\epsilon_1(b_1) \leq \epsilon_2(b_2)$. Then via the identification of elements with subsets of the $\circ$–region we conclude $(\epsilon_1(b_1))^\circ \leq (\epsilon_2(b_2))^\circ$ from which also $\epsilon_1^s(b_1^\circ) \leq \epsilon_2^s(b_2^\circ)$. Thus there exists a a_0 such that $b_1^\circ \leq \zeta_1(a_0)$ and $\zeta_2(a_0) \leq b_2^\circ$. Intersecting with α we get $b_1^\circ \leq \zeta_1(a_0) \cap \alpha$ and $\zeta_2(a_0) \cap \alpha \leq b_2^\circ$. Now, put $b_0 := a_0 \cap \alpha$. Then $\zeta_1(b_0) = \zeta_1(a_0) \cap \alpha$ as well as $\zeta_2(b_0) = \zeta_2(a_0)$, from which follows that $b_1 \leq \epsilon_1(b_0)$ and $\epsilon_2(b_0) \leq b_2$.

Assume that Θ has local interpolation. Put $\vdash\, := \,\vdash_{\Theta^s}$. Assume $\varphi \vdash \psi$. Then (a) $\alpha \to \varphi \vdash \alpha \to \psi$. Moreover, we also have (b) $\beta \to \varphi \vdash \beta \to \psi$ and (c) $\omega \to \varphi \vdash \omega \to \psi$. (b) can be reformulated into (b′) $\alpha \to \blacklozenge_\beta\varphi \vdash \alpha \to \blacklozenge_\beta\psi$. The cases (a), and (b′) receive similar treatment. Take (a). We have $\varphi_s \vdash_\Theta \psi_s$. There exists by assumption on Θ an χ such that $var(\chi) \subseteq var(\varphi_s) \cap var(\psi_s)$ and $\varphi_s \vdash_\Theta \chi \vdash_\Theta \chi_s$. Then $\alpha \to (\varphi_s)^s \vdash \alpha \to \chi^s \vdash \alpha \to (\varphi_s)^s$. Now take the substitution σ as defined above. Then

$$(\varphi \dashv\vdash)\alpha \to \sigma((\varphi_s)^s) \vdash \alpha \to \sigma(\chi^s) \vdash \alpha \to \sigma((\psi_s)^s)(\dashv\vdash \psi) \,.$$

Put $\zeta := \sigma(\chi^s)$. Then we have $var(\alpha \to \zeta) \subseteq var(\varphi) \cap var(\psi)$ and $\alpha \to \varphi \vdash \alpha \to \zeta \vdash \alpha \to \psi$. Similarly, we find a formula η such that $\alpha \to \blacklozenge_\beta\varphi \vdash \alpha \to \eta \vdash \alpha \to \blacklozenge_\beta\psi$. Hence $\beta \to \varphi \vdash \beta \to \blacklozenge_\alpha\eta \vdash \beta \to \psi$. For (c) we can appeal to the fact that classical logic has interpolation to find a θ such that $\alpha \to \blacklozenge_\omega\varphi \vdash \alpha \to \blacklozenge_\omega\theta \vdash \alpha \to \blacklozenge_\omega\psi$. (For in the scope of $\blacklozenge_\omega$, φ and ψ reduce to nonmodal formulae.) Then put $\lambda := (\alpha \to \zeta) \wedge (\beta \to \blacklozenge_\alpha\eta) \wedge (\omega \to \theta)$. It follows that $\varphi \vdash \lambda \vdash \psi$. Thus Θ^s has interpolation. Exactly the same proof can be used for global interpolation (no use of the deduction theorem has been made). So, *local interpolation* and *global interpolation* have been shown to be invariant under simulation.

Theorem 6.7.12 (Simulation Theorem). *The simulation map $\Lambda \mapsto \Lambda^s$ is an isomorphism from the lattice of bimodal normal logics onto the interval $[\boldsymbol{Sim}, \mathsf{Th}\boxed{\bullet}]$ in $\mathcal{E}\,\boldsymbol{K}_1$ preserving and reflecting the following properties of logics.*

- *finite, strongly recursive, recursive axiomatizability,*
- *codimension n,*
- *tabularity,*
- *local and global decidability,*
- *local and global $\mathcal{C}$–computability ($\mathcal{C}$–hardness, $\mathcal{C}$–completeness),*
- *local and global completeness,*
- *local and global finite model property,*
- *local and global interpolation,*
- *strong and weak compactness,*
- *g–, df–, r–, d– and c–persistence,*
- *being a Sahlqvist logic of (special) rank n.*

Exercise 216. Show the remaining claims in the Simulation Theorem.

Exercise 217. Show by direct calculation that $(-)^s$ is a $\bigsqcup$–homomorphism.

Exercise 218. Using the simulation, we can simulate the fusion operator inside $\mathcal{E}\,\mathbf{K}_1$ as follows. Define $\Lambda\widehat{\otimes}\Theta := (\Lambda \otimes \Theta)^s$. By the transfer results $\Lambda\widehat{\otimes}\Theta$ has finite model property (is complete etc.) iff Λ and Θ are. Now suppose that Λ and Θ are logics with finite model property such that their join, i. e. $\Lambda \sqcup \Theta$, fails to have the finite model property. Show that $\Lambda\widehat{\otimes}\Theta$ is a logic with finite model property such that adding $(\Box p \leftrightarrow \blacksquare p)^s$ fails to have the finite model property. Clearly, $\mathbf{Sim} \oplus (\Box p \leftrightarrow \blacksquare p)^s$ has the finite model property. Thus we have shown that if finite model property is not preserved under joins, there is a Sahlqvist logic of special rank 1 which fails to preserve joins. We will later refine this result somewhat; but already it shows that very simple axioms can induce rather drastic effects if added to a logic, on the condition that such effects can occur at all.

6.8. Simulation and Transfer — Some Generalizations

We have proved the simulation and transfer theorems only for the case of two monomodal logics. This suggests two dimensions of generalization. First, we can generalize to the simulation and transfer of several logics with one operator each, and of two logics with several operators; of course, the two can be combined. Let us begin with the *fusion*. Obviously, if we can generalize the transfer theorems to the case of two logics with several operators, we have — by induction — a transfer theorem for the independent fusion of n monomodal logics. The case of infinitely many logics needs to be discussed separately. In fact, it does give rise to a number of exceptions which must be carefully noted.

Now let us study the case of two polymodal logics Λ and Θ. The constructions of an *ersatz* is straightforwardly generalized. The only difficulty we face is that the proof of *Consistency Reduction* makes use of Makinson's Theorem. Recall, namely, that in the model building procedure we want to insert certain small models at an internal set of nodes. What we therefore need is that $\Lambda \otimes \Theta$ allows for at least one finite model, if for example transfer of finite model property is desired. This is true if only Λ and Θ both have a finite model. Namely, the following construction yields frames for the fusion. Take a frame $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ for $\mathbf{K}_\kappa$ and a frame $\mathfrak{G} = \langle \mathfrak{g}, \mathbb{G}\rangle$ for $\mathbf{K}_\lambda$. We define the $\mathbf{K}_{\kappa+\lambda}$–frame $\mathfrak{F} \otimes \mathfrak{G}$ as follows. The underlying set is $f \times g$. For $\alpha < \kappa$, $\langle v_1, w_1\rangle \lhd^{\otimes}_{\alpha} \langle v_2, w_2\rangle$ iff $w_1 = w_2$ and $v_1 \lhd_\alpha v_2$. For $\beta < \lambda$ we put $\langle v_1, w_1\rangle \lhd^{\otimes}_{\kappa+\beta} \langle v_2, w_2\rangle$ iff $v_1 = v_2$ and $w_1 \lhd_\beta w_2$. Finally, $\mathbb{F}\otimes\mathbb{G}$ is the set of all unions of sets $a \otimes b = \{\langle v, w\rangle : v \in a, w \in b\}$, where $a \in \mathbb{F}$ and $b \in \mathbb{G}$. It is easy to check that this is indeed a frame. For if $\alpha < \kappa$ then $\blacklozenge_\alpha(a \otimes b) = (\blacklozenge_\alpha a) \otimes b$ and if $\beta < \lambda$ then $\blacklozenge_{\kappa+\beta}(a \otimes b) = a \otimes (\blacklozenge_\beta b)$. The frame $\mathfrak{F} \otimes \mathfrak{G}$ is called the **tensor product** of $\mathfrak{F}$ and $\mathfrak{G}$.

Lemma 6.8.1. *Let $\mathfrak{F}$ be a Λ–frame and $\mathfrak{G}$ be a Θ–frame. Then $\mathfrak{F}\otimes\mathfrak{G}$ is a $\Lambda\otimes\Theta$–frame.*

PROOF. The projections of a set a are defined by

$$\begin{array}{rcl} \llcorner a \lrcorner_1 & := & \{v : (\exists w)(\langle v, w\rangle \in a)\} \\ \llcorner a \lrcorner_2 & := & \{w : (\exists v)(\langle v, w\rangle \in a)\} \end{array}$$

Let β be a valuation on $\mathfrak{F} \otimes \mathfrak{G}$. Define a valuation $\gamma_1(p) := \{v : \llcorner \beta(p) \lrcorner_1\}$, and $\gamma_2(p) := \{w : \llcorner \beta(p) \lrcorner_2\}$. By induction it is shown that if φ uses only modalities $\Box_\alpha$, $\alpha < \kappa$, then $\llcorner \overline{\beta}(\varphi) \lrcorner_1 = \overline{\gamma_1}(\varphi)$ and if φ uses only modalities $\kappa + \beta$ then $\llcorner \overline{\beta}(\varphi) \lrcorner_2 = \overline{\gamma_2}(\varphi)$. Hence, if φ is a theorem of Λ, it is a theorem of $\mathfrak{F} \otimes \mathfrak{G}$, and if it is a theorem of Θ, then it is a theorem of $\Lambda \otimes \Theta$, under suitable identification of modalities. This proves the theorem. ◻

Let us note that if we have infinitely many logics, this method does not allow to find a finite model for their fusion. However, suppose that there is a number m such that all logics have a frame of size at most m (e. g. if they are all monomodal) then we can construct a finite model for the infinite fusion. The basic idea is that there are only finitely many choices of relations, so rather than taking the infinite product as above, we will collapse a lot of relations into a single one. Let us first observe that if Λ is a logic with infinitely many operators which has a frame $\mathfrak{F}$ of size $\leq n$, then the theory of $\mathfrak{F}$ contains many axioms of the form $\Diamond_\alpha p \leftrightarrow \Diamond_\beta p$, $\alpha, \beta < \kappa$, saying that the relations α and β are identical. It is then possible to regard $\mathsf{Th}\,\mathfrak{F}$ as a finite operator theory, and to compress $\mathfrak{F}$ into a frame containing finitely many relations. Call an **unindexed frame** a pair $\langle f, R\rangle$ where R is a *set* of binary relations over f. The map $\langle f, \langle \lhd_\alpha : \alpha < \kappa\rangle\rangle \mapsto \langle f, \{\lhd_\alpha : \alpha < \kappa\}\rangle$ is called the **compression map**. Given $n := \sharp f$ there are finitely many unindexed frames of size n, for there are at most 2^{n^2} different binary relations, and a compressed frame contains a subset of them. Notice also that if $\mathfrak{f}$ and $\mathfrak{g}$ are finite frames of size n, $\mathfrak{f}$ a frame for Λ and $\mathfrak{g}$ a frame for Θ, then there exists an unindexed frame of size n for $\Lambda \otimes \Theta$. For we can assume the underlying sets to be identical. Then we interpret $\lhd_\alpha$, $\alpha < \kappa$ by $\lhd_\alpha$ and $\lhd_{\kappa+\beta}$ by $\lhd_\beta$, and then compress. Now, for the final result, take a set of logics Λ_i, $i \in I$, such that there is a number n such that each Λ_i has a frame of size $\leq n$. Then take the unindexed tensor product. It is finite. Now uncompress the frame, assigning each modality its relation.

LEMMA 6.8.2. *Let Λ_i, $i \in I$, be an indexed family of polymodal logics. Let n be a number such that every Λ_i has a model of size $\leq n$. Then $\bigotimes_{i \in I} \Lambda_i$ has a model of size $\leq n!$.*

DEFINITION 6.8.3. *A property $\mathcal{P}$ of logics is said to be **preserved under fusion** if for each family Λ_i, $i \in I$, of consistent logics, $\bigotimes_{i \in I} \Lambda_i$ has $\mathcal{P}$ whenever each Λ_i has $\mathcal{P}$. If we can infer $\mathcal{P}$ for the factors from that of the fusion we say $\mathcal{P}$ is **reflected under fusion**. $\mathcal{P}$ **transfers under fusion** iff it is both preserved and reflected under fusion.*

THEOREM 6.8.4 (Transfer Theorem). *The following properties are transferred under fusion.*

- *finite and recursive axiomatizability, provided the indexed collection is finite,*
- *g–, df–, ti–, r–, d– and c–persistence,*
- *local and global decidability,*
- *local and global finite model property, under the condition that there exists a uniform bound for the smallest model of the factors,*
- *local and global completeness,*
- *weak and strong compactness,*
- *local interpolation, local Halldén completeness.*

Of course, at some points things can be improved. Finite axiomatizability is reflected under fusion even when I is infinite and the same holds for recursive axiomatizability. Tabularity transfers exactly when almost all factors are trivial.

Now let us turn to simulation. We can generalize in two directions. First, we can simulate an extension of $\mathbf{K}_{\kappa+\kappa}$ by using only the modalities of $\mathbf{K}_\kappa$. (Notice that we cannot simply simulate $\mathbf{K}_{\kappa+\lambda}$ for different κ and λ. But if $\kappa \neq \lambda$, it is possible to adjoin to one of the logics a number of trivial operators to make the reduction work in this case.) In that case we double up the frame and pick one modality to play the complicated role of encoding the double structure. Notice that the reduction can only be iterated a finite number of times, unlike the fusion. But we can also wrap up the modalities in one step as follows. For n modalities, f^n consists of n different copies f^i, $i < n$, plus $n-1$ extra points $*^0, \dots, *^{n-2}$. Given $x \in f$ we write x^i for the corresponding point in f^i. Then the relation $\lhd_j$ is coded among the x^j. The points $*^i$ are related by $*^j \leqslant *^k$ iff $k = j - 1$. Furthermore we have $x^j \leqslant *^k$ iff $j = k$. This serves to distinguish the x^j from the x^k, $k \neq j$. Finally, for $j \neq k$ we put $x^j \leqslant y^k$ iff $x = y$. Notice that $x^j \leqslant x^j$ iff $x \lhd_j x$. Then the formulae α_j and ω_j are defined as follows.

$$\begin{array}{lcll}
\omega(j) & := & \boxminus^{j+1}\bot \wedge \diamondminus^{i}\top & j < n-1 \\
\alpha(j) & := & \neg\omega(j+1) \wedge \diamondminus\,\omega(j) & j < n-2 \\
\alpha(n-2) & := & \diamondminus\,\omega(n-2) & \\
\alpha(n-1) & := & \bigwedge_{j<n-1} \neg\omega(j) \wedge \neg\alpha(j) &
\end{array}$$

As before, the logic of simulation structures can be axiomatized. It is the best strategy to define the operators $\boxminus_{\alpha(j)}$ and $\boxminus_{\omega(j)}$. These operators all satisfy $\mathbf{alt}_1$. Moreover, we have $\alpha(j) \to \diamondminus_{\alpha(i)}\top$ for $i \neq j$. In total, we have the following postulates.

$$\begin{array}{ll}
\alpha(i) \to \diamondminus\,\alpha(j) & i \neq j \\
\alpha(i) \wedge \diamondminus_{\alpha(j)} p. \to .\ \boxminus_{\alpha(j)}\, p & i \neq j \\
\alpha(i) \to \diamondminus\,\omega(i) & \\
\alpha(i) \to \neg\diamondminus\,\omega(j) & i \neq j \\
\alpha(i) \wedge \diamondminus_{\alpha(j)} p. \to \diamondminus_{\alpha(k)} \diamondminus_{\alpha(j)} p & i \neq k,\ j \neq k \\
\diamondminus^{\leq 4}\omega(j) \wedge p. \to .\ \boxminus^{\leq 4}(\omega(j) \to p) &
\end{array}$$

for $j \neq k$. This defines $\mathbf{Sim}(n)$. The special postulates ensure that we can move from x^j to x^k in one step. The price to be paid for this is that the special rank of $\mathbf{Sim}(n)$ is 1 for $n > 0$. There is a simulation which avoids this (see the exercises), but it has other disadvantages. We put $\mathbf{Inc}(n) := \mathbf{K.alt}_1 \oplus \boxminus^n \bot$, which is the same as $(\mathbf{K}_n \oplus \bot)^s$.

Theorem 6.8.5 (Simulation Theorem). *The simulation map $\Lambda \mapsto \Lambda^s$ is an isomorphism from the lattice of normal n–modal logics onto the interval $[\boldsymbol{Sim}(n), \boldsymbol{Inc}(n)]$ preserving and reflecting the following properties.*

- *finite, strongly recursive and recursive axiomatizability,*
- *tabularity,*
- *local and global decidability,*
- *local and global $\mathcal{C}$–computability ($\mathcal{C}$–hardness, $\mathcal{C}$–completeness),*
- *local and global completeness,*
- *local and global finite model property,*
- *local and global interpolation,*
- *strong and weak compactness,*
- *g–, df–, r–, d–, and c–persistence,*
- *being a Sahlqvist logic of (special) rank n,*
- *being weakly transitive,*
- *being of bounded alternativity.*

To close this chapter we will generalize the conservativity result of this chapter to consequence relations. Apart from being interesting in its own right, this will also sharpen our understanding of the methods used previously. As before, the situation we look at is the fusion of two monomodal consequence relations and the simulation of a bimodal consequence relation by a monomodal consequence relation. The results are general, however. The simplification is made only to avoid baroque notation and to be able to concentrate on the essentials of the method. As is to be expected, the proofs of some theorems are more involved. In particular, to show that the fusion of two consequence relations is conservative requires other methods because an analogue of Makinson's Theorem is missing. (See Section 3.9.) For the purpose of the next definition recall that if R is a set of rules, $\vdash^R$ denotes the least modal consequence relation containing R. We have the two translations $\tau_\Box$ and $\tau_\blacksquare$ from $\mathcal{P}_1$ into $\mathcal{P}_2$, defined in Section 6.2. $\tau_\Box$ replaces $\boxminus$ by $\Box$ and $\tau_\blacksquare$ replaces $\boxminus$ by $\blacksquare$.

Definition 6.8.6. *Let $\vdash_1 \; = \; \vdash^R$ and $\vdash_2 \; = \; \vdash^S$ be two monomodal consequence relations. Put $T := \tau_\Box[R] \cup \tau_\blacksquare[S]$. Then $\vdash_1 \otimes \vdash_2$ is a bimodal consequence relation defined by $\vdash_1 \otimes \vdash_2 := \vdash^T$. $\vdash_1 \otimes \vdash_2$ is called the **fusion** of $\vdash_1$ and $\vdash_2$.*

Let $\vdash$ be a bimodal consequence relation. Define two consequence relations, $(\vdash)_\Box$ and $(\vdash)_\blacksquare$ by

$$\begin{aligned} \vdash_\Box &:= \tau_\Box^{-1}[\vdash] \\ \vdash_\blacksquare &:= \tau_\blacksquare^{-1}[\vdash] \end{aligned}$$

We call $\vdash_\Box$ and $\vdash_\blacksquare$ the **white** and **black reduct** of $\vdash$, respectively.

Proposition 6.8.7. *Let $\vdash_1$ and $\vdash_2$ be monomodal consequence relations. Then $\vdash_1 \otimes \vdash_2$ is the least consequence relation such that $(\vdash_1 \otimes \vdash_2)_\Box = \vdash_1$ and $(\vdash_1 \otimes \vdash_2)_\blacksquare = \vdash_2$.*

Our first main result is the analogue of Theorem 6.2.3.

Theorem 6.8.8. *Let either $\vdash_1$ be inconsistent or else $\vdash_2$ be consistent. Then*

$$(\vdash_1 \otimes \vdash_2)_\Box = \vdash_1$$

First, let us deal with the easy case. If $\vdash_1$ is inconsistent then so is $\vdash_1 \otimes \vdash_2$, and the fusion is clearly conservative. So, let us from now on assume that $\vdash_1$ is consistent. In that case, $\vdash_2$ is also assumed to be consistent. Now suppose $\mathfrak{M} = \langle \mathfrak{A}, D\rangle$ and $\mathfrak{N} = \langle \mathfrak{B}, E\rangle$ are matrices such that $\vdash_1 \subseteq \vdash_\mathfrak{M}$ and $\vdash_2 \subseteq \vdash_\mathfrak{N}$. Then construct a bimodal matrix $\mathfrak{M} \otimes \mathfrak{N}$ as follows. It is based on the tensor product of boolean algebras. This tensor product is formed from so–called **basic tensors**. These are pairs $\langle x, y\rangle \in A \times B$, denoted by $x \otimes y$. The (boolean) tensor algebra is the boolean algebra generated by the basic tensors with respect to the following rules.

$$\begin{array}{rcl}
(x_1 \otimes y_1) \cap (x_2 \otimes y_2) & := & (x_1 \cap x_2) \otimes (y_1 \cap y_2) \\
x \otimes 0 & := & 0 \otimes 0 \\
0 \otimes y & := & 0 \otimes 0 \\
(x_1 \otimes y) \cup (x_2 \otimes y) & := & (x_1 \cup x_2) \otimes y \\
(x \otimes y_1) \cup (x \otimes y_2) & := & x \otimes (y_1 \cup y_2) \\
-(x \otimes y) & := & (-x) \otimes y \cup x \otimes (-y) \cup (-x) \otimes (-y)
\end{array}$$

It can be shown that an element of $\mathfrak{A} \otimes \mathfrak{B}$ is of the form $\bigcup_{i<n} x_i \otimes y_i$, where $x_i > 0$ as well as $y_i > 0$ for all $i < n$. (Moreover, if $n = 0$, then this disjunction denotes the tensor $0 \otimes 0$.) For a basic tensor $x \otimes y = 0 \otimes 0$ iff $x = 0$ or $y = 0$. This defines the **tensor product** of boolean algebras.

Lemma 6.8.9. *The following holds.*

(A.) $x_1 \otimes y_1 = x_2 \otimes y_2$ *iff*
  (1.) $0 \in \{x_1, x_2\}$ *and* $0 \in \{y_1, y_2\}$ *or*
  (2.) $x_1 = x_2$ *and* $y_1 = y_2$.

(B.) $x_1 \otimes y_1 \leq x_2 \otimes y_2$ *iff*
  (1.) $x_1 = 0$ *or*
  (2.) $y_1 = 0$ *or*
  (3.) x_1, x_2, y_1 *and* y_2 *are all distinct from 0 and* $x_1 \leq x_2$ *and* $y_1 \leq y_2$.

Proof. Clearly, (A.) follows from the construction. Therefore we need to show (B.). If $x_1 \leq x_2$ and $y_1 \leq y_2$, then

$$(x_1 \otimes y_1) \cap (x_2 \otimes y_2) = (x_1 \cap x_2) \otimes (y_1 \cap y_2) = x_1 \otimes y_1 \ .$$

So, $x_1 \otimes y_1 \leq x_2 \otimes y_2$. Now let $x_1 = 0$ or $y_1 = 0$. Then $x_1 \otimes y_1 = 0 \otimes 0 \leq x_2 \otimes y_2$. Let all elements be distinct from zero. Then $x_1 \otimes y_1 \leq x_2 \otimes y_2$ implies $(x_1 \otimes y_1) \cap (x_2 \otimes y_2) =$

$x_1 \otimes y_1$. So, $(x_1 \cap x_2) \otimes (y_1 \cap y_2) = x_1 \otimes y_1$. Now, $x_1 \neq 0$ and $y_1 \neq 0$, hence $x_1 \cap x_2 = x_1$ and $y_1 \cap y_2 = y_1$. This implies $x_2 \neq 0$ and $y_2 \neq 0$, and moreover $x_1 \leq x_2$ and $y_1 \leq y_2$. ◻

The bottom element is $0 \otimes 0$, and the top element is $1 \otimes 1$. The action of $\Box$ and $\blacksquare$ is defined via their duals, $\Diamond$ and $\blacklozenge$, in the following way.

$$\begin{array}{lcl} \Diamond(\bigcup_{i<n} x_i \otimes y_i) & := & \bigcup_{i<n} \Diamond(x_i \otimes y_i) \\ \Diamond(x \otimes y) & := & (\Diamond x) \otimes y \\ \blacklozenge(\bigcup_{i<n} x_i \otimes y_i) & := & \bigcup_{i<n} \blacklozenge(x_i \otimes y_i) \\ \blacklozenge(x \otimes y) & := & x \otimes (\Diamond y) \end{array}$$

This defines now the **fusion** of two monomodal algebras. Finally, let $D \otimes E$ denote the least filter containing the elements $x \otimes y$, $x \in D$ and $y \in E$. It can be shown that this is likewise the filter generated by $x \otimes 1$, $x \in D$, and $1 \otimes y$, $y \in E$. Finally, put

$$\mathfrak{M} \otimes \mathfrak{N} := \langle \mathfrak{A} \otimes \mathfrak{B}, D \otimes E \rangle$$

The first fact to note is

Lemma 6.8.10. *Assume that* $0 \notin D$ *and* $0 \notin E$. *Then* $x \otimes y \in D \otimes E$ *iff* $x \in D$ *and* $y \in E$.

Proof. From right to left holds by definition of $D \otimes E$. From left to right, assume that $x \otimes y \in D \otimes E$. Then there is $x' \in D$ and $y' \in E$ such that $x \otimes y \geq x' \otimes y'$. Since $x' > 0$ and $y' > 0$ by assumption on D and E, $x \geq x'$ and $y \geq y'$ from which $x \in D$ and $y \in E$, since D and E are filters, and $x \otimes y \in D \otimes E$, by definition of $D \otimes E$. ◻

Lemma 6.8.11. *The map* $x \mapsto x \otimes 1$ *is an in injective homomorphism from* $\mathfrak{A}$ *into* $\mathfrak{A} \otimes \mathfrak{B} \restriction \{1, -, \cap, \Box\}$. *Dually, the map* $y \mapsto 1 \otimes y$ *is an injective homomorphism from* $\mathfrak{B}$ *into* $\mathfrak{A} \otimes \mathfrak{B} \restriction \{1, -, \cap, \blacksquare\}$.

This is left as an exercise.

Lemma 6.8.12. *Assume that* $\mathfrak{N} = \langle \mathfrak{B}, E \rangle$ *is a monomodal matrix and* $0 \notin E$. *Then* $(\vdash_{\mathfrak{M} \otimes \mathfrak{N}})_\Box \subseteq \vdash_{\mathfrak{M}}$.

Proof. Assume that $\mathfrak{M} = \langle \mathfrak{A}, D \rangle$. Let $\rho = \langle \Delta, \varphi \rangle \notin \vdash_{\mathfrak{M}}$. Then $0 \notin D$. We want to show that $\tau_\Box[\rho] := \langle \tau_\Box[\Delta], \tau_\Box(\varphi) \rangle \notin \vdash_{\mathfrak{M} \otimes \mathfrak{N}}$. To that end, let β be a valuation such that $\overline{\beta}[\Delta] \subseteq D$ but $\overline{\beta}(\varphi) \notin D$. Then $\gamma(p) := \beta(p) \otimes 1$ is a valuation into $\mathfrak{M} \otimes \mathfrak{N}$. Moreover, by Lemma 6.8.11, for any formula $\chi \in \mathcal{P}_1$, $\overline{\gamma}(\tau_\Box(\chi)) = \overline{\beta}(\chi) \otimes 1$. It follows with Lemma 6.8.10 that $\overline{\gamma}(\tau_\Box(\chi)) \in D \otimes E$ iff $\overline{\beta}(\chi) \in D$. Hence, $\overline{\gamma}[\tau_\Box[\Delta]] \subseteq D \otimes E$, and $\overline{\gamma}(\tau_\Box(\varphi)) \notin D \otimes E$. So, $\tau_\Box[\rho] \notin \vdash_{\mathfrak{M} \otimes \mathfrak{N}}$. ◻

Recall from Section 3.5 the following notation. Given a modal algebra $\mathfrak{B}$, $\mathfrak{B}_z$ denotes the algebra of elements $\leq z$. This is called the relativization of $\mathfrak{B}$.

Lemma 6.8.13. *The map* $r_z : \mathfrak{x} \to \mathfrak{x} \cap (1 \otimes z)$ *is a homomorphism with respect to the operations* 1, $-$, $\cap$, $\Box$.

Proof. Let $\mathfrak{C} := \mathfrak{A}\otimes\mathfrak{B}_z$. We have $1^{\mathfrak{A}\otimes\mathfrak{B}} = 1\otimes 1$, and $1^{\mathfrak{C}} = 1\otimes z = (1\otimes 1)\cap(1\otimes z)$. It is easy to check that r_z commutes with $\cap$ and $\cup$. The latter is useful in order to derive that $r_z(-\mathfrak{x}) = -r_z(\mathfrak{x})$. Namely, we can reduce this now to the case where $\mathfrak{x} = x\otimes y$. In that case,

$$\begin{aligned} r_z(-(x\otimes y)) &= r_z((-x)\otimes y \cup x\otimes(-y)\cup(-x)\otimes(-y)) \\ &= (-x)\otimes(y\cap z)\cup x\otimes((-y)\cap z)\cup(-x)\otimes(-y)\cap z \\ &= -^{\mathfrak{C}}(x\otimes(y\cap z)) \\ &= -^{\mathfrak{C}} r_z(x\otimes y) \end{aligned}$$

Finally,

$$\begin{aligned} r_z(\Diamond(x\otimes y)) &= r_z((\blacklozenge x)\otimes y) \\ &= (\blacklozenge x)\otimes(y\cap z) \\ &= \Diamond(x\otimes(y\cap z)) \\ &= \Diamond(r_z(x\otimes y)) \end{aligned}$$

∎

Lemma 6.8.14. $(\vdash_{\mathfrak{M}\otimes\mathfrak{N}})_\Box \supseteq \vdash_{\mathfrak{M}}$.

Proof. Let $\rho = \langle\Delta,\varphi\rangle$ be a rule of monomodal logic. Assume that $\tau_\Box[\rho] \notin \vdash_{\mathfrak{M}\otimes\mathfrak{N}}$ where $\langle\mathfrak{A}, D\rangle$ and $\mathfrak{N} = \langle\mathfrak{B}, E\rangle$. We have to show that $\rho \notin \vdash_{\mathfrak{M}}$. To that end, note that there is a valuation γ into $\mathfrak{M}\otimes\mathfrak{N}$ such that $\overline{\gamma}[\Delta] \subseteq D\otimes E$ but $\overline{\gamma}(\varphi) \notin D\otimes E$. Let the variables occurring in Δ and φ be p_i, $i < k$. We may assume that there exist elements c_j, $j < m$, of $\mathfrak{B}$ such that $c_j > 0$ for all $j < m$, $c_i \cap c_j = 0$ iff $i = j$ and for certain $a^i_j \in A$,

$$\gamma(p_i) = \bigcup_{j<m} a^i_j \otimes c_j$$

(Of course, some or all of the a^i_j may be zero.) Now, put

$$\beta_j(p_i) := a^i_j$$

Then $\gamma(p_i)\cap c_j = \beta_j(p_i)\otimes c_j$. It follows that $\overline{\gamma}(\tau_\Box(\chi))\cap c_j = \overline{\beta}(\chi)\otimes c_j$ for every $\chi \in \Delta\cup\{\varphi\}$. Finally, observe that $r_{c_j}[D\otimes E] = D\otimes E_{c_j}$, where E_{c_j} is the filter generated by $y\cap c_j$, $y\in E$. Notice that it may well happen that $E_{c_j} = B$, for example if $-c_j \in E$. Now, for $\chi\in\Delta\cup\{\varphi\}$, $\overline{\gamma}(\tau_\Box(\chi))\cap c_j = \overline{\beta_j}(\chi)\otimes c_j \in D\otimes E_{c_j}$ iff $\overline{\beta_j}(\chi)\in D$. Now, $\overline{\gamma}[\tau_\Box[\Delta]] \subseteq D\otimes E$ implies that $\overline{\beta_j}[\Delta]\subseteq D$ for all $j < m$ and $\overline{\gamma}(\tau_\Box(\varphi)) \notin D\otimes E$ implies that for some $j < m$, $\overline{\beta_j}(\chi) \notin D$. Hence $\rho \notin \vdash_{\mathfrak{M}}$. ∎

Corollary 6.8.15. *Let $\mathfrak{N} = \langle\mathfrak{B}, E\rangle$ be a matrix such that $0 \notin E$. Then* $(\vdash_{\mathfrak{M}\otimes\mathfrak{N}})_\Box = \vdash_{\mathfrak{N}}$.

Corollary 6.8.16. *Let $\mathfrak{M}$ be a matrix for $\vdash_1$ and $\mathfrak{N}$ a matrix for $\vdash_2$. Then $\mathfrak{M}\otimes\mathfrak{N}$ is a matrix for $\vdash_1 \otimes \vdash_2$.*

We are now in a position to prove Theorem 6.8.8. Assume that $\vdash_1$ is consistent. Then $\vdash_2$ is consistent, according to the assumptions of the theorem. Hence there exists a matrix $\mathfrak{N} = \langle \mathfrak{B}, E\rangle$ such that $\vdash_2 \supseteq \vdash_{\mathfrak{N}}$. Now assume that $\rho \notin \vdash_1$. Then there exists a matrix $\mathfrak{N}$ such that $\rho \notin \vdash_{\mathfrak{N}}$. Then $\tau_\Box[\rho] \notin \vdash_{\mathfrak{M}\otimes\mathfrak{N}}$. Thus, $\tau_\Box[\rho] \notin \vdash_1 \otimes \vdash_2$, since $\mathfrak{M} \otimes \mathfrak{N}$ is a matrix for the fusion. It follows that $\rho \notin (\vdash_1 \otimes \vdash_2)_\Box$. Hence $(\vdash_1 \otimes \vdash_2)_\Box \subseteq \vdash_1$. The converse inclusion is straightforward from the definition. The theorem is proved. As an immediate consequence we obtain from it the fact that the fusion of logics is directly connected with the related consequence relations.

Theorem 6.8.17. *Let Λ and Θ be monomodal logics. Then*

$$\vdash_{\Lambda\otimes\Theta} = \vdash_\Lambda \otimes \vdash_\Theta \qquad \Vdash_{\Lambda\otimes\Theta} = \Vdash_\Lambda \otimes \Vdash_\Theta$$

Exercise 219. Prove Theorem 6.8.17.

Exercise 220. Is Halldén–completeness reflected under simulations?

Exercise 221. Propose a different simulation by putting $x^j \leqslant x^k$ iff $j = k$ (and $x \lhd_j x$) or $k = j+1 \pmod{n}$). Show that in this simulation the special rank of the simulating logic is not increased even in the case it is zero.

Exercise 222. Discuss the tradeoffs of the various simulations in terms of the alternativity bounds versus transitivity bounds.

Exercise 223. Show that tabularity transfers under fusion exactly when all but one factor only have frames of size 1.

Exercise 224. Recall that a logic is α–compact if for every consistent set based on $< \alpha$ variables there is a model based on a frame for that logic. Show that for finite α, Λ is α–compact only if $\Lambda \otimes \mathbf{K}$ is 1–compact.

Exercise 225. Show that a logic is 1–compact iff it has at least one Kripke–frame.

Exercise 226. Call a property $\mathcal{P}$ **operator compact** if it holds of an infinite fusion of logics iff it holds of all finite sub–fusions. That is to say, $\bigotimes_{i\in I} \Lambda_i$ has $\mathcal{P}$ iff $\bigotimes_{i\in J} \Lambda_i$ has $\mathcal{P}$ for all finite $J \subseteq I$. Find out which properties are operator compact.

CHAPTER 7

Lattices of Modal Logics

7.1. The Relevance of Studying Lattices of Logics

We have seen in an earlier chapter that the lattices of (normal) modal logics are — at least from a lattice theoretic point of view — quite well–behaved. However, a typical fact of lattice theory is that the abstract properties of a lattice give very little insight into the actual structure of a lattice. We always need a lot of special information about it. We will see that there is no hope of deducing strong results about logics using these abstract properties; however, there is a mixture between abstract theory and concrete work with modal logic that yields surprisingly deep insights into both the structure of the lattices and the properties of the logics. The main tool will be that of a *splitting*, developed within logic by V. A. Jankov [**108, 109**] for intuitionistic logic and in modal logic especially by Wim Blok [**25**] and Wolfgang Rautenberg [**168, 170**]. These authors have also paid explicit attention to the structure of lattices in their work. Elsewhere, lattice theory had only a minor role to play. The motivation for introducing the concept of a splitting has been solely the quest for studying the lattice of normal monomodal logics as an object. However, it soon appeared that there are deep connections between splitting results and intrinsic properties of logics. Here we will develop the theory with hindsight, paying attention to the interplay between the structure of the lattices of extensions of a logic and properties of the logics in that lattice.

The objects of study are extensions of the logic $\mathbf{K}_\kappa$ (or extensions thereof), the minimal logic for Kripke–frames with κ operators. Most results hold only if κ is **finite**, but they will be explicitly marked — as we have done before on similar occasions. $\mathcal{E}\,\mathbf{K}_\kappa$ is the image of $\langle \wp(\mathcal{P}_\kappa), \bigcap, \bigcup \rangle$ under the map $\alpha : X \mapsto \mathbf{K}_\kappa \oplus X$. This map commutes with infinite joins but generally not with meets. (For example, let $X_1 := \{p\}$ and $X_2 := \{q\}$ with $p \neq q$. Then $X_1 \cap X_2 = \varnothing$ but obviously $\mathbf{K} \oplus X_1 = \mathbf{K} \oplus X_2 = \mathbf{K} \oplus \bot$.) The structure of the lattice is fully determined by α. However, since $\wp(\mathcal{P}_\kappa)$ is uncountable, we must restrict ourselves to finite sets, or occasionally also recursive or recursively enumerable sets. These restrictions occasionally cause complications; the set of finitely axiomatizable logics is not closed under intersection, except for weakly transitive logics, and generally not closed under infinitary unions. The set of (strongly) recursively axiomatizable logics are closed under unions and meets, but closed under infinitary unions only under special circumstances. Mostly, we will consider finitely axiomatizable logics. Even here the situation is not so favourable. We will show in Chapter 9 that for given (finite!) sets X, Y it is undecidable whether $\mathbf{K}_\kappa \oplus X = \mathbf{K}_\kappa \oplus Y$. So, not too much should be expected from this approach.

In general we will be concerned with decidable subsets of $\mathcal{E}\,\mathbf{K}_\kappa$. A subset $E \subseteq \mathcal{E}\,\mathbf{K}_\kappa$ is called **decidable** if for given finite X, the problem '$\alpha(X) \in E$' is decidable. α is a closure operator on $\mathcal{P}_\kappa$. A closed element is of the form $\alpha(X)$. Take two elements, $\alpha(X)$ and $\alpha(Y)$. Suppose that $\alpha(X)$ is decidable and Y finite. Then it is decidable whether or not $\alpha(Y) \subseteq \alpha(X)$. Namely, we only have to test all elements of Y whether they are in $\alpha(X)$. And since '$\varphi \in \alpha(X)$' is decidable, so is therefore '$\alpha(Y) \subseteq \alpha(X)$'.

Proposition 7.1.1. ($\kappa < \aleph_1$.) *$\mathbf{K}_\kappa \oplus X$ is decidable iff $\{\Lambda : \Lambda \subseteq \mathbf{K}_\kappa \oplus X\}$ is decidable iff for all finite Y it is decidable whether or not $\mathbf{K}_\kappa \oplus Y \subseteq \mathbf{K}_\kappa \oplus X$.*

We remark here that to say that $\alpha(X)$ decidable is not the same as to say that $\{\alpha(X)\}$ is decidable. Now take the lattice $\mathcal{E}\,\mathbf{S4.3}$. We will see later that all extensions are finitely axiomatizable and decidable. Hence, we can decide for two logics $\mathbf{S4.3} \oplus X$ and $\mathbf{S4.3} \oplus Y$, where X and Y are finite, whether $\mathbf{S4.3} \oplus X \subseteq \mathbf{S4.3} \oplus Y$ and hence whether $\mathbf{S4.3} \oplus X = \mathbf{S4.3} \oplus Y$. The same holds for the lattices of extensions of $\mathbf{K.alt}_1$ and $\mathbf{K5}$. In general, however, not only are there non–finitely axiomatizable logics, there are also finitely axiomatizable, undecidable logics. In general, therefore, intrinsic properties of logical systems and properties of logics in the lattice $\mathcal{E}\,\mathbf{K}_\kappa$ are unrelated, with some important exceptions. One such exception are *splitting pairs*. We say that $\langle \mathbf{K}_\kappa \oplus X, \mathbf{K}_\kappa \oplus Z\rangle$ is a *splitting pair* in the lattice $\mathcal{E}\,\mathbf{K}_\kappa$ if the lattice is the disjoint union of the principal ideal generated by $\mathbf{K}_\kappa \oplus X$, and the principal filter generated by $\mathbf{K}_\kappa \oplus Z$. Suppose that $\langle \mathbf{K}_\kappa \oplus X, \mathbf{K}_\kappa \oplus Z\rangle$ is a splitting pair and both $\mathbf{K}_\kappa \oplus X$ and $\mathbf{K}_\kappa \oplus Z$ are decidable. Then for each finite set Y, '$\mathbf{K}_\kappa \oplus Y = \mathbf{K}_\kappa \oplus X$' is decidable. For let Y be given; and let Y be finite. Then, by the decidability of '$\mathbf{K}_\kappa \oplus Z$', for every $\varphi \in Y$, '$\mathbf{K}_\kappa \oplus \varphi \subseteq \mathbf{K}_\kappa \oplus Z$' is decidable. Hence '$\mathbf{K}_\kappa \oplus Y \subseteq \mathbf{K}_\kappa \oplus Z$' is decidable. Moreover, '$\mathbf{K}_\kappa \oplus Y \not\supseteq \mathbf{K}_\kappa \oplus Z$' is equivalent to '$\mathbf{K}_\kappa \oplus Y \subseteq \mathbf{K}_\kappa \oplus X$'. The latter is decidable. So, splitting pairs are ideal tools for getting a grip on the structure of the lattice. Yet, it turns out that such pairs are quite rare. In general, to have sufficiently many such pairs the base logic must be strengthened to include at least an axiom of weak transitivity. There are many interesting logics that are weakly transitive, especially in the case of a single operator. But naturally arising logics in several operators will most likely not be weakly transitive. (However, one can always add a universal modality to make a logic weakly transitive.) This is on the one hand a rather negative fact, but on the other hand even the absence of splittings reveals something about the structure of the lattice.

7.2. Splittings and other Lattice Concepts

Let $\langle L, \top, \bot, \sqcap, \sqcup\rangle$ be a bounded lattice. An element x is called $\sqcap$**–irreducible** if $x \neq \top$ and for every pair y, z of elements such that $y \sqcap z = x$ either $y = x$ or $z = x$. Dually, we say that x is $\sqcup$**–irreducible** if $x \neq \bot$ and for every pair of elements y and z such that $x = y \sqcup z$ we have either $x = y$ or $x = z$. An element x is called $\sqcap$**–prime** or **meet prime** ($\sqcup$**–prime** or **join prime**) if $x \neq \top$ ($x \neq \bot$) and for every pair y and z of elements such that $y \sqcap z \leq x$ ($y \sqcup z \geq x$) we have either $y \leq x$ or $z \leq x$ (either $y \geq x$ or $z \geq x$). The following is an easy fact about lattices.

Proposition 7.2.1. *Let $\mathfrak{L} = \langle L, \top, \bot, \sqcap, \sqcup\rangle$ be a bounded lattice and $x \in L$. If x is $\sqcap$–prime, x is $\sqcap$–irreducible. If x $\sqcup$–prime it is also $\sqcup$–irreducible. Suppose that $\mathfrak{L}$ is distributive. Then x is $\sqcap$–irreducible iff x is $\sqcap$–prime and x is $\sqcup$–irreducible iff x is $\sqcup$–prime.*

In finite lattices meet–irreducibility is equivalent to having exactly one upper cover, and join–irreducibility is equivalent to having exactly one lower cover. Now, in a finite lattice, every element is the meet of meet–irreducible elements, and every element is the join of join–irreducible elements. This follows from the fact that for every pair of elements $x < y$ there exists a maximal element p such that $p \geq x$ but $p \ngeq y$. p must be meet–irreducible, for every element strictly larger than p is above y, and if $p = u \sqcap v$ for some $u, v > p$ then $u, v \geq y$ and hence $p = u \sqcap v \geq y$, a contradiction. The reader may now check that any finite distributive lattice is isomorphic to the lattice of upward closed sets of meet–irreducibles, and — dually — also isomorphic to the lattice of downward closed sets of join–irreducibles. This is the general idea behind the topological representation that we will develop later.

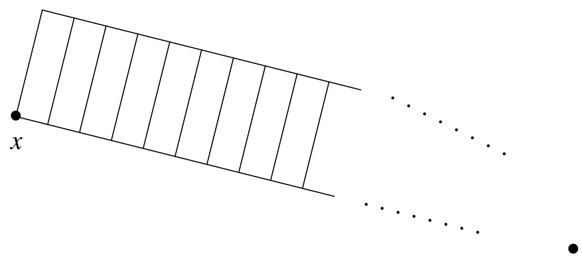
In a complete lattice we can define infinitary versions of these notions as well. Call x $\bigsqcap$**–irreducible** or **strictly meet–irreducible** if whenever $x = \bigsqcap\langle y_i : i \in I\rangle$ we have $x = y_i$ for some $i \in I$. Call an element x $\bigsqcap$**–prime** or **strictly meet–prime** if $x \geq \bigsqcap\langle y_i : i \in I\rangle$ implies $x \geq y_i$ for some $i \in I$. And analogously $\bigsqcup$**–irreducibility** or **strict join–irreducibility** and $\bigsqcup$**–primeness** or **strict join–primeness** are defined. Again, primeness is stronger than irreducibility. A set $Y = \langle y_i : i \in I\rangle$ is called a **subreduction** of x if $\bigsqcap Y \leq x$ while $y_i \nleq x$ for all $i \in I$.

Proposition 7.2.2. *Let $\mathfrak{L}$ be a complete and distributive lattice. If $\mathfrak{L}$ is upper continuous then every $\bigsqcup$–irreducible element is $\bigsqcup$–prime; if $\mathfrak{L}$ is lower continuous then every $\bigsqcap$–irreducible element is also $\bigsqcap$–prime.*

Proof. Surely, the two statements are dual to each other, so let us prove only the first. Let x be $\bigsqcup$–irreducible and let $x \leq \bigsqcup\langle y_i : i \in I\rangle$. Then $x = x \sqcap \bigsqcup\langle y_i : i \in I\rangle$. Since $x \sqcap \bigsqcup\langle y_i : i \in I\rangle = \bigsqcup\langle x \sqcap y_i : i \in I\rangle$ we have $x = \bigsqcup\langle x \sqcap y_i : i \in I\rangle$, so by the irreducibility of x, $x = x \sqcap y_i$ for some $i \in I$. Thus for some $i \in I$, $x \leq y_i$, as required. ∎

The converse of these statements fails to hold. Just consider the lattice $1 + \mathbb{Z} \times \mathbb{Z} + 1$, obtained by adjoining to the product of the linear lattice $\langle \mathbb{Z}, min, max\rangle$

FIGURE 7.1.



with itself a new bottom and a new top element. This lattice is not continuous, but distributive. No element is even simply meet– or join–irreducible. So all strictly join–irreducible elements are also strictly join–prime. Lattices of modal logics are upper continuous, so strictly join–irreducible logics are also strictly join–prime. To give the reader an impression of lattices which are distributive but fail one of the large distributive laws, consider the lattice in Figure 7.1. The bottom element is the intersection both of the lower descending chain and the upper descending chain. Hence x is $\bigsqcap$–irreducible but not $\bigsqcap$–prime. Other examples are the lattices of open subsets of the topological spaces $\mathbb{R}^n$, and the lattice $\mathcal{E}\,\mathbf{K.alt}_1$ (see Chapter 7.6). The following example illustrates that $\mathcal{E}\,\mathbf{K.alt}_1$ is not lower continuous. The theory of the one–point reflexive frame, $\boxed{\circ}$, contains the theory of the infinite frame $\langle \omega, \prec \rangle$ where $i \prec j$ iff $j = i + 1$. For the map $n \mapsto 0$ is a contraction of $\langle \omega, \prec \rangle$ onto $\boxed{\circ}$. The theory of $\langle \omega, \prec \rangle$ is the intersection of the theories of $\langle n, \prec \rangle$, by the fact that the n–transits of the respective roots are isomorphic. So $\{\mathsf{Th}\,\langle n, \prec \rangle : n \in \omega\}$ is a subreduction of the theory of $\boxed{\circ}$. Strongly irreducible elements have a nice characterization from a lattice theoretic point of view. Given an element $x \in L$ we say that y is an **upper cover** of x if $x < y$ but for no element z we have $x < z < y$. x is called a **lower cover** of y if y is an upper cover of x. We write $x \prec y$ to say that x is a lower cover of y (or that y is an upper cover of x). The following is not hard to see.

PROPOSITION 7.2.3 (Blok). *In a complete lattice, x is strongly join–irreducible iff it has an upper cover y and for all $z > x$ we have $z \geq y$. Dually, x is strongly meet–irreducible iff it has a lower cover y and for all $z < x$ we have $z \leq y$.*

For finitely axiomatizable logics we can show they are never the limit of an ascending sequence of logics (see [**23**]).

THEOREM 7.2.4. *Let Λ be finitely axiomatizable. Then for every logic $\Theta \subsetneq \Lambda$ there exists a lower cover Θ^o of Λ such that $\Theta \subseteq \Theta^o$.*

Proof. Consider the following property $\mathcal{P}$. X has $\mathcal{P}$ iff $\Theta \oplus X \not\supseteq \Lambda$. This is a property of finite character. For $\Lambda = \Theta \oplus E$ for some finite set E. Thus if $\Theta \oplus X \supseteq \Theta \oplus E$ then there exists a finite subset $X_0 \subseteq X$ such that $E \subseteq \Theta \oplus X_0$. Therefore $\mathcal{P}$ has finite character. So $\varnothing$ is contained in a maximal set X^o, by Tukey's Lemma. Put $\Theta^o := \Theta \oplus X^o$. By definition of X^o, Θ^o is a lower cover of Λ. For consider an extension $\Theta^o \oplus Y$. If $\Theta^o \oplus Y \not\supseteq \Lambda$ then $X^o \cup Y$ must have $\mathcal{P}$, which by the maximality of X^o means that $Y \subseteq X^o$. Thus $\Theta^o \oplus Y = \Theta^o$, as required. ◻

Similarly the following theorem is proved. For the purpose of stating the theorem, a pair x/y is called a **quotient** if $x \geq y$. The quotient x/y is **prime** if $x > y$ and for no z, $x > z > x$.

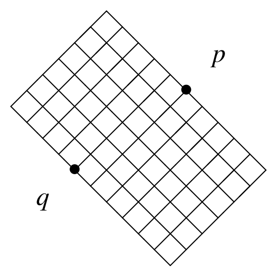
Theorem 7.2.5. *Let $\Lambda_o < \Lambda^o$. Then the interval $[\Lambda_o, \Lambda^o]$ contains a prime quotient.*

Proof. It is not hard to see that the interval contains a logic $\Theta \neq \Theta_o$ which is finitely axiomatizable over Λ_o. Now show as above that there exists a lower cover Θ_o for Θ. ◻

The lattice of rational (or real) numbers in the closed interval $[0, 1]$, with *min* and *max* as operations, contains no prime quotients. So this is a nontrivial property of lattices of logics. Logics which are not finitely axiomatizable may also have lower covers. Examples are the logics $\iota(M)$ of Chapter 2.6, for infinite M. The covers are not necessarily unique. Moreover, even if Λ has a lower cover Λ_o it may happen that there exist logics Θ such that $\Theta \subsetneq \Lambda$ but $\Theta \not\subseteq \Lambda_o$.

The notion of a **sublattice** and homomorphic image of a lattice is defined just as all the other algebraic concepts. (However, notice that our lattices usually have infinitary operations. So the concepts must be extended to infinitary operations; we trust that the reader understands how this is done.) If we just consider lattices as objects, we can afford to be vague as to whether or not we consider infinitary operations or the top and bottom elements as being primitve operations, because we can define them from the others if necessary. When we consider homomorphisms of lattices, however, this makes a great difference. A homomorphism which is faithful to the finitary operations need not be faithful to the infinitary ones. (This is a point where the notion of a category defined in Section 4.4 is helpful. Taking the class of complete lattices with maps respecting the finitary operations results in a different category than taking the class of complete lattices together with maps preserving the infinitary operations.) Let us say, therefore, that a lattice only has the finitary operations. Recall that locales are structures $\langle L, \sqcap, \bigsqcup \rangle$ where $\sqcap$ and $\bigsqcup$ are the finitary meet and infinitary join. (See Section 3.) Notions of subobject and homomorphism differ with respect to the top and bottom element, depending on whether they are present in the signature. If they are they must be preserved under embeddings. In boolean algebras, where the top and bottom are definable, a subalgebra must contain the top and bottom of the original algebra. For lattices this need not be so. For our purposes, we are quite happy not to have the top and the bottom in the signature.

Figure 7.2.



Recall from Section 1.1 the notation $\uparrow X$ and $\downarrow X$. An upper cone is a **filter** if it is closed under finite intersections; a lower cone is called an **ideal** if it is closed under finite unions. A filter (an ideal) is **principal** if it is of the form $\uparrow\{x\}$ ($\downarrow\{x\}$) for some x. We usually write $\downarrow x$ instead of $\downarrow\{x\}$ and $\uparrow x$ instead of $\uparrow\{y\}$. If F is a principal filter and $F = \uparrow x = \uparrow y$ then $x = y$. Similarly for ideals. Filters and ideals are special sorts of sublattices. If they are principal, then they are also complete sublattices, or, for that matter, sublocales.

Now we come to the main definition of this section, that of a *splitting*, shown in Figure 7.2.

Definition 7.2.6. *A pair $\langle p, q\rangle$ of elements is said to* ***split*** *a lattice $\mathfrak{L}$ if for every element $x \in L$ we have $x \leq p$ or $x \geq q$, but not both. Equivalently, $\langle p, q\rangle$ is a splitting if L is the disjoint union of the ideal $\downarrow p$ and the filter $\uparrow q$. If that is so, p is said to* ***split*** *$\mathfrak{L}$, and q is said to* ***co−split*** *$\mathfrak{L}$. p and q are* ***splitting companions*** *of each other. We write $\mathfrak{L}/p$ or $\bot/p$ for the splitting companion of p.*

The situation is depicted in Figure 7.2. We give some examples from modal logic.

Example. $\mathbf{K}_1/\mathsf{Th}\,\boxed{\bullet} = \mathbf{K}_1.\mathbf{D}$. For let $\Lambda \nsubseteq \mathsf{Th}\,\boxed{\bullet}$. Then $\Box\bot \notin \Lambda$, and so $\Diamond\top \in \Lambda$.

Example. $\mathbf{S4}/\mathsf{Th}\,\boxed{\circ\!\longrightarrow\!\circ} = \mathbf{S5}$. $\boxed{\circ\!\longrightarrow\!\circ}$ consists of two points, 0 and 1, and $\lhd = \{\langle 0, 0\rangle, \langle 0, 1\rangle, \langle 1, 1\rangle\}$. Assume that Λ is an extension of $\mathbf{S4}$ such that $\Lambda \nsubseteq \mathbf{S5}$. We have to show that $\mathfrak{h}$ is a frame for Λ. So, by assumption, $\Diamond(p \wedge \Diamond\Box\neg p) \vee \Diamond(\neg p \wedge \Diamond\Box p)$ is consistent with Λ and therefore there exists a model $\langle\mathfrak{F}, \beta, x\rangle \vDash p; \Diamond\Box\neg p$. Put

$$V := \overline{\beta}(\Box(p \to \Box\Diamond p) \wedge \Box(\neg p \to \Box\Diamond\neg p))$$

V is an internal set of $\mathfrak{F}$ and nonempty. It is a generated subset. Furthermore, every point not in V has a sucessor in V. To see this, define the **span** of $y \in f$ to be the set of formulae $\varphi \in \{p, \neg p\}$ such that there exists a $z \rhd y$ with $\langle\mathfrak{F}, \beta, z\rangle \vDash \varphi$. A point y is of **minimal span** if every successor of y has the same span as y. It is easy to see that every point sees a point of minimal span. V is the set of points of minimal span. Let

W be the set of successors of x which are not in V. W is not empty, for it contains x. (By the fact that x has span $\{p, \neg p\}$, and has a successor of span $\{\neg p\}$, $x \notin V$.) Define π by $\pi(y) := 0$ if $y \in W$ and $\pi(y) := 1$ if $y \in V$. π is a contraction onto $\mathfrak{h}$ of the subframe generated by x in $\mathfrak{F}$.

Example. (See [**241**] for details.) Let Θ be the tense logic of the rational numbers $\langle \mathbb{Q}, < \rangle$ and Λ the tense logic of the real numbers $\langle \mathbb{R}, < \rangle$. Then $\Lambda = \Theta/\mathsf{Th}\,\boxed{\circ \longrightarrow \circ}$. For let $\langle f, \lhd \rangle$ be a rooted frame for Θ. Then $\lhd$ is a linear dense order without endpoints, though not necessarily irreflexive. Suppose that there is a tense p–morphism onto $\boxed{\circ \longrightarrow \circ}$. Then there are sets $A, B \subseteq f$ such that for every $x \in A$ and every $y \in B$: $x \lhd y$. Moreover, every element in A has a successor in A, and every element in B has a predecessor in B. Such pair of sets is called a *gap*. Hence, the frames for Λ are those frames for Θ which contain no gaps. In particular, $\langle \mathbb{Q}, < \rangle$ is not a frame for Λ.

The following theorem gives a full characterisation of splitting elements and splitting pairs.

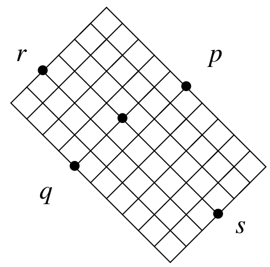
Theorem 7.2.7. *An element p splits a complete lattice iff p is $\bigsqcap$–prime. Dually, an element q co–splits a complete lattice iff q is $\bigsqcup$–prime. If p splits $\mathfrak{L}$, the companion is uniquely determined, similarly for co–splitting elements. The map $p \mapsto p_\star$ which sends a strictly meet–prime element to its splitting companion is an isomorphism from the poset of $\bigsqcap$–prime elements onto the poset of $\bigsqcup$–elements. Its inverse is the map $q \mapsto q^\star$ which sends a co–splitting element to its splitting companion.*

Proof. First, if p splits $\mathfrak{L}$, then the complement of $\downarrow p$ is a principal filter $\uparrow q$ for some q. q is certainly unique. Moreover, consider $p \geq \bigsqcap \langle y_i : i \in I \rangle$. Then, $\bigsqcap \langle y_i : i \in I \rangle \notin \uparrow q$. This can only be the case if for some $i \in I$, $y_i \notin \uparrow q$, which means for that i, $y_i \leq p$. Thus p is $\bigsqcap$–prime. Dually, we can show that if q co–splits the lattice, then q is $\bigsqcup$–prime. Conversely, assume that p is $\bigsqcap$–prime. Take $q := \bigsqcap \langle y : y \nleq p \rangle$. We have $q \nleq p$. Moreover, if $x \nleq p$ then by definition of q, $x \geq q$. And if $x \geq q$, then $x \leq p$ cannot hold. So, we have a splitting pair $\langle p, q \rangle$. Dually for $\bigsqcup$–prime elements. Now for the last claim. Take as a map the map that sends a splitting element to its companion. This map is injective; for if $p_1 \neq p_2$ then $\uparrow p_1 \neq \uparrow p_2$. Then also $\downarrow (p_1)_\star \neq \downarrow (p_2)_\star$, showing $(p_1)_\star \neq (p_2)_\star$. The map is surjective, because each $\bigsqcup$–prime element has a companion. Now consider $p_1 \leq p_2$, and let $\langle p_1, q_1 \rangle$ and $\langle p_2, q_2 \rangle$ be splitting pairs. Then $\downarrow p_1 \subseteq \downarrow p_2$, whence $\uparrow q_1 \supseteq \uparrow q_2$, from which $q_1 \leq q_2$. ∎

If we have a splitting pair, then there arise two natural sublattices, namely $\downarrow p$ and $\uparrow q$. Generally, in modal logic, we are interested in lattices of extensions, so the lattices $\uparrow q$ are of special importance. It is generally the case that if $\langle \widehat{p}, \widehat{q} \rangle$ is a splitting pair of $\mathfrak{L}$ and $\widehat{p} \geq q$ then $\widehat{p}$ splits the lattice $\mathfrak{L}/p$. (In case $\widehat{p} \ngeq q$ the induced splitting is trivial.) The splitting companion is $q \sqcup \widehat{q}$. So we have the following equation.

$$(\mathfrak{L}/p_1)/p_2 = \mathfrak{L}/p_1 \sqcup \mathfrak{L}/p_2$$

FIGURE 7.3. Iterated Splitting



The iterated splitting does therefore not depend on the order in which we split. Thus for a possibly infinite set N of prime elements of $\mathfrak{L}$ we define

$$\mathfrak{L}/N := \bigsqcup \langle \mathfrak{L}/p : p \in N \rangle$$

In Figure 7.3 the situation is shown where p and r are being split in succession, with q the splitting companion of p and s the splitting companion of r. In case we have a complete boolean algebra and a splitting pair $\langle p, q \rangle$, p is a coatom and q an atom. (Only coatoms are $\sqcap$–irreducible, and only atoms are $\sqcup$–irreducible.)

THEOREM 7.2.8. *A complete lattice $\mathfrak{L}$ is isomorphic to $\mathbf{2} \times \mathfrak{K}$ for some complete lattice $\mathfrak{K}$ iff there exists a splitting pair $\langle p, q \rangle$ such that p is a coatom and q is an atom.*

PROOF. Assume $\mathfrak{L} \cong \mathbf{2} \times \mathfrak{K}$. We may actually assume that $\mathfrak{L} = \mathbf{2} \times \mathfrak{K}$. Let b be the bottom element of $\mathfrak{K}$, and t the top element of $\mathfrak{K}$. (These elements exist in $\mathfrak{K}$ since it is a complete lattice.) Put $t^* := \langle 0, t \rangle$ and $b_* := \langle 1, b \rangle$. The pair $\langle t^*, b_* \rangle$ is a splitting pair of $\mathfrak{L}$. For assume that $\langle i, y \rangle \nleq t^* = \langle 0, t \rangle$. Then $i \nleq 0$, since $y \leq t$ by choice of t. Hence $i = 1$. But then $\langle i, y \rangle \geq \langle 1, b \rangle = b_*$. And conversely. Now assume that $\langle p, q \rangle$ is a splitting pair and that p is a coatom, while q is an atom in $\mathfrak{L}$. We show that the map $x \mapsto x \sqcup q$ is a bijection between the ideal $\downarrow p$ and the filter $\uparrow q$ with inverse $y \mapsto y \sqcap p$. For notice that under the assumptions, $p \sqcap q = \bot$ and $p \sqcup q = \top$. (For $q \nleq p$ and so $p \sqcap q < q$. Therefore $p \sqcap q = \bot$, for q is an atom. Likewise, $p \sqcup q > p$ and so $p \sqcup q = \top$, since p is a coatom.) Now for $x \leq p$ and $y \geq q$ we have

$$(x \sqcup q) \sqcap p = (x \sqcap p) \sqcup (q \sqcap p) = x \sqcap p = x$$

$$(y \sqcap p) \sqcup q = (y \sqcup q) \sqcap (p \sqcup q) = y \sqcup q = y$$

This shows that both maps are bijective. Moreover, they are isotonic and hence isomorphisms. $\Box$

It follows, for example, that a finite boolean algebra is isomorphic to a direct product of some copies of **2**. However, in general, in a splitting pair $\langle p, q\rangle$ neither is p a coatom, nor is q an atom. An example is the six element chain $\mathbf{6} = \langle 6, min, max\rangle$. In this lattice $\langle 2, 3\rangle$ is a splitting pair, but 2 is not an atom and 3 not a coatom. The reader may verify that also the splitting $\langle \mathsf{Th}\,\boxed{\bullet}, \mathbf{K.D}\rangle$ is an example.

Theorem 7.2.9 and Corollary 7.2.10 of the published version are false. Theorem 7.2.9 states that if $\langle \Theta, \Lambda\rangle$ is a splitting of the lattice of extensions of Σ then $\langle \vdash^m_\Theta, \vdash_\Lambda\rangle$ is a splitting of $\mathcal{E}(\vdash_\Sigma$. A counterexample was given by Emil Jerabek. Let $\Sigma = \mathbf{K}$, $\Theta = \mathbf{K} \oplus \Box\bot$ and $\Lambda = \mathbf{K.D}$. The consequence $\vdash^m_{\mathbf{K}}$ contains the rule $\langle\{\Box p\}, p\rangle$, but this rule is not in $\vdash^m_{\mathbf{K}\oplus\Box\bot}$. So the former is not contained in the latter. On the other hand, it has the same tautologies as $\vdash_{\mathbf{K}}$ and so it does not contain $\vdash_{\mathbf{K.D}}$. It does not matter whether we start with a splitting of the lattice of normal logics or of quasinormal logics. Observe that the same counterexample can be used for quasinormal logics, choosing $\mathbf{K} + \Box\bot$.

Exercise 227. Show Proposition 7.2.1.

Exercise 228. Let $\mathfrak{L}$ be a lattice, not necessarily complete. Show that an element $\neq \bot, \top$ is join–irreducible (meet–irreducible) iff it has exactly one lower cover (exactly one upper cover).

Exercise 229. Let $\mathfrak{L}$ be a complete lattice. Show that if p is $\bigsqcap$–irreducible in $\mathfrak{L}$ ($\bigsqcap$–prime) then it is also $\bigsqcap$–irreducible ($\bigsqcap$–prime) in any complete sublattice containing it.

Exercise 230. Show with a specific example that there are lattices $\mathfrak{L}$ and elements p such that p is not $\bigsqcap$–prime in $\mathfrak{L}$ but $\bigsqcap$–prime in the lattice $\mathfrak{L}/N$, for some set of $\bigsqcap$–prime elements N. Thus the notion of being a splitting element is not stable under iterated splittings.

Exercise 231. A logic Λ is called **essentially 1–axiomatizable** over Θ if for every axiomatization $\Lambda = \Theta(X)$ over Θ there exists a $\varphi \in X$ such that $\Lambda = \Theta(\varphi)$. Show that a modal logic Λ co–splits $\mathcal{E}\Theta$ iff it is essentially 1–axiomatizable over Θ.

7.3. Irreducible and Prime Logics

In this section we will investigate the possibility of characterizing irreducibility and primeness of logics algebraically.

Theorem 7.3.1. *$\Lambda \in \mathcal{E}\,\Theta$ is $\bigsqcap$–irreducible only if $\Lambda = \mathsf{Th}\,\mathfrak{A}$ for a subdirectly irreducible $\mathfrak{A}$.*

Proof. Surely we have $\Lambda = \mathsf{Th}\,\mathfrak{A}$ for some $\mathfrak{A}$. Suppose that $\mathfrak{A}$ is not subdirectly irreducible. Then, by Theorem 4.1.3, $\mathfrak{A}$ is a subdirect product of $\langle \mathfrak{A}_i : i \in I\rangle$, where every $\mathfrak{A}_i$ is subdirectly irreducible. Then we have $\mathsf{Th}\,\mathfrak{A} = \bigsqcap\langle\mathsf{Th}\,\mathfrak{A}_i : i \in I\rangle$. Hence there is a $\mathfrak{A}_i$ such that $\mathsf{Th}\,\mathfrak{A} = \mathsf{Th}\,\mathfrak{A}_i = \Lambda$. ◻

The converse of Theorem 7.3.1 is generally false. For example, take the **S4.3**–frame $\mathfrak{o} := \langle\omega + 1, \geq\rangle$. This frame is generated by ω and therefore the algebra $\mathfrak{Ma}(\mathfrak{o})$ of subsets of that frame is subdirectly irreducible. However, $\mathsf{Th}\,\mathfrak{Ma}(\mathfrak{o}) = \mathbf{Grz.3}$ has the finite model property and is therefore not irreducible in $\mathcal{E}\,\mathbf{K}$. A useful result is this.

Proposition 7.3.2. *Let Λ be $\bigsqcap$–irreducible. Then $\Lambda = \mathsf{Th}\,\mathfrak{A}$ for an $\mathfrak{A}$ which is finitely generated.*

Proof. Let $\Lambda = \mathsf{Th}\,\mathfrak{A}$. We have $\mathfrak{A} \nvDash \varphi$ iff there is a valuation β and a U such that $\langle\mathfrak{A}, \beta, U\rangle \vDash \neg\varphi$. Take $\mathfrak{C}$ to be the subalgebra generated by $\beta(p)$, $p \in var(\varphi)$. Let $V := U \cap C = \{a \in U : a \in C\}$. We claim that V is an ultrafilter on C. For V is obviously upward closed in C. Moreover, it is closed under intersection, since C is closed under intersection and U is as well. Also, if $x \in C$, then also $-x \in C$. Now either $x \in U$ or $-x \in U$. Hence either $x \in V$ or $-x \in V$. So, V is an ultrafilter in C. Since $\mathfrak{C}$ is a subalgebra of $\mathfrak{A}$, we have $\mathsf{Th}\,\mathfrak{C} \supseteq \mathsf{Th}\,\mathfrak{A}$. Also, $\langle\mathfrak{C}, \beta, V\rangle \vDash \varphi$, since $\overline{\beta}(\neg\varphi) \in C$ and $\overline{\beta}(\neg\varphi) \in U$. We conclude that if there is a model based on $\mathfrak{A}$, there is a model based on a finitely generated subalgebra of $\mathfrak{A}$. So

$$\mathsf{Th}\,\mathfrak{A} = \bigcap\langle\mathsf{Th}\,\mathfrak{C} : \mathfrak{C} \rightarrowtail \mathfrak{A}, \mathfrak{C} \text{ finitely generated}\rangle$$

Consequently, if $\mathsf{Th}\,\mathfrak{A}$ is $\bigsqcap$–irreducible, there must be a finitely generated $\mathfrak{C} \rightarrowtail \mathfrak{A}$ such that $\mathsf{Th}\,\mathfrak{C} = \mathsf{Th}\,\mathfrak{A}$. ◻

Theorem 7.3.3. *Λ is irreducible in $\mathcal{E}\,\Theta$ if and only if $\Lambda = \mathsf{Th}\,\mathfrak{A}$ for a subdirectly irreducible algebra such that $\mathsf{Th}\,\mathfrak{A}$ is prime in $\mathcal{E}\mathsf{Th}\,\mathfrak{A}$.*

Proof. Observe that in a lattice an element x is $\bigsqcap$–irreducible iff it is $\bigsqcap$–prime in $\uparrow x$. ◻

Thus, in order to find algebras whose logics are irreducible we need to answer the question of which algebras are splitting algebras. This is by no means trivial or more easy, so the last theorem is really of little practical significance. By analogy, an algebra $\mathfrak{A}$ or a frame $\mathfrak{F}$ is called **prime** in $\mathcal{E}\Theta$ or $\mathsf{Alg}\,\Theta$ if $\mathsf{Th}\,\mathfrak{A}$ ($\mathsf{Th}\,\mathfrak{F}$) is prime in $\mathcal{E}\Theta$. For the statement of our next theorems we shall introduce the notion of a presentation of an algebra. We shall first give a general definition and then specialize to modal algebras.

Definition 7.3.4. *Let $\mathcal{V}$ be a variety of Ω–algebras and $\mathfrak{A} \in \mathcal{V}$. A pair $\langle X, E\rangle$ is called a **presentation of** $\mathfrak{A}$ if (i.) X is a set, $E \subseteq Tm_\Omega(X) \times Tm_\Omega(X)$, and (ii.) $\mathfrak{Fr}_{\mathcal{V}}(X)/\Theta(E) \cong \mathfrak{A}$, where $\Theta(E)$ is the least congruence in $\mathfrak{Fr}_{\mathcal{V}}(X)$ containing E. A presentation is **finite** if X and E are both finite. $\mathfrak{A}$ is called **finitely presentable in** $\mathcal{V}$ if $\mathfrak{A}$ has a finite presentation in $\mathcal{V}$.*

Since every algebra is the homomorphic image of a free algebra in a variety, every algebra of a variety is presentable in it. Simply let $h : \mathfrak{Tm}_\Omega(X) \twoheadrightarrow \mathfrak{A}$ be a surjective homomorphism. Then put $E := ker(h)$. The pair $\langle X, E\rangle$ is a presentation of $\mathfrak{A}$.

Now fix an isomorphism i. Then $i \circ h_{\Theta(E)} : X \to A$. So, we may actually assume that X consists of elements of A. Then $\mathfrak{A}$ is generated by X, and E is a set of equations such that the minimal congruence containing E is the kernel of the canonical map $\mathfrak{Fr}_{\mathcal{V}}(X) \twoheadrightarrow \mathfrak{A}$. In the present context, we speak of logics rather than varieties, and of open filters rather than congruences. We write $\langle\Delta\rangle$ for the open filter generated by Δ.

Definition 7.3.5. *Let $\mathfrak{A}$ be a κ–modal algebra, Θ a κ–modal logic. A* ***presentation*** *of $\mathfrak{A}$ over Θ is a pair $\langle X, \Delta\rangle$ where $X \subseteq A$, is a set of variables and Δ a set of terms over X such that the map $pr : \mathfrak{Fr}_\Theta(X)/\langle\Delta\rangle \to \mathfrak{A}$ induced by the identity assignment $pr(a) := a$ is an isomorphism. $\mathfrak{A}$ is called* ***finitely presentable*** *if both X and Δ can be chosen finite. We call Δ a* ***diagram*** *of $\mathfrak{A}$ over Θ. Δ is not uniquely determined. We also say that $\mathfrak{A}$ is α–****presentable*** *if $\sharp X = \alpha$.*

If Δ can be chosen finite, we write δ to denote the set Δ as well as $\bigwedge \Delta$. Moreover, we may identify $\mathfrak{Can}_\Theta(X)$ and $\mathfrak{Can}_\Theta(\alpha)$. In order to appreciate this definition, consider the dual frames $\mathfrak{A}^+$ and $\mathfrak{Can}_\Theta(\alpha)$. Of course, since $\mathfrak{A}$ is α–generated there is a homomorphism $h : \mathfrak{Fr}_\Theta(\alpha) \twoheadrightarrow \mathfrak{A}$. Thus, $h^+ : \mathfrak{A}^+ \rightarrowtail \mathfrak{Can}_\Theta(\alpha)$. So, $\mathfrak{A}^+$ is a generated subframe of the α–canonical frame for Θ. It is clear that there exists a possibly infinite Δ such that $\mathfrak{Fr}_\Theta(\alpha)/\langle\Delta\rangle \cong \mathfrak{A}$. We also call this a **diagram**. Thus, viewing the canonical frame as equivalence classes of formulae, there is a set of formulae characterizing the notion *being in the generated subframe underlying* $\mathfrak{A}^+$ generated by α. $\mathfrak{A}$ is finitely presentable if there is a finite set of this kind. Now, recall that even if Θ is the modal theory of $\mathfrak{A}$, there are many ways in which $\mathfrak{A}^+$ lies embedded in the canonical Θ–frame, just because there are many ways to generate $\mathfrak{A}$ by α elements. The diagram of $\mathfrak{A}$ completely determines the map $h : \mathfrak{Fr}_\Theta(\alpha) \twoheadrightarrow \mathfrak{A}$ — at least up to isomorphisms of $\mathfrak{A}$, which is the best we can hope for anyway. Now fix h, let Δ be the associated diagram, and $\mathfrak{H}$ is the h^+–image of the frame underlying $\mathfrak{A}^+$. Thus in the canonical frame, a world x satisfies this diagram, $x \in \boxtimes^\omega\Delta$, iff x is in $\mathfrak{H}$.

If $\mathfrak{A}$ is finite and the number κ of primitive modalities is finite then for every $k \in \omega$, $\mathfrak{A}$ is k–presentable iff it is k–generated. In particular, $\mathfrak{A}$ is k–presentable for $k := \sharp A$:

Proposition 7.3.6. ($\kappa < \aleph_0$.) *Let $\mathfrak{A}$ be finite. Define*

$$\begin{aligned}\delta(\mathfrak{A}) := \quad & \bigwedge\langle p_a \wedge p_b \leftrightarrow p_{a\cap b} : a, b \in A\rangle \\ & \wedge \bigwedge\langle \neg p_a \leftrightarrow p_{-a} : a \in A\rangle \\ & \wedge \bigwedge\langle \Box_i p_a \leftrightarrow p_{\blacksquare_i a} : a \in A, i < \kappa\rangle\end{aligned}$$

Then $\delta(\mathfrak{A})$ is a diagram for $\mathfrak{A}$ over Θ.

Proof. Consider $\epsilon : \mathfrak{Fr}_\Theta(var(\delta)) \to \mathfrak{A} : p_a \mapsto a$. ϵ is surjective and factors through pr, that is, $\epsilon = pr \circ \kappa$ for a homomorphism κ. Since $pr \circ \kappa = \epsilon$ is surjective,

pr is surjective. It remains to show that pr is injective. This is done by showing that $\kappa(\varphi) = \kappa(\psi)$ iff $\varphi \leftrightarrow \psi \in \langle\delta\rangle$ and for every φ there is a $a \in A$ such that $\varphi \leftrightarrow p_a \in \langle\delta\rangle$. This is proved by induction on φ. Hence $\sharp im(\kappa) = \sharp A$ and since A is finite, pr is injective. □

This diagram uses as many variables as there are elements in the algebra. However, it is possible to use far less variables. First, a finite algebra is isomorphic to the set algebra over a finite Kripke–frame $\mathfrak{f}$. Given $\mathfrak{f}$ we may replace the diagram by

$$\begin{aligned}\delta(\mathfrak{f}) := \quad & \bigvee\langle p_v : v \in f\rangle \\ & \wedge\bigwedge\langle p_v \to \neg p_w : v \neq w\rangle \\ & \wedge\bigwedge\langle p_v \to \Diamond_j p_w : v \lhd_j w, j < \kappa\rangle \\ & \wedge\bigwedge\langle p_v \to \neg\Diamond_j p_w : v \not\lhd_j w, j < \kappa\rangle\end{aligned}$$

This is clearly an equivalent characterization of $\mathfrak{A}$. Notice that by passing from variables for sets of worlds to variables for worlds we have used $k := \sharp f$ variables where previously we had 2^k. We can compress this diagram further. Let $\ell := \ulcorner {}^2 log\, k \urcorner$, the least integer such that $2^\ell \geq k$. Then there exists an injection $s : f \to \wp(\ell)$. Now take variables q_i, $i < \ell$. For $S \subseteq \{0, \dots, \ell - 1\}$ define

$$\chi(S) := \bigwedge_{i \in S} q_i \wedge \bigwedge_{i \notin S} \neg q_i$$

Then the diagram for $\mathfrak{f}$ can be replaced by the following formula

$$\begin{aligned}\delta(\mathfrak{f}) := \quad & \bigvee\langle \chi(s(v)) : v \in f\rangle \\ & \wedge\bigwedge\langle \chi(s(v)) \to \Diamond_j \chi(s(w)) : v \lhd_j w, j < \kappa\rangle \\ & \wedge\bigwedge\langle \chi(s(v)) \to \neg\Diamond_j \chi(s(w)) : v \not\lhd_j w, j < \kappa\rangle\end{aligned}$$

Thus the number of variables needed is at most logarithmic in the size of the underlying frame. It is clear that a sharp bound is the number of generators of $\mathfrak{A}$. We have shown, however, that this number is at most doubly logarithmic in $\sharp A$.

Definition 7.3.7. *Let $\mathfrak{A}$ be subdirectly irreducible. Let Δ be a diagram of $\mathfrak{A}$ over Θ, $\mathfrak{B} \in \mathsf{Alg}\,\Theta$. For a compound modality $\boxplus$ and a finite subset $\delta \subseteq \Delta$ we say that $\mathfrak{B}$ is $\boxplus\delta$–**consistent with** $\mathfrak{A}$, if a valuation $\beta : var[\Delta] \to B$ and an ultrafilter U exists such that $\langle\mathfrak{B}, \beta, U\rangle \vDash \neg p_c; \boxplus\delta$, where $p_c \in var[\Delta]$, c being an opremum of $\mathfrak{A}$. $\mathfrak{B}$ is called ω–**consistent with** $\mathfrak{A}$ if a valuation $\beta : var[\Delta] \to B$ exists satisfying $\langle\mathfrak{B}, \beta, U\rangle \vDash \neg p_c; \{\boxplus\delta : \boxplus$ compound$, \delta \in \Delta\}$. If $\mathfrak{B}$ is $\boxplus\delta$–consistent with $\mathfrak{A}$ for every compound modality $\boxplus$ and finite subset $\delta \subseteq \Delta$ then $\mathfrak{B}$ is said to be **weakly consistent with** $\mathfrak{A}$.*

Theorem 7.3.8. *Let $\mathfrak{A}$ be subdirectly irreducible. Then the following assertions are equivalent for all $\mathfrak{B} \in \mathsf{Alg}\,\Theta$:*

(i) *$\mathfrak{B}$ is ω–consistent with $\mathfrak{A}$.*
(ii) $\mathfrak{A} \in \mathsf{SH}(\mathfrak{B})$.
(iii) $\mathfrak{A} \in \mathsf{HS}(\mathfrak{B})$.

FIGURE 7.4.

$$\begin{tikzcd}
& var[\Delta] \arrow[dl, "\beta"'] \arrow[d, "\gamma"] & \\
\mathfrak{B} \arrow[r, "\epsilon"] & \mathfrak{C} & \mathfrak{A} \arrow[l, "\zeta"'] \\
& \mathfrak{Fr}_\Theta(var[\Delta]) \arrow[ul, "\overline{\beta}"] \arrow[u, "\overline{\gamma}"] \arrow[ur, "\eta"'] &
\end{tikzcd}$$

PROOF. Let Δ be a diagram of $\mathfrak{A}$ over Θ. Assume first (i). Let then $\beta : var[\Delta] \to B$ and U be such that $\overline{\beta}(\neg p_c) \in U$ and for all compound modalities $\boxplus$ and formulae $\delta \in \Delta$ we have $\overline{\beta}(\boxplus\delta) \in U$. Let F be the open filter generated by the $\overline{\beta}(\boxplus\delta)$. Consider the induced mapping $\epsilon : \mathfrak{B} \to \mathfrak{B}/F =: \mathfrak{C}$ and put $\gamma := \epsilon \circ \beta : var[\Delta] \to C$. Then $\overline{\gamma}(\boxplus\delta) = 1$ for all $\boxplus$ and $\delta \in \Delta$. The homomorphism $\overline{\gamma}$ factors through $\eta : \mathfrak{Fr}_\Theta(var[\Delta]) \to \mathfrak{A}$, for $\mathfrak{A} \cong \mathfrak{Fr}_\Theta(var[\Delta])/\langle\Delta\rangle$. (See Figure 7.4.) $\mathfrak{A}$ is subdirectly irreducible and has a minimal nontrivial congruence relation, which is generated by c. To show that the induced mapping ζ is injective it is therefore sufficient to show that $\zeta(c) \neq 1$ — or, equivalently — that $\zeta(-c) \neq 0$. But $\zeta(-c) = \zeta \circ \eta(\neg p_c) = \overline{\gamma}(\neg p_c) = \epsilon \circ \overline{\beta}(\neg p_c) > 0$, because for every $a \in F$ there exists a $\boxplus$ and a finite subset $\delta \subseteq \Delta$ such that $a \geq \overline{\beta}(\boxplus\delta)$. Then $\overline{\beta}(\neg p_c) \cap a \geq \overline{\beta}(\neg p_c) \cap \overline{\beta}(\boxplus\delta) = \overline{\beta}(\neg p_c \wedge \boxplus\delta) \neq 0$. Hence ζ is injective and $\mathfrak{A} \in \mathsf{S}(\mathfrak{C}) \subseteq \mathsf{SH}(\mathfrak{B})$. This shows (ii). That (ii) implies (iii) we have seen in Chapter 1.3. Finally, assume (iii) holds. We will show (i). Let $\mathfrak{A} \in \mathsf{HS}(\mathfrak{B})$. Then there is a $\mathfrak{C}$, an injective $\epsilon : \mathfrak{C} \rightarrowtail \mathfrak{B}$ and a surjection $\rho : \mathfrak{C} \twoheadrightarrow \mathfrak{A}$. (See Figure 7.5.) Here, $\mathfrak{F} := \mathfrak{Fr}_\Theta(var[\Delta])$ and κ is the canonical mapping. $\mathfrak{F}$ is free and thus κ can be lifted over ρ to $\gamma : \mathfrak{F} \to \mathfrak{C}$. Then $\beta := \epsilon \circ \gamma$. For every $n \in \omega$, $\kappa(\neg p_c \wedge \boxplus\delta) = \kappa(\neg p_c) > 0$ and since ϵ is injective, $\beta(\neg p_c \wedge \boxplus\delta) = \epsilon \circ \gamma(\neg p_c \wedge \boxplus\delta) > 0$ for every $\boxplus\delta$. Hence $\mathfrak{B}$ is ω–consistent with $\mathfrak{A}$; (i) is shown. ◻

LEMMA 7.3.9. *Let $\mathfrak{A}$ be subdirectly irreducible. If $\mathfrak{B}$ is weakly consistent with $\mathfrak{A}$ then there is a $\mathfrak{S} \in \mathsf{Up}(\mathfrak{B})$ which is ω–consistent with $\mathfrak{A}$.*

PROOF. Let S be the set of formulae $\boxplus\delta$ such that δ is a finite subset of Δ and $\boxplus$ is a compound modality. (We have $\boxplus = \Box^s$ for some finite set s of sequences over κ.) For each $\sigma \in S$ we have a valuation $\beta_\sigma : var[\Delta] \to B$ and an ultrafilter U_σ such that

FIGURE 7.5.

$$\begin{tikzcd}
& \mathfrak{C} \arrow[r, "\epsilon"] \arrow[dl, "\rho"'] & \mathfrak{B} \\
\mathfrak{A} & & \\
& \mathfrak{F} \arrow[uu, "\gamma"] \arrow[uur, "\beta"'] \arrow[ul, "\kappa"] &
\end{tikzcd}$$

$\overline{\beta}_\sigma(\neg p_c \wedge \sigma) \in U_\sigma$. Let s be a finite set of sequences and δ a finite subset of Δ.

$$E(s, \delta) := \{\Box^t \delta' : s \subseteq t, \delta \subseteq \delta'\}$$

Next put

$$\mathbb{E} := \{E(s, \delta) : s \subset \kappa^*, \delta \subseteq \Delta, s, \delta \text{ finite}\}$$

$\mathbb{E}$ has the finite intersection property, for we have

$$E(s_1, \delta_1) \cap E(s_2, \delta_2) = E(s_1 \cup s_2, \delta_1 \cup \delta_2) .$$

Therefore, $\mathbb{E}$ is contained in an ultrafilter, say V. Now define $\mathfrak{C} := \prod_V \mathfrak{B}$, $\gamma := \prod_V \beta_\sigma$. Then $\overline{\gamma}(\neg p_c \wedge \boxplus\delta) > 0$ for every $\boxplus\delta \in S$. The set

$$\{\overline{\gamma}(\neg p_c \wedge \Box^s \delta) : s \subseteq \kappa^*, \delta \subseteq \Delta\}$$

has the finite intersection property and is contained in an ultrafilter. Let it be U. Then $\langle \mathfrak{C}, \gamma, U\rangle \vDash \neg p_c; S$. Hence $\mathfrak{C}$ is ω–consistent with $\mathfrak{A}$. ◻

Putting our results together we get

THEOREM 7.3.10. *Let $\mathfrak{A}$ be subdirectly irreducible. Then the following assertions are equivalent for every $\mathfrak{B}$:*

(i) $\mathfrak{B}$ *is weakly consistent with* $\mathfrak{A}$.
(ii) $\mathfrak{A} \in \mathsf{SHUp}(\mathfrak{B})$.
(iii) $\mathfrak{A} \in \mathsf{HSP}(\mathfrak{B})$.

PROOF. (i) $\Rightarrow$ (ii). Assume (i). By Lemma 7.3.9 there is a $\mathfrak{C} \in \mathsf{Up}(\mathfrak{B})$ which is ω–consistent with $\mathfrak{A}$. Hence by Theorem 7.3.8 $\mathfrak{A} \in \mathsf{SH}(\mathfrak{C}) \subseteq \mathsf{HSUp}(\mathfrak{B})$. (ii) $\Rightarrow$ (i). If $\mathfrak{B}$ is not weakly consistent with $\mathfrak{A}$ then there is a $\boxplus\delta$ such that $\mathfrak{B} \vDash \boxplus\delta \to p_c$. It follows that $\mathfrak{S} \vDash \boxplus\delta \to p_c$ for any $\mathfrak{S} \in \mathsf{Up}(\mathfrak{B})$. Hence $\mathfrak{S}$ is not ω–consistent with $\mathfrak{A}$ and thus $\mathfrak{A} \notin \mathsf{SH}(\mathfrak{C})$. Since this is valid for all $\mathfrak{S} \in \mathsf{Up}(\mathfrak{B})$, $\mathfrak{A} \notin \mathsf{SHUp}(\mathfrak{B})$. Now (ii) implies (iii) because $\mathsf{HSUp}(\mathfrak{B}) \subseteq \mathsf{HSP}(\mathfrak{B})$, and (iii) implies (i). ◻

Notice that we did not actually use Jónsson's Lemma, but derived it. That this is possible was first observed in [**234**]. Notice that in the course of the definitions we have made crucial use of the fact that $\mathfrak{A}$ is subdirectly irreducible. The previous theorem shows quite clearly why this must be so. For if not, all varieties are of the form $\mathsf{HSUp}(\mathcal{K})$ for a class $\mathcal{K}$. The following Splitting Theorem can now be proved. The original version appeared in Rautenberg [**170**] but only for finite algebras in weakly transitive logics. In Kracht [**120**] it was generalized to the case where $\mathfrak{A}$ is finitely presentable. The following version is fully general and was proved in Wolter [**234**].

Theorem 7.3.11 (Splitting Theorem). *Let $\mathfrak{A}$ be a subdirectly irreducible modal algebra with diagram Δ. Then the following are equivalent.*

1. $\mathsf{Th}\,\mathfrak{A}$ *is prime in* $\mathcal{E}\,\Theta$.
2. *There is a compound modality* $\boxplus$ *and a finite* $\delta \subseteq \Delta$ *such that for all* $\mathfrak{B} \in \mathsf{Alg}\,\Theta$*:*

   (†) *If* $\mathfrak{B}$ *is* $\boxplus\delta$*–consistent with* $\mathfrak{A}$ *then* $\mathfrak{B}$ *is weakly consistent with* $\mathfrak{A}$.
3. $\{\mathfrak{B} : \mathfrak{A} \notin \mathsf{HSP}\mathfrak{B}\}$ *is a variety.*
4. $\{\mathfrak{B} : \mathfrak{A} \notin \mathsf{HSP}\mathfrak{B}\}$ *is closed under ultraproducts.*

Moreover, if $\mathfrak{A}$ *and* $\boxplus\delta$ *fulfill* (†) *we have* $\mathcal{E}\Theta/\mathfrak{A} = \Theta \oplus \boxplus\delta \to p_c$.

Proof. Clearly, the first and the last are equivalent, for if we have a splitting, the corresponding splitting partner axiomatizes a variety, the variety of all algebras whose theory is not contained in the theory of $\mathfrak{A}$, or those algebras in whose variety $\mathfrak{A}$ is not contained. Thus, we have to show the equivalence of the last three. Let S be the set of all $\boxplus\delta$. Assume that (†) fails for all $\sigma \in S$. For every $\sigma \in S$ there is a $\mathfrak{B}_\sigma$ which is σ–consistent with $\mathfrak{A}$ but not weakly consistent with $\mathfrak{A}$. Hence by the preceding theorem, for every $\sigma \in S$, $\mathfrak{A} \notin \mathsf{HSP}(\mathfrak{B}_\sigma)$ that is $\mathsf{Th}\,\mathfrak{B}_\sigma \not\supseteq \mathsf{Th}\,\mathfrak{A}$. However, a suitable ultraproduct $\mathfrak{C} = \prod_{\sigma\in S} \mathfrak{B}_\sigma$ is weakly consistent with $\mathfrak{A}$, and hence $\mathsf{Th}\,\mathfrak{A} \supseteq \mathsf{Th}\,\mathfrak{C}$, from which follows that $\{\mathfrak{B} : \mathfrak{A} \notin \mathsf{HSP}\mathfrak{B}\}$ is not closed under ultraproducts and hence also not a variety. Also, $\mathsf{Th}\,\mathfrak{A}$ is not prime in $\mathcal{E}\Theta$. Let now (†) be fulfilled by some $\boxplus\delta$. Then $\{\mathfrak{B} : \mathfrak{A} \notin \mathsf{HSP}\mathfrak{B}\} = \{\mathfrak{B} : \mathfrak{B} \vDash \boxplus\delta \to p_c\}$ and hence it is a variety, and so closed under ultraproducts. Thus $\mathsf{Th}\,\mathfrak{A}$ is prime, with splitting companion $\Theta \oplus \boxplus\delta \to p_c$. ∎

There are two important subcases of the Splitting Theorem, which are each quite characteristic. One concerns the case of weakly transitive logics, the other that of cycle–free frames. Recall that weakly transitive logics are characterized by the existence of a strongest compound modality. Furthermore, each finite subdirectly irreducible algebra has a diagram. Hence (†) is easy to satisfy in this case.

Corollary 7.3.12 (Rautenberg). ($\kappa < \aleph_0$.) *Let* Θ *be weakly transitive. Then every finite subdirectly irreducible* Θ*–algebra splits* $\mathcal{E}\Theta$.

Let us say that a finite algebra is **cycle–free** iff $\mathfrak{A} \vDash \boxtimes^m \bot$ for some $m \in \omega$. It can be seen that $\mathfrak{A}$ is cycle–free iff the corresponding Kripke–structure contains no cycles. If $\mathfrak{A} \nvDash \boxtimes^{\leq m-1} \bot$ then $\boxtimes^{\leq m-1} \bot$ is an opremum of $\mathfrak{A}$.

Corollary 7.3.13 (Blok). ($\kappa < \aleph_0$.) *Let $\mathfrak{A}$ be finite subdirectly irreducible and cycle–free. Then $\mathfrak{A}$ is prime in every variety containing it. Then there exists a finite number m such that $\mathcal{E}\,\Theta/\mathfrak{A} = \Theta \oplus \boxtimes^{\leq m+1}\delta(\mathfrak{A}) \to \neg \boxtimes^m \bot$.*

Proof. Let $\boxtimes^m \bot$ be an opremum of $\mathfrak{A}$; let $\delta := \delta(\mathfrak{A})$ be a diagram of $\mathfrak{A}$. Then $\boxtimes^{\leq m+1}\delta$ satisfies (†) in Theorem 7.3.11. For if $\langle \mathfrak{B}, \beta, U\rangle \vDash \boxtimes^{\leq m+1}\delta \wedge \boxtimes^m \bot$ then for all $r > 0$ $\langle \mathfrak{B}, \beta, U\rangle \vDash \boxtimes^{\leq m+r}\delta \wedge \boxtimes^m \bot$ because $1 = \overline{\beta}(\boxtimes^m \bot \to \boxtimes^m \boxtimes^r \delta) = \overline{\beta}(\boxtimes^m \bot \to \boxtimes^{m+r}\delta)$ for all $r \geq 0$. So for all $r > 0$, $\overline{\beta}(\boxtimes^{\leq m}\delta \wedge \boxtimes^m \bot) \leq \overline{\beta}(\boxtimes^m \bot \wedge \boxtimes^{\leq m} \boxtimes^r \delta)$. ∎

Exercise 232. Show that the number of variables needed to axiomatize $\mathcal{E}\Lambda/\mathfrak{A}$ over Λ is equal to the number of elements needed to generate $\mathfrak{A}$.

Exercise 233. Let $\Theta = \bigsqcup_{i \in I} \mathcal{E}\,\Lambda/\mathfrak{B}_i$ where $\mathfrak{B}_i$ are splitting algebras. Suppose that the $\mathsf{Th}\,\mathfrak{B}_i$ are incomparable. Then the minimum number n of variables such that Θ is n–axiomatizable is equal to the minimum number k such that every $\mathfrak{B}_i$ is k–generated.

Exercise 234. A variety is said to have the **congruence extension property** (CEP) if for every algebra $\mathfrak{A}$ and every subalgebra $\mathfrak{B} \leq \mathfrak{B}$ the following holds. Every congruence Θ on $\mathfrak{B}$ is the restriction to B of a congruence on $\mathfrak{A}$. Hence, in a variety with CEP we have $\mathsf{HS}\,\mathfrak{A} = \mathsf{SH}\,\mathfrak{A}$ for all $\mathfrak{A}$. Show that Abelian groups have the CEP, but groups in general do not.

Exercise 235. Show that the variety of κ–modal algebras has CEP for any κ.

***Exercise 236.** Show that the algebra $\mathfrak{A}$ of the frame $\boxed{\circ \longrightarrow \circ}$ splits the lattice of extensions of **K**.**alt**$_3$.**B**.**T** of reflexive, symmetric frames with at most three successors. However, show that in general it does not follow that if $\mathfrak{A} \in \mathsf{SHUp}\,\mathfrak{B}$ then also $\mathfrak{A} \in \mathsf{SH}\,\mathfrak{B}$. The latter is in fact in many cases true, and a counterexamples are rather difficult to construct. For example, $\mathfrak{B}$ must in any case be infinite. (Can you show this?)

7.4. Duality Theory for Upper Continuous Lattices

We have seen that lattices of modal logics are locales also called upper continuous lattices and that they are generated by their join–compact elements, because by Theorem 2.9.8 the latter coincide with the finitely axiomatizable logics. In this chapter we want to go deeper into the structure theory of such locales and see what further properties of the lattices of logics we can derive. There is a duality theory for locales (see [**110**]). A topological space always gives rise to an upper continuous lattice, also called a locale. Namely, let $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ be a topological space over the set X with open sets $\mathbb{X}$. Then let $\Omega(\mathfrak{X}) := \langle \mathbb{X}, \cap, \bigcup\rangle$. By definition of a topological space all finite intersections of open sets are open, and all unions of open sets are open; so, $\Omega(\mathfrak{X})$ is a locale. Let $f : \mathfrak{X} \to \mathfrak{Y}$ be a continuous map. Then let $\Omega(f)$ be the the restriction of $f^{-1} : 2^Y \to 2^X : A \mapsto f^{-1}[A]$ to $\mathbb{X}$. By continuity of f this is a map from $\mathbb{X}$ to $\mathbb{Y}$. It is not hard to see that $\Omega(f) : \Omega(\mathfrak{Y}) \to \Omega(\mathfrak{X})$, i. e. that $\Omega(f)$ is a homomorphism of locales. It is easily seen that Ω is a contravariant functor from the category of topological spaces into the category of upper–continuous distributive lattices. Thus Ω can also be construed as a covariant functor from the category of topological spaces to the category of locales.

To make the exposition analogous to that of Stone–Duality, let us switch from the category of locales to the dual category, that of frames. A **point** of a frame is a surjective frame–morphism $p : \mathfrak{L} \twoheadrightarrow \mathbf{2}$. To have a surjection onto $\mathbf{2}$ means that L can be split into two sets, $F := p^{-1}(1)$ and $I := p^{-1}(0)$, where F is a filter and I an ideal, and I is closed under arbitrary joins. For if $p(a_s) = 0$ for all $s \in S$, then $p(\bigsqcup_{s\in S} a_s) = \bigsqcup_{s\in S} p(a_s) = 0$. Thus $I = \downarrow x$ for some x. Moreover, x must be $\sqcap$–irreducible (since F is closed under intersection). So, x is $\sqcap$–prime. This characterization is exact. For if x is $\sqcap$–irreducible then $\downarrow x$ is an ideal and $L - \downarrow x$ is a filter. Namely, from $y \nleq x$ and $y \leq z$ we deduce $z \nleq x$, and from $y, z \nleq x$ also $y \sqcap z \nleq x$. Let $pt(\mathfrak{L})$ denote the set of points of $\mathfrak{L}$. There is a bijection between the set of points and the $\sqcap$–irreducible elements of $\mathfrak{L}$, defined by $x \mapsto p_x$, where p_x is defined by $p_x(y) = 0$ iff $y \leq x$. On $pt(\mathfrak{L})$ we take as open sets the sets of the form $\widehat{x} = \{p : p(x) = 1\}$. So, since every p is of the form p_y for some y, we have $\widehat{x} = \{p_y : y \ngeq x\}$. Now put $\mathfrak{Spc}(\mathfrak{L}) = \langle pt(\mathfrak{L}), \{\widehat{x} : x \in L\}\rangle$. In place of homomorphisms we might also take the meet–irreducible elements as elements of $\mathfrak{Spc}(\mathfrak{L})$. $\mathfrak{Spc}(\mathfrak{L})$ is a topological space, by the next theorem, if we add that $\varnothing = \widehat{\top}$ and $pt(\mathfrak{L}) = \widehat{\bot}$.

Lemma 7.4.1. *The map $\widehat{x} = \{p : p(x) = 1\}$ commutes with arbitrary joins and finite meets.*

Proof. Let $z = \bigsqcup_{i\in I} x_i$. $p \in \widehat{z}$ iff $p(z) = 1$ iff $p(x_i) = 1$ for at least one $i \in I$ iff $p \in \widehat{x_i}$ for at least one $i \in I$. So $\widehat{z} = \bigcup_{i\in I} \widehat{x_i}$. Furthermore, $p \in \widehat{x_1 \sqcap x_2}$ iff $p(x_1 \sqcap x_2) = 1$ iff $p(x_1) = 1$ and $p(x_2) = 1$ iff $p \in \widehat{x_1} \cap \widehat{x_2}$. Hence $\widehat{x_1 \sqcap x_2} = \widehat{x_1} \cap \widehat{x_2}$. ∎

The sets of the form $\uparrow x$ are closed sets. Now consider a homomorphism $h : \mathfrak{L} \to \mathfrak{M}$. Then the map $p \mapsto p \circ h$ maps a point $p : \mathfrak{M} \twoheadrightarrow \mathbf{2}$ of $\mathfrak{M}$ onto a point of $\mathfrak{L}$. Thus put $\mathfrak{Spc}(h) : p \mapsto p \circ h$.

Theorem 7.4.2. *The map $\mathfrak{Spc}$ mapping a locale $\mathfrak{L}$ onto $\mathfrak{Spc}(\mathfrak{L})$ and a homomorphism $h : \mathfrak{L} \to \mathfrak{M}$ onto $\mathfrak{Spc}(h) : p \mapsto p \circ h$ is a covariant functor from the category of locales to the category of topological spaces and continuous maps.*

Proposition 7.4.3. *Let Loc^{op} be the opposite category of the category Loc of locales and homomorphisms; and let Top be the category of topological spaces. Then Ω considered as a functor $\mathsf{Top} \to \mathsf{Loc}^{op}$ is left adjoined to $\mathfrak{Spc}$ considered as a functor $\mathsf{Loc}^{op} \to \mathsf{Top}$.*

The proof of this fact is an exercise. Interesting for us are the unit and counit of this adjunction. If we have a locale, the function $x \mapsto \widehat{x}$ is a canonical map $\mathfrak{L} \to \Omega(\mathfrak{Spc}(\mathfrak{L}))$. This map is surjective, but in general not injective, see the exercises below. Likewise, given a topological space $\mathfrak{X}$, we have a map $\mathfrak{X} \to \mathfrak{Spc}(\Omega(\mathfrak{X}))$. This map is surjective, but in general not injective. For a point x, the set $\{x\}$ is join–irreducible and so its complement is meet–irreducible. Hence it gives rise to a point p_x of $\Omega(\mathfrak{X})$. There may exist x and y such that $x \neq y$ and $p_x = p_y$. This motivates the following definition.

Definition 7.4.4. *A locale $\mathfrak{L}$ is called **spatial** if the map $x \mapsto \widehat{x}$ is a bijection from $\mathfrak{L}$ onto $\Omega(\mathfrak{Spc}(\mathfrak{L}))$. A space $\mathfrak{X}$ is called **sober** if the map $x \mapsto p_x$ is a bijection from $\mathfrak{X}$ onto $\mathfrak{Spc}(\Omega(\mathfrak{X}))$.*

Theorem 7.4.5. *A locale is spatial iff every element is the meet of meet–irreducible elements.*

Proof. Let $\mathfrak{L}$ be spatial. Then there is a canonical map $x \mapsto \widehat{x}$. This is surjective iff $\mathfrak{L}$ is spatial. ◻

A characterization of sober spaces is harder to obtain. There are various reasons why a space can fail to be sober. First of all, it can happen that in $\mathfrak{X}$ two different elements, say x and y, are contained in the same open sets. Then the locale of open sets is isomorphic to the locale of the space $\mathfrak{Y}$ which differs from $\mathfrak{X}$ only in that x has been taken away. ($\mathfrak{Y}$ is the image of $\mathfrak{X}$ by a continuous map.) Therefore we define the following.

Definition 7.4.6. *Let $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ be a topological space. $\mathfrak{X}$ is a **T_0–space** if for every pair of elements $x, y \in X$ if $x \neq y$ there exists an open set O such that $\sharp(O \cap \{x, y\}) = 1$.*

The Sierpiński–space is a T_0–space, though not a T_2–space. Now we define a relation $\leq$ on points of a T_0–space by $x \leq y \Leftrightarrow \overline{\{x\}} \supseteq \overline{\{y\}}$. We call $\leq$ the specialization order. (In [**110**] the converse ordering is considered. The order defined here has the advantage to make the sets $\uparrow x$ closed and coincide with the upper set in $\Omega(\mathfrak{X})$.)

Proposition 7.4.7. *A topological space is a T_0–space iff the specialization order is a partial order.*

Proof. Let $\mathfrak{X}$ be a T_0–space. Since $\overline{\{x\}} \supseteq \overline{\{x\}}$ we have $x \leq x$. Moreover, if $x \leq y$ and $y \leq z$ then $\overline{\{x\}} \supseteq \overline{\{y\}}$ and $\overline{\{y\}} \supseteq \overline{\{z\}}$, from which $\overline{\{x\}} \supseteq \overline{\{z\}}$, or $x \leq z$. Hence $\leq$ is a partial order. Now assume that $\mathfrak{X}$ is not a T_0–space. Then there exist x, y, $x \neq y$ such that for all open sets O, either $O \cap \{x, y\} = \varnothing$ or $\{x, y\} \subseteq O$. Hence, for all closed sets C, either $\{x, y\} \subseteq C$ or $C \cap \{x, y\} = \varnothing$. It follows that $y \in \overline{\{x\}}$ and $x \in \overline{\{y\}}$. So, $x \leq y$ and $y \leq x$, that is, $\leq$ is not a partial order. ◻

Lemma 7.4.8. *A space is T_0 iff the map $x \mapsto p_x$ is injective.*

The proof of this lemma is left as an exercise.

Theorem 7.4.9. *For any locale, the space* $\mathfrak{Spc}(\mathfrak{L})$ *is sober.*

Proof. Let $\mathfrak{L}$ be a locale and A a meet–irreducible element of $\Omega(\mathfrak{Spc}(\mathfrak{L}))$. Then A is of the form $\widehat{x}$ for some $x \in L$. Since the map $y \mapsto \widehat{y}$ preserves infinite joins, there is a largest such x. We show that x is $\sqcap$–irreducible. For otherwise, there are y and z such that $x = y \sqcap z$ and so $\widehat{x} = \widehat{y} \cap \widehat{z}$, a contradiction to the assumption that A is meet–irreducible. Now let p_x be the point defined by x, x $\sqcap$–irreducible. Then

$$\overline{\{p_x\}} = pt(\mathfrak{L}) - \widehat{x}\,.$$

For $\widehat{x} = \{p_y : p_y(x) = 1\} = \{p_y : x \nleq y\}$. Thus, $\widehat{x} = \{p_y : y \notin \uparrow x\}$. Hence, $pt(\mathfrak{L}) - \widehat{x} = \{p_y : y \geq x\}$. This is a closed set. Moreover, it is the least closed set containing x. Thus the canonical map $x \mapsto \widehat{x}$ is surjective. Moreover, it is injective, since for different points p, q we cannot have $p(x) = q(x)$ for all x. Thus $\widehat{x}$ contains just one of p and q and so $\mathfrak{Spc}(\mathfrak{L})$ is a T_0–space, and so by the previous lemma the canonical map is injective. ◻

The specialization ordering on $\mathfrak{Spc}(\mathfrak{L})$ can be determined easily in $\mathfrak{L}$. For notice that we have

$$\begin{array}{rcl} p_x \leq p_y & \Leftrightarrow & \overline{\{p_x\}} \supseteq \overline{\{p_y\}} \\ & \Leftrightarrow & \widehat{x} \subseteq \widehat{y} \\ & \Leftrightarrow & \{q : q(y) = 1\} \subseteq \{q : q(x) = 1\} \\ & \Leftrightarrow & \{z : p_z(y) = 1\} \subseteq \{z : p_z(x) = 1\} \\ & \Leftrightarrow & \{z : z \ngeq y\} \subseteq \{z : z \ngeq x\} \\ & \Leftrightarrow & x \leq y\,. \end{array}$$

The interest in this duality of sober spaces with spatial locales for our purposes lies in the possibility to describe the lattices of modal logics as certain sober spaces. The way to approach the structure of a sober space is by first studying the specialization ordering and then looking at the topology defined over it. However, some care is needed. Just as with boolean algebras, the space of irreducible elements alone cannot provide a complete description of the lattice. For notice that in many cases there are at least two topologies for a given specialization order. Namely, given $\langle X, \leq\rangle$ let $Y(X, \leq)$ be the set of all lower closed sets, that is, sets of the form $\downarrow S$ for some S. This is the finest topology we can define. Also, let $\Phi(X, \leq)$ be the smallest topology that contains $\varnothing$ and all sets of the form

$$X - (\uparrow x_1 \cup \uparrow x_2 \cup \ldots \cup \uparrow x_n)$$

$\Phi(X, \leq)$ is called the **weak topology** and $Y(X, \leq)$ the **Alexandrov topology**.

Theorem 7.4.10. *Let* $\mathfrak{X} = \langle X, \mathbb{X}\rangle$ *be a* T_0*–space with specialization order* $\leq$*. Then*

$$\Phi(X, \leq) \subseteq \mathbb{X} \subseteq Y(X, \leq)$$

Proof. First, all sets of the form $\uparrow x$ for a single element x must be closed sets. Moreover, $\uparrow x = \overline{\{x\}}$. For $y \in \overline{\{x\}}$ iff every closed set containing x also contains y iff $\overline{\{y\}} \subseteq \overline{\{x\}}$ iff $y \geq x$. This shows the first inequality. The second follows, for $Y(X, \leq)$ contains all sets which are lower closed. ∎

Theorem 7.4.11. *Let $\mathfrak{L}$ be a spatial locale. $\mathfrak{L}$ is continuous iff the topology on $\mathfrak{Spc}(\mathfrak{L})$ with respect to the specialization order is the Alexandrov topology.*

Proof. A locale is continuous iff every $\bigsqcap$–irreducible element is $\bigsqcap$–prime. Now let x be $\bigsqcap$–irreducible. Assume that x is not $\bigsqcap$–prime. Then we can find a sequence y_i of elements such that $\bigsqcap y_i \leq x$ while for all i we have $y_i \nleq x$. Since we have a spatial locale, we can choose the y_i to be $\bigsqcap$–irreducible. Now, $y_i \leq x$ iff $\overline{\{p_x\}} \not\supseteq \overline{\{p_{y_i}\}}$ iff $p_{y_i} \notin \overline{\{p_x\}}$. By the next theorem, $\lim \overline{y}$ exists, since $\mathfrak{Spc}(\mathfrak{L})$ is sober. We have $\lim \overline{y} \leq x$ by assumption. Hence $\lim \overline{y} \in \overline{\{p_x\}}$. This shows that the set $\bigcup \overline{\{p_{y_i}\}}$ is not closed, even though it is upward closed. Taking complements, we see that we have a downward closed set which is not open. Hence the topology is not equal to the finest topology. For the other direction, just reason backwards. ∎

It is a rather intricate matter to say exactly what specialization orders admit a sober topology. We will prove a rather useful theorem, stating that the specialization order on a sober space must be closed under lower limits.

Theorem 7.4.12. *Let $\mathfrak{X}$ be sober and $\overline{x} = \langle x_i : i \in \omega \rangle$ be a descending chain of points. Then* $\lim \overline{x}$ *exists in $\mathfrak{X}$.*

Proof. Let $\langle x_i : i \in \omega \rangle$ be a descending sequence of points. Then the sequence $\langle \overline{\{x_i\}} : i \in \omega \rangle$ is ascending, that is, $\overline{\{x_i\}} \subseteq \overline{\{x_j\}}$ if $i \leq j$. Let T be the closure of its union. Assume that $T = S_1 \cup S_2$ for some closed sets S_1 and S_2. Then almost all x_i are in S_1 or in S_2. Hence, by directedness, either all x_i are in S_1 or all x_i are in S_2. Thus $T = S_1$ or $T = S_2$, showing T to be indecomposable. By sobriety, $T = \overline{\{y\}}$ for some y. It is easy to see that $y = \lim \overline{x}$. ∎

Such elements corresponding to limits of chains are $\sqcap$–irreducible elements which are $\bigsqcap$–reducible. We have seen that in lattices of logics a $\bigsqcap$–irreducible logic is the logic of a subdirectly irreducible algebra. By the decomposition theorem, every logic is the intersection of the $\mathsf{Th}\,\mathfrak{A}$ for some subdirectly irreducible $\mathfrak{A}$. We have seen, however, that we are not entitled to conclude that $\mathsf{Th}\,\mathfrak{A}$ is $\bigsqcap$–irreducible if $\mathfrak{A}$ is subdirectly irreducible. Nevertheless, the following theorem can be established.

Theorem 7.4.13. *In the lattices $\mathcal{E}\,\boldsymbol{K}_\kappa$, every logic is the intersection of $\bigsqcap$–irreducible logics.*

Proof. Let Θ be a logic, and $\varphi \notin \Theta$. Then let $S_\varphi = \{\Lambda \supseteq \Theta : \varphi \notin \Lambda\}$. S_φ is not empty, and is closed under direct upper limits. For if for a chain $\Lambda_n \in S_\varphi$, $\Lambda_n \supseteq \Lambda_{n+1}$ we have $\lim \Lambda_n \notin S_\varphi$, then there exists an n_0 such that $\varphi \in \Lambda_{n_0}$. Hence we conclude that there must be a maximal element in S_φ, which we denote by S^*_φ. (This element is not necessarily unique.) Now every proper extension of S^*_φ contains φ, while S^*_φ does not. Hence S^*_φ is $\bigsqcap$–irreducible. Now $\Theta = \bigsqcap\langle S^*_\varphi : \varphi \notin \Theta\rangle$. ◻

Corollary 7.4.14. *For any modal logic* Λ, $\mathcal{E}\,\Lambda$ *is spatial.*

We can use Theorem 7.4.13 for a sharper variant of the representation theorem. Above we have seen that in a sober space the limit of a descending chains always exists. This limit is not $\bigsqcap$–irreducible. Under mild assumptions on the space it is redundant in the representation. Let $Irr(\mathfrak{L})$ be the set of $\bigsqcap$–irreducibles; and $\mathfrak{Irr}(\mathfrak{L}) := \langle Irr(\mathfrak{L}), \leq\rangle$. Then define the topology as before, namely putting $\check{x} = \{p \in Irr(\mathfrak{L}) : p(x) = 1\} = \widehat{x} \cap Irr(\mathfrak{L})$. Let $\mathfrak{ISpc}(\mathfrak{L}) = \langle Irr(\mathfrak{L}), \{\check{x} : x \in L\}\rangle$.

Definition 7.4.15. *A topological space is called a* $\boldsymbol{T_D}$*–**space** if for every point* x *the set* $\{x\}$ *is relatively open in* $\overline{\{x\}}$.

Theorem 7.4.16. *The natural map* $x \mapsto \check{x} : \mathfrak{L} \to \Omega(\mathfrak{ISpc}(\mathfrak{L}))$ *is an isomorphism iff every element is an intersection of* $\bigsqcap$*–irreducible elements.*

We obtain that every lattice of extensions is representable by a $\mathbf{T}_D$–space. Let us make this explicit. With a logic Λ we associate a locale $\mathcal{E}\,\Lambda$, and two spaces. The first is $\mathfrak{Spc}(\mathcal{E}\,\Lambda)$. It consists of the $\sqcap$–irreducible logics, ordered by set inclusion. The topology has as closed sets the sets of the form $\uparrow\Theta$, where $\Theta \supseteq \Lambda$. The second space is $\mathfrak{ISpc}(\mathcal{E}\,\Lambda)$. Its members are the $\bigsqcap$–irreducible logics ordered by set inclusion. Again, the closed sets are the sets of the form $\uparrow\Theta$ for $\Theta \supseteq \Lambda$. It is immediately clear that $\mathfrak{ISpc}(\mathcal{E}\,\Lambda)$ is nothing but the relativization of the space $\mathfrak{Spc}(\mathcal{E}\,\Lambda)$ to the set of $\bigsqcap$–irreducible elements, and so uniquely defined. We have proved in addition that given $\mathfrak{ISpc}(\mathcal{E}\,\Lambda)$, the space $\mathfrak{Spc}(\mathcal{E}\,\Lambda)$ is uniquely determined. It can in fact be constructed. (See the exercises.)

Exercise 237. Show Proposition 7.4.3.

Exercise 238. Prove Lemma 7.4.8.

Exercise 239. Take the lattice $\mathfrak{V} = 1 + (\mathbb{Z} \times \mathbf{2}) + 1$, which is isomorphic to the direct product of $\mathbb{Z}$ with the two–element chain plus a bottom and a top element. Show that $\mathfrak{V}$ is a locale, but not spatial.

Exercise 240. For a topological space $\mathfrak{X}$, the **soberification** is the space $\mathfrak{Spc}(\Omega(\mathfrak{X}))$. Show that for lattices of modal logics, $\mathfrak{Spc}(\mathfrak{L})$ is the soberification of $\mathfrak{ISpc}(\mathfrak{L})$. (This shows how $\mathfrak{Spc}(\Omega(\mathfrak{X}))$ can be recovered from $\mathfrak{ISpc}(\mathfrak{X})$.)

Exercise 241. Show with a counterexample that $\Phi(X, \leq)$ does not necessarily contain *only* sets of the form

$$X - (\uparrow x_1 \cup \uparrow x_2 \cup \ldots \cup \uparrow x_n)$$

(in addition to $\varnothing$). Thus depending on the properties of $\leq$, the topology $\Phi(X, \leq)$ may contain more sets.

7.5. Some Consequences of the Duality Theory

Lattices of (normal) extensions of a logic are algebraic, with the compact elements being the finitely axiomatizable logics; we have seen also that in the lattice $\mathcal{E}\,\Lambda$ every element is the intersection of $\bigsqcap$–irreducible elements. The finitely axiomatizable logics are closed under finite union, just as the compact elements. An infinite join of finitely axiomatizable logics need not be finitely axiomatizable again. Likewise, the finite meet of finitely axiomatizable logics need not be finitely axiomatizable. However, in the case of weakly transitive logics, any finite intersection of finitely axiomatizable logics is again finitely axiomatizable.

Definition 7.5.1. *A locale is* ***coherent*** *if (i) every element is the join of compact elements and (ii) the meet of two compact elements is again compact.*

Coherent locales allow a stronger representation theorem. Let $\mathfrak{L}$ be a locale, $K(\mathfrak{L})$ be the set of compact elements. They form a lattice $\mathfrak{K}(\mathfrak{L}) := \langle K(\mathfrak{L}), \sqcap, \sqcup\rangle$, by definition of a coherent locale. Given $\mathfrak{K}(\mathfrak{L})$, $\mathfrak{L}$ is uniquely identified by the fact that it is the lattice of ideals of $\mathfrak{K}(\mathfrak{L})$.

Lemma 7.5.2. *A locale is coherent iff it is isomorphic to the locale of ideals of a distributive lattice.*

Proof. Let $\mathfrak{L}$ be coherent. Denote by $Id(\mathfrak{K}(\mathfrak{L}))$ the set of ideals in $\mathfrak{K}(\mathfrak{L})$; moreover, let $\mathfrak{I}\mathfrak{K}(\mathfrak{L}) := \langle Id(\mathfrak{K}(\mathfrak{L})), \cap, \bigcup'\rangle$. Here, if I_c, $c \in C$, is a family of ideals, $\bigcup'_{c\in C} I_c$ is the least ideal containing $\bigcup_{c\in C} I_c$. This is a locale; the compact elements are of the form $\downarrow S$ where S is finite. For $x \in L$ let $x^\star$ be defined by $x^\star := \{y \in K(\mathfrak{L}) : y \leq x\}$. We show that this map is an isomorphism from $\mathfrak{L}$ onto $\mathfrak{I}\mathfrak{K}(\mathfrak{L})$. First, $x^\star$ is clearly an ideal. Moreover, if $x \leq y$ then $x^\star \subseteq y^\star$. Conversely, let $I \subseteq \mathfrak{K}(\mathfrak{L})$ be an ideal; then put $I_\star := \bigsqcup I$. If $I \subseteq J$ then $I_\star \leq J_\star$. Both maps are order preserving; it is therefore enough to show that one is the inverse of the other. We show that $x = (x^\star)_\star$ and that $I = (I_\star)^\star$. For the first, observe that $x \geq (x^\star)_\star$ generally holds in a lattice. Since by assumption every element is the union of join compact elements, we also have $(x^\star)_\star \geq x$. This gives $x = (x^\star)_\star$. For the second claim, let y be compact and I an ideal. Suppose that $y \in I$. Then $y \leq \bigsqcup I = I_\star$ and so $y \in (I_\star)^\star$. Conversely, if $y \in (I_\star)^\star$ then $y \leq I_\star = \bigsqcup I$. By compactness, there exists a finite subset $I_0 \subseteq I$ such that $y \leq \bigsqcup I_0$. Since I is closed under finite joins, $\bigsqcup I_0 \in I$, and so $y \in I$. Hence I and $(I_\star)^\star$ contain the same compact elements. So they are identical subsets of $K(\mathfrak{L})$. Now let $\mathfrak{D}$ be a distributive lattice. Then $\mathfrak{I}(\mathfrak{D}) = \langle Id(\mathfrak{D}), \cap, \bigcup'\rangle$ is coherent. An ideal is compact iff it is principal, that is, of the form $\downarrow y$ for some y. Since $\downarrow y \cap \downarrow z = \downarrow(y \sqcap z)$, principal ideals are closed under meet. Furthermore, suppose that I is an ideal. Then $I = \bigcup'_{x\in I} \downarrow x$. So $\mathfrak{I}(\mathfrak{D})$ is a coherent locale. ◻

If we have a lattice homomorphism $\mathfrak{K}(\mathfrak{L}) \to \mathfrak{K}(\mathfrak{M})$ then this map can be extended uniquely to a homomorphism of locales $\mathfrak{L} \to \mathfrak{M}$. Not all locale homomorphisms arise this way, and so not all locale maps derive from lattice homomorphisms. Hence call a map $f : \mathfrak{L} \to \mathfrak{M}$ **coherent** if it maps compact element into compact elements.

Theorem 7.5.3. *The category* DLat *of distributive lattices and lattice homomorphisms is dual to the category* CohLoc *of coherent locales with coherent maps.*

Now if Λ is weakly transitive, the intersection of two finitely axiomatizable logics is again finitely axiomatizable. Now, a logic is compact in $\mathcal{E}\,\Lambda$ iff it is finitely axiomatizable over Λ. We conclude the following theorem.

Proposition 7.5.4. *Let Λ be weakly transitive. Then $\mathcal{E}\,\Lambda$ is coherent.*

The converse need not hold. In Chapter 9 we will see that $\mathcal{E}\,\mathbf{K.alt}_1$ is coherent (because every logic in this lattice is finitely axiomatizable) but the logic is not weakly transitive.

The next property of lattices has already made an appearance earlier, namely *continuity*. It is connected with a natural question about axiomatizability of logics.

Definition 7.5.5. *Let $\mathfrak{L}$ be a complete lattice. A set $X \subseteq L$ is a* **generating set** *if for every member of L is the join of a subset of X. $\mathfrak{L}$ is said to have a* **basis** *if there exists a least generating set. Moreover, X is a* **strong basis** *for $\mathfrak{L}$ if every element has a nonredundant representation, that is, for each x there exists a minimal $Y \subseteq X$ such that $x = \bigsqcup Y$.*

Theorem 7.5.6. *Let $\mathfrak{L}$ be a locale. $\mathfrak{L}$ has a basis iff (i) $\mathfrak{L}$ is continuous and (ii) every element is the meet of $\bigsqcap$–irreducible elements. $\mathfrak{L}$ has a strong basis iff it has a basis and there exists no infinite properly ascending chain of $\bigsqcap$–prime elements.*

Proof. Assume that (i) and (ii) hold. Let I be the set of $\bigsqcap$–irreducible elements. Since $\mathfrak{L}$ is continuous, I is also the set of $\bigsqcap$–prime elements. Hence every element of I splits $\mathfrak{L}$. Take an element $x \in L$. Let $J := \uparrow x \cap I$. Then put $K := I - J$ and $x_0 := \mathfrak{L}/K = \bigsqcup_{y \in K} \mathfrak{L}/y$. Since for no $y \in K$ it holds that $x \leq y$, we have $x \geq \mathfrak{L}/y$; and hence $x \geq x_0$. We also have $x_0 \leq x$ since $x_0 \leq q$ for all $q \in J$, and so also $x_0 \leq \bigsqcap J = x$. Hence x is a union of elements of the form $\mathfrak{L}/y$, y $\bigsqcap$–irreducible. We have to show that the set X of $\bigsqcup$–irreducible elements is minimal. The elements of X are also $\bigsqcup$–prime and hence compact. Moreover, let $Y \subsetneq X$. Say, $y \in X - Y$. Then for no subset $U \subseteq Y$ we can have $\bigsqcup U = y$, since y is $\bigsqcup$–irreducible. Thus X is a basis. Now let $\mathfrak{L}$ have a basis, X. Let $x \in X$. Assume that $x = \bigsqcup Y$ for some $Y = \{y_i : i \in I\}$. By assumption, each y_i is the join of some set X_i, $X_i \subseteq X$. Put $X_0 := \bigcup_{i \in I} X_i$. Then $x = \bigsqcup X_0$. Suppose $x \notin X_0$. Then $X - \{x\}$ is a generating set, contradicting our assumption on X. Consequently, X consists of $\bigsqcup$–irreducible elements. $\mathfrak{L}$ is a locale, and so X consists of the $\bigsqcup$–prime elements. Hence, each element $y \in X$ has a splitting companion $y_\star$. Now take $x \in L$. Put $U := (X - \uparrow x)$. Consider the element $x^0 := \bigsqcap_{y \in U} y_*$. Clearly, $x \leq x^0$; for if $y \in U$ then $x \nleq y$, so $y_\star \geq x$. Now, if $x \neq x^0$ there exists a $y \in X$ such that $y \leq x^0$ but $y \nleq x$. Then $x \geq y_\star$. Hence $x^0 \geq y_\star$. Contradiction. So, $x = x^0$. Hence, (ii) is fulfilled. Moreover, each element is the meet of prime elements. This implies that (i) holds. For assume that u is $\bigsqcap$–irreducible and the intersection of a set of $\bigsqcap$–prime elements. Then this set is one–membered, and u is therefore also $\bigsqcap$–prime.

Now let us turn to strong bases. Assume that $\mathfrak{L}$ has a base and let $\langle x_i : i \in \omega\rangle$ be a strictly ascending chain of $\bigsqcap$–prime elements. Then $\langle \mathfrak{L}/x_i : i \in \omega\rangle$ is a strictly ascending chain of $\bigsqcup$–prime elements. Let y be its limit. Then there exists no minimal set Y of $\bigsqcup$–primes whose join is y. Thus $\mathfrak{L}$ fails to have a strong basis. Assume on the other hand that there exists no such chain. Then for each upper closed set P of $\bigsqcap$–prime elements the set $max(P) := \{x \in P : (\forall y \in P)(y \geq x \Rightarrow y = x)\}$ is well–defined and is the least set Q such that $\downarrow Q = P$. ◻

Proposition 7.5.7. *Let $\mathfrak{L}$ be a locale with a strong basis. Then the elements of $\mathfrak{L}$ are in one–to–one correspondence with antichains in $\mathfrak{I}\mathfrak{S}\mathfrak{pc}(\mathfrak{L})$.*

Proof. By Theorem 7.2.7 the map $p \mapsto p^\star := \mathfrak{L}/p$ induces an order isomorphism from the poset of $\bigsqcap$–primes onto the poset of $\bigsqcup$–primes, whose inverse is $q \mapsto q_\star$. Let x be an element. Then let Y be a minimal set of $\bigsqcup$–primes whose join is x. Then Y is an antichain. Then $Y_\star$ is an antichain of $\bigsqcap$–primes. Put $x_o := \bigsqcap Y_\star$. Conversely, given an antichain Z of $\bigsqcup$–primes, let $Z^o := \bigsqcup Z^\star$. Then $x = (x_o)^o$ as well as $(Z^o)_o$, so this is a bijection. ◻

Theorem 7.5.8. *Let Λ be a modal logic. Then $\mathcal{E}\,\Lambda$ has a basis iff $\mathcal{E}\,\Lambda$ is continuous.*

Since continuous lattices are the exception in modal logic, most extension lattices do not have a basis. We can sharpen the previous theorem somewhat to obtain stricter conditions on continuity.

Corollary 7.5.9. *Let Λ be weakly transitive and have the finite model property. Then the following are equivalent.*

1. *$\mathcal{E}\,\Lambda$ has a basis.*
2. *$\mathcal{E}\,\Lambda$ has a strong basis.*
3. *Every extension of Λ has the finite model property.*
4. *Every extension of Λ is the join of co–splitting logics.*
5. *Every join of co–splitting logics has the finite model property.*

Proof. The most interesting part is perhaps the fact that if we have a basis, then we already have a strong basis. But this follows, because there is no infinite strictly ascending chain of $\bigsqcap$–irreducibles. For the $\bigsqcap$–irreducible logics are the logics of finite frames, and for finite frame $\mathfrak{f}$, $\mathfrak{g}$ we have that $\mathsf{Th}\,\mathfrak{g} \supseteq \mathsf{Th}\,\mathfrak{f}$ implies $\sharp g \leq \sharp f$. If Λ is weakly transitive then every finite subdirect irreducible algebra induces a splitting. On the other hand, if Λ has the finite model property, then no more elements can induce splittings. The equivalence now follows directly. ◻

Corollary 7.5.10. *Let $\mathsf{Alg}\,\Lambda$ be locally finite. Then $\mathcal{E}\,\Lambda$ is continuous.*

Proof. Since $\mathsf{Alg}\,\Lambda$ is locally finite, every extension of Λ has the finite model property (see Theorem 4.8.7). Moreover, the one–generated Λ–algebra, $\mathfrak{Fr}_\Lambda(\{p\})$, is finite. Now consider the elements $\boxplus p$, $\boxplus$ a compound modality. Up to equivalence there exist only finitely many of them. So a largest compound modality exists. So, Λ is weakly transitive and has the finite model property. By Corollary 7.5.9, $\mathcal{E}\,\Lambda$ has a basis, and by Theorem 7.5.8 it is therefore continuous. □

The converse does not hold. The lattice $\mathcal{E}\,$**S4.3** is continuous but **S4.3** fails to be locally finite. Now, finally, consider the question of finite axiomatizability. In lattices which have a basis, this can be decided rather easily. Given a set S, a relation $\prec$ is called a **well–partial order** (**wpo**) if it is a partial order and there are no infinite strictly descending chains, and no infinite antichains. There is a rather famous theorem by J. B. Kruskal which says that the set of finite trees under the embedding–ordering is a wpo ([**135**]).

Theorem 7.5.11. *Let $\mathfrak{L}$ be locale with a basis. Then the following are equivalent.*

1. $\geq$ *is a well partial order on* $\mathfrak{I}\mathfrak{S}\mathfrak{pc}(\mathfrak{L})$.
2. $\mathfrak{L}$ *has a strong basis and no infinite antichains exist in* $\mathfrak{I}\mathfrak{S}\mathfrak{pc}(\mathfrak{L})$.
3. *Every element is a finite union of* $\bigsqcup$*–primes.*
4. $\mathfrak{L} \cong \mathfrak{K}(\mathfrak{L})$.
5. *Every strictly ascending chain in* $\mathfrak{L}$ *is finite.*

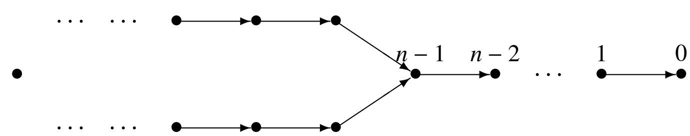
Proof. If $\mathfrak{L}$ has a basis it is continuous and so the properties on the poset of $\bigsqcap$–irreducible elements can equivalently be checked on the poset of $\bigsqcup$–irreducible elements. Let (1.) be the case. Then (2.) holds by Theorem 7.5.6. Now consider an element x of $\mathfrak{L}$ and let Y be a minimal set of $\bigsqcup$–primes such that $x = \bigsqcup Y$. Then Y is an antichain, and so Y is finite. Now, if every element is a finite union of join–primes, then every element is compact. Now let $\langle x_i : i \in \omega\rangle$ be an infinite strictly ascending chain. Then its limit cannot be compact. Finally, assume that there are no infinite strictly ascending chains in $\mathfrak{L}$. Then there are no strictly ascending chains in $Irr(\mathfrak{L})$. Furthermore, if X is an infinite antichain, then we can choose an infinite ascending chain of subsets of X corresponding to an infinite ascending chain of elements in $\mathfrak{L}$. Hence $\geq$ is a wpo, as required. □

Corollary 7.5.12. ($\kappa \leq \aleph_0$.) *Let $\mathcal{E}\,\Lambda$ have a strong basis. Then the following are equivalent.*

1. *Every extension of $\mathcal{E}\,\Lambda$ is finitely axiomatizable.*
2. *$\mathcal{E}\,\Lambda$ is finite or countably infinite.*
3. *There exists no infinite set of incomparable splitting logics.*

There are lattices of modal logics in which there are infinite ascending chains of irreducible elements and infinite antichains. The latter has been shown in [**61**] (see exercises below). Another example is the logics of Chapter 2.6. Here we will produce an infinite ascending chain of irreducible elements; a first proof of this fact was given by Blok [**23**]. Our example will allow to prove a number of very interesting negative facts about modal logics in general. Let $\mathfrak{F}$ and $\mathfrak{G}$ be frames. Then let $\mathfrak{F} \otimes \mathfrak{G}$

FIGURE 7.6. $\bullet \otimes (\underline{\omega} \oplus \underline{\omega}) \otimes \underline{n}$



be the following frame.

$$\begin{array}{lcl} h & := & f \times \{0\} \cup g \times \{1\} \\ \lhd^h & := & \quad \{\langle\langle x, 0\rangle, \langle y, 0\rangle\rangle : x, y \in f, x \lhd^f y\} \\ & & \cup \quad \{\langle\langle x, 1\rangle, \langle y, 1\rangle\rangle : x, y \in g, x \lhd^g y\} \\ & & \cup \quad \{\langle\langle x, 0\rangle, \langle y, 1\rangle\rangle : x \in f, y \in g\} \\ \mathbb{H} & := & \{a \times \{0\} \cup b \times \{1\} : a \in \mathbb{F}, b \in \mathbb{G}\} \\ \mathfrak{F} \otimes \mathfrak{G} & := & \langle h, \lhd^h, \mathbb{H}\rangle \end{array}$$

Moreover, if α is an ordinal number, let $\underline{\alpha} := \langle \alpha, \ni\rangle$. We are interested in the logic of the frames of the form $\bullet \otimes (\underline{\alpha} \oplus \underline{\beta}) \otimes \underline{\gamma}$ for infinite α and β. In Figure 7.6 the frame $\bullet \otimes (\underline{\omega} \oplus \underline{\omega}) \otimes \underline{n}$ is shown.

Consider the following formulae.

$$\begin{array}{lcl} \varphi_0 & := & p_0 \wedge \neg p_1 \wedge \Box(\neg p_0 \wedge \neg p_1) \\ \varphi_1 & := & \neg p_0 \wedge p_1 \wedge \Box(\neg p_0 \wedge \neg p_1) \end{array}$$

The logic $\mathbf{G}.\Omega_2$ is defined as follows.

$$\begin{array}{ll} \mathbf{G}.\Omega_2 := \mathbf{K}.\mathbf{G} & \oplus\, \Diamond p_0 \wedge \Diamond p_1 \wedge \Diamond p_2 \to \bigvee_{i<j<3} \Diamond(p_i \wedge p_j) \vee \bigvee_{i \neq i} \Diamond(p_i \wedge \Diamond p_j) \\ & \oplus\, \Diamond\varphi_0 \wedge \Diamond\varphi_1 \to \Diamond\Diamond\varphi_0 \wedge \Diamond\Diamond\varphi_1 \\ & \oplus\, \Diamond\varphi_0 \wedge \Diamond\varphi_1 \to \Box\neg(\Diamond\varphi_0 \wedge \Diamond\varphi_1) \end{array}$$

THEOREM 7.5.13. *Every extension of* $\mathbf{G}.\Omega_2$ *is complete with respect to frames of the form (i)* $\bullet \otimes (\underline{\omega} \oplus \underline{\omega}) \otimes \underline{\alpha}$, $\alpha \leq \omega$, *or (ii)* $\underline{\alpha}$, $\alpha \leq \omega$.

The Theorem 7.5.13 is proved as follows. Every extension of $\mathbf{G}.\Omega_2$ is complete with respect to simple noetherian frames, by Theorems 8.6.14 and 8.6.15 of Section 8.6. Moreover, it is easy to see that the Kripke–frames underlying the reduced canonical frames for $\mathbf{G}.\Omega_2$ have the structure $\bullet \otimes (\underline{\alpha} \oplus \underline{\beta}) \otimes \underline{\gamma}$, for certain ordinal numbers α, β and γ. Furthermore, if α and β are nonzero, they must be infinite. Let $\Lambda \supseteq \mathbf{G}.\Omega_2$, and let $\varphi \notin \Lambda$. Then there exists a model $\langle \mathfrak{F}, \beta, x\rangle \vDash \neg\varphi$ based on a generated subframe of a reduced weak canonical frame. Let $y \in f$; define $C(y) := \{\chi \in sf(\varphi) : y \vDash \chi\}$. Call y φ–**maximal** if for every $z \rhd y$ such that $C(z) = C(y)$

also $z \lhd y$. Let D be the set of points which are φ–maximal. D has $\leq 2 \cdot \sharp sf(\varphi)$ points. Consider the subframe $\mathfrak{D}$ of $\mathfrak{F}$ based on D. It is cofinal, that is, it contains all points of depth 0. Let $\gamma(p) := \beta(p) \cap D$. There exists a weak successor $y \in D$ of x and $\langle \mathfrak{D}, \beta, y \rangle \vDash \varphi$. $\mathfrak{D}$ can be partitioned into four possibly empty pairwise disjoint frames $\mathfrak{W}$, $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{Z}$, each linearly ordered by $\lhd$ such that $\mathfrak{D} \cong \mathfrak{Z} \oslash (\mathfrak{A} \oplus \mathfrak{B}) \oslash \mathfrak{W}$. Observe that if $\mathfrak{A}$ is empty, we can choose $\mathfrak{B}$ in such a way that it is empty, too.
Case 1. $\mathfrak{A}$ or $\mathfrak{B}$ is empty. Then $\mathfrak{D}$ is linear. This case is rather straightforward.
Case 2. Both $\mathfrak{A}$ and $\mathfrak{B}$ are nonempty. Then $\mathfrak{Z}$ is nonempty and so $\mathfrak{Z} \cong \boxed{\bullet}$. Let $\mathfrak{G} = \mathfrak{Z} \oslash ((\underline{\alpha}_1 \oslash \mathfrak{A}) \oplus (\underline{\alpha}_2 \oslash \mathfrak{B})) \oslash (\underline{\alpha}_3 \oslash \mathfrak{W})$ for certain ordinal numbers α_1, α_2 and α_3. $\mathfrak{G}$ is a $\mathbf{G}.\Omega_2$–frame. Let $\mathfrak{E}$ be the subframe of $\mathfrak{G}$ based on the union $W \cup A \cup B \cup Z$. It is possible to extend the valuation γ on $\mathfrak{D}$ to a valuation δ on $\mathfrak{F}$ such that (i) $\delta(p) \cap E = \gamma(p)$, (ii) each $\delta(p)$ is a finite union of intervals and singletons, (iii) $\langle \mathfrak{F}, \delta, y \rangle \vDash \neg\varphi$. (Namely, let $x \in \delta(p)$ for $x \in \alpha_1$ iff for the root y_1 of $\mathfrak{A}$, $y \in \gamma(p)$. Let $x \in \delta(p)$ for $x \in \alpha_2$ iff for the root y_2 of $\mathfrak{B}$, $y \in \gamma(p)$, and let $x \in \delta(p)$ for $x \in \alpha_3$ iff for the root y of $\mathfrak{W}$, $y_3 \in \gamma(p)$.)

Now, the size of $\mathfrak{D}$ depends on φ. However, it is always finite. Hence we obtain that the logic of $\bullet \oslash (\underline{\alpha} \oplus \underline{\beta}) \oslash \underline{\gamma}$, where $\alpha \leq \beta$ is identical to the logic of $\bullet \oslash (\underline{\alpha'} \oplus \underline{\beta'}) \oslash \underline{\gamma'}$, where $\alpha' \leq \beta'$ if (i) $\alpha = \alpha'$, $\beta = \beta'$ or $\alpha = \alpha'$ and $\beta, \beta' \geq \omega$ or $\alpha, \alpha', \beta, \beta' \geq \omega$, and (ii) $\gamma = \gamma'$ or $\gamma, \gamma' \geq \omega$. This completes the proof of Theorem 7.5.13. We deduce the following.

Theorem 7.5.14. $\mathcal{E}\ \boldsymbol{G}.\Omega_2 \cong \omega + \mathit{2} + \omega^{op}$.

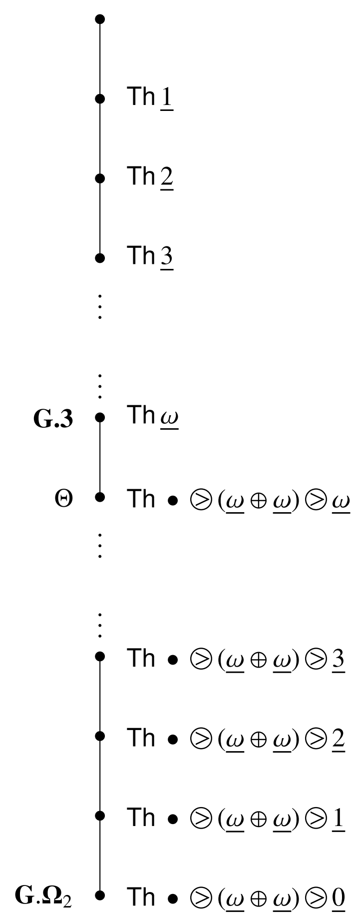
Proof. It is a matter of direct verification (using some formulae) that the logics $\mathsf{Th}\,\underline{\gamma}$, $\gamma \leq \omega$, as well as $\mathsf{Th}\ \bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{\gamma}$, $\gamma \leq \omega$, are pairwise distinct. Now, the theorem is established by the following facts.

1. $\underline{m} \rightarrowtail \underline{n}$ iff $m \leq n$.
2. $\bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{m} \twoheadrightarrow \bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{n}$ iff $m \geq n$.
3. $\mathsf{Th}\,(\bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{n}) \subseteq \mathsf{Th}\,(\bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{\omega})$.
4. $\underline{n} \rightarrowtail \underline{\omega} \rightarrowtail \bullet \oslash (\underline{\omega} \oplus \underline{\omega}) \oslash \underline{\omega}$.

This ends the proof. ∎

Call a set Δ of formulae **independent** if for every $\delta \in \Delta$ we have $\delta \notin \mathbf{K} \oplus (\Delta - \{\delta\})$. (For example, a basis is an independent set.) A logic Λ is **independently axiomatizable** if there exists an independent set Δ such that $\Lambda = \mathbf{K} \oplus \Delta$. Every finitely axiomatizable logic is independently axiomatizable. It has been shown in Chagrov and Zakharyaschev [**42**] that there exists a logic which is not independently axiomatizable. Furthermore, Kracht [**125**] gives an example of a logic which is not finitely axiomatizable, but all its proper extensions are. Such a logic is called **pre–finitely axiomatizable**. Here is a logic that has both properties.

FIGURE 7.7. $\mathcal{E}\,\mathbf{G}.\Omega_2$



THEOREM 7.5.15. *The logic of the frame $\bullet \circledcirc (\underline{\omega} \oplus \underline{\omega}) \circledcirc \underline{\omega}$ is pre–finitely axiomatizable. It splits the lattice of extensions of $\boldsymbol{G}.\Omega_2$. Moreover, it is not axiomatizable by a set of independent formulae.*

PROOF. The first two claims are immediate. For the last, let $\Theta := \mathsf{Th}\ \bullet \circledcirc (\underline{\omega} \oplus \underline{\omega}) \circledcirc \underline{\omega}$. Let Δ be a set of formulae such that $\Theta = \mathbf{K} \oplus \Delta$. $\mathbf{G}.\Omega_2$ is finitely axiomatizable. Hence there exists a finite set $\Delta_0 \subseteq \Delta$ such that $\mathbf{G}.\Omega_2 \subseteq \mathbf{K} \oplus \Delta_0 \subsetneq \Theta$. Let

$\delta, \delta' \in \Delta - \Delta_0$ be two different formulae. Then either $\mathbf{K} \oplus \Delta_0 \oplus \delta \subseteq \mathbf{K} \oplus \Delta_0 \oplus \delta'$ or $\mathbf{K} \oplus \Delta_0 \oplus \delta' \subseteq \mathbf{K} \oplus \Delta_0 \oplus \delta$. Hence the set Δ is not independent. ◻

Moreover, with this logic we have an example of an algebra that induces a splitting but is not finitely presentable. This shows that the generality of Theorem 7.3.11 is really needed.

Theorem 7.5.16. *Let $\mathfrak{A}$ be the algebra generated by the singleton sets of $\bullet \otimes (\underline{\omega} \oplus \underline{\omega}) \otimes \underline{\omega}$ is not finitely presentable. Its logic splits $\mathcal{E}$ $\boldsymbol{G}.\Omega_2$.*

Proof. Take the freely 2–generated $\mathbf{G}.\Omega_2$–algebra, with the generators a and b, and consider the open filter F generated by the set E.

$$\begin{array}{rl} E := & \{a \to -b \cap \blacksquare(-a \cap -b), b \to -a \cap \blacksquare(-a \cap -b)\} \\ \cup & \{(a \cup b) \to \blacklozenge(\blacksquare^{n+1}0 \cap -\blacksquare^{n}0) : n \in \omega\} \end{array}$$

We will show that $\mathfrak{A} \cong \mathfrak{Fr}_\Theta(\{a, b\})/F$ and that for every finite subset E_0 of E and filter F_0 generated by E_0, $\mathfrak{A}$ is actually not isomorphic to the quotient $\mathfrak{Fr}_\Theta(\{a, b\})/F_0$. The second claim is easy. For if $E_0 \subseteq E$ is finite, for some n, $(a \cup b) \to \blacklozenge(\blacksquare^{n+1}0 \cap -\blacksquare^{n}0) \notin E_0$. It is consistent to add $(a \cup b) \cap \blacksquare^{n+1}0$, that is to say, adding that formula does not generate the trivial filter. So, no finite subset is enough to generate the filter of E. Now we show the first claim. Let us look at the reduced 2–canonical frame rather than the 2–canonical frame. (See Section 8.6 for a definition.) Let W be the set of (noneliminable) points satisfying $-a \cap -b \cap \blacksquare(-a \cap -b)$. This set is the set of 0–definable points. It is not hard to see that it is linearly ordered by $\lhd$. Moreover, it is not finite, by choice of E. Moreover, we can show that every world of infinite depth is eliminable, and so that $\langle W, \lhd\rangle$ is isomorphic to $\underline{\omega}$. To see that, it is enough to show that for every formula $\varphi(p, q)$, $\varphi(a, b)$ is either finite or cofinite. This is shown by induction on φ. Now, let u be of depth ω in W, and let $\psi(a, b)$ hold at a point of infinite depth in W. We may assume that ψ is strictly simple (see Section 1.7 for a definition). Any set $(\Diamond\chi)(a, b)$ containing u is cofinite, and so is every set $(\Box\chi)(a, b)$ containing u. Now, a nonmodal formula $\mu(a, b)$ containing u holds at all points of W. So, u is eliminable. However, the frame consists only of noneliminable points. This shows that $\langle W, \lhd\rangle \cong \underline{\omega}$. Now, every world of $-W$ must see all worlds of W. Moreover, it must be in $a \cup b \cup \blacklozenge a \cup \blacklozenge b$. Let A be the set of worlds in $(a \cup \blacklozenge a) \cap -\blacklozenge b$, B the set of worlds in $(b \cup \blacklozenge b) \cap -\blacklozenge a$. Finally, $Z := -(W \cup A \cup B)$. Every world of Z is in $\blacklozenge a$ and in $\blacklozenge b$. For otherwise it contains $\blacksquare - a$ or $\blacksquare - b$. Suppose it contains both; then it is in W. Suppose it contains only one, say $\blacksquare - b$. Then it does not contain $\blacksquare - a$, so it contains $\blacklozenge a$. But then it is a member of A. Contradiction. By the axioms of $\mathbf{G}.\Omega_2$, if $x \in Z$, then x has no successor in Z. So, Z has one member only and it generates the frame. Both A and B are linearly ordered. This follows from two facts. (i) No member of A sees a member of B, and no member of B sees a member of A. (ii) There are no three incomparable points. Both A and B must be infinite, by the axioms of $\mathbf{G}.\Omega_2$. Hence, by the same argument as for W, they are isomorphic to $\underline{\omega}$. ◻

Exercise 242. Show that a coherent locale is spatial.

Exercise 243. Let $\mathbf{Grz}_3$ be the logic of **Grz**–structures of depth 3. Show that $\mathbf{Grz}_3$ is locally finite but that $\mathcal{E}\,\mathbf{Grz}_3$ is uncountable. *Hint.* The frames for $\mathbf{Grz}_3$ are simply characterized by the fact that they are a poset in which no strictly ascending chain of length 4 exists.

Exercise 244. Show that there exists a logic Θ such that all extensions of Θ are finitely axiomatizable over Θ but not all extensions are finitely axiomatizable over **K**.

7.6. Properties of Logical Calculi and Related Lattice Properties

In this section we will map out the distribution of logics that have a certain property in terms of their closure under lattice operations. Some easy facts are the following. Given a class $\mathfrak{X}$ of frames, the logics which are $\mathfrak{X}$–complete are closed under (infinite) intersection. The set of logics which are $\mathfrak{X}$–persistent are closed under (infinite) union. Moreover, it can be shown that logics which are $\mathfrak{X}$–elementary for some modal class $\mathfrak{X}$ of frames form a sublocale in the locale $\mathcal{E}\,\Lambda$. We will show here a special case, the most important one, namely that of the Sahlqvist logics (see [**127**]). The general fact is proved the same way. Recall that $\mathfrak{Sq}_n$ denotes the class of logics axiomatizable by a set of Sahlqvist axioms of Sahlqvist rank n.

Theorem 7.6.1. ($\kappa < \aleph_0$.) *The logics $\mathfrak{Sq}_n$ form a sublocale of the locale $\mathcal{E}\,\mathbf{K}_\kappa$ of κ–modal logics.*

Proof. The only thing which is not straightforward is the closure under meet. To this end take two elementary Sahlqvist formulae of rank n, $(\forall x)\alpha(x)$and $(\forall x)\beta(x)$. We want to show that $(\forall x)\alpha(x) \vee (\forall x)\beta(x)$ is again Sahlqvist of rank n. Define formulae γ_k, $k \in \omega$, by

$$\gamma_k := (\forall x)([\bigwedge_{j\leq k}(\forall y \rhd^j x)\alpha(y)] \vee [\bigwedge_{j\leq k}(\forall y \rhd^j x)\beta(y)])$$

where $x \lhd^j y$ iff there exists a path of length j from x to y. Observe now that

$$\mathfrak{F} \vDash (\forall x)\alpha(x) \vee (\forall x)\beta(x) \text{ iff } \mathfrak{F} \vDash \{\gamma_k : k \in \omega\}$$

For if $\mathfrak{F} \vDash (\forall x)\alpha(x) \vee (\forall x)\beta(x)$ then $\mathfrak{F} \vDash (\forall x)\alpha(x)$ or $\mathfrak{F} \vDash (\forall x)\beta(x)$. Assume withouth loss of generality the first. Thus $\mathfrak{F} \vDash (\forall x)(\forall y \rhd^j x)\alpha(y)$ for all j and consequently $\mathfrak{F} \vDash \gamma_k$ for all k. For the converse assume $\mathfrak{F}$ satisfies all γ_k. Take a world $w \in \mathfrak{F}$. Then for an infinite number of $k \in \omega$ we have either

$$\mathfrak{F} \vDash \{(\forall y \rhd^j x)\alpha(y)[w] : j \leq k\}$$

or

$$\mathfrak{F} \vDash \{(\forall y \rhd^j x)\beta(y)[w] : j \leq k\}$$

Let the first be the case. Denote by $\mathfrak{G}$ be the subframe generated by w in $\mathfrak{F}$. Then $\mathfrak{G} \vDash (\forall x)\alpha(x)$. Consequently,

$$\mathfrak{G} \vDash (\forall x)\alpha(x) \vee (\forall x)\beta(x)$$

This holds for all generated subframes, and so it holds for $\mathfrak{F}$ as well, since $\alpha(x), \beta(x)$ are restricted. Secondly, for every k, γ_k is Sahlqvist of rank n. Two cases need to be distinguished. First case is $n = 0$. Then γ_k is constant and so Sahlqvist of rank 0. Second case $n > 0$. Since the formula begins with a universal quantifier and $(\forall y \rhd^j x)$ is a chain of universal quantifiers, the rank of γ_k is the maximum of the ranks of α and β, hence at most n. ◻

As a particular application of the splitting theorem we will prove that there are effective means to axiomatize tabular logics and that the tabular logics form a filter in the lattice of logics. This is by no means trivial to show, and requires some advanced methods, despite the seemingly simple character of the generating algebra. In general, these results have been shown to hold in congruence distributive varieties by Baker in [**2**]. However, his proof uses abstract algebra, whereas here we can make use of geometric tools. Moreover, some results can be sharpened. Crucial to the analysis are two families of logics, the family of weakly transitive logics and logics of bounded alternativity. Recall from Theorem 3.2.12 that logics of bounded alternativity are canonical. Consider now a logic which is m–transitive and satisfies $\mathbf{alt}_n$ for some $m, n \in \omega$. Then each rooted subframe of the canonical frame has at most $\frac{n^{m+1}-1}{n-1}$ many points. For, by induction, the kth wave consists of at most n^k many points and the transit of a point is the union of the k–waves for $k \leq m$, by m–transitivity. So, there are only finitely many non–distinct frames in the canonical frame.

Theorem 7.6.2. *A logic is tabular iff it is of bounded alternativity and weakly transitive.*

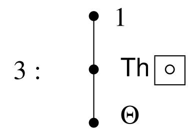
Proof. Let Θ be tabular, say, $\Theta = \mathsf{Th}\,\mathfrak{f}$ for some finite $\mathfrak{f}$. Let $n = \sharp f$. Then Θ is n–transitive, and each point has at most n successors. Conversely, let Θ be weakly transitive and of finite alternativity. Then Θ is canonical, hence complete. Each rooted subframe for Θ is finite and bounded in size by $1 + n + n^2 + \ldots + n^m$. So there are up to isomorphic copies finitely many rooted frames for Θ. Their disjoint union, $\mathfrak{g}$, is finite and $\Theta = \mathsf{Th}\,\mathfrak{g}$. ◻

Corollary 7.6.3. *The tabular logics form a filter in the lattice of modal logics.*

How can we axiomatize the logic of a single frame $\mathfrak{f}$ if it is finite? First, we know that there is an axiom of weak transitivity satisfied by $\mathfrak{f}$, say $\mathbf{trs}_m$. Hence, we know that all finite frames for $\mathsf{Th}\,\mathfrak{f}$ induce a splitting of the lattice of extensions of $\mathbf{K}.\mathbf{trs}_m$. Moreover, there is an axiom $\mathbf{alt}_n$ satisfied by $\mathfrak{f}$, so in fact all rooted frames are finite. Hence we have the following representation.

$$\mathsf{Th}\,\mathfrak{f} = \mathbf{K}.\mathbf{alt}_m.\mathbf{trs}_n/N$$

Figure 7.8.



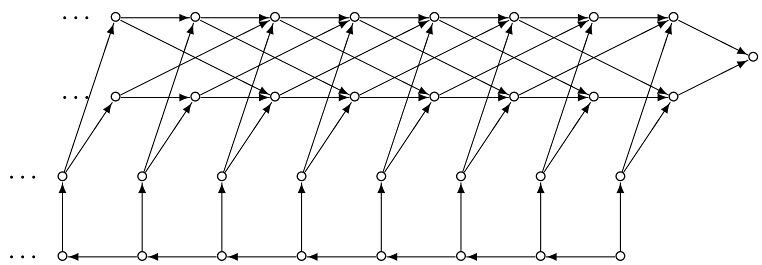
where N is the set of logics of rooted frames $\mathfrak{g}$ which are not generated subframes of p–morphic images of $\mathfrak{f}$ and which are frames for $\mathbf{K.alt}_m.\mathbf{trs}_n$. There are finitely many of them.

Proposition 7.6.4. *Tabular logics are finitely axiomatizable and of finite codimension.*

The converse does not hold. We give an example in monomodal logic, namely the logic of the veiled recession frame. It is of codimension 2. Its lattice of extensions is depicted in Figure 7.8; it is isomorphic to $\langle 3, \leq\rangle$. Let $\mathfrak{r}$ denote the frame defined by $\mathfrak{r} := \langle \omega, \lhd\rangle$ where $n \lhd m$ iff $n \leq m+1$ (see [**25**]). Let $\mathfrak{R}$ be the algebra of the finite and cofinite subsets of ω and $\Theta := \mathsf{Th}\,\mathfrak{B} = \mathbf{K}(\Box\Diamond p \to \Diamond\Box^2 p,\ \Box p \to p,\ \Diamond p \wedge \Box(p \to \Box p) \to p)$. $\mathfrak{R}$ is 1–generated; simply take $\{0\} \subseteq \omega$. It is not entirely simple to see that the logic of the veiled recession frame is $\bigsqcap$–prime in its own lattice. However, consider the structure of the frame $\mathfrak{Fr}_\Theta(1)$. It is constructed by taking a product of $\mathfrak{R}$, each factor representing a possible valuation of p into the algebra $\mathbb{R}$. The canonical value of p in that frame is then the sequence of the values $\beta(p)$. Crucial is the fact that a valuation sends p to a non–trivial set, that is, a set $\neq \varnothing$ and $\neq r$, iff it does not send $\Diamond p \wedge \neg p$ to $\varnothing$. Notice on the other hand that a set generates $\mathbb{R}$ exactly if it is nontrivial. (This will be an exercise.) Now consider an extension $\Lambda \supsetneq \Theta$. Then there is a nontrivial map $h : \mathfrak{Fr}_\Theta(1) \twoheadrightarrow \mathfrak{Fr}_\Lambda(1)$. By the fact that Θ is the logic of the veiled recession frame, $\mathfrak{Fr}_\Lambda(1)$ must not contain any generated component looking like the recession frame. Hence, all components have been killed by h, that is, $h(\Diamond p \wedge \neg p) = \varnothing$, since the latter formula defines exactly the generated components looking like veiled recession frames. To put this differently, $h(\Diamond p \to p) = 1$, that is, $\Diamond p \to p \in \Lambda$, which means that Λ contains the logic of the one–point reflexive frames. Hence $\mathfrak{R}$ is prime in $\mathcal{E}\,\Theta$ and induces a splitting $\mathcal{E}\,\Theta/\Theta = \mathsf{Th}\,\boxed{\circ}$. It is no coincidence that the counterexample is a logic of depth 2, by the Theorem 2.9.9. We will show later that with two operators there are $2^{\aleph_0}$ many logics of codimension 1.

Our last example concerns the problem of closure under union of completeness properties. We have seen that tabular logics are closed under union. Blok [**21**] has first given an example of two logics which have the finite model property such that their join does not. Here we will show that there are complete logics such that their join is incomplete. The example is based on [**62**], one of the first examples of an incomplete logic. (The article by Fine appeared in the same edition of the journal

FIGURE 7.9.



where THOMASON published his [**207**]. The two results have been independently obtained.) Consider the frame of Figure 7.9. The frame is denoted by $\mathfrak{f}$. The first three rows of points are of finite depth, the lowest consists of points of infinite depth. Notice that the points of infinite depth form an infinite ascending chain. We refer to the first two rows as the **spine** of the frame, the points of the third row are called **needles**. The lowest row is the **head**. Consider the algebra $\mathbb{F}$ generated by the finite subsets of the spine. This set contains all finite sets of needles. It can be shown inductively that each element of $\mathbb{F}$ has a maximal point. That is, if $a \in \mathbb{F}$ then for every x there exists a $y \in a$ such that $x \lhd y$ and for all $y \lhd z \in a$ we have $z = y$. This holds of the generating sets, and if it holds of set a, b then of $-a$, of $\blacklozenge a$, and of $a \cup b$. Hence $\mathfrak{F} \vDash$ **Grz**. On the other hand $\mathfrak{f} \nvDash$ **Grz**.

THEOREM 7.6.5 (Fine). *The logic of $\mathfrak{F}$ is incomplete.*

PROOF. Consider a formula φ such that φ can be satisfied in $\mathfrak{F}$ under a valuation β at a point w_0 iff there exist points z_1, z_2, z_3, y and x such that (i.) $w_0 \lhd x \lhd y \lhd z_1; z_2; z_3$ and $z_2 \lhd z_3$, $z_1 \ntriangleleft z_3$, $z_3 \ntriangleleft z_1$, $x \ntriangleleft w_0$, $y \ntriangleleft x$, (ii.) for all v such that $z_2 \lhd v$ either $v = z_2$ or $v = z_3$, (iii.) for all v such that $y \lhd v$ either $y = v$ or $z_1 \lhd v$ or $z_2 \lhd v$. We leave it to the reader to construct such a formula φ. Then $\langle \mathfrak{F}, \beta, w_0 \rangle \vDash \varphi$ iff w_0 is in the head of the frame. Moreover, if $\mathfrak{G}$ is a frame for $\mathsf{Th}\,\mathfrak{F}$ such that $\mathfrak{G} \nvDash \neg\varphi$ then its underlying Kripke–frame contains an infinite strictly ascending chain of points and so the Kripke–frame on which it is based is not a frame for $\mathsf{Th}\,\mathfrak{F}$. Hence there exists no model for φ based on a Kripke–frame. Thus $\mathsf{Th}\,\mathfrak{F}$ is incomplete. ∎

COROLLARY 7.6.6. *The logics $\mathsf{Th}\,\mathfrak{f}$ and **Grz** are complete. Their union is incomplete, however.*

We summarize the facts in Table 1. We include here not only facts about closure under finite union and finite intersection but also about upper and lower limits. Not all facts have been shown so far. That the finite model property and completeness

TABLE 1. Closure of Properties of Logics

	sup	inf	$\sqcup$	$\sqcap$
tabularity	yes	no	yes	yes
finite model property	no	yes	no	yes
completeness	no	yes	no	yes
compactness	?	no	no	?
canonicity	yes	no	yes	yes
Δ–elementarity	yes	no	yes	yes
finite axiomatizability	no	no	yes	?
decidability	no	no	no	yes
subframe logic	yes	yes	yes	yes
Halldén–completeness	yes	no	no	no
interpolation	yes	no	no	no

may be lost under suprema is left as an exercise. FRANK WOLTER in [**240**] gives an example of two compact logics whose join is not compact. **G.3** is the infimum of logics which are compact, Δ–elementary and canonical, but **G.3** is neither. Nonpreservation of decidability under join is shown in Section 9.4. Nonpreservation of decidability under suprema is straightforward. Now consider the frames $\mathfrak{p}_n = \langle n+1, \lhd\rangle$ where $i \lhd j$ iff $i = j = 0$ or $i < j$. Let $P := \{\mathfrak{p}_n : 0 < n \in \omega\}$. These frames are frames for **K4.3**. Clearly, the logic of finitely many such frames is decidable (it is tabular), but also the logic of any cofinite set R. For this logic is simply **K4.3**$/Q$, where $Q = P - R$. So, consider the map κ sending each subset of P to the logic of its frames. (This is similar to $\iota(M)$ of Section 2.6.) Clearly, the splitting formula for $\mathfrak{p}_n$ is satisfiable in $\kappa(M)$ iff $n \in M$. So, $\kappa(M)$ is undecidable for a nonrecursive set M and hence it is not finitely axiomatizable. This logic is the supremum of finitely axiomatizable logics with finite model property. These logics are decidable. Yet, $\kappa(M)$ is not decidable. The logics $\kappa(M)$ are also useful in showing that Δ–elementary logics as well as finitely axiomatizable logics are not closed under infima. From Theorem 9.4.6 one can easily deduce that Halldén–complete logics are not closed under meets. Halldén–complete logics are closed under suprema just like logics with interpolation. The proof is rather straightforward. Logics with interpolation are not closed under intersection or meet. This follows from the classification of logics containing **Grz** with interpolation (see [**148**]). That subframe logics are closed under all these operations has been shown in Section 3.5.

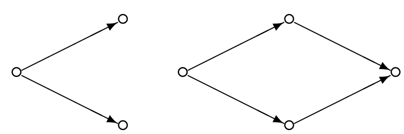
There is also the notion of a *bounded property*. A logic Λ is said to *bound* a property $\mathcal{P}$ if Λ does not possess $\mathcal{P}$ but all its proper extensions do. For facts about bounded properties see Section 8.8.

Exercise 245. Show Corollary 7.6.3.

Exercise 246. Show that a set in the veiled recession frame generates the algebra of finite and cofinite sets iff it is not empty and not the full set. *Hint.* Proceed as follows. Show that there must exist a set of the form $[n) = \{m : n \leq m\}$. Then all $[k)$ for $k \geq n$ exist. Then we have almost all singletons $\{k\}$. Now observe that $\blacklozenge\{n\} = \{n-1, n, n+1\}$ and show that we get all other singletons as well.

Exercise 247. Let $\mathfrak{f}$ be a finite frame and let $\mathsf{Th}\,\mathfrak{f}$ split $\mathcal{E}\Theta$. Show that there is a formula in $\ulcorner {}^2log(\sharp f + 1)\urcorner$ variables axiomatizing $\Theta/\mathfrak{f}$. *Hint.* Use a different axiomatization than developed above. Instead of adding $\mathbf{alt}_n$ add an axiom saying that the frames must have at most $\sharp f$ points.

Exercise 248. Using the fact that **S4.3** is axiomatized over **S4** by splitting the following two frames, show that there is no axiom for **.3** using one variable only. However, show that **S4**.**3**.t can be axiomatized over by axioms using a single variable.



Exercise 249. Let Θ have the finite model property. Show that any logic in $\mathcal{E}\,\Theta$ has a lower cover.

Exercise 250. Construct a formula φ satisfying the requirements of the proof of Theorem 7.6.5. Supply the remaining details of the proof!

****Exercise 251.** Show that there is an ascending chain of logics which have the finite model property such that their supremum fails to have the finite model property. *Hint.* Put $P_n := \{\mathfrak{p}_m : 0 < m \leq n\}$ with $\mathfrak{p}_n$ as above. Next put $\Theta_n := \mathbf{K4}/P_n$ and $\Theta_\omega := \mathbf{K4}/P$. Clearly, Θ_ω is the supremum of the Θ_n. Show that (a) Θ_ω has the same finite models as **G**, (b) $\mathbf{K4}/\mathfrak{p}_n$ is axiomatizable by a constant formula. Deduce that Θ_n has the finite model property for all n. (c) $\Theta \neq \mathbf{G}$.

Exercise 252. (Continuing the previous exercise.) Show that Θ_ω has interpolation. (So, there exist logics with interpolation but without the finite model property.)

Exercise 253. Show that there exist logics with interpolation which are undecidable.

Exercise 254. Let Θ be the modal logic of frames $\langle N, \lhd_1, \lhd_2, \lhd_3, \lhd_4, \lhd_5\rangle$, where $\langle N, \lhd_1\rangle$ is isomorphic to the set of natural numbers and the successor function, $\lhd_2 =$

$\lhd_1^{\smile}$, $\lhd_3 = \lhd_1^+$, $\lhd_4 = \lhd_2^+$, and finally $\lhd_5$ a subset of the diagonal of N. Show that Θ has up to isomorphism only frames of the form specified. Consider extensions of the form

$$\Theta_m := \Theta \oplus \{\Box_2^n \bot \to \Box_5 \bot : n < m\} \ .$$

Let $\Theta_\omega := sup\{\Theta_m : m \in \omega\}$. Show that Θ_ω is consistent and has no frames. Clearly, all Θ_m are complete. Conclude that completeness is not preserved under suprema.

Exercise 255. Show that 0–axiomatizability is preserved under join and meet, and under suprema. (It is not known whether 0–axiomatizability is preserved under infima.)

7.7. Splittings of the Lattices of Modal Logics and Completeness

We will start the investigation of splittings in the lattices of modal logics by studying the lattice of extensions of the minimal modal logic. We will see that for a logic Λ in order to split the lattice $\mathcal{E}\,\mathbf{K}_\kappa$, Λ must be the logic of a finite cycle–free frame. There are several questions which come to mind. First, what happens if we try to iterate the construction? Are there possibilities to split the lattices by algebras which did not previously induce a splitting? Secondly, what interesting properties do the resulting splitting logics have? We will give quite complete answers to these questions. First of all, notice that by the unravelling technique we know that if a formula φ is consistent with $\mathbf{K}_\kappa$ then it has a model based on a cycle–free frame, in fact a totally intransitive tree. This shows that *only* cycle–free frames can split $\mathcal{E}\,\mathbf{K}_\kappa$.

Proposition 7.7.1 (Blok). ($\kappa < \aleph_0$.) *A logic splits $\mathcal{E}\,\boldsymbol{K}_\kappa$ iff it is the logic of a finite rooted cycle–free frame.*

Proof. $\mathbf{K}_\kappa = \bigcap\langle \mathsf{Th}\,\mathfrak{f} : \mathfrak{f}$ finite cycle–free$\rangle$. So, an element Θ is prime in $\mathcal{E}\,\mathbf{K}_\kappa$ only if $\Theta \supseteq \mathsf{Th}\,\mathfrak{f}$, for some finite $\mathfrak{f}$, that is, $\Theta = \mathsf{Th}\mathfrak{h}$ for some $\mathfrak{h}$ which is a p–morphic image of some generated subframe of $\mathfrak{f}$. If $\mathfrak{f}$ is cycle–free, so is $\mathfrak{h}$. On the other hand, by Corollary 7.3.13 finite rooted cycle–frames induce a splitting of $\mathcal{E}\,\mathbf{K}_\kappa$. ∎

We can derive the following fact about the structure of the lattice of extensions from [**23**].

Theorem 7.7.2 (Blok). *$\mathcal{E}\,\boldsymbol{K}_\kappa$ has no elements of finite nonzero dimension. Hence $\mathcal{E}\,\boldsymbol{K}_\kappa$ is atomless.*

Proof. Suppose otherwise. Then there is an atom Λ in the lattice. Atoms are $\bigsqcup$–irreducible and so $\bigsqcup$–prime, by the fact that lattices of extensions are upper continuous. Hence they are co–splitting, and so there must be a splitting companion of Λ. Call it Θ. We know from the previous theorem that $\Theta = \mathsf{Th}\,\mathfrak{f}$ for some finite, rooted and cycle–free frame. It is easy to see that there is a splitting logic $\Theta' \subsetneq \Theta$. Namely, if $\mathfrak{f} = \langle f, \lhd\rangle$ with w_0 the root, let $u \notin f$ and put $\lhd' := \lhd \cup \{\langle u, w_0\rangle\}$. Finally, $\mathfrak{g} := \langle f \cup \{u\}, \lhd'\rangle$ and $\Theta' := \mathsf{Th}\,\mathfrak{g}$. $\mathfrak{g}$ is rooted and cycle-free and so Θ' splits

$\mathcal{E}\,\mathbf{K}_\kappa$. It has a splitting partner Λ' for which $\Lambda' \subsetneq \Lambda$. But then $\Lambda' = \mathbf{K}_\kappa$, which is impossible. ◻

We will show now another theorem by Blok which says that any iterated splitting of $\mathbf{K}_\kappa$ has the finite model property. However, our method is markedly different from that of Blok, keeping with the spirit of the technique of constructive reduction. Moreover, the original proofs only deal with monomodal logics and are redone here for polymodal logics.

Theorem 7.7.3. ($\kappa < \aleph_0$.) *Adding an axiom of the form $\boxtimes^k\bot \to \varphi$ preserves decidability, global completeness, and the global finite model property.*

Proof. The proof is based on the observation that for any $m, n \in \omega$ there is a finite set $S(m,n)$ of substitutions form the set of formulae over p_i, $i < m$, into the set of fromulae over p_i, $i < n$, such that for any finite set of generators $\{a_0, \dots, a_{n-1}\}$ for the set algebra $\mathbb{F}$ of a refined frame $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ for the valuation $\beta : p_i \mapsto a_i$, $i < n$, any formula φ in the variables p_i, $i < m$, and any point $x \in f$

$$(\dagger)\ \ \langle \mathfrak{F}, \beta, x\rangle \vDash \{\boxtimes^k\bot \to \varphi^\sigma : \sigma \in S(m,n)\} \qquad \Leftrightarrow \qquad \mathfrak{F} \vDash \boxtimes^k\bot \to \varphi$$

For then it holds that for all χ, ψ based on the sentence letters $p_0, \dots, p_{n-1}$

$$\boxtimes^\omega \chi \vdash_{\Lambda(\boxtimes^k\bot\to\varphi)} \psi \ \Leftrightarrow\ \boxtimes^\omega\chi; \boxtimes^\omega\Big(\bigwedge_{\sigma\in S(m,n)} \boxtimes^k\bot \to \varphi^\sigma\Big) \vdash_\Lambda \psi$$

From this we can deduce that if Λ is globally complete, so is now the logic $\Lambda \oplus \boxtimes^k\bot \to \varphi$ and if Λ has the global finite model property, so does also the logic $\Lambda \oplus \boxtimes^k\bot \to \varphi$. For a proof of this consequence from ($\dagger$) just check all models on rooted refined frames $\mathfrak{F}$ where the underlying set algebra is generated by the values of $\beta(p_0), \dots, \beta(p_{n-1})$. It is enough to show the theorem in the class of refined frames.

Now for the proof of ($\dagger$). From right to left holds for any set $S(m,n)$. So the difficult part is from left to right. We begin by constructing the $S(m,n)$. Consider the subframe $\mathfrak{C}$ based on the set C of all points x such that $\langle \mathfrak{F}, x\rangle \vDash \boxtimes^k\bot$. $\mathfrak{C}$ is a generated subframe of $\mathfrak{F}$ and hence refined since $\mathfrak{F}$ is. By Theorem 2.7.14, C is finite, bounded in size by a function depending only on n (and k). Hence $\mathfrak{C}$ is full. Consider now the induced valuation on $\mathfrak{C}$, also denoted by β. It is possible to show that any set $T \subseteq C$ can be presented as the extension of $\tau_T(a_0, \dots, a_{n-1})$ under β for a suitable τ_T which is of modal degree $\leq 2k$. Collect in $S(m,n)$ all substitutions $\sigma : p_i \mapsto \tau_i(p_0, \dots, p_{n-1})$, $i < m$, for formulas of depth $\leq 2k$. $S(m,n)$ is finite. We show ($\Rightarrow$) of ($\dagger$) with these sets. To that end, assume that $\mathfrak{F} \nvDash \boxtimes^k\bot \to \varphi$ for some φ such that $var(\varphi) = \{p_i : i < m\}$. Then there exist γ and x such that

$$\langle \mathfrak{F}, \gamma, x\rangle \vDash \boxtimes^k\bot \wedge \neg\varphi$$

Then $x \in C$ and so we have by the fact that $\mathfrak{C}$ is a generated subframe

$$\langle \mathfrak{C}, \gamma, x\rangle \vDash \neg\varphi\ .$$

There exist $\tau_i(\vec{p})$, $i < m$, such that $\gamma(p_i) = \overline{\beta}(\tau_i(\vec{p}))$. It follows that $\overline{\gamma}(\varphi) = \overline{\beta}(\varphi[\tau_i(\vec{p})/p_i : i < m])$. Put $\sigma : p_i \mapsto \tau_i(\vec{p})$, $i < m$. Then $\overline{\gamma}(\varphi) = \overline{\beta}(\varphi^\sigma)$. Therefore

$$\langle \mathfrak{C}, \beta, x\rangle \vDash \neg\varphi^\sigma \ .$$

And so

$$\langle \mathfrak{F}, \beta, x\rangle \vDash \boxtimes^k \bot \wedge \neg\varphi^\sigma \ .$$

This demonstrates (†). ◻

Since $\mathbf{K}_\kappa$ has the global finite model property we conclude that splitting finitely many frames does not disturb the global finite model property. However, we can conclude the following.

Theorem 7.7.4 (Blok). ($\kappa < \aleph_0$.) *All splittings $\mathcal{E}\, \boldsymbol{K}_\kappa/N$ where N is a set of finite rooted cycle–free frames have the local finite model property.*

Proof. The proof will be performed for the case $\kappa = 1$. The generalization to arbitrary finite κ is straightforward but somewhat tedious. Let N be a set of finite rooted cycle–free frames. Suppose that N is finite. By Theorem 7.3.13 the splitting formula is of the form $\boxtimes^{\leq m+1}\delta(\mathfrak{A}) \rightarrow \neg \boxtimes^m \bot$ for some m and $\mathfrak{A}$. By repated use of Theorem 7.7.3, one can show that $\mathcal{E}\,\mathbf{K}_\kappa/N$ has the global finite model property. Now let N be infinite. Put $\Lambda := \mathcal{E}\,\mathbf{K}_\kappa/N$. Suppose that φ is a formula and $d = dp(\varphi)$. By induction on d we prove that every Λ–consistent formula of degree $\leq d$ has a Λ–model based on a frame of depth $\leq d$. This certainly holds for $d = 0$. Let $N(d)$ be the set of frames of depth $\leq d$ contained in N whose powerset algebra is generable by at most $\sharp var(\varphi)$ elements. This set is finite by Theorem 2.7.14. It is enough to show that if φ is Λ–consistent then it has a model based on a frame of depth $\leq d$ not having a generated subframe that is contractible to a member of $N(d)$. Suppose that φ is consistent with Λ. Then it is consistent with $\mathbf{K}_\kappa$ as well. So there exists a finite Kripke–frame $\mathfrak{f}$ such that $\mathfrak{f} \nvDash \neg\varphi$. Two cases have to be considered.
Case 1. $\varphi \vdash \boxtimes^d\bot$ (where $\vdash\, = \,\vdash_{\mathbf{K}_\kappa}$). Then any model for φ is based on a cycle–free frame of depth $\leq d$. Now φ is consistent with Λ and so it is consistent with $\mathbf{K}_\kappa/N(d)$. Thus it has a finite model based on a frame $\mathfrak{g}$ by the fact that $N(d)$ is finite. $\mathfrak{g}$ must be a frame which is not reducible to any member in $N(d)$, hence $\mathfrak{g} \notin N$, since $\mathfrak{g}$ is of depth $\leq d$.
Case 2. $\varphi \nvdash \boxtimes^d\bot$. Then we can build a finite model of $\varphi \wedge \neg \boxtimes^d \bot$ with root w_0. Let φ be in normal form and $\varphi = \bigvee_{i<n} \varphi_i$. Let $\varphi_i = \mu_i \wedge \Box\chi_i \wedge \bigwedge_{j<m_i} \Diamond\psi^i_j$. Pick $i < n$ such that φ_i is Λ–consistent. Let $\langle \mathfrak{g}_j, \beta_j, y_j\rangle \vDash \psi^i_j \wedge \chi_i$, $\mathfrak{g}_j = \langle g_j, \lhd_j\rangle$. Assume that the g_j are disjoint, of depth $\leq d-1$ and rooted at y_j; moreover, let $w_0 \notin g_i$ for all i. By inductive hypothesis we may assume that no $\mathfrak{g}_j$ has a generated subframe which can be contracted to a member of N. Now put $h := \{w_0\} \cup \bigcup_{j<m_i} g_j$, $\lhd := \bigcup_{j<m_i} \lhd_j \cup \{\langle w_0, y_j\rangle : j < m_i\}$, $\mathfrak{h} := \langle h, \lhd\rangle$. By choice of i, μ_i is consistent in

boolean logic. Let δ be a valuation into $\mathbf{2}$ such that $\overline{\delta}(\mu_i) = 1$. Put

$$\gamma(p) := \begin{cases} \bigcup_{j<m_i} \beta_j(p) \cup \{w_0\} & \text{if } \delta(p) = 1 \\ \bigcup_{j<m_i} \beta_j(p) & \text{if } \delta(p) = 0 \end{cases}$$

Then $\langle \mathfrak{h}, \gamma, w_0 \rangle \vDash \varphi_i$ as can be verified immediately. Suppose that $\mathfrak{h}$ has a generated subframe $\mathfrak{k}$ which can be contracted to a member of N. Then, by construction of $\mathfrak{h}$, $\mathfrak{k} = \mathfrak{h}$. So, $\mathfrak{h}$ is cycle–free. Suppose therefore that for every $i < n$ the construction must yield a cycle–free frame. Then $\varphi \vdash \boxtimes^m \bot$ for some m. It is not hard to show that $m \leq d$. This contradicts our assumptions, however. Hence, there is an $i < n$ and $\mathfrak{g}_j$, $j < n_i$ such that the frame $\mathfrak{h}$ constructed from the $\mathfrak{g}_j$ has a cycle. (So, one of the $\mathfrak{g}_j$ contains a cycle.) Then $\mathfrak{h}$ is not contractible to a member of N. ◻

Here now is another result about lattices of extensions, which is rather peculiar in nature.

Theorem 7.7.5. *The lattice $\mathcal{E}\,\boldsymbol{K}.t$ is isomorphic to a direct product $\mathit{2} \times \mathfrak{L}$, where $\mathfrak{L}$ has no splittings.*

Proof. First of all, $\mathcal{E}\,\mathbf{K}.t$ has a splitting $\langle \Lambda, \Theta \rangle$, that is induced by the one–point irreflexive frame, by the fact that it splits $\mathcal{E}\,\mathbf{K}_2$, of which $\mathbf{K}.t$ is a member. But notice that Θ has finite model property by Theorem 3.6.1, because it is axiomatized by the constant formula $\Diamond\top \vee \Diamond\top$. The frames for Θ are all frames which are different from the one point irreflexive frame. Suppose now that $\Theta' \subsetneq \Theta$. Then there must be a finite frame for Θ' which is not a frame for Θ. The only choice is the one–point irreflexive frame; and then $\Theta' = \mathbf{K}.t$. Thus Θ is an atom, and Λ is a coatom. Now Theorem 7.2.8 yields the decomposition into $\mathbf{2} \times \mathfrak{L}$. Any splitting of $\mathfrak{L}$ induces a splitting of $\mathbf{2} \times \mathfrak{L}$, so we are done if we can show that no other splittings exist. This is left as an exercise. ◻

We have seen in the preceding chapter that there are quite strong incomplete logics, so completeness is actually not a guaranteed property of modal logics, as has been believed until the early seventies, before counterexamples have been produced by S. K. Thomason [**207**] and Kit Fine [**62**], later also by Johan van Benthem [**9**]. Nevertheless, despite the fact that there are incomplete logics one might still believe that the phenomenon of incompleteness is somewhat marginal. To get an insight into the whereabouts of incomplete logics Kit Fine has proposed in [**61**] to study the degree of incompleteness of a logic. This is defined to be the set of all logics sharing the same *Kripke–frames* with a given logic. Of course, only one of these logics can be complete, so the cardinality of this set gives an indication of how many incomplete logics close to a given logic exist.

Definition 7.7.6. *The* ***Fine–spectrum*** *of a logic Λ is the set $sp(\Lambda) := \{\Theta : \mathsf{Krp}(\Theta) = \mathsf{Krp}(\Lambda)\}$. The cardinality of the Fine–spectrum of Λ is called the* ***degree of incompleteness of*** *Λ. If Λ has degree of incompleteness 1 then it is called* ***intrinsically complete*** *or* ***strictly complete****.*

We can give a lattice theoretic analogue of this definition. Suppose we are given a set $P \subseteq L$ in a lattice. Then the **P–spectrum** of an element is the set

$$sp_P(x) := \{y : (\forall p \in P)(x \leq p \Leftrightarrow y \leq p)\}.$$

With P being the set of complete logics (or the logics of rooted frames), we get our definition of the Fine–spectrum. It is clear that each P–spectrum has a maximal element, namely $\bigsqcap\langle p : x \leq p \in P\rangle$. In general, there is however no least element. Fine–spectra do not always contain a minimal element. Here is an example. (See also Section 7.9.)

We are interested in the tense theory of ω^{op}. Put $\Theta := \mathbf{G.3}.t.\mathbf{K4.3}.\mathbf{D}^{-}$. It is not hard to show that the Kripke–frames for Θ are of the form α^{op}, where α is a limit ordinal. For the axioms of **G.3** guarantee $\lhd$ to be a conversely well–founded relation. Adding $\mathbf{K4.3}^{-}$ forces the relation to be a well–order. Finally, adding $\mathbf{D}^{-}$ α is forced to be a limit ordinal. The theory of ω^{op} contains Θ. What we will show is that this theory is a maximal consistent and complete logic in the lattice of tense logics, and that it is the intersection of countably many logics extending it. Since the extensions have no frame, they are incomplete, and are all in the Fine–spectrum of the inconsistent tense logic. Hence, in tense logic the Fine–spectrum of the inconsistent logic is not an interval.

Let us define for subsets A, B of ω, $A \sim B$ iff for almost all $n \in \omega$, $n \in A$ iff $n \in B$. We say that in this case A and B are **almost equal**. It is not hard to show that $\sim$ is a congruence in the algebra of sets over ω, with the operations

$$\begin{array}{rcl} \blacksquare\!\!-A & := & \{n : (\forall m < n)(m \in A)\} \\ -\!\!\blacksquare A & := & \{n : (\forall m > n)(m \in A)\} \end{array}$$

Furthermore, let $\mathbb{O}$ be the set of finite and cofinite subsets of ω. $\mathbb{O}$ is closed under all operations, and therefore $\langle\omega, >, \mathbb{O}\rangle$ is a general frame. $\mathbb{O}$ contains exactly the sets which are almost zero or almost one. For the purpose of the next theorem, a **partition** of a set M is a subset $\mathbb{X}$ of $\wp(M)$ such that (a) $\varnothing \notin \mathbb{X}$, (b) $M = \bigcup\mathbb{X}$, and (c) for any $S, T \in \mathbb{Z}$, if $S \neq T$ then $S \cap T = \varnothing$.

Lemma 7.7.7. *Let $A \subseteq \wp(\omega)$ be a finite partition of ω. Then the least algebra containing A is the set of all sets which are almost identical to a union of elements of A.*

Lemma 7.7.8. *Let $G := \{S_i : i < n\}$ be a finite set of subsets of ω. Then there exists $H = \{T_j : j \leq k\}$, which is a partition of H of cardinality $\leq 2^n$ such that G and H generate the same algebra of sets.*

Lemma 7.7.9. *Let $\mathbb{Q}$ be an n–generated set of sets. Then $\langle\omega, >, \mathbb{Q}\rangle$ satisfies $\chi(2^n)$, where $\chi(k)$ is the following formula.*

$$\begin{array}{ll} & \bigvee_{i<k} \Diamond\Box\neg p_i \\ \vee & \bigvee_{i<j<k} \Diamond\Box(p_i \leftrightarrow p_j) \\ \vee & \Diamond\Box \bigvee_{i<k} p_i \end{array}$$

Proof. Let β be a valuation and let $a_i := \beta(p_i)$. Pick some point $n \in \omega$. We have to show that $\langle\langle\omega, >, \mathbb{Q}\rangle, k, \beta\rangle \vDash \chi(2^n)$. To that end, assume that $k \nvDash \bigvee_{i<k} \Diamond\Box\neg p_i$. Then all a_i are infinite. Assume furthermore that $k \nvDash \bigvee_{i<j<k} \Diamond\Box(p_i \leftrightarrow p_j)$. Then for no different i, j we have $a_i \sim a_j$. Hence no two sets a_i and a_j are almost equal. We know by Lemma 7.7.8 that in an n–generated algebra there are at most 2^n sets such that no two are almost equal. Moreover, their union is almost ω. Hence $k \vDash \Diamond\Box \bigvee_{i<k} p_i$. That had to be shown. □

Lemma 7.7.10. *No logic containing $\Theta \oplus \chi(k)$ has a nonempty Kripke–frame.*

Proof. Let $\mathfrak{f}$ be a (nonempty) Kripke–frame for $\Theta \oplus \chi(k)$. Then $\mathfrak{f}$ is isomorphic to a converse well–order α^{op}, where α is a limit ordinal, and so not finite. But then $\chi(k)$ is clearly not valid in $\mathfrak{f}$. □

Theorem 7.7.11. *The tense logic of ω^{op} is the downward limit of logics which have no frame.*

Proof. Suppose that $\omega^{op} \nvDash \neg\varphi$. Then there exists a model $\langle\omega^{op}, \beta, k\rangle \vDash \varphi$. Now let $\mathbb{Q}$ be the algebra generated by $\beta(p)$, $p \in var(\varphi)$. Clearly, we have $\langle\omega, >, \mathbb{Q}\rangle \nvDash \neg\varphi$. $\mathbb{Q}$ is finitely generated. Hence $\langle\omega, >, \mathbb{Q}\rangle$ satisfies an axiom $\chi(k)$ for some k. Hence φ is consistent in $\mathsf{Th}\,\Omega \oplus \chi(k)$. So,

$$\mathsf{Th}\,\omega^{op} = \bigcap_{k\in\omega} \mathsf{Th}\,\omega^{op} \oplus \chi(k)$$

This concludes the proof of the theorem. □

Corollary 7.7.12. *The spectrum of the inconsistent tense logic is not an interval.*

Theorem 7.7.13. *The spectrum of the monomodal logic $\mathsf{Th}\,\boxed{\bullet}$ is not an interval.*

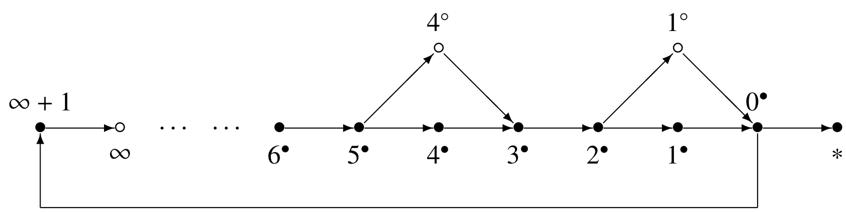
This last theorem follows immediately from the Simulation Theorem. The situation is different if we specialize on the set $\mathbb{P}$ of *prime* or splitting elements. The $\mathbb{P}$–spectrum of x is called the **prime–spectrum** and denoted by $psp(x)$.

Proposition 7.7.14. *Let $\mathfrak{L}$ be a lattice. For every $x \in L$ there are x_0 and x^0 such that $psp(x) = [x_0, x^0]$. x^0 is $\mathbb{P}$–complete and x_0 is a union of co–splitting elements. Consequently, if $x = x_0$ and x is $\mathbb{P}$–complete then x is strictly $\mathbb{P}$–complete.*

Proof. Put $x^0 := \bigsqcap\langle p : p \in \mathbb{P}, p \geq x\rangle$ and $x_0 := \bigsqcup\langle \mathfrak{L}/p : p \in \mathbb{P}, p \ngeq x\rangle$. Suppose that $z \in psp(x)$. Then clearly $z \leq x^0$. We have to show that $x_0 \leq z$. Now for $p \in \mathbb{P}$ we have $p \ngeq x \Leftrightarrow p \ngeq z$ whence $\mathfrak{L}/p \leq x \Leftrightarrow \mathfrak{L}/p \leq z$. Therefore $z \geq \bigsqcup\langle\mathfrak{L}/q : p \in \mathbb{P}, p \ngeq x\rangle = x_0$. Conversely, if $x_0 \leq z \leq x^0$ then $p \geq z \Leftrightarrow p \geq x$ for all $p \in \mathbb{P}$ and so $z \in psp(x)$. □

This shows among other that the splittings of $\mathbf{K}_\kappa$ by cycle–free frames are intrinsically complete. Thus incomplete logics cannot be 'close' to unions of co–splitting logics. The difficult question remains as to where these logics are. We will spend the

FIGURE 7.10. $\mathfrak{P}_M$, $M = \{1, 4, \ldots\}$



rest of this section and the whole next section to prove one of the most beautiful theorems in modal logic. We will call it *Blok's Alternative*, because it says that logics have only two choices for their degree of incompleteness: 1 or $2^{\aleph_0}$.

THEOREM 7.7.15 (Blok's Alternative). *A logic containing* $\boldsymbol{K}_1$ *is intrinsically complete iff it is an iterated splitting of* $\boldsymbol{K}_1$. *Otherwise, the logic has degree of incompleteness* $2^{\aleph_0}$.

The proof is done in two steps. The first step consists in showing that the coatoms have degree of incompleteness $2^{\aleph_0}$. In the second step we use the frames produced in the first step to show that in fact *any* non–splitting logic has degree of incompleteness $2^{\aleph_0}$. Let us begin by defining the frames $\mathfrak{P}_M$, where $M \subseteq \omega - \{0\}$. The set of worlds of $\mathfrak{P}_M$ is the (disjoint) union of the sets $\{n^\bullet : n \in \omega\}$, $\{*, \infty, \infty + 1\}$ and $\{n^\circ : n \in M\}$. We put

$$x \lhd y \Leftrightarrow \begin{cases} & x \in \{m^\bullet, m^\circ\}, y \in \{n^\bullet, n^\circ\}, m > n, \\ \text{or} & x = y = m^\circ, \\ \text{or} & x = 0^\bullet, y \in \{*, \infty + 1\}, \\ \text{or} & x \in \{\infty, \infty + 1\}, y = \infty, \\ \text{or} & x = \infty, y \in \{n^\bullet, n^\circ\}. \end{cases}$$

This defines the frame $\mathfrak{p}_M$. The frame corresponding to $M = \{1, 4, \ldots\}$ is shown in the picture. $\mathfrak{p}_M$ is intuitively obtained as follows. We have the set of natural numbers ordered by $\lhd = >$. Each point in that set is referred to as $n^\bullet$. Moreover, if $n \in M$ we also have a reflexive companion n°. Notice that the set $\{*\}$ is the only nontrivial generated subframe. We let $\mathbb{P}_M$ be the algebra of all finite sets not containing ∞ and of all cofinite sets containing ∞. This set is closed under all operations. Indeed, if we have a finite set S not containing ∞, $*$ or $\infty + 1$, or a cofinite set then $\blacklozenge S$ is cofinite, and contains ∞. Furthermore, we have $\blacklozenge\{*\} = \blacklozenge\{\infty + 1\} = \{0^\bullet\}$ which is finite and does not contain ∞. If S is cofinite and contains ∞, so does $\blacklozenge S$. Thus this is well–defined and so $\mathfrak{P}_M := \langle \mathfrak{p}_M, \mathbb{P}_M \rangle$ is a frame.

Lemma 7.7.16. *$\mathfrak{P}_M$ is 0–generated.*

Proof. We will show two things. (1.) $\{*\}$ is definable, and (2.) all other sets are definable from $\{*\}$. (1.) is easy. We have $\blacksquare\varnothing = \{*\}$. For (2.) we need to do some work. First, $\blacksquare\{*\} = \{0^\bullet\}$. Moreover, $\{\infty + 1\} = \blacklozenge^{\leq 2}\{0^\bullet\} - \blacklozenge^{\leq 1}\{0^\bullet\}$. Now define by induction on n the polynomials $i_n(p)$, $p_n^\bullet(p)$, $p_n^\circ(p)$.

$$\begin{array}{lcl} i_n(p) & := & \bigvee_{k\leq n}(p_k^\bullet(p) \vee p_k^\circ(p)) \\ p_0^\bullet(p) & := & p \\ p_0^\circ(p) & := & \bot \\ p_{n+1}^\bullet(p) & := & \Box i_n(p) \\ p_{n+1}^\circ(p) & := & \Diamond p_n^\bullet(p) \wedge \neg p_{n+1}^\bullet(p) \wedge \neg\Diamond p_{n+1}^\bullet(p) \end{array}$$

Then $i_0(p) = p$. By induction it is verified that $p_n^\bullet(\{0^\bullet\}) = \{n^\bullet\}$ and that $p_n^\circ(\{0^\bullet\}) = \{n^\circ\}$ if $n \in M$, and $= \varnothing$ otherwise. Hence $i_n(\{0^\bullet) = \{k^\bullet : k \leq n\} \cup \{k^\circ : k \leq n, k \in M\}$. Finally, $\{0^\bullet\} = \blacksquare\varnothing$, so we can define all these sets if we replace p by $\Box\bot$. So all singleton sets except for $\{\infty\}$ are 0–definable, and all finite sets not containing ∞ and their complements are also 0–definable. ∎

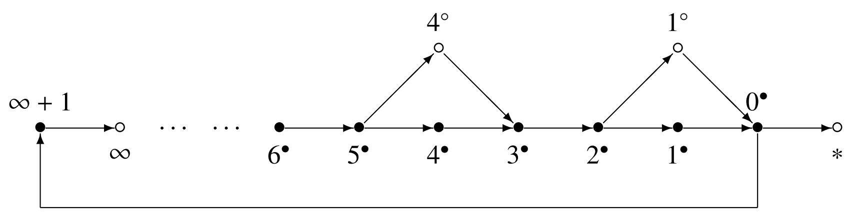
Notice that the definition of the polynomials does not depend on the set $M \subseteq \omega$. We will make use of that in the next section. The next goal is to show that the logics of these frames are all of codimension 2. To see this we can use the splitting theorem. The algebra is subdirectly irreducible and the logic it generates is 4–transitive. Hence we are done if we can show this algebra to be finitely presentable. But this is entirely obvious, since we have just shown that it is 0–generated, and so isomorphic to $\mathfrak{Fr}_\Lambda(0)$. Hence, we can take as a diagram simply $\top$. To get an extension of this logic we simply add the formula $\Box\bot$, for $\Box\bot$ is an opremum! Therefore, $\mathsf{Th}\,\mathfrak{P}_M$ has codimension 2. Finally, for different sets M, the logics of the frames are distinct, simply by the fact that the points n° are definable by constant formulae. We have proved now not only that there are uncountably many incomplete logics with Kripke–frame $\boxed{\bullet}$, but also that the logic of the irreflexive point is co–covered by them.

Definition 7.7.17. *Let $\mathfrak{L}$ be a lattice, and $x \in L$. Define the* ***co–covering number*** *of x to be the cardinality of the set $\{y : y \prec x\}$.*

Theorem 7.7.18 (Blok). *The logic of the irreflexive point is co–covered by $2^{\aleph_0}$ incomplete logics. Hence it has co–covering number $2^{\aleph_0}$ and degree of incompleteness $2^{\aleph_0}$.*

Next we need to deal with the frame $\boxed{\circ}$. The solution will be quite similar. Namely, instead of the frame $\mathfrak{p}_M$ we define the frame $\mathfrak{q}_M$ consisting of the same set,

FIGURE 7.11. $\mathfrak{Q}_M$, $M = \{1, 4, \ldots\}$



and the relation $\lhd$ differs minimally in that we now put $* \lhd *$. The rest is the same. We put $\mathbb{Q}_M := \mathbb{P}_M$ and this defines $\mathfrak{Q}_M$.

LEMMA 7.7.19. *$\mathfrak{Q}_M$ is 1–generated. Moreover, a set generates $\mathfrak{Q}_M$ iff it is $\neq \varnothing$ and $\neq q_M$.*

PROOF. Since the frame is identical on $-\{*\}$, we can use (2.) of the previous proof to show that $\mathbb{Q}_M$ is generated from $\{*\}$ and so 1–generated. For the second claim it suffices to show that $\{*\}$ is definable from any other nontrivial set. We claim that $-\{*\} = \blacklozenge^{\leq 4} b - \blacksquare^{\leq 4} b$ iff $b \neq \varnothing$ and $-b \neq \varnothing$. Namely, suppose that $b \neq \varnothing$ as well as $-b \neq \varnothing$. Then, because $b \neq \varnothing$, $\blacklozenge^{\leq 4} b \supseteq -\{*\}$. For if $x \in b$ for some x then by 4–transitivity all elements that can at all reach x are in $\blacklozenge^{\leq 4} b$. Hence either $* \in b$ and then $\blacklozenge^{\leq 4} b = 1$ or $* \notin b$ and then $\blacklozenge^{\leq 4} b = -\{*\}$. **Case 1.** $* \in b$. Then there exists an $x \neq *$ such that $x \notin b$. Consequently, $\blacksquare^{\leq 4} b = \{*\}$. **Case 2.** $* \notin b$. Then $\blacklozenge^{\leq 4} b = -\{*\}$, and $\blacksquare^{\leq 4} b = \varnothing$, as required. ◻

The algebra underlying $\mathfrak{Q}_M$ has two subalgebras, the two element algebra and itself. It has three homomorphic images, itself, the trivial algebra and the algebra of $\boxed{\circ}$, the latter corresponding to the subframe generated by $\{*\}$. Now take a logic $\Theta \supseteq \Lambda_M := \mathsf{Th}\,\mathfrak{P}_M$. We have $h : \mathfrak{Fr}_{\Lambda_M}(1) \twoheadrightarrow \mathfrak{Fr}_{\Theta}(1)$. We will use an argument similar to the one before. However, this time we do not have such a simple structure for the free frame. Namely, it consists of countably many copies of $\mathfrak{Q}_M$, each corresponding to a different generating set. The generating sets are exactly the nontrivial sets. We now reason as with the veiled recession frame of Section 7.6.

LEMMA 7.7.20. *Let S be an internal set of $\mathfrak{Q}_M$ and $C := S \cap \blacklozenge - S$. $C = \varnothing$ iff $S = \varnothing$, $S = \{*\}$ or $S = q_M$.*

PROOF. Suppose that $S = \varnothing$. Then $C = \varnothing$. Suppose on the other hand that $-S = \varnothing$. Then $C \subseteq \blacklozenge\varnothing = \varnothing$. Finally, if $S = \{*\}$, then $* \notin \blacklozenge - S$, and so $C = \varnothing$

as well. Now assume that neither of the three is the case. Let S be cofinite. Then it contains ∞. $-S$ is finite, and $\blacklozenge - S$ contains ∞ if $-S$ contains at least n° or $n^\bullet$ for some n. In that case, $\infty \in C$. Now let $-S \subseteq \{\infty + 1, *\}$. In this case $0^\bullet \in C$. This finishes the case where S is cofinite. So, let now S be finite. It does not contain ∞, so $\infty + 1 \in \blacklozenge - S$. Hence if $\infty + 1 \in S$ we are done. So, assume from now on $\infty + 1 \notin S$. Let n be the smallest number such that $S \cap \{n^\bullet, n^\circ\} \neq \varnothing$. $n := -1$ if such number does not exist. If $n > 0$, $0^\bullet \in -S$ and so $\{n^\bullet, n^\circ\} \subseteq \blacklozenge - S$. In that case, $C \neq \varnothing$. Now let $n = 0$. Then, since $\infty + 1 \in -S$, $0^\bullet \in \blacklozenge - S$, and so $0^\bullet \in C$. Finally, let $n = -1$. Then $S = \{*\}$. But that was excluded. ◻

So, $h(p \to \Box p \wedge \neg p \to \Box\neg p) = 1$ iff $h(p \to \Box p) = 1$ and $h(\neg p \to \Box\neg p) = 1$ iff $h(p) = \varnothing$ or $h(\neg p) = \varnothing$ iff $h(p)$ does not generate the full algebra of internal sets iff $h(p)$ generates a subalgebra isomorphic to the algebra of subsets of $\boxed{\circ}$. Hence if $p \to \Box p \notin \Theta$, $\mathfrak{Fr}_\Theta(1)$ contains a generated subframe isomorphic to $\mathfrak{Q}_M$, so that $\Theta = \Lambda_M$. The following is now proved.

Theorem 7.7.21 (Blok). *The logic of the reflexive point is co–covered by $2^{\aleph_0}$ incomplete logics. Hence it has co–covering number $2^{\aleph_0}$ and degree of incompleteness $2^{\aleph_0}$.*

Exercise 256. Show that there are $2^{\aleph_0}$ iterated splittings of $\mathbf{K}_1$.

Exercise 257. Show that the frame underlying the freely 0–generated $\mathbf{K}_1$–algebra has cardinality $2^{\aleph_0}$.

Exercise 258. Show that $\mathsf{Th}\,\boxed{\bullet}$ is co–covered by $2^{\aleph_0}$ logics in the lattice $\mathcal{E}\mathbf{K.trs}_4$. *Hint.* This should be entirely easy.

Exercise 259. Describe the structure of the canonical frame for one variable in $\mathsf{Th}\,\mathfrak{Q}_M$.

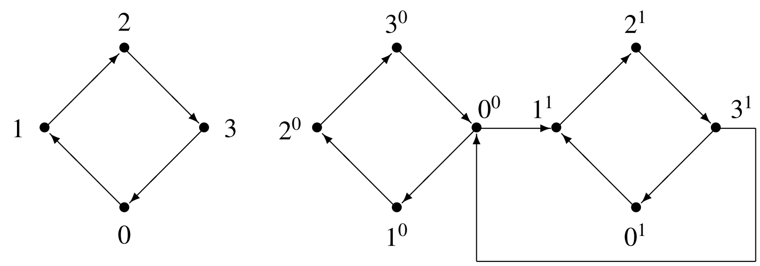
*__Exercise 260.__ Show that the lattice of tense logics has exactly one splitting. *Hint.* Show first that $\mathbf{K}.t$ is complete with respect to finite frames which contain no forward cycle, and hence no backward cycle. Now let $\mathfrak{f}$ be such a frame and let it have more than one point. Create a sequence of frames $\mathfrak{f}_n$ such that (i) they contain a reflexive point and so $\mathsf{Th}\,\mathfrak{f}_n \not\subseteq \mathsf{Th}\,\mathfrak{f}$ and (ii) that $\bigcap \mathsf{Th}\,\mathfrak{f}_n \subseteq \mathsf{Th}\,\mathfrak{f}$. (See [**123**].)

Exercise 261. Let $\mathfrak{A}$ be a finite n–generated modal κ–algebra. Then $\mathsf{Th}\,\mathfrak{A}$ can be axiomatized by formulae containing at most $n + 1$ variables.

7.8. Blok's Alternative

Now that we have shown for the two coatoms that they are maximally incomplete, we proceed to a proof that in fact all consistent logics which are not joins of

Figure 7.12.



co–splittings of $\mathcal{E}\,\mathbf{K}_1$ have degree of incompleteness $2^{\aleph_0}$. Suppose that Θ is a consistent logic and not identical to a logic $\mathbf{K}_1/N$, N a set of cycle–free frames. Then Θ properly includes Θ_0, the splitting of $\mathbf{K}_1$ by all cycle–free frames not belonging to $\mathsf{Krp}\,\Theta$. Since Θ_0 has the finite model property, there must be a finite frame $\mathfrak{f}$ such that $\mathfrak{f} \vDash \Theta_0$ but $\mathfrak{f} \nvDash \Theta$. By the construction of Θ_0, $\mathfrak{f}$ cannot be cycle–free. It is by playing with $\mathfrak{f}$ that we obtain a set of frames $\mathfrak{G}_M$, where $M \subseteq \omega$ and $\mathfrak{G}_M$ is not a frame for Θ, but all Kripke–frames for $\mathsf{Th}\,\mathfrak{G}_M$ are frames for Θ, so that $\Theta \cap \mathsf{Th}\,\mathfrak{G}_M$ is an incomplete logic different from Θ.

Consider now the frame $\mathfrak{f}$. The desired frames $\mathfrak{G}_M$ will be produced in two stages. First, we know that $\mathfrak{f}$ contains a cycle, say $c_0 \lhd c_1 \lhd \ldots \lhd c_\gamma = c_0$. Put $C := \{c_i : i < \gamma\}$. We assume the cycle to be minimal, that is, $c_i \lhd c_j$ iff $j \equiv i + 1 \pmod{\gamma}$. The proof of the existence of such a cycle is an exercise. We produce a variant of $\mathfrak{f}$, $\mathfrak{f}^n$, by blowing up this cycle. This variant is defined as follows. We put $f^n := (f - C) \cup C \times n$. (Here, the union is assumed to be disjoint.) We write x^i rather than $\langle x, i\rangle$ for $x \in C$ and $i < n$. Also we write $C^n := \{c^i_j : i < n, j < \gamma\}$. So, f^n consists in replacing C by n copies, so that a point $x \in C$ is now split into $x^0, \ldots, x^{n-1}$. $\blacktriangleleft$ is defined as follows.

$$x \blacktriangleleft y \text{ iff } \begin{cases} \quad\ (1) & x, y \notin C^n \text{ and } x \lhd y \\ \text{or } (2) & x = c^i_j, y = c^i_k \text{ and } k \equiv j + 1 \pmod{\gamma}, \\ \text{or } (3) & x = c^i_0, y = c^{i+1}_1, \\ \text{or } (4) & x = c^i_j, y \notin C^n, c_j \lhd y, \\ \text{or } (5) & x \notin C^n, y = c^0_j, \text{ and } x \lhd c_j \end{cases}$$

Figure 7.12 shows a duplication of a minimal four cycle according to the previous construction.

Proposition 7.8.1. *Define π by $\pi(c^i_j) := c_j$ and $\pi(y) := y$ for $y \notin C^n$. Then $\pi : \mathfrak{f}^n \twoheadrightarrow \mathfrak{f}$.*

Proof. This is a straightforward checking of cases. The first p–morphism condition is proved as follows. Assume $x \blacktriangleleft y$. Then if (1) holds, also $x \lhd y$; moreover,

$x = \pi(x)$ and $y = \pi(y)$. If (2) holds, then $x = c^i_j$ and $y = c^i_k$ with $k \equiv j + 1 \pmod{\gamma}$. Then $\pi(x) = c_j$ and $\pi(y) = c_k$. By the fact that C is a cycle, $c_j \lhd c_k$. Similarly for (3). Now let $x = c^i_j \blacktriangleleft y$, $y \notin C^n$. Then $\pi(x) = c_j \lhd y = \pi(y)$. Similarly for (5). This concludes the proof of the first condition. For the second condition assume $\pi(x) \lhd u$. **Case 1.** $x \notin C^n$. Then $\pi(x) = x$. Now either **(1a)** $u \notin C$; in this case $x \blacktriangleleft u$ and $\pi(u) = u$. Or **(1b)** $u \in C$; in this case $\pi(c^0_j) = u$ for some c_j such that $x \lhd c_j$. Then by (5) $x \lhd c^0_j$. **Case 2.** $x \in C^n$, $x = c^i_j$, for some i, j. Then either **(2a)** $u \notin C$ in which case we have $u = \pi(u)$ and $x \blacktriangleleft u$ by (4); or **(2b)** $u \in C$. In this case we must have $u = c_k$ for $k \equiv j + 1 \pmod{\gamma}$, by the fact that the cycle was chosen to be minimal. Hence if we put $y = c^i_k$ then $x \blacktriangleleft y$ by (3), and $\pi(y) = u$. ◻

We also note the following. If $\iota : \mathfrak{f} \rightarrowtail \mathfrak{g}$, then $\iota[C]$ is a cycle of $\mathfrak{g}$, and $\mathfrak{g}^n$ is defined in an analogous way. Then there is a p–morphism $\iota^n : \mathfrak{f}^n \rightarrowtail \mathfrak{g}^n$ defined by $\iota^n(x) := \iota(x)$ if $x \notin C$ and $\iota^n(c^i_j) := \iota(c_j)^i$ for $c_j \in C$. We call ι^n the **lifting** of ι. Likewise, a p–morphism $\pi : \mathfrak{f} \twoheadrightarrow \mathfrak{g}$ which is injective on C defines a lifting $\pi^n : \mathfrak{f}^n \twoheadrightarrow \mathfrak{g}^n$, with $\mathfrak{g}^n$ analogously defined.

Take two pointed frames $\langle \mathfrak{F}, x\rangle$ and $\langle \mathfrak{G}, y\rangle$. Assume that the underlying sets of worlds are disjoint. Then the connected sum, denoted by $\mathfrak{F} \veebar^x_y \mathfrak{G}$ — or, if the context allows this, by $\mathfrak{F} \veebar \mathfrak{G}$ — is defined as follows.

$$\begin{array}{lcl} \lhd_h & := & \lhd_f \cup \lhd_g \cup \{\langle x, y\rangle\} \\ \mathbb{H} & := & \{a \cup b : a \in \mathbb{F}, b \in \mathbb{G}\} \\ \mathfrak{F} \veebar^x_y \mathfrak{G} & := & \langle f \cup g, \lhd_h, \mathbb{H}\rangle \end{array}$$

It is easy to see that whenever $\{x\}$ is internal in $\mathfrak{F}$, $\mathfrak{F} \veebar^x_y \mathfrak{G}$ is a frame. For $\blacklozenge_h(a \cup b) = \blacklozenge_f a \cup \blacklozenge_g b$ if $y \notin b$ and $\blacklozenge_h(a \cup b) = \blacklozenge_f a \cup \blacklozenge_g b \cup \{x\}$ else.

In what is to follow we will fix $\mathfrak{G}$ to be $\mathfrak{P}_M$, $M \subseteq \omega - \{0\}$, and $\mathfrak{F}$ will be the full frame corresponding to a frame of the form $\mathfrak{f}^n$ for a finite frame with a cycle C. Given this choice, the construction always yields a frame. For this cycle we assume to have fixed an enumeration $C := \{c_i : i < n\}$ and so defines a frame of the form $\mathfrak{f}^n$. Now define $\mathfrak{F} \veebar \mathfrak{P}_M$ to be $\mathfrak{F} \veebar^x_{\infty+1} \mathfrak{P}_M$, where $x = c^{n-1}_1$. Notice that the construction depends on several parameters, the number n, the distinguished cycle, the order of the cycle, and the set M. Let us agree to call a point of $\mathfrak{F} \veebar \mathfrak{P}_M$ **bad** if it is in g, and **good** otherwise. There are only finitely many good points, since $\mathfrak{F}$ is finite. The following theorem holds quite generally for arbitrary $\mathfrak{G}$ in place of $\mathfrak{P}_M$ (if glued at the same point x).

Lemma 7.8.2. *Suppose that $\mathfrak{f} \nvDash \varphi$ and that $d > dp(\varphi)$. Then $\mathfrak{f}^d \veebar \mathfrak{P}_M \nvDash \varphi$.*

Proof. Suppose that $\langle \mathfrak{f}, \beta, x\rangle \vDash \neg\varphi$. Fix the p–morphism $\pi : \mathfrak{f}^d \twoheadrightarrow \mathfrak{f}$ of Lemma 7.8.1. If $x \in C$ choose $y := x$, else $y := x^0$. Take the valuation $\gamma(p) := \pi^{-1}[\beta(p)]$. Then by Proposition 7.8.1, since $d > dp(\varphi)$ we have $\langle \mathfrak{f}^d, \gamma, y\rangle \vDash \neg\varphi$. Now let δ be any valuation on $\mathfrak{f}^d \veebar \mathfrak{G}$ such that $\delta(p) \cap f^d = \gamma(p)$. We claim that $\langle \mathfrak{f}^d \veebar \mathfrak{G}, \delta, y\rangle \vDash \neg\varphi$. For that purpose it is enough to show that the $d - 1$–transit of y in $\mathfrak{f}^d \veebar \mathfrak{G}$ contains no bad points. For then it is identical to the $d-1$–transit of x in $\mathfrak{f}^d$. Since $\blacktriangleleft \cap(\mathfrak{f}^d \times \mathfrak{f}^d) = \lhd^d$, the two transits are isomorphic as frames and the conclusion follows. So, suppose there is a chain $\langle y_i : i < d\rangle$ such that $y_0 = y$ and $y_i \lhd y_{i+1}$, $i < d-1$. If it contains a bad point, by construction of $\mathfrak{f}^d \veebar \mathfrak{P}_M$ there is a $i_0 < d-1$ such that $y_{i_0} = c_1^{d-1}$. Then there must be a $i_1 < i_0$ such that $y_{i_1} = c_0^{d-2}$, a $i_2 < i_1$ such that $y_{i_2} = c_0^{d-3}$ etc. It follows finally that there is a i_{d-1} such that $y_{i_{d-1}} = c_0^0$. But this means that $i_{d-1} < 0$, contradiction. ☐

Let us get some more insight into the structure of $\mathfrak{f}^d \veebar \mathfrak{P}_M$. First, $\mathfrak{P}_M$ is a generated subframe of $\mathfrak{f}^d \veebar \mathfrak{P}_M$. Next, look at the polynomials p_n°, $p_n^\bullet$ and i_n. Since $\mathfrak{P}_M$ is a generated subframe of $\mathfrak{f}^d \veebar \mathfrak{P}_M$, $p_n^\bullet(\Diamond\top \wedge \Box^2\bot)$ is satisfiable in $\mathfrak{f}^d \veebar \mathfrak{P}_M$, and if $n \in M$, $p_n^\circ(\Diamond\top \wedge \Box^2\bot)$ is also satisfiable in that frame. However, if $n \notin M$, then $p_n^\circ(\Box^2\bot)$ is not satisfiable at n°. Moreover, $p_n^\bullet(\Diamond\top \wedge \Box^2\bot)$ is satisfiable only at $n^\bullet$ and $p_n^\circ(\Diamond\top \wedge \Box^2\bot)$ only at n° if $n > \sharp f^d$. From that we deduce

Proposition 7.8.3. *Let $d \in \omega$ and M, N be subsets of ω. Suppose that there exists a $n \in (M-N) \cup (M-N)$ such that $\sharp f^d < n$. $\mathsf{Th}\, \mathfrak{f}^d \veebar \mathfrak{P}_M \not\subseteq \mathsf{Th}\, \mathfrak{f}^d \veebar \mathfrak{P}_N$.*

Proof. Suppose $M \neq N$. Then there exists a n such that $n \in N-M$ or $n \in M-N$. Assume the first. Then $p_n^\circ(\Diamond\top \wedge \Box^2\bot)$ is not satisfiable in $\mathfrak{f}^d \veebar \mathfrak{P}_M$ since $n > \sharp f^d$. Hence $\neg p_n^\circ(\Diamond\top\wedge\Box^2\bot) \in \mathsf{Th}\, \mathfrak{f}^d \veebar \mathfrak{P}_M$. However, $p_n^\circ(\Diamond\top\wedge\Box^2\bot)$ is satisfiable in $\mathfrak{f}^d \veebar \mathfrak{P}_N$. Now assume $n \in M-N$. Then $p_{n+1}^\bullet(\Diamond\top \wedge \Box^2\bot) \to \Diamond p_n^\circ(\Box^2\bot)$ is a theorem of $\mathfrak{f}^d \veebar \mathfrak{P}_M$ but not of $\mathfrak{f}^d \veebar \mathfrak{P}_N$. ☐

Lemma 7.8.4. *Let $\langle \mathfrak{f}, y\rangle$ be a finite pointed frame. Assume that $\mathbb{H}$ is a subalgebra of the algebra of internal sets of $\mathfrak{f} \veebar^y_{\infty+1} \mathfrak{P}_M$. Then either $\mathbb{H}$ is finite or there is a p–morphism $\pi : \mathfrak{f} \veebar^y_{\infty+1} \mathfrak{P}_M \twoheadrightarrow \mathfrak{K}$ for some atomic frame $\mathfrak{K}$, such that $\mathbb{H}$ is the π–preimage of $\mathbb{K}$. Moreover, if $\{*\} \in \mathbb{H}$, then $\mathfrak{K} \cong \mathfrak{g} \veebar^{\pi(y)}_{\infty+1} \mathfrak{P}_M$ for some contraction $\pi : \mathfrak{f} \twoheadrightarrow \mathfrak{g}$.*

Proof. Assume that $\mathbb{H}$ is infinite. Then it contains a nontrivial subset of $\mathfrak{P}_M$. By the results of the previous section, the trace algebra induced by $\mathbb{H}$ on $\mathfrak{P}_M$ is the entire algebra $\mathbb{P}_M$. Now put $x \sim y$ iff for all $a \in \mathbb{H}$, $x \in a$ iff $y \in a$. If x is bad and $\neq *$, then $x \sim y$ only if $y = x$. Hence, $[x]_\sim := \{y : x \sim y\}$ is finite. $\mathfrak{f} \veebar \mathfrak{P}_M$ is atomic; therefore there is a set a_x for every x such that $[x]_\sim = a_x$. Now $x \lhd y$ iff $x \in \blacklozenge a_y$ and from this follows easily that $\sim$ is a net, and we have an induced p–morphism. If $\{*\} \in \mathbb{H}$, then $[*]_\sim = \{*\}$. It is easy to see that the induced map comes from a p–morphism $\pi : \mathfrak{f}^d \twoheadrightarrow \mathfrak{g}^d$. ☐

Lemma 7.8.5. *Let $\langle \mathfrak{f}, y\rangle$ be a finite pointed frame and $\mathfrak{G}$ generated subframe of $\mathfrak{f} \veebar^y_{\infty+1} \mathfrak{P}_M$. Then*

(a) $\mathfrak{G} \cong \mathfrak{g} \veebar^y_{\infty+1} \mathfrak{P}_M$, *where $\mathfrak{g} \leq \mathfrak{f}$ and $y \in g$, or*
(b) $\mathfrak{G} \cong \mathfrak{P}_M$, *or*
(c) $\mathfrak{G} \cong \boxed{\bullet}$, *or*
(d) $\mathfrak{G}$ *is a generated subframe of $\mathfrak{f}$ and $y \notin h$.*

With $\langle \mathfrak{f}, y\rangle$ given, let $\mathfrak{f} \barwedge^{y}_{\infty+1} \mathfrak{P}_M$ or simply $\mathfrak{f} \barwedge \mathfrak{P}_M$ denote the frame which results from $\mathfrak{f} \veebar^{y}_{\infty+1} \mathfrak{P}_M$ by factoring out the net $\sim$ for which $[*]_\sim = \{*, z\}$ for some $z \in f$ and $[x]_\sim = \{x\}$ for all $x \notin \{*, z\}$.

Lemma 7.8.6. *Let $\langle \mathfrak{f}, y\rangle$ be a finite frame. Let Λ be an extension of $\mathsf{Th}\, \mathfrak{f} \veebar^{y}_{\infty+1} \mathfrak{P}_M$. Then Λ is the intersection of logics of the form*

1. $\mathsf{Th}\, \mathfrak{g} \veebar \mathfrak{P}_M$ *or* $\mathsf{Th}\, \mathfrak{g} \barwedge \mathfrak{P}_M$*, where* $\mathfrak{g}$ *is a contractum of some generated subframe of* $\mathfrak{f}$ *containing y,*
2. $\mathsf{Th}\, \mathfrak{P}_M$*,*
3. $\mathsf{Th}\, \boxed{\bullet}$ *and*
4. $\mathsf{Th}\, \mathfrak{g}$*, where* $\mathfrak{g}$ *is a contractum of a subframe of* $\mathfrak{f}$ *not containing y.*

Proof. The proof is by induction on $\sharp f$. We assume that the theorem holds for all frames $\mathfrak{g}$ such that $\sharp g < \sharp f$. Let $\mathfrak{F} = \mathfrak{f} \veebar^{y}_{\infty+1} \mathfrak{P}_M$ or $\mathfrak{F} = \barwedge^{y}_{\infty+1}\mathfrak{P}_M$. There exists a formula ξ such that $\langle \mathfrak{F}, \beta, x\rangle \vDash \xi$ iff the transit of x is $\mathfrak{F}$, and the map $\overline{\beta}$ is onto. Before showing the existence of ξ, let us see how it proves the theorem. Consider an extension Λ of $\mathsf{Th}\, \mathfrak{F}$. It is proper iff it contains $\neg\xi$. Now the algebra $\mathfrak{Fr}_\Lambda(n)$ is a subdirect product of all algebras generated by valuations into $\mathfrak{F}$ satisfying $\neg\xi$. By Lemma 7.8.4 and Lemma 7.8.5 these algebras correspond to frames of the form (a) with g being of smaller cardinality than f, (b) or (c). By induction hypothesis the desired conclusion follows.

Now for the existence of ξ. We perform the argument for $\mathfrak{F} = \mathfrak{f} \veebar \mathfrak{P}_M$. First, $\mathfrak{F}$ is d–transitive with $d := \sharp f + 4$. For each $x \in f$ let p_x be a variable. Add p_* as a variable. Let p_0 be the variable of the root of $\mathfrak{f}$. Define now $\psi := \Diamond^3 p^\bullet_n(\Box p_*) \wedge \Box\Box\bot$ for $n > \sharp f$.

Lemma 7.8.7. *Let $\langle \mathfrak{F}, \beta, x\rangle \vDash \psi$ for some β. Then $x = 0^\bullet$.*

The proof of this lemma is an exercise. Now put

$$\begin{array}{lcl} \rho(n^\bullet) & := & p^\bullet_n(\psi), \\ \rho(n^\circ) & := & p^\circ_n(\psi), \\ \rho(\infty+1) & := & \Diamond\Diamond\rho(0^\bullet) \wedge \neg\Diamond\rho(0^\bullet)\ . \end{array}$$

By Lemma 7.8.7, $\langle \mathfrak{F}, \beta, y\rangle \vDash \rho(\alpha)$, $\alpha \in p_M$, iff $x = \alpha$.

$$\begin{array}{rrl} \delta := & & p_* \to \Box\bot \\ & \wedge & p_y \to \Diamond\rho(\infty+1) \\ & \wedge \bigvee & \langle p_x : x \in f \cup \{*\}\rangle \vee \Diamond^{\leq 2}\rho(0^\bullet) \\ & \wedge \bigwedge & \langle p_x \to \neg p_{x'} : x \neq x'; x, x' \in f \cup \{*\}\rangle \\ & \wedge \bigwedge & \langle p_x \to \Diamond p_{x'} : x \lhd x'; x, x' \in f \cup \{*\}\rangle \\ & \wedge \bigwedge & \langle p_x \to \neg\Diamond p_{x'} : x \ntriangleleft x'; x, x' \in f \cup \{*\}\rangle \end{array}$$

$$\xi := \quad p_0 \wedge \Box^{\leq d}\delta$$

Suppose that $\langle\mathfrak{F},\beta,x\rangle \vDash \xi$. Then $x \vDash p_0$. It can be shown chasing successors that $x \vDash \Diamond^{\leq d} p_y$, and so $x \vDash \Diamond^{\leq d}\Diamond\rho(\infty+1)$. Since $\rho(\infty+1)$ can only be satisfied at $\infty+1$, we have first of all that the transit of x contains $\mathfrak{P}_M$. Moreover, $\overline{\beta}(\rho(\infty+1)) = \{\infty+1\}$, and so the trace algebra of $im[\overline{\beta}]$ relative to $\mathfrak{P}_M$ is $\mathbb{P}_M$. Using this one can show that the transit of x must contain exactly $\sharp f$ many points outside of $\mathfrak{P}_M$. Finally, it is shown that $\beta(p_x) = \{o(x)\}$ for some function $o : f \to f$, and that o is an isomorphism. It follows that x generates the whole frame, and that the algebra induced by $\overline{\beta}$ is the entire algebra of internal sets. This concludes the proof. ∎

Lemma 7.8.8. *Let $\langle\mathfrak{f}, y\rangle$ be a finite frame. Then* $\mathsf{Th}\,\mathfrak{f} \veebar \mathfrak{P}_M$ *has finite codimension in* $\mathcal{E}\,\boldsymbol{K}$.

The previous results also hold if $\mathfrak{P}_M$ is replaced by $\mathfrak{Q}_M$. However, some adaptations have to be made. For example, ψ is now defined by

$$\psi := \Diamond^3 p_n^{\bullet}(\Box p_*) \wedge \Box p_* \ .$$

Furthermore, in δ we need to put $p_* \to \Box p_*$ in place of $p_* \to \Box\bot$. The analogue of Lemma 7.8.7 does not hold with this definition. The entire argument goes through nevertheless, with some slight complication. We will not spell out the full details. This is what we need to prove the next theorem.

Theorem 7.8.9 (Blok). *Let Θ be consistent and not equal to a splitting $\boldsymbol{K}_1/N$ by a set of cycle–free frames. Then Θ is co–covered by $2^{\aleph_0}$ incomplete logics.*

Proof. If Θ is consistent, then $\Theta \subseteq \mathsf{Th}\boxed{\bullet}$ or $\Theta \subseteq \mathsf{Th}\boxed{\circ}$, by Makinson's Theorem. Suppose that $\Theta \subseteq \mathsf{Th}\boxed{\bullet}$. Take a minimal $\mathfrak{f}$ such that $\mathfrak{f}$ is not a frame for Θ. Then there exists a formula φ such that $\mathfrak{f} \nvDash \varphi$ but $\varphi \in \Theta$. Put $d := dp(\varphi) + 1$. Let $M \subseteq \omega - \{0\}$. Then by Lemma 7.8.2, $\mathfrak{f}^d \veebar \mathfrak{P}_M \nvDash \varphi$, so $\Lambda_M := \mathsf{Th}\,\mathfrak{f}^d \veebar \mathfrak{P}_M \not\supseteq \Theta$. Let Λ^o_M be a maximal extension of Λ_M not containing Θ. By Lemma 7.8.6, the fact that $\Theta \subseteq \mathsf{Th}\boxed{\bullet}$ and the choice of $\mathfrak{f}$, Λ^o_M is incomplete. Moreover, it is $\sqcap$–irreducible, and so Λ^o_M is of the form $\mathsf{Th}\,\mathfrak{g} \veebar \mathfrak{P}_M$, $\mathsf{Th}\,\mathfrak{g} \barwedge \mathfrak{P}_M$ or $\mathsf{Th}\,\mathfrak{P}_M$. $\Theta \sqcap \Lambda^o_M$ co–covers Θ. For if $\Theta \supsetneq \Theta' \supsetneq \Theta \sqcap \Lambda^o_M$ then $\Theta' \sqcup \Lambda^o_M \supseteq \Theta$, since $\Theta' \sqcup \Lambda^o_M$ properly contains Λ^o_M. On the other hand $\Theta = (\Theta' \sqcup \Lambda^o_M) \sqcap \Theta = (\Lambda^o_M \sqcap \Theta) \sqcup \Theta' = \Theta' \not\subseteq \Theta$. Contradiction. Λ^o_M is incomplete. and by minimality of $\mathfrak{f}$ all finite frames of Λ_M are frames for Θ. It remains to be seen that there exist $2^{\aleph_0}$ many different such logics. To that effect note that by Proposition 7.8.3 that there is a number c such that if $(N - M) \cup (M - N)$ contains an $n \geq c$ then $\mathsf{Th}\,\mathfrak{g} \veebar \mathfrak{P}_M$ and $\mathsf{Th}\,\mathfrak{g} \veebar \mathfrak{P}_N$ are incomparable, and similarly $\mathsf{Th}\,\mathfrak{g} \barwedge \mathfrak{P}_M$ and $\mathsf{Th}\,\mathfrak{g} \barwedge \mathfrak{P}_N$ are incomparable. Hence, for such sets Λ^o_M and Λ^o_N are incomparable. Moreover, $\Theta \cap \Lambda^o_M \neq \Theta \cap \Lambda^o_N$. Both are co–covers of Θ. Hence, any set N such that $N \cap \{0, \ldots, c-1\} = M \cap \{0, \ldots, c-1\}$ would have sufficed equally well. There are $2^{\aleph_0}$ many such sets. This concludes the proof of the case $\Theta \subseteq \mathsf{Th}\boxed{\bullet}$. If $\Theta \not\subseteq \mathsf{Th}\boxed{\bullet}$, use $\mathfrak{Q}_M$ is place of $\mathfrak{P}_M$. ∎

Little remains to be done to complete the proof of Blok's Alternative. Namely, we need to show that there arise no new splittings except for $\mathbf{K.D} = \mathbf{K}/\boxed{\bullet}$, which has a splitting frame $\boxed{\circ}$, yielding the inconsistent logic as an iterated splitting of **K**. Since the inconsistent logic was always included in the theorems, they hold for that case as well. Hence, the following theorem concludes the proof.

Theorem 7.8.10 (Blok). *Let* $\Lambda = \boldsymbol{K}_1/N$ *for a set* N *of cycle–free frames. Then either* $\boxed{\bullet} \in N$ *and* $\mathsf{Th}\,\boxed{\circ}$ *splits* $\mathcal{E}\Lambda$ *or* $\boxed{\bullet} \notin N$ *and* Θ *splits* $\mathcal{E}\Lambda$ *iff it splits* $\mathcal{E}\boldsymbol{K}_1$.

Proof. Let $\boxed{\bullet} \notin N$. Then $\boxed{\bullet}$ is a frame for Λ. Assume that $\mathfrak{f}$ is not cycle–free and minimally so, that is, for every frame $\mathfrak{g}$ to which $\mathfrak{f}$ reduces either $\mathfrak{g} = \mathfrak{f}$ or $\mathfrak{g}$ is cycle–free. Then

$$\prod \langle \mathsf{Th}\,\mathfrak{f}^d \veebar \mathfrak{P}_M : d \in \omega \rangle \subseteq \mathsf{Th}\,\mathfrak{f}$$

But for no d $\mathsf{Th}\,\mathfrak{f}^d \veebar \mathfrak{P}_M \subseteq \mathsf{Th}\,\mathfrak{f}$. Hence $\mathsf{Th}\,\mathfrak{f}$ is not prime in $\mathcal{E}\Lambda$. ◻

Exercise 262. Show that if a frame contains a cycle, it contains a minimal cycle.

Exercise 263. Let $\mathfrak{L}$ be a distributive lattice, $x, y \in L$. Show that the interval $[x \sqcap y, y]$ is isomorphic to the interval $[x, x \sqcup y]$. (In fact, distributivity is stronger than necessary.)

Exercise 264. Show Lemma 7.8.7. *Hint.* Show first that $\langle \mathfrak{F}, \beta, x\rangle \vDash p_n^\bullet(p)$ implies that there is a chain $\langle x_i : i < n+1\rangle$ such that $x = x_0$, $x_i \lhd x_j$ iff $i < j$, and that all x_i are different.

7.9. The Lattice of Tense Logics

Tense logic is on the one hand an important area of modal logic when it comes to its applications. On the other hand it has proved to be influential also in the theoretical development of modal logic. It was here that many counterexamples to general conjectures have been found. Moreover, it has been shown that the rich theory of extensions of **K4** is not the effect of it being operator transitive, but rather of the combination of it being operator transitive and weakly transitive. The first implies the second when there is one basic operator, but already in tense logic this fails to be the case. The consequences of this will be seen below. We begin this section by showing that in the lattice of tense logic there exist incomplete logics of codimension 1. The first example is due to S. K. Thomason [**206**]. (See also Section 7.7.) We present it in the form given in Rautenberg [**169**]. Let

$$\Lambda := \mathbf{G.3}.t \oplus \mathbf{K4.3.D}^{-} \oplus \Box\Diamond p \to \Diamond\Box p\ .$$

A Kripke–frame for $\mathbf{G.3}.t \oplus \mathbf{K4.3}^{-}$ is nothing but a converse well–order λ^{op}. A Kripke–frame for $\mathbf{G.3}.t \oplus \mathbf{K4.3.D}^{-}$ is a converse well–order λ^{op} where λ is a limit–ordinal. It is easy to show, however, that no such frame can be a frame for $\Box\Diamond p \to \Diamond\Box p$. Λ therefore has no Kripke–frames. But it is not inconsistent. For example, let Ω be the frame based on ω with the finite and cofinite subsets as internal sets. Then Ω is a frame for Λ. It can be shown that $\Lambda = \mathsf{Th}\,\Omega$.

Theorem 7.9.1 (Thomason). *The lattice of tense logics has an incomplete logic of codimension 1.*

Already, this allows to deduce that the lattice $\mathcal{E}\,\mathbf{K}_1$ has an incomplete logic of codimension 2. So, Theorem 2.9.9 is the best possible result. One can play several variations on that theme. It is possible to produce $2^{\aleph_0}$ many incomplete logics of any finite codimension in $\mathcal{E}\,\mathbf{K}.t$. It that way many results of Section 7.8 can be shown using tense logic.

The previous section has illustrated that we can deduce important facts about logics if we concentrate on certain sublattices. These sublattices also had the best properties we could expect, namely to be continuous sublattices. The general situation is far from being that good. As an illustration we will study some basic properties of the lattice of tense logics. It is easy to see that there are embeddings of the lattice of modal logic into the lattice of tense logics, and that there are also maps in the opposite direction. However, these maps usually behave well only with respect to one of the two operations. The general situation is as follows. We have a pair of modal languages, $\mathcal{L}_1$ and $\mathcal{L}_2$, and $\mathcal{L}_1$ is a sublanguage of $\mathcal{L}_2$. Then a logic Θ over $\mathcal{L}_1$ has a natural least extension Θ^e in the lattice of $\mathcal{L}_2$–logics. Moreover, each $\mathcal{L}_2$–logic gives rise to an $\mathcal{L}_1$–logic by restricting to the language $\mathcal{L}_1$. Since we are always dealing with normal modal logics, this is satisfied.

We will study the extension lattices of $\mathbf{K}.t$, $\mathbf{K4}.t$ and $\mathbf{S4}.t$. The method is uniform in all three cases and can be transferred to numerous other logics. Each of these logics has the finite model property and therefore only finite algebras can induce splittings. Thus we can concentrate on the Kripke–frames of those algebras. Here is an important fact about tense algebras.

Proposition 7.9.2. *For finite tense algebras $\mathfrak{A}$ the following are equivalent.*

1. $\mathfrak{A}$ *is subdirectly irreducible.*
2. $\mathfrak{A}$ *is directly irreducible.*
3. $\mathfrak{A}^+$ *is connected.*
4. $\mathfrak{A}$ *is simple.*

This can easily be established by using the duality theory. It can also be obtained using the methods of Section 4.3. Notice that this theorem fails for infinite algebras; a counterexample will be given below in the exercises. This is due to the fact that there exist logics which are not weakly transitive. Notice also that $\mathbf{K4}.t$ and $\mathbf{S4}.t$ while being operator transitive are not weakly transitive. This has to be borne in mind.

For the definition of subreducing frames we use the following construction. Take two frames $\mathfrak{f}$, $\mathfrak{g}$ and let $x \in f$, $y \in g$. Then let $\mathfrak{f}\,{}^x\wr_y\,\mathfrak{g}$ denote the frame $\langle f\,{}^x\wr_y\, g, \lhd\rangle$ where $f\,{}^x\wr_y\, g = f \times \{0\} \cup g \times \{1\} - \{\langle y, 1\rangle\}$ and $\lhd$ is defined by

(i) $\langle a, 0\rangle \lhd \langle b, 0\rangle$ iff $a \lhd_{\mathfrak{f}} b$
(ii) $\langle a, 1\rangle \lhd \langle b, 1\rangle$ iff $a \lhd_{\mathfrak{g}} b$
(iii) $\langle x, 0\rangle \lhd \langle b, 1\rangle$ iff $y \lhd_{\mathfrak{g}} b$

FIGURE 7.13.

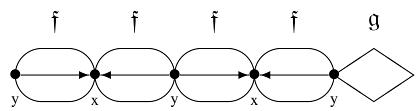
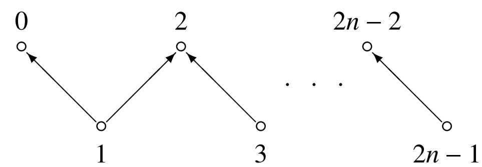


FIGURE 7.14. The garland $\mathfrak{Gl}_{2n-1}$



This is well–defined whenever $x \lhd_{\mathfrak{f}} x \Leftrightarrow y \lhd_{\mathfrak{g}} y$. If $\mathfrak{f}$, $\mathfrak{g}$ are transitive, $\lhd$ will be taken to be the transitive closure of the relation defined above. We call $\mathfrak{f}\,{}^{x}\!\wr_{y}\,\mathfrak{g}$ a **book** and $\mathfrak{f}$ and $\mathfrak{g}$ the **pages**. When the choice of the points is clear we write $\mathfrak{f} \wr \mathfrak{g}$ instead of $\mathfrak{f}\,{}^{x}\!\wr_{y}\,\mathfrak{g}$. With two points $x \lhd y \in f$ fixed we define ${}^{n}\mathfrak{f}$, $n \geq 1$, by

$$\begin{aligned} {}^{1}\mathfrak{f} &:= \mathfrak{f} \\ {}^{2k+1}\mathfrak{f} &:= {}^{2k}\mathfrak{f}\,{}^{x}\!\wr_{x}\,\mathfrak{f} \\ {}^{2k+2}\mathfrak{f} &:= {}^{2k+1}\mathfrak{f}\,{}^{y}\!\wr_{y}\,\mathfrak{f} \end{aligned}$$

We distinguish the elements of different pages in ${}^{n}\mathfrak{f}$ by indices $0, \ldots, n-1$. The map $\varphi : x_i \mapsto x$ is a p–morphism; for if $x_i \lhd y_j$ then either $i = j$ and thus $x \lhd y$ by (i) and (ii) or $i + 1 = j$ and $x \lhd y$ by (iii). Now if $\varphi(x_i) \lhd y$ we have $y = \varphi(y_i)$ and $x_i \lhd y_i$. Likewise for $x \lhd \varphi(y_i)$. The same can be shown in the transitive case. By this we see that any map $\psi : {}^{n}\mathfrak{f} \wr \mathfrak{g} \rightsquigarrow \mathfrak{f}$ satisfying $\psi(x_i) = x$ is $n-1$–localic with respect to any point x_0 of the first page. This suggests that by taking a suitable $\mathfrak{g}$ so that there is no p–morphism from ${}^{n}\mathfrak{f} \wr \mathfrak{g}$ to $\mathfrak{f}$ for any $n \in \omega$, $\{\mathsf{Th}\,{}^{n}\mathfrak{f} \wr \mathfrak{g} : n \in \omega\}$ is a subreduction of $\mathsf{Th}\,\mathfrak{f}$. A particularly important class of frames for our purposes and for illustration of the books let us introduce the *garlands*. A garland is a zigzag frame shown in Figure 7.14. Formally, we define $\mathfrak{gl}_n$ as a frame $\langle n+1, \lhd\rangle$ where $i \lhd j$ iff $i = j$ or i is odd and $j = i \pm 1$. Thus, $\mathfrak{g}_n$ has $3n$ arrows and $n+1$ points. Note that a garland is isomorphic to a book where each page is the frame $\boxed{\circ\!\longrightarrow\!\circ}$. Garlands can be characterized modally. A frame $\mathfrak{f} = \langle g, \lhd\rangle$ is called **meager** if there are no two points $x \lhd y \lhd x$. A connected frame $\mathfrak{f}$ is a garland iff it is reflexive, meager, of alternativity 3 in both directions, and of depth 2. Thus $\mathbf{Ga} := \mathbf{Grz}.t.\mathbf{alt}_3^{+}.\mathbf{alt}_3^{-}$ as the reader may verify. We first prove an important

Lemma 7.9.3. *Suppose that* $\mathfrak{f}$ *is not a cluster. Then* $\mathsf{Th}\,\mathfrak{g}$ *splits* $\mathcal{E}\,\boldsymbol{K}.t$*,* $\mathcal{E}\,\boldsymbol{K4}.t$ *and* $\mathcal{E}\,\boldsymbol{S4}.t$ *only if* $\mathfrak{g}$ *is a garland.*

Proof. $\mathfrak{f}$ is finite and connected and $\sharp f = n$. Since $\mathfrak{f}$ is not a cluster, there are points x, y such that $x \lhd y \ntrianglelefteq x$. Also $\mathfrak{f} = \mathfrak{Tr}^n_{\mathfrak{f}}(x)$. Now consider the frame ${}^m\mathfrak{f} \wr_0 \mathfrak{gl}_{2n+8}$. This is well defined in case x and y are both reflexive points. In case one of them is not reflexive we take ${}^m\mathfrak{f} \wr_0 \mathfrak{gl}^\circ_{2n+8}$ instead, where $\mathfrak{gl}^\circ_{2n+8}$ is identical to $\mathfrak{gl}_{2n+8}$ except that $0 \ntriangleleft 0$. We have to show that if there is a p–morphism (in both relations) $\pi : {}^m\mathfrak{f} \wr_0 \mathfrak{gl}^{(\circ)}_{2n+8} \to \mathfrak{f}$ then (a) $\mathfrak{f}$ is reflexive and transitive, (b) $\mathfrak{f}$ is of alternativity ≤ 3, (c) $\mathfrak{f}$ is meager and (d) $\mathfrak{f}$ is of depth ≤ 2. For then $\mathfrak{g}$ splits the lattice only if it satisfies (a) – (d) and thus is a garland. The details of the proof are left to the reader. The crucial fact to be shown is that if $x \in f$ there exists a $z \in gl_{2n+8}$ such that $p(z) = x$. For then $\pi[\mathfrak{gl}_{2n+8}] = \mathfrak{f}$, from which the rest follows. ◻

This considerably reduces the class of possible splitting frames. However, we will also show that most of the garlands and clusters cannot split any of these logics. We do this by establishing a lemma on splittings of $\mathcal{E}$ **Ga**.

Lemma 7.9.4. $\mathsf{Th}\,\mathfrak{gl}_n$ *splits* $\mathcal{E}\,\boldsymbol{Ga}$ *iff* $n \leq 1$.

Once this lemma is proved it follows that $\mathfrak{gl}_n$ cannot split $\mathcal{E}$ **K**.t, $\mathcal{E}$ **K4**.t nor $\mathcal{E}$ **S4**.t for $n > 1$ since all these lattices contain $\mathcal{E}$ **Ga**.

Proposition 7.9.5. *There is a p–morphism* $\pi : \mathfrak{gl}_m \to \mathfrak{gl}_n$ *iff* $n = 0$*,* $m = \omega$ *or* n *divides* m*.*

Proof. ($\Rightarrow$) Suppose $\pi : \mathfrak{gl}_m \to \mathfrak{gl}_n$ is a p–morphism. Assume $n > 0$. Write $i \equiv j$ for $\pi(i) = \pi(j)$. We now have the following

Claim 7.9.6. *On the condition that* $n > 0$*, if* $i \equiv j$ *then* $i \equiv j (mod\ 2)$*. Moreover, if* $i \equiv j$ *then* $i - 1 \equiv j - 1$ *or* $\equiv j + 1$ *and* $i + 1 \equiv j - 1$ *or* $\equiv j + 1$*, whenever these points exist.*

For suppose $i \equiv j$ and that i is even and j is odd. Then $j \lhd j-1, j$ and if $j+1 \leq m$ then also $j \lhd j+1$. By the p–morphism condition for $\lhd$ and the fact that $i \lhd k$ iff $k = i$ we get $j - 1 \equiv i$ and $j + 1 \equiv i$. Similarly, if $i > 0$ then by the p–morphism condition for $\rhd$ we have $i - 1 \equiv j$ and if $i < m$ also $i + 1 \equiv j$. Continuing this argument we get $k \equiv \ell$ for all k, ℓ and hence $n = 0$, which we have excluded. Now let again $i \equiv j$. Then if both are even, $i-1, i, i+1 \lhd i$ and $j-1, j, j+1 \lhd j$ whenever these points exist. By the second p–morphism condition, either $i-1 \equiv j-1$ or $i-1 \equiv j$ or $i-1 \equiv j+1$. But since j is even and $i-1$ is odd, $i-1 \equiv j$ cannot hold. Likewise, $i+1 \equiv j$ as well as $i \equiv j-1, j+1$ cannot occur. This proves Claim 7.9.6.

In order to prove that m is a multiple of n we look at subsets C of $\mathfrak{gl}_m$ which are connected and on which $\pi \restriction C$ is injective. Such sets are called **partial sections**. If $\pi \restriction C$ is also surjective, in other words, if $\sharp C = n + 1$ then C is called **section**. We now prove

Claim 7.9.7. *If C is a partial section and $\sharp C > 1$ then C is contained in exactly one section.*

To see this observe first that since C is connected and π is a $\lhd$–homomorphism, $\pi[C]$ is connected as well. Therefore, $C \subseteq \{0, \dots, m\}$ is an interval as is $\pi[C] \subseteq \{0, \dots, n\}$. If $i, i+1 \in C$ and $\pi : i \mapsto k$ then by Claim 7.9.6, $\pi(i+1) = k+1$ or $\pi(i+1) = k-1$. If $\pi : i+1 \mapsto k+1$ then $\pi : i+2 \mapsto k+2$, for by the same argument $\pi : i+2 \mapsto k, k+2$ but $\pi(i+2) = k$ contradicts the injectivity of $\pi \upharpoonright C$. But if $\pi : i+1 \mapsto k-1$ then similarly $\pi : i+2 \mapsto k-2$. So, by induction, either $\pi \upharpoonright C$ is a strictly increasing or strictly decreasing function. Now if $C = \{i, \dots, j\}$ and $\pi[C] = \{k, \dots, \ell\}$ and π is monotone increasing then if $k > 0$, $\pi(i) = k > 0$ we get by the second p–morphism condition that either $i-1$ or $i+1$ is in the preimage of $k-1$. But $\pi(i+1) = k+1$. Thus $i > 0$ and $\pi(i-1) = k-1$. So we add $i-1$ to C. Likewise, $\pi(j) = \ell$ and if $\ell < n$ then $j < m$ and we add $j+1$ to C. Similarly, if π is decreasing. This proves Claim 7.9.7.

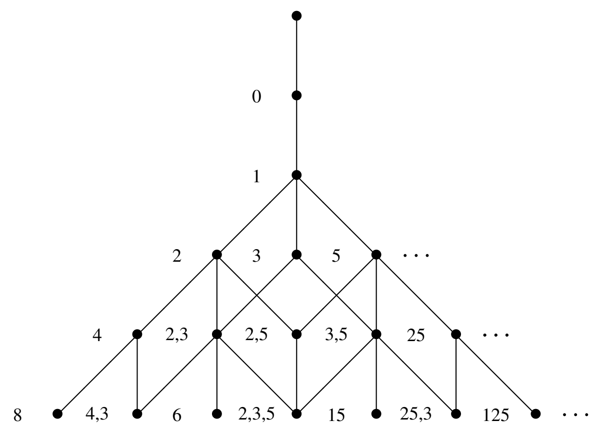
Now $\mathfrak{gl}_m$ contains exactly m subsets of the form $\{i, i+1\}$. π is injective on each of them and they are all contained in one and only one section. Each section contains $n+1$ points and thus n subsets $\{j, j+1\}$. Hence n divides m or $m = \omega$.
($\Leftarrow$) If $n = 0$ take the constant map $j \mapsto 0$. If $n > 0$, $\mathfrak{g}_m$ must be covered by sections as follows. If S, T are sections then $S = T$ or $\sharp(S \cap T) \leq 1$. Each section is an interval of $n+1$ points and each pair $\{i, i+1\}$ is in exactly one section. Hence the sections are $S_k = \{nk, nk+1, \dots, n(k+1)\}$. On each section π is bijective. Suppose that π is increasing on S_i. Then $\pi(n(k+1)) = n$. Thus π must be decreasing on S_{i+1} and vice versa. Thus let π be increasing on all even numbered sections S_{2i} and decreasing on all odd numbered sections S_{2i+1}. Thus $\pi(i) = s$ iff $i = 2kn + s$ or $i = 2(k+1)n - s$ for some k. We show that π defined this way is a p–morphism. We have $i \lhd i$ and $\pi(i) \lhd \pi(i)$. Moreover, if $i \lhd i+1$ or $i \lhd i-1$ then i is odd. Now if $\pi(i) = s$ then either $i = 2kn + s$ or $i = (2k+2)n - s$. In both cases s is odd as well and $s \lhd s+1, s-1$ and $\{s-1, s+1\} = \pi[\{i-1, i+1\}]$. Similarly if i is even. Now suppose $\pi(i) \lhd \ell$. If $i \in S$, $\pi(i) \lhd \ell$, then take $s \in \pi^{-1}(\ell) \cap S$. s is unique. Since $\pi \upharpoonright S : \langle S, \lhd \rangle \to \mathfrak{gl}_n$ is an isomorphism, $i \lhd s$ as well. Similarly if $\ell \lhd \pi(i)$. ◻

With this result in our hands we can probe quite deeply into the structure of $\mathcal{E}\,\mathbf{Ga}$ and also prove the desired theorem. We have that $\mathbf{Ga} = \mathsf{Th}\,\mathfrak{gl}_\omega$ since $\mathsf{Th}\,\mathfrak{gl}_\omega \supseteq \mathsf{Th}\,\mathfrak{gl}_n$ for every n. Each logic containing $\mathbf{Ga}$ must be complete. This is due to the fact that logics of bounded alternative are complete in general. The $\bigsqcap$–irreducible elements are the $\mathsf{Th}\,\mathfrak{gl}_n$ for $n \in \omega$. Every proper extension of $\mathbf{Ga}$ which is not trivial is therefore an intersection $\bigsqcap \langle \mathsf{Th}\,\mathfrak{gl}_n : n \in F \rangle$ where $F \subseteq \omega$ is finite. For if F is infinite we immediately have

$$\bigsqcap \langle \mathsf{Th}\,\mathfrak{g}_n : n \in F \rangle = \mathbf{Ga}\,,$$

since $\mathfrak{gl}_\omega$ is contained in $\mathsf{Up}\,\mathfrak{gl}_n$ for a non–trivial ultrafilter U on F. (This can also be shown without the use of ultrafilters. This is left as an exercise.) Therefore every proper extension of $\mathbf{Ga}$ is tabular while $\mathbf{Ga}$ itself is not.

FIGURE 7.15. $\mathcal{E}$ **Ga**



THEOREM 7.9.8. ***Ga*** *is pretabular.*

It is now straightforward to map out the structure of the locale $\mathcal{E}$ **Ga**. We will establish here the structure of the corresponding T_D–space. Recall that the latter consists of the $\bigsqcap$–irreducible logics as points, and the closed sets correspond to sets of the form $\uparrow\Lambda$ where Λ is a logic. Thus, by the results obtained above, the points are logics $\mathsf{Th}\,\mathfrak{g}_n$, $n \in \omega$. Let $\mu := \langle \omega - \{0\}, |\rangle^{op}$ where $m \mid n$ iff m divides n. Now let $\mu + 1$ be the poset obtained by adding a new top element. Recall that $\Phi(\mu + 1)$ denotes the weak topology on $\mu + 1$. In this topology, an upper set is closed if it is (i) the entire space or (ii) a finite union of sets of the form $\uparrow x$.

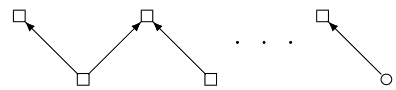
THEOREM 7.9.9. $\mathfrak{I}\mathfrak{Spc}(\mathcal{E}\ \boldsymbol{Ga}) \cong \Phi(\mu + 1)$.

The corresponding space $\mathfrak{Spc}(\mathcal{E}\,\mathbf{Ga})$ can easily be determined. It has a new element at the bottom (corresponding to **Ga** itself, which is $\sqcap$–irreducible but not $\bigsqcap$–irreducible). The closed sets are those sets which are finite (and hence do not contain **Ga**) or contain **Ga** (and hence all other elements). Thus the upper part of $\mathcal{E}$ **Ga** is depicted in Figure 7.15. To the left of each node we have written numbers n such that the node is the intersection of the logics of the corresponding $\mathfrak{gl}_n$.

PROPOSITION 7.9.10. *For $n \geq 3$ there are infinitely many logics of codimension n in $\mathcal{E}$* ***Ga****.*

The proof of Lemma 7.9.4 is now easy. Clearly, both $\mathfrak{gl}_0$ and $\mathfrak{gl}_1$ split the lattice. But for $n > 1$ observe that the sequence $\langle \mathfrak{gl}_p : p \text{ prime}, p > n\rangle$ subreduces $\mathfrak{gl}_n$. As we have noted, this implies also that none of the garlands $\mathfrak{gl}_n$ split $\mathcal{E}\,\mathbf{S4}.t$ unless $n \leq 1$. It will turn out soon that we cannot improve this result for $\mathcal{E}\,\mathbf{S4}.t$. But for $\mathcal{E}\mathbf{K}.t$ and

Figure 7.16. $\mathfrak{h}_k$



$\mathcal{E}\,\mathbf{K4}.t$ even these cases are ruled out. Look at the sequence $\langle \mathfrak{gl}_n^\circ : n \in \omega \rangle$ where $\mathfrak{gl}_n^\circ$ differs from $\mathfrak{gl}_n$ in that $n \ntriangleleft n$. The maps π and ρ defined by $\pi : \mathfrak{gl}_n^\circ \rightsquigarrow \mathfrak{gl}_1 : j \mapsto j$ (mod 2) and $\rho : \mathfrak{gl}_n^\circ \rightsquigarrow \mathfrak{gl}_0 : j \mapsto 0$ are n–localic with respect to 0. Thus this sequence subreduces both frames in $\mathcal{E}\,\mathbf{K4}.t$ and in $\mathcal{E}\,\mathbf{K}.t$.

Lemma 7.9.11. *For no n, $\mathsf{Th}\,\mathfrak{gl}_n$ splits $\mathcal{E}\,\boldsymbol{K}.t$, $\mathcal{E}\,\boldsymbol{K4}.t$.*

It now remains to treat the clusters. Here the situation is quite similar to the situation of the garlands.

Lemma 7.9.12. *Suppose $\mathfrak{g}$ is a cluster. Then $\mathsf{Th}\,\mathfrak{g}$ splits $\mathcal{E}\,\boldsymbol{K}.t$ and $\mathcal{E}\,\boldsymbol{K4}.t$ only if $\mathfrak{g} \cong \boxed{\bullet}$ and $\mathcal{E}\,\boldsymbol{S4}.t$ only if $\mathfrak{g} \cong \boxed{\circ}$.*

Proof. Let $n := \sharp g > 1$ and $\mathfrak{h}_k = \langle h_k, \triangleleft \rangle$ with

$$h_k = \{0, \dots, k\} \times \{1, \dots, n\} - \{\langle k, n \rangle\}$$

and $\langle i, j \rangle \triangleleft \langle i', j' \rangle$ iff (i) i is odd, $i' = i + 1$ or $i - 1$ or (ii) i is even and $i' = i$. This can be visualized by Here, $\square$ denotes a cluster with n points and $\circ$ a cluster with 1 point. There is no p–morphism from $\mathfrak{h}_k$ into $\mathfrak{g}$ as there is no way to map a point belonging to an $n - 1$–point cluster onto a n–point cluster.

Now look at the k–transit of $\langle 0, 0 \rangle$ in $\mathfrak{h}_k$; call it $\mathfrak{e}$. Let e be its underlying set. Every point in e is contained in an n–point cluster since $\langle i, j \rangle \in e$ iff $i < k$. Thus there is a p–morphism $\pi : \mathfrak{e} \to \mathfrak{g}$. Extend π to a map $p^+ : \mathfrak{h}_k \rightsquigarrow \mathfrak{g}$. p^+ is k–localic with respect to $\langle 0, i \rangle$ for every i. Hence $\mathfrak{h}_k$ is k–consistent with $\mathfrak{g}$. It follows that $\langle \mathsf{Th}\,\mathfrak{h}_k : k \in \omega \rangle$ is a subreduction of $\mathsf{Th}\,\mathfrak{g}$. $\square$

Now we have collected all the material we need to prove the splitting theorems. Notice that a splitting frame for any of these logics can only be one–point cluster or a two–point garland. We will now show that the frames not excluded by the above lemmata are indeed splitting frames.

Theorem 7.9.13. *Λ splits the lattice $\mathcal{E}\boldsymbol{S4}.t$ iff $\Lambda = \mathsf{Th}\,\mathfrak{gl}_1 = \mathsf{Th}\,\boxed{\circ \longrightarrow \circ}$ or $\Lambda = \mathsf{Th}\,\mathfrak{gl}_0 = \mathsf{Th}\,\boxed{\circ}$.*

Proof. ($\Leftarrow$) The nontrivial part is $\mathfrak{gl}_1$. We will show $\mathcal{E}\,\mathbf{S4}.t/\mathfrak{gl}_1 = \mathbf{S5}.t$ by proving that (†) of the Splitting Theorem holds for $m = 1$. Therefore let $\mathfrak{A}$ be an algebra satisfying $\mathsf{Th}\,\mathfrak{A} \nsupseteq \mathbf{S5}t$. Then there is a set $c \in A$ of $\mathfrak{A}$ such that $0 < c \cap \blacklozenge\blacksquare - c$. Consequently, in the underlying Kripke–frame there are two points $s \triangleleft t$ such that

$s \in c$ and $t \in \blacksquare\!\!- -\, c$ whence $t \ntriangleleft s$. Now we have $\delta(\mathfrak{gl}_1) = (p_a \leftrightarrow \neg p_b) \wedge \Diamond\!\!- p_b \wedge -\!\!\Diamond p_a \wedge (\Box\!\!- p_b \leftrightarrow p_b) \wedge (-\!\!\Box p_a \leftrightarrow p_a)$. Suppose that under these circumstances we can construct valuations $\alpha_n : \{p_a, p_b\} \to A$ such that $s \in \overline{\alpha}_n(p_a \wedge \Box^{\leq n}\delta(\mathfrak{gl}_1))$. Then (†) of the Splitting Theorem, 7.3.11 (2), $\mathfrak{gl}_1$ splits $\mathcal{E}\,\mathbf{S4}.t$. To construct the α_n we define inductively subsets a_n, b_n in $\mathfrak{A}$ as follows:

$$
\begin{array}{llll}
(0) & b_0 & := & \blacksquare\!\!- -\, c \\
 & a_0 & := & -b_0 \\
(A) & a_{2k+1} & := & -b_{2k+1} \\
 & b_{2k+1} & := & b_{2k} \cap -\!\!\blacklozenge\, a_{2k} \\
(B) & a_{2k+2} & := & a_{2k+1} \cap \blacklozenge\!\!- b_{2k+1} \\
 & a_{2k+2} & := & -a_{2k+1}
\end{array}
$$

Furthermore define $T_0 := \{s\}$, $T_{2k+1} := \blacklozenge\!\!- T_{2k}$, $T_{2k+2} := -\!\!\blacklozenge\, T_{2k+1}$. The T_n are not necessarily internal. Since $T_{2k} \subseteq -\!\!\blacksquare\ \blacklozenge\!\!- T_{2k} = -\!\!\blacksquare\, T_{2k+1}$ and furthermore $T_{2k+1} \subseteq \blacksquare\!\!-\, -\!\!\blacklozenge\, T_{2k+1} = \blacksquare\!\!- T_{2k+1}$ it follows that $T_n \subseteq \boxtimes^{\leq 1} T_{n+2}$. For $\boxtimes^{\leq 1} T_{2k+2} := \blacksquare\!\!- T_{2k+2} \cap -\!\!\blacksquare\, T_{2k+2} \supseteq T_{2k+1} \cap -\!\!\blacksquare\, T_{2k+1} \supseteq T_{2k}$ and dually for odd n. Consequently, $s \in \boxtimes^{\leq n} T_{n+1}$. We now verify the following claims:

$$
\begin{array}{llll}
(I) & a_n \cap b_n & = & \varnothing \\
 & a_n \cup b_n & = & 1 \\
(II) & a_n & = & -\!\!\blacksquare\, a_n \\
 & b_n & = & \blacksquare\!\!- b_n \\
(III) & T_n \cap a_n & = & T_n \cap a_{n+1} \\
 & T_n \cap b_n & = & T_n \cap b_{n+1} \\
(IV) & T_n & \subseteq & \blacklozenge\!\!- b_n \cap -\!\!\blacklozenge\, a_n
\end{array}
$$

(I) holds by construction. (II) – (IV) are verified by induction; for (II) we only need to show $a_n \subseteq -\!\!\blacksquare\, a_n$ and $b_n \subseteq \blacksquare\!\!- b_n$. By symmetry of (A) and (B) we may restrict to (A). $b_{2k+1} = b_{2k} \cap -\!\!\blacklozenge\, a_{2k} = \blacksquare\!\!- b_{2k} \cap -\!\!\blacklozenge\, a_{2k} \subseteq \blacksquare\!\!- b_{2k} \cap \blacksquare\!\!-\, -\!\!\blacklozenge\, -\!\!\blacklozenge\, a_{2k} = \blacksquare\!\!- (b_{2k} \cap -\!\!\blacklozenge\, a_{2k}) = \blacksquare\!\!- b_{2k+1}$, $a_{2k+1} = -b_{2k+1} = -\blacksquare\!\!- b_{2k+1} = \blacklozenge\!\!- a_{2k+1} \subseteq -\!\!\blacksquare\ \blacklozenge\!\!-\ \blacklozenge\!\!- a_{2k+1} = -\!\!\blacksquare\ \blacklozenge\!\!- a_{2k+1} = -\!\!\blacksquare\ -\ \blacksquare\!\!- b_{2k+1} = -\!\!\blacksquare\ - b_{2k+1} = -\!\!\blacksquare\, a_{2k+1}$. This shows (II). For (III) we now observe that $T_{2k} \cap b_{2k+1} = T_{2k} \cap b_{2k} \cap -\!\!\blacklozenge\, a_{2k} = T_{2k} \cap b_{2k}$ since $T_{2k} \subseteq -\!\!\blacklozenge\, a_{2k}$ by (IV). $T_{2k} \cap a_{2k+1} = T_{2k} \cap a_{2k}$ immediately follows. To prove (IV) we observe that if $y \in T_{2k+1}$ there is a $x \in T_{2k}$ such that $x \rhd y$. By induction hypothesis we have $x \in \blacklozenge\!\!- b_{2k}$ and therefore there is a $z \rhd x$ such that $z \in T_{2k} \cap b_{2k}$. Now $z \in b_{2k} \cap -\!\!\blacklozenge\, a_{2k} = b_{2k+1}$ and, as $y \lhd x \lhd z$, $y \in \blacklozenge\!\!- b_{2k+1}$. To show $y \in -\!\!\blacklozenge\, a_{2k+1}$ we distinguish two cases: (α) $y \in a_{2k+1}$ and (β) $y \in b_{2k+1}$. In case (α) we immediately have $y \in -\!\!\blacklozenge\, a_{2k+1}$ and in case (β) we have $y \in -\!\!\blacklozenge\, a_{2k}$. But as

$$-\!\!\blacklozenge\, a_{2k+1} = -\!\!\blacklozenge\, (a_{2k} \cup -\!\!\blacksquare\, b_{2k}) \supseteq -\!\!\blacklozenge\, a_{2k}$$

we also have $y \in -\!\!\blacklozenge\, a_{2k+1}$.

Now we put $\alpha_n : p_a \mapsto a_{n+1}, p_b \mapsto b_{n+1}$. It remains to be shown that $s \in \overline{\alpha}_n(p_a \wedge \boxtimes^{\leq n}\delta(\mathfrak{gl}_1))$. Notice that (I) – (IV) together yield $T_{n+1} \subseteq \overline{\alpha}_n(\delta(\mathfrak{gl}_1))$ whence $\{s\} \subseteq \boxtimes^{\leq n} T_{n+1} \subseteq \overline{\alpha}_n(\boxtimes^{\leq n}\delta(\mathfrak{gl}_1))$. And since by (III) and the fact that $s \in T_{n+1}$ we have $s \in \alpha_n(p_a)$, everything is proved. □

Theorem 7.9.14. *The following holds.*

1. Θ *splits* $\mathcal{E}$ $\boldsymbol{K}.t$ *iff* $\Theta = \mathsf{Th}\boxed{\bullet}$.
2. Θ *splits* $\mathcal{E}$ $\boldsymbol{K4}.t$ *iff* $\Theta = \mathsf{Th}\boxed{\bullet}$.

Proof. $\boxed{\bullet}$ is cycle–free and therefore splits $\mathcal{E}$ $\boldsymbol{K}.t$. A fortiori it splits $\mathcal{E}$ $\boldsymbol{K4}.t$. □

Notes on this section. The results of this section have mainly been established in [**123**]. The connections between monomodal logics and their minimal tense extensions have been investigated thoroughly in a series of papers by Frank Wolter. He has shown that many properties are lost in passing from a monomodal logic to its minimal tense extension; these are among other the finite model property and completeness. (See [**238**], [**239**].) Wolter has also shown that the minimal tense extension of a modal logic need not be conservative. For completeness results in tense logic see also [**242**] and [**236**].

Exercise 265. Complete the proof of Lemma 7.9.3.

Exercise 266. Show Proposition 7.9.2.

Exercise 267. Show that **S4**.**2**.t, where **S4**.**2** = **S4** $\oplus\ \Diamond\Box p \to \Box\Diamond p$, is 2–transitive. Hence show that every finite connected frame splits $\mathcal{E}$**S4**.**3**.t and **S5**.t.

Exercise 268. Show that $\mathcal{E}$ **S4**.t has exactly two elements of codimension 2.

Exercise 269. Let $\mathfrak{Gl}_\omega$ the frame consisting of $\mathfrak{gl}_\omega$ and the finite and cofinite sets as internal sets. Describe the bidual $(\mathfrak{Gl}_\omega)_+{}^+$. Show that its algebra of sets is subdirectly irreducible but that the frame is not connected.

Exercise 270. Show without the use of ultrafilters that $\bigcap_{n\in M} \mathfrak{gl}_n = \mathbf{Ga}$ whenever $M \subseteq \omega$ is infinite.

Part 3

Case Studies

CHAPTER 8

Extensions of K4

8.1. The Global Structure of $\mathcal{E}$K4

The first general results in modal logic were obtained for transitive logics. Moreover, transitive logics have for a long time been in the focus of interest of modal logicians since the motivating applications very often yielded transitive logics. Moreover, as we shall see, the structure of **K4**–frames is also by far easier than the the structure of frames for nontransitive frames. The first general result about a large classes of logics rather than individual logics was undeniably Robert Bull's [**34**] and the sequel [**60**] by Kit Fine. These papers gave an exhaustive theory of logics containing **S4.3**. However, **S4.3** is a very strong logic, and the result therefore did not cover many important transitive logics. The starting point of a general theory of transitive logics was the paper [**63**] by Kit Fine in which it was shown that the addition of an axiom of finite width makes transitive logics complete. The paper also contained useful results concerning weak canonical frames. The paper [**66**] was another breakthrough. It introduced the notion of a subframe logic. It was shown that all transitive subframe logics have the finite model property. Moreover, characterizations were given of canonical subframe logics. Many important logics turn out to be subframe logics. These results were discovered around the same time by Michael Zakharyaschev, who introduced the still more general notion of a *cofinal subframe logic*. All transitive cofinal subframe logics were shown to have the finite model property. Moreover, each transitive logic is axiomatizable by a set of so–called *canonical formulae*. These are formulae which characterize the frames of the logic by some geometric condition. This allows to deal with transitive logics by means of geometric conditions on frames rather than axioms. These results have established a novel way of thinking about transitive logics, and have also influenced the general theory of modal logic through the concepts of subframe and cofinal subframe logic.

Before we plunge into the theory of **K4**–frames, we will introduce some notation and recall some easy facts about **K4**–frames and p–morphisms between them. In transitive frames we say that y is a **strong** or **strict successor** of x if $x \lhd y \ntriangleleft x$ and that y is a **weak successor** of x if $x \lhd y$ or $x = y$. We write $x \vec{\lhd} y$ if y is a strong successor of x and $x \unlhd y$ if y is a weak successor of x. A cluster is a maximal set C of points such that given $x, y \in C$ then x is a weak successor of y. Thus, for elements x, y of a cluster C, if $x \neq y$ then $x \lhd y$. Consequently, if $\sharp C > 1$ then $\lhd \cap C^2 = C^2$.

On the other hand, if $\sharp C = 1$, say $C = \{x\}$, then either $x \lhd x$ (and so $\lhd \cap C^2 = C^2$) or $x \ntriangleleft x$ (and so $\lhd \cap C^2 = \varnothing$). C is called **degenerate** if it contains a single point and $x \ntriangleleft x$, and **improper** if $C = \{x\}$ with $x \lhd x$. If $\sharp C > 1$, C is called **proper**. The **type** of a cluster C is the cardinality of the set $\{y : x \lhd y\}$ for some $x \in C$. This is clearly independent of the choice of x. If the type is at least 2 then the cluster is proper, if the type is 1 the cluster is improper, and the cluster is degenerate if the type is 0. Instead of saying that a cluster has type 0 we also say that it has type $\varnothing$. The cluster containing a given point x is denoted by $C(x)$. If C and D are clusters, and for some $x \in C$ and $y \in D$ we have $x \lhd y$, then in fact for all $x \in C$ and all $y \in D$ holds that $x \lhd y$. This justifies the notation $C \lhd D$ for clusters. Finally, if C and D are distinct clusters and $C \lhd D$, then all points of D are strong successors of all points in C. Consequently, we may represent a transitive frame as a triple $\langle p, \blacktriangleleft, \nu\rangle$ where $\langle p, \blacktriangleleft\rangle$ is a partial order and ν a function from p to ω. ($\langle p, \blacktriangleleft\rangle$ is the ordering of clusters, and $p(x)$ is the type of the cluster represented by x.) Let $\mathfrak{f} = \langle f, \lhd\rangle$ be a Kripke–frame. Recall that the depth of a point x, in symbols $dp(x)$, is an ordinal number defined by

$$dp(x) := \{dp(y) : x \vec{\lhd} y\}$$

This requires that $dp(y)$ is defined for all strong successors y of x. In particular, x is of depth 0 if for every successor y of x, $y \lhd x$. In other words, x has no strong successors. We then say that the cluster of x is **final**. Notice that the definition assigns depth only to points for which there is no infinite ascending chain of strong successors, or to rephrase this, if the partial order of clusters from x onwards has no ascending chains. The definition as given also extends to infinite ordinals; for example, a point is of depth ω if all strong successors are of finite depth and, moreover, for each $n \in \omega$ there is at least a point of depth n. It is left as an exercise to show that if x is of depth α and $\alpha > \beta$ then there is a strong successor of depth β. We call the points of a given depth α the **α–slice** of the frame.

Definition 8.1.1. *Let $\mathfrak{f}$ be a Kripke–frame and α an ordinal number. Then $\mathfrak{f}^{<\alpha}$ denotes the subframe of points of depth $< \alpha$. Likewise, $\mathfrak{f}^{>\alpha}$ and $\mathfrak{f}^{\alpha}$ denote the subframes of points of depth $> \alpha$ and $= \alpha$, respectively. $\mathfrak{F}^{<\alpha}$ is the subframe induced by $\mathfrak{F}$ on the Kripke–frame $\mathfrak{f}^{<\alpha}$. Likewise for $\mathfrak{F}^{>\alpha}$, if that is well–defined.*

Definition 8.1.2. *Let $\mathfrak{F}$ be a frame. The **depth of** $\mathfrak{F}$ is the set of all ordinals α such that there exist points of depth α. A frame is said to satisfy the **ascending chain condition** (**acc**) iff there exists no infinite, strictly ascending chain of points. Such frames are also called **noetherian**.*

Notice that a frame $\mathfrak{F}$ is noetherian iff $\mathfrak{F} = \mathfrak{F}^{<\alpha}$ for some α. (In fact, any ordinal greater or equal to the depth of $\mathfrak{F}$ suffices if $\mathfrak{F}$ is noetherian.)

We begin by recalling the frame constructor $\oplus$, standing for the disjoint union. It takes two frames and places them side by side. We say that it places the frames **parallel** to each other. Now we introduce a second constructor, $\circledcirc$. Rather than placing frames parallel to each other it places the first before the second. To be precise, let $\mathfrak{f} = \langle f, \lhd_f \rangle$ and $\mathfrak{g} = \langle g, \lhd_g \rangle$ with f and g disjoint. Then

$$\mathfrak{f} \oslash \mathfrak{g} := \langle f \cup g, \lhd_f \cup f \times g \cup \lhd_g \rangle$$

Thus, the underlying set is the disjoint union of the sets of the individual frames. And $x \lhd y$ iff either x and y are in f and $x \lhd_f y$ or x in f and y in g (and no other condition) or x and y are both in g and $x \lhd_g y$. Now if $I = \langle I, < \rangle$ is an ordered set and $\mathfrak{f}_i = \langle f, \lhd_i \rangle$, $i \in I$, an indexed family of frames on disjoint sets then

$$\oslash_{i \in I} \mathfrak{f}_i := \langle \bigcup_I f_i, \bigcup_I \lhd_i \cup \bigcup_{i<j} f_i \times f_j \rangle$$

Proposition 8.1.3 (Decomposition). *Let $\mathfrak{f} = \mathfrak{g} \oslash \mathfrak{h}$. (1.) A generated subframe of $\mathfrak{f}$ is of the form $\mathfrak{d} \oslash \mathfrak{h}$ for a $\mathfrak{d} \rightarrowtail \mathfrak{g}$ or of the form $\mathfrak{d}$ for some $\mathfrak{d} \rightarrowtail \mathfrak{h}$. (2.) If $p : \mathfrak{g} \twoheadrightarrow \widehat{\mathfrak{g}}$ and $q : \mathfrak{h} \twoheadrightarrow \widehat{\mathfrak{h}}$ are contractions, so is $p \oslash q : \mathfrak{g} \oslash \mathfrak{h} \twoheadrightarrow \widehat{\mathfrak{g}} \oslash \widehat{\mathfrak{h}}$.*

Call a p–morphism **lateral** if it is depth preserving. A contraction map $\pi : \mathfrak{g} \twoheadrightarrow \mathfrak{h}$, $\mathfrak{g}$, $\mathfrak{h}$ finite, is **minimal** if in any decomposition $\pi = \rho \circ \rho'$ either ρ or ρ' is an isomorphism. Any contraction of a finite frame can be decomposed into a sequence of minimal contractions and there is only a small number of such contractions. First of all, let us take a point v of minimal depth which is identified via π with another point and let w be of minimal depth such that $\pi(v) = \pi(w)$. Now two cases arise. Either $dp(w) = dp(v)$ or $dp(w) > dp(v)$. In the latter case it is easily seen that there exists a pair w and v' such that $v' \in C(v)$ and $dp(w) = dp(v') + 1$. We may assume that $v' = v$. A non–lateral p–morphism collapses at least a point with an immediate successor. In that case, the successor cannot be irreflexive. Furthermore, if $w \lhd v$ then $C(w)$ is mapped onto $C(v)$. So we have the situation $\pi : \mathbf{m} \oslash \mathbf{n} \twoheadrightarrow \mathbf{k}$. Then the restriction of π to the cluster $\mathbf{n}$ is surjective. Hence $k \leq n$. Moreover, $k < n$ implies that we can decompose π in the following way:

$$\mathbf{m} \oslash \mathbf{n} \twoheadrightarrow \mathbf{m} \oslash \mathbf{k} \twoheadrightarrow \mathbf{n} \oslash \mathbf{k} .$$

This is left for the reader to verify. So, by minimality and the fact that we have a non–lateral contraction, $k = n$. Likewise, $m > n$ cannot occur, because then we can decompose π into

$$\mathbf{m} \oslash \mathbf{n} \twoheadrightarrow \mathbf{n} \oslash \mathbf{n} \twoheadrightarrow \mathbf{n} .$$

So $m \leq n$. $\pi : \mathbf{m} \oslash \mathbf{n} \twoheadrightarrow \mathbf{n}$ is actually indecomposable iff it is injective on both clusters. If $m \leq n$ such a contraction can be constructed.

Now we come to the case where two points of identical depth are identified. Let d be the minimal depth of points on which p is not injective. It is clear that if π is a contraction, and we let $\widehat{\pi}$ be defined by $\widehat{\pi}(w) := w$ if w is of depth $\neq d$ and $\widehat{\pi}(w) = \pi(w)$ else, $\widehat{\pi}$ is a contraction as well, and π factors through $\widehat{\pi}$. Thus by minimality π is lateral. Let v and w be two distinct points such that $\pi(v) = \pi(w)$. Then two cases arise. **1. Case.** $C(v) = C(w)$. Then π is of type $\mathbf{m} \twoheadrightarrow \mathbf{n}$ for $m > n$. **2. Case.** $C(v) \neq C(w)$. Then π is of the type $\pi : \mathbf{m} \oplus \mathbf{n} \twoheadrightarrow \mathbf{k}$. Then either all three are degenerate or all three are nondegenerate. Again we find that π must be injective on both clusters, hence $m = n$, and that $k < n$ contradicts minimality. So, $\pi : \mathbf{n} \oplus \mathbf{n} \twoheadrightarrow \mathbf{n}$, and π is injective on the individual clusters.

Theorem 8.1.4. *π is a minimal contraction iff it is of the following types.* ***Lateral.*** *(a.) $\boldsymbol{n{+}1} \twoheadrightarrow \boldsymbol{n}$, ≥ 1, (b.) $\varnothing \oplus \varnothing \twoheadrightarrow \varnothing$ or (c.) $\boldsymbol{n} \oplus \boldsymbol{n} \twoheadrightarrow \boldsymbol{n}$.* ***Nonlateral.*** *$\varnothing \otimes \boldsymbol{n} \twoheadrightarrow \boldsymbol{n}$ or $\boldsymbol{m} \otimes \boldsymbol{n} \twoheadrightarrow \boldsymbol{n}$ with $m \leq n$. In all cases, π is injective on the two individual subclusters.*

Recall the notion of a local subframe. We know that if $\mathfrak{G}$ is a local subframe of $\mathfrak{F}$ then any contraction on $\mathfrak{G}$ can be extended to a contraction on $\mathfrak{F}$ which is the identity on all points not in $\mathfrak{F}$. For example, clusters are always local in a Kripke–frame. Notice that this is so no matter how they are embedded into the frame. So, clusters can always be contracted. Moreover, there are additional situations which are useful to know. For example, in a finite frame, if x has a single immediate successor, y, and y is reflexive, then the set $\{x, y\}$ is local. For if $x \lhd z$ then either $x = z$ or $y \unlhd z$ by the fact that y is the only immediate successor, and so $y \lhd z$, since y is reflexive. And $y \lhd z$ implies $x \lhd z$. Thus the set is local, and there is a contraction reducing it to a single point in the total frame. We call x in this situation a **unifier** (for y). In general, x is a unifier for a set X if X is the set of all immediate successors of x.

Say f (or $\mathfrak{f}$) is **totally local** if it is local and for all $x \in g - f$ and all $y, z \in f$ we have $x \lhd y$ iff $x \lhd z$. In case $\mathfrak{f}$ is totally local, we write sometimes $\mathfrak{g}[\mathfrak{f}]$ to denote the given, totally local occurrence of $\mathfrak{f}$. A replacement of $\mathfrak{f}$ by another frame $\mathfrak{h}$ is then unambigously defined. We define $\mathfrak{g}[\mathfrak{h}/\mathfrak{f}]$ to be a frame with underlying set $g - f + h$ and relation $x \widehat{\lhd} z$ iff (1.) neither x and z are in $\mathfrak{h}$ and $x \lhd z$ or (2.) $x \in g$ and $z \in h$ or $x \in h$ and $z \in g$ or (3.) both x, z in h and then $x \lhd z$ (in $\mathfrak{h}$).

From a theoretical viewpoint it has shown useful to introduce four parameters to classify transitive logics.

Depth. Let $\mathfrak{F}$ be a refined transitive frame. The **depth** of $\mathfrak{F}$ is the set of all α such that there exists a point of depth α in $\mathfrak{F}$. We write $dp(\mathfrak{F})$ to denote the depth of $\mathfrak{F}$. The **depth** of $\Lambda \supseteq \mathbf{K4}$ is the supremum of all $dp(\mathfrak{F})$, where $\mathfrak{F}$ is a refined Λ–frame – if such a supremum exists. Otherwise Λ is said to be of unbounded depth. The logic of frames of depth $\leq n$ can be axiomatized for finite n. Namely, put

$$\begin{array}{rcll} dp(>n) & := & p_0 & \wedge \Box^+ \bigwedge_{i<j<n+1} p_i \to \neg p_j \\ & & & \wedge \Box^+ \bigwedge_{i<j<n+1} p_i \to \Diamond p_j \\ & & & \wedge \Box^+ \bigwedge_{i<j<n+1} p_j \to \neg \Diamond p_i \end{array}$$

(Here the convention $\Box^+\varphi := \varphi \wedge \Box\varphi$ is used.) $dp(> n)$ is satisfiable in $\mathfrak{F}$ if there is a chain $\langle x_i : i < n \rangle$ of length n such that $x_i \ntriangleleft x_{i+1}$ for every $i < n - 1$; we call such a chain **properly ascending**. Clearly, if no such chain exists, $\mathfrak{F}$ is of depth $\leq n$, and conversely. Let $dp(\leq n) := \neg dp(> n)$. This axiom is also denoted by J_n in the literature. Then $\mathbf{K4}.dp(\leq n)$ is the logic of frames of depth $\leq n$.

Width. A set $\{x_0, \ldots, x_k\}$ of points is an **antichain** of length $k + 1$ if for no two distinct points x_i, x_j we have $x_i \lhd x_j$. A refined frame $\mathfrak{F}$ has width n if there is a point

x and an antichain Y of length n such that x sees every member of Y. We denote by $wd(\mathfrak{F})$ the width of $\mathfrak{F}$. The width of Λ is the supremum of all widths of refined Λ–frames (if such a supremum exists) and denoted by $wd(\Lambda)$. If a logic is of width n for some $n < \omega$ it is of **finite width**. A logic is of **width n** if it satisfies the axiom

$$wd(\leq n) := \bigwedge_{i=0}^{n} \Diamond p_i. \rightarrow . \bigvee_{i \neq j} \Diamond(p_i \wedge p_j) \vee \bigvee_{i \neq j} \Diamond(p_i \wedge \Diamond p_j)$$

This axiom is also denoted by I_n in the literature.

Tightness. Let $\mathfrak{F}$ be a refined transitive frame. Let x be a point and $x \lhd y$. Moreover, let $\mathfrak{G}$ be a subframe such that no point of $\mathfrak{G}$ is comparable with x. Then if $\mathfrak{G}$ is of depth α, $\mathfrak{F}$ is said to be of **tightness** α. We write $ti(\mathfrak{F}) = \alpha$ in that case. The **tightness** of $\Lambda \supseteq$ **K4** is the supremum of all $ti(\mathfrak{F})$, where $\mathfrak{F}$ is a refined Λ–frame, if such a supremum exists. Consider the following formula.

$$\begin{array}{rcll} ti(> n) & := & p_0 & \wedge\Box^+ \bigwedge_{i<j<n+2} p_i \rightarrow \neg p_j \\ & & & \wedge\Box^+ p_0 \rightarrow \Diamond p_{n+1} \\ & & & \wedge\Box^+ \bigwedge_{i<j<n+1} p_i \rightarrow \Diamond p_j \\ & & & \wedge\Box^+ \bigwedge_{j<i<n+2} p_i \rightarrow \neg\Diamond p_j \\ & & & \wedge\Box^+ \bigwedge_{i<n+1} p_{n+1} \rightarrow \neg\Diamond p_i \\ & & & \wedge\Box^+ \bigwedge_{0<i<n+1} p_i \rightarrow \neg\Diamond p_{n+1} \end{array}$$

This formula is satisfiable in a refined frame iff there is a point seeing a properly ascending chain of length $k + 1$ and an incomparable point. Let $ti(\leq k) := \neg ti(> k)$. Then $\mathfrak{F} \vDash ti(\leq k)$ iff $\mathfrak{F}$ does not have tightness $> k$. So, **K4**.$ti(\leq k)$ is the logic of frames of tightness $\leq k$.

Fatness. For a cardinal number α, the cluster $\underline{\alpha}$ is the cluster containing α many points. We also write **n** for $\underline{n}$, $n \in \omega$. We call the type of a cluster C also its **fatness** and denote it by $ft(C)$. The fatness of a refined frame, $ft(\mathfrak{F})$ is the supremum of all $ft(x)$, where x is a point of $\mathfrak{F}$. The fatness of a logic Λ is the supremum of the set of all $ft(\mathfrak{F})$, $\mathfrak{F}$ a refined frame for Λ. Equivalently, a frame fails to be of fatness n if the following formula is satisfiable

$$\begin{array}{rl} ft(> n) := p_0 & \wedge\Box^+ \bigwedge_{i<j<n+1} p_i \rightarrow \neg p_j \\ & \wedge\Box^+ \bigwedge_{i<j<n+1} p_i \rightarrow \Diamond p_j \end{array}$$

The logic of frames of fatness $\varnothing$ is the logic **G** $= \Diamond p \rightarrow \Diamond(p \wedge \Box\neg p)$.

The significance of the choice of these logics will become clearer later on. Let us say here only that the logics of finite depth **K4**.$dp(\leq k)$ form an infinite descending chain in the lattice $\mathcal{E}$(**K4**); likewise the logics **K4**.I_k (alias **K4**.$wd(\leq k)$) and **K4**.$ti(\leq k)$ as well as **K4**.$ft(\leq k)$. Moreover, the axiom of *linearity*, known as **.3** is the same as I_1 and the same as $wd(\leq 1)$. **Grz** is the logic of reflexive frames of fatness 1, **G** the logic of frames of fatness $\varnothing$. The last two statements need a rigorous proof, but with the methods developed in this chapter this is actually very easy.

Exercise 271. Prove Proposition 8.1.3.

Exercise 272. Show that $\mathbf{m} \otimes \mathbf{n} \twoheadrightarrow \mathbf{1} \otimes \mathbf{n} \twoheadrightarrow \mathbf{n}$.

Exercise 273. Let $\mathfrak{f}$ be a finite reflexive frame. Show that there is a p–morphism $\mathfrak{f} \twoheadrightarrow \mathfrak{g}$ such that $\sharp g = \sharp f - 1$. So there always exists a truly minimal p–morphism.

Exercise 274. Let $\mathfrak{f} = \langle f, \lhd \rangle$ be a frame of depth ≥ 2. Define $x \shortparallel y$ by $x \not\lhd y$ and $y \not\lhd x$. We call $\langle f, \shortparallel \rangle$ the **antiframe** of $\mathfrak{f}$. Show that if $\mathfrak{f}$ is $\otimes$–decomposable the antiframe $\langle f, \shortparallel \rangle$ is disconnected. Show that the converse need not hold.

8.2. The Structure of Finitely Generated K4–Frames

Fundamental for the study of the lattice $\mathcal{E}\,\mathbf{K4}$ is the fact that the structure of the finitely generated algebras is rather well–behaved. Looking at the underlying Kripke–frame for a refined frame, we can say that all these algebras contain an upper part, consisting of points of finite depth, and that each point of infinite depth is 'behind' this upper part. Moreover, this upper part is atomic, and each level is finite. It should be said that the infinitely generated algebras are not as nice as that, so the restriction to finitely generated algebras is indeed necessary. But we know that logics are complete with respect to such frames, so we do not loose anything by specializing to such frames. Before we begin let us note a simple fact.

Theorem 8.2.1. *Let $\mathfrak{F}$ be a transitive refined frame and $\mathbb{F}$ be n–generated. Then the clusters have type $\varnothing$ or $\boldsymbol{t}$ for $t \leq 2^n$.*

Proof. Let $A = \{a_i : i < n\}$ be a generating set and let x and y be two points in a cluster. By induction on the sets generated from A it is shown that $x \in a_i \Leftrightarrow y \in a_i$ for all $i < n$ implies that $x \in b \Leftrightarrow y \in b$ for all $b \in \mathbb{F}$. Hence a cluster can have at most 2^n distinct points. ◻

This idea of *induction on the sets generated from A* can be made precise in the following way. Let us take a general frame $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$ for **K4** which is refined, such that the algebra of subsets is n–generated. Let $\{a_0, \dots, a_{n-1}\}$ be a generating set. Let $\mathbb{P}_n := \{p_i : i < n\}$. Put $v : \mathbb{P}_n \to \mathbb{F} : p_i \mapsto a_i$, $i < n$. Then for any $b \in \mathbb{F}$ there is a φ with $var(\varphi) \subseteq \mathbb{P}_n$ and $b = \overline{v}(\varphi)$. Therefore we can prove a property of all internal sets by induction on the constitution of formulae φ such that $var(\varphi) \subseteq \mathbb{P}_n$. If $\overline{v}(\varphi) = b$ and $dp(\varphi) = k$, we say that b is **k–definable**. (Notice that 0–definable was also defined to mean that the set is definable by a constant formula. Throughout this chapter we will not use the term 0–definable in that latter sense.) Notice finally, that if $\mathfrak{F}$ is n–generated then any generated subframe is n–generated by the same set of generators. Since $\mathfrak{F}$ is refined, $\mathfrak{f}$ is transitive. We will now show that $\mathfrak{f}$ contains a generated subframe $\mathfrak{f}^{<\omega}$ such that each point of $\mathfrak{f}^{<\omega}$ is of finite depth and, moreover, any point not in $\mathfrak{f}^{<\omega}$ sees a point of arbitrary finite depth. Such a structure is called *top–heavy*.

Definition 8.2.2. *A frame $\mathfrak{F}$ is called **top–heavy** if every point not in $\mathfrak{F}^{<\omega}$ has a successor of depth n for any $n < \omega$.*

$\mathfrak{f}^{<\omega}$ will be defined inductively. The generated subframe of $\mathfrak{F}$ induced by $\mathfrak{f}^{<\omega}$ will be denoted by $\mathfrak{F}^{<\omega}$.

Lemma 8.2.3. *Let α be an ordinal and $\mathfrak{f}$ be a Kripke–frame. Then the subframe $\mathfrak{f}^{<\alpha}$ consisting of all points of depth $< \alpha$ is a generated subframe of $\mathfrak{f}$.*

Proof. Let x be of depth β. $dp(x)$ is well–defined only if all strong successors have a depth. By definition, if $x \lhd y$ then either y is not a strong successor, in which case y belongs to the cluster of x and is of equal depth, since x and y have the same successors and the same strong successors. Or y is a strong successor and has a depth according to the definition of the depth. ◻

Lemma 8.2.4. *Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$ be n–generated and refined. Then the following holds.*

1. *There are at most $2^n(2^{2^n-1} + 1)$ points of depth 0.*
2. *$\mathfrak{F}^{<1}$ is a full finite frame.*
3. *Every set $a \subseteq \mathfrak{f}^{<1}$ is 2–definable.*
4. *Every point has a weak successor of depth 0.*

Proof. Because $\mathfrak{F}^{<1}$ is a generated subframe of $\mathfrak{F}$, it is refined and generated by the a_i. So if (1.) holds, (3.) holds as well, since finite refined frames are full. Let us start then by attacking (2.). For $S \subseteq \{0, 1, \dots, n-1\}$ define

$$\chi(S) := \bigwedge_{i \in S} p_i \wedge \bigvee_{i \notin S} \neg p_i$$

Put $A_S := \overline{v}(\chi(S))$. Every node is in exactly one of the sets A_S. If $x \in A_S$, $\chi(S)$ is called the **atom** of x. Let $\mathcal{C}$ be system of subsets of $\{0, 1, \dots, n-1\}$. Define

$$\begin{array}{lcl} \mathcal{C}\text{–}span & := & \bigwedge\langle \Diamond\chi(S) : S \in \mathcal{C}\rangle \wedge \bigwedge\langle \neg\Diamond\chi(S) : S \notin \mathcal{C}\rangle \\ mspan & := & \left\{ \begin{array}{ll} & \Box\bot \\ \vee & \bigvee\langle \chi(S) \wedge \mathcal{C}\text{–}span \wedge \bigwedge_{\mathcal{D} \subsetneq \mathcal{C}} \neg\Diamond\mathcal{D}\text{–}span : S \in \mathcal{C}\rangle \end{array} \right. \end{array}$$

We say that x is of **span** $\mathcal{C}$ if $x \in \overline{v}(\mathcal{C}\text{–}span)$. We say that x is of **miminal span** if it is in $\overline{v}(mspan)$. So, x is of minimal span iff it is without successor or it has successors, among which one has the same atom as x and all successors of x have the same span as x. The set of points of span $\mathcal{C}$ is 1–definable, the set of minimal span is 2–definable.

(†) Let x and z be points of identical span and identical atom. Assume that x and z are of minimal span. Then $x = z$.

Let the span of x by $\mathcal{C}$. We show by induction on φ that $x \in \overline{v}(\varphi)$ iff $z \in \overline{v}(\varphi)$. It follows that $x \in b$ iff $z \in b$ for all $b \in \mathbb{F}$ and by refinedness $x = z$. Now for the inductive proof. Since x and z have the same atom, $x \in \overline{v}(p_i)$ iff $z \in \overline{v}(p_i)$ for all

$i < n$. The induction steps for $\neg$ and $\wedge$ are straightforward. Now let $\varphi = \Diamond\eta$. Assume that $x \in \overline{v}(\varphi)$. Then there exists a y such that $x \lhd y$ and $y \in \overline{v}(\eta)$. Now y has minimal span (as it succeeds a point of minimal span). Hence its span is equal to the span of x. Let $\chi(S)$ be the atom of y. Then $\Diamond\chi(S)$ is a conjunct of $\mathcal{C}$–*span*. But z has span $\mathcal{C}$, and so there is a point u such that $z \lhd u$ and u has atom $\chi(S)$. Also u has span $\mathcal{C}$. By induction hypothesis, $u \in \overline{v}(\eta)$. Hence $z \in \overline{v}(\Diamond\eta) = \overline{v}(\varphi)$.

It follows now that a point x is of depth 0 iff it is of minimal span. From left to right is clear. Now assume that x is of minimal span. If it has no successors, the case is clear. Now let x have successors. Let y be a successor of x. By assumption, the span of y equals the span of x, so there is a successor z of y satisfying the same atom as x, by definition on the formula *mspan*. By (†), $z = x$, since z and x have the same span. It follows that $y \lhd x$. So, x is of depth 0.

Now we come to (1.). What we need to specify for a point of depth 0 is its span, which is then minimal, and hence defines the size of the cluster as well as the atoms which are represented in it, unless the point is successorless, in which case its own atom suffices for a characterization. There are at most 2^n degenerate clusters. If the cluster is nondegenerate, it has size between 1 and 2^n, for each size k there is a choice of $\binom{2^n}{k}$ atoms. Summing this we get

$$\sharp\mathfrak{f}^1 \leq 2^n + \sum_{k=1}^{2^n} k\binom{2^n}{k} = 2^n + 2^n \sum_{k=0}^{2^n-1} \binom{2^n-1}{k} = 2^n(2^{2^n-1}+1)$$

Moreover, each point in that frame is 2–definable. Hence each subset is. This shows (3.).

Next we show (4.). Put $Z_0(x) := \{S : x \in \blacklozenge A_S\}$. $Z_0(x)$ is finite. Hence there is a weak successor y of x such that for all $z \rhd y$, $Z_0(z) = Z_0(y)$. Two cases arise. **Case 1.** y is reflexive. Then y is of minimal span and so of depth 0. **Case 2.** y is irreflexive. Then if y has no successors it is of minimal span and so of depth 0. Finally, if y has successors, then let $z \rhd y$. Since $Z_0(z) = Z_0(y)$ and since the atom of z is in $Z_0(y)$, the atom of z is in $Z_0(z)$. It follows that z is of minimal span. ◻

Let the function $\delta(n, k)$ be defined inductively as follows

$$\begin{array}{lcl} \delta(n,1) & := & 2^n(2^{2^n-1}+1) \\ \delta(n,k+1) & := & \delta(n,1)(2^{\delta(n,k)}-1) \end{array}$$

Theorem 8.2.5. *Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ be an n–generated refined* ***K4****–frame. Then the following holds*

1. *There are at most $\delta(n, k+1)$ points of depth $\leq k$.*
2. *$\mathfrak{F}^{<k+1}$ is a full finite frame.*
3. *Every subset of $\mathfrak{F}^{<k+1}$ is $4k+2$–definable.*
4. *Every point in $\mathfrak{F}$ not of depth $< k$ has a weak successor of depth k.*

Proof. The case $k = 0$ is covered by the previous lemma. Now we do the induction step for k. Again, (2.) follows from (1.) and (1.) follows from (3.). For each point x of depth $< k$ there exists a formula $\gamma(x)$ in the variables $\mathbb{P}_n$ such that $x \in \overline{v}(\gamma(x))$. Let us consider the set P of formulae $\gamma(x)$ where $x \in f^{<k}$. We know by induction hypothesis that $\sharp P \leq \delta(n,k)$ and that for each x, $\gamma(x)$ if of degree $\leq 4k-2$. Moreover, the disjunction of all $\gamma(x)$ is a formula of degree $4k-2$ defining the points of depth $< k$. Denote it by $dp(<k)$. Let $Q \subseteq P$. Say that a point in the frame is of **width** Q if $x \in \overline{v}(Q\text{–}width)$ where

$$Q\text{–}width := \neg dp(<k) \wedge \bigwedge \langle \Diamond C : C \in Q \rangle \wedge \bigwedge \langle \neg\Diamond C : C \in P - Q \rangle$$

Clearly, a point in the frame is of width Q iff it is not of depth $< k$ and sees exactly the points defined by Q. Notice that the width can never be empty, in fact must contain at least $k-1$ formulae. Furthermore, a point is of **minimal width** if it satisfies

$$mwidth := \bigvee_{Q \subseteq P} (Q\text{–}width \wedge \bigwedge_{R \subsetneq Q} \neg\Diamond R\text{–}width)$$

A point is of minimal width if it is of width Q for some Q and no successor can be of lesser width (unless it is of depth $< k$ in which case it has no width). It is easy to see that the points of minimal width together with the points of depth $< k$ form a generated subframe. However, this is not yet the desired subframe. For now we have to start the same procedure as above, selecting points of minimal span. As in the previous lemma, each point has an atom A_S, $S \subseteq \{0, 1, \dots, n-1\}$. A point is of **span** $\mathcal{C}$ *within being of minimal width*, $C \subseteq 2^n$, if it satisfies the formula

$$\mathcal{C}\text{–}span := \bigwedge \langle \Diamond(mwidth \wedge \chi(S)) : S \in \mathcal{C} \rangle \wedge \bigwedge \langle \neg\Diamond(mwidth \wedge \chi(S)) : S \notin \mathcal{C} \rangle$$

It is easy to check that a point is of $\mathcal{C}$–span within being of minimal width iff it has exactly the successors with atoms $\chi(S)$ of minimal width. Finally, the formula

$$mspan := \Box dp(<k) \vee \bigvee (\langle \chi(S) \wedge \mathcal{C}\text{–}span \wedge \bigwedge_{\mathcal{D} \subsetneq \mathcal{C}} \neg\Diamond\mathcal{D}\text{–}span : \chi(S) \in \mathcal{C} \subseteq 2^n)$$

defines the points of minimal span within being of minimal width. We claim it defines the slice of points of depth k. It is easy to see that if a point is of depth k it must be of minimal width and within that of minimal span. For the converse, however, we prove the following.

(‡) Let x, z be of minimal width and of minimal span within being of minimal width. Assume that x and z are of equal width, equal span and equal atom. Then $x = z$.

By induction on the constitution of the formula φ in the variables $\mathbb{P}_n$ we show that for points x and z satisfying the assumptions of (‡), $x \in \overline{v}(\varphi)$ iff $z \in \overline{v}(\varphi)$. This suffices for a proof. Now for variables this is so by the fact that x and z have the same atom. The steps for $\wedge$ and $\neg$ are straightforward. Now assume that $\varphi = \Diamond\eta$ and that $x \in \overline{v}(\varphi)$. Let x have span Q and width $\mathcal{C}$. Then there exists a y such that $x \lhd y$ and $y \in \overline{v}(\eta)$. Suppose that y is of depth $< k$. Then $\Diamond\gamma(y)$ is a conjunct of Q–*span*. So, z has a successor satisfying $\gamma(y)$. But only y satisfies $\gamma(y)$. Therefore $z \lhd y$. Now assume that y is not of depth $< k$. Then it has span Q and width $\mathcal{C}$. Let $\chi(S)$ be the atom of

y. Now x has a successor of span Q, width $\mathcal{C}$ and atom $\chi(S)$. Then z has a successor u with span Q, width $\mathcal{C}$ and atom $\chi(S)$. By induction hypothesis, $u \in \overline{v}(\eta)$. Hence $z \in \overline{v}(\Diamond\eta) = \overline{v}(\varphi)$.

Next we shall show that the points satisfying *mspan* are exactly the points of depth k. Clearly, if a point is of depth k it satisfies *mspan*. So, assume that x satisfies *mspan*. If every successor of x is of depth $< k$ we are done. So let this not be case. Then the atom of x is in the span of every successor of x which is not of depth $< k$. Moreover, no successor of x which is not of depth $< k$ has lesser width than x. So, let $x \lhd y$ and y not of depth $< k$. Then y has the same width and the same span as x and the atom of x is in the span of y, which means that there exists a successor u of y not of depth $< k$ which has the same atom as x. By (‡), $u = x$ and so $y \lhd x$. This shows that x is of depth k.

The formula defining the points is of degree $4k + 2$, since at most four modal operators are stacked on top of formulae defining points of depth $< k$. Counting the points, we find that if there are $\delta(n, k)$ points of depth $< k$, there can be at most $2^{\delta(n,k)} - 1$ nonempty subsets. For each subset C, there is a (possibly empty) set of points having minimal width C. Within that set, we have counted in effect the number of points of depth 0. This shows (1.). Finally (4.). If x is not of depth $< k$, then it has a width. If that width is not minimal, there is a successor v of minimal width, since the width is finite. As in the proof of the previous theorem we see that v has a weak successor w of minimal span within being of minimal width, by the same argument. A weak successor of w is a successor of x. If the width of x is minimal, then argue with $x = w$. This completes the proof. ◻

Let us remark that the bound for the number of points of depth 0 is exact. There is a frame (the frame underlying the free algebra) which has $\delta(n, 0)$ many points. However, for greater depth the bound is too large. This is so because not for every set S of points of depth $< k$ there exists a point of width exactly S. For if x sees a point $y \in S$ and $y \lhd z$ then $x \lhd z$ as well, so suitable S are only those sets which are closed under successors. However, to count their number is more intricate and not of immediate interest for us. As a first consequence we will show that the upper part of the lattice $\mathcal{E}\,\mathbf{K4}$ is rather well–behaved. Recall from Section 4.8 the Theorem 4.8.7 on locally finite varieties.

Theorem 8.2.6 (Segerberg). *Let $\Lambda \supseteq \boldsymbol{K4} \oplus dp(\leq n)$. The variety of Λ–algebras is locally finite. Consequently, Λ has the finite model property.*

Theorem 8.2.7 (Maksimova). *An extension of $\boldsymbol{K4}$ is of finite codimension iff it is tabular.*

Proof. The direction from right to left follows from Proposition 7.6.4. Now assume that Λ has finite condimension. Consider the sequence of logics $\mathsf{Th}(\mathfrak{Fr}_\Lambda(n))$, $n \in \omega$. Clearly, $\Lambda \subseteq \mathsf{Th}(\mathfrak{Fr}_\Lambda(n))$ for all n, and $\Lambda = \mathsf{Th}(\mathfrak{Fr}_\Lambda(n))$. Furthermore, $\mathsf{Th}(\mathfrak{Fr}_\Lambda(n+1)) \subseteq \mathsf{Th}(\mathfrak{Fr}_\Lambda(n))$. If equality holds then we have $\Lambda = \mathsf{Th}(\mathfrak{Fr}_\Lambda(n))$. These facts together show that since Λ has finite codimension there is an n_0 such that

$\Lambda = \mathsf{Th}(\mathfrak{Fr}_\Lambda(n_0))$. Now consider the extensions $\Lambda(d) := \mathsf{Th}(\mathfrak{Fr}^{<d}_\Lambda(n_0))$ for $d < \omega$, where $\mathfrak{Fr}^{<d}_\Lambda(n_0)$ is the subframe of points of depth $< d$. Again, $\Lambda = \bigcap_{d \in \omega} \Lambda(d)$ and $\Lambda(d+1) \subseteq \Lambda(d)$. If equality holds, $\Lambda = \Lambda(d)$. So there exists a d_0 such that $\Lambda = \Lambda(d_0) = \mathsf{Th}(\mathfrak{Fr}^{<d_0}_\Lambda(n_0))$. This frame is finite. Hence Λ is tabular. ◻

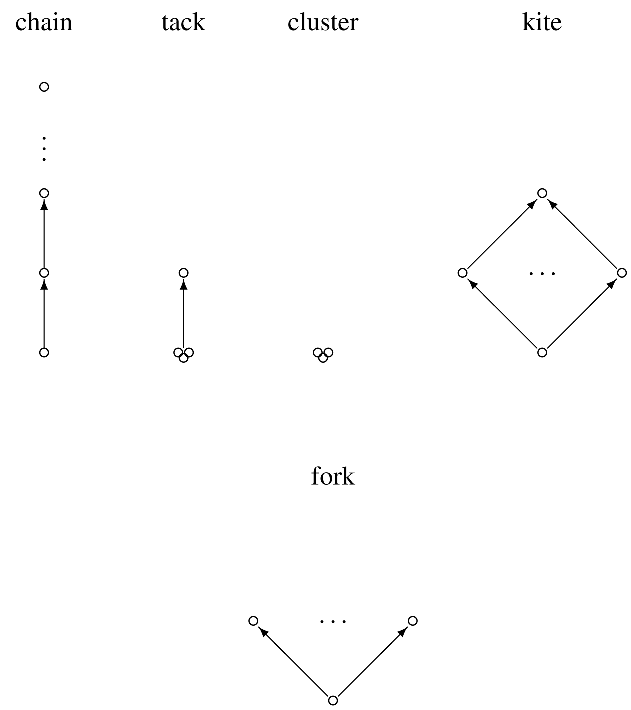
Thus, if one of the four parameters, the depth, is finite, we have more or less good control over the situation. A logic is called **pretabular** if it is not tabular, but all of its proper extensions are. By abstract arguments one can show that any logic which is not tabular must be contained in a pretabular logic. In the mid–seventies it was established by Leo Esakia and V. Meskhi in [**57**] that $\mathcal{E}$**S4** has exactly five such logics. (This fact has been proved independently also by Wolfgang Rautenberg and Larisa Maksimova.) The proof is a real classic in the theory of **K4**. We prove it by playing with the fundamental parameters of *depth*, *fatness* and *width*. Let Π be a pretabular logic in $\mathcal{E}$ **S4**. If its width, fatness and depth is finite, Π is tabular. So, one of the parameters is infinite. Suppose, then, that the depth of Π is not finite. Then as one can easily see, the n–element chains are models of that logic. This follows from the following useful fact.

Proposition 8.2.8. *Let $\mathfrak{f}$ be a noetherian Kripke–frame for $\boldsymbol{S4}$ of depth α. The map $x \mapsto dp(x)$ is a p–morphism onto $\langle\alpha, \geq\rangle$. Let $\mathfrak{f}$ be a noetherian Kripke–frame for $\boldsymbol{G}$ of depth α. The map $x \mapsto dp(x)$ is a p–morphism onto $\langle\alpha, >\rangle$.*

Proof. Let $\mathfrak{f}$ be a Kripke–frame for **S4**. Call the map which sends x to its depth π. Clearly, if $x \lhd y$ then $\pi(x) \geq \pi(y)$, by definition of depth. Next, assume that $\pi(x) \geq \beta$. Then x is of depth at least β. It follows by definition of depth that there exists a successor of x which has depth β. Likewise for **G**. ◻

Namely, let Π be of unbounded depth. By the fact that frames for Π are top–heavy, for each $d < \omega$ there exists a Π–frame $\mathfrak{f}$ of depth d. Then, by the previous theorem, $\langle d, \geq\rangle$ is also a Π–frame. The logic of all finite chains is **Grz.3**. It is also the logic of the infinite chain $\langle\omega, \geq\rangle$. It is clearly not tabular, but each proper extension is of finite depth; since the logic is of fatness 1 and of width 1, every extension is tabular. This concludes the case of infinite depth. Now we may assume that the depth of Π is finite. Let now the fatness be unbounded. We distinguish two cases. **Case 1.** The depth of prefinal clusters is unbounded. Notice that if $\mathbf{n}$ is a prefinal cluster, take the subframe $\mathfrak{g}$ generated by that cluster. Then $\mathfrak{g} \twoheadrightarrow \mathbf{n} \otimes \circ$. So, all these frames can be mapped onto a frame of the form $\mathbf{n} \twoheadrightarrow \circ$, called **tacks** in [**57**]. Let Π be the logic of the structure $f_{\aleph_0} = \underline{\aleph_0} \otimes \circ$. Π has depth 2 and width 2 but infinite fatness. It is the logic of $\underline{\aleph_0} \otimes \circ$. Then Π is not tabular. But every proper extension must be of finite fatness and hence tabular. **Case 2.** The depth of final clusters is unbounded. Then Π = **S5**, the logic of the lonely clusters, called **clots**. For in a refined frame with a final cluster of size n we find a generated subframe isomorphic to the cluster $\mathbf{n}$. So, $\Pi \subseteq$ **S5**. On the other hand, **S5** is not tabular, since it is not of finite fatness. Yet, every proper extension must be.

FIGURE 8.1. The five pretabular extensions of **S4**



Now we have reduced the investigation to logics of finite depth and finite fatness. Since Π is not tabular its width must be unbounded. Now, if the depth and fatness is bounded then the number of immediate incomparable successors of a point is also not bounded. (If it were, let Π be of depth d and fatness f, and let points have at most q immediate incomparable successors. Then there are at most $1 + q + q^2 + \ldots + q^{d-1}$ clusters; each cluster has at most f points. Hence Π is tabular.) So for any n there is a refined frame $\mathfrak{F}_n$ and a point x with n incomparable immediate successors. Take the subframe $\mathfrak{G}_n$ generated by x. Then Π is the logic of the $\mathfrak{G}_n$, as is not hard to see. Furthermore, we may assume that $\mathfrak{G}_n$ are of fatness 1. (Otherwise, take the set of all frames $\mathfrak{H}_n$ resulting from $\mathfrak{G}_n$ by collapsing each cluster to $\mathbf{1}$. This set is of unbounded width, and so not tabular. Since Π is pretabular, Π is the logic of the $\mathfrak{H}_n$.) Let $g_n = \{x\} \cup I_n \cup N_n \cup Z_n$, where I_n is the set of immediate successors of x of depth 0, N_n the set of immediate successors of x not of depth 1, and Z_n the set of the remaining points. **Case 1.** The set $\{\sharp I_n : n \in \omega\}$ is unbounded. Let $\mathfrak{k}_n := \langle k_n, \lhd_n \rangle$, where

$$\begin{array}{rcl} k_n & := & \{x\} \cup I_n \cup \{y\} \\ \lhd_n & := & \{\langle w, w\rangle : w \in k_n\} \cup \{\langle x, w\rangle : w \in k_n\} \,. \end{array}$$

Then $\mathfrak{b}_n \cong \mathbf{1} \otimes \bigoplus_{i \leq n} \mathbf{1}$ with $n = \sharp I_n + 1$; this frame is also called the n–**fan**. So $\mathfrak{k}_n$ is a $\sharp I_n + 1$–fan. Let $\rho_n : g_n \to k_n$ be defined by $\rho_n \restriction \{x\} \cup I_n = id$, and $\rho_n(w) = y$ for $w \in N_n \cup Z$. This is a p–morphism. Now Π is included in the logic of all n–fans. The logic of n–fans is not tabular, but a proper extension is, since it must be of finite width. This exhausts the first case. **Case 2.** $\{\sharp I_n : n \in \omega\}$ is bounded. Then $\{\sharp N_n : n \in \omega\}$ is unbounded. Put $\mathfrak{d}_n := \mathbf{1} \otimes (\bigoplus_{1 \leq n} \mathbf{1}) \otimes \mathbf{1}$, and call it the n–**top** or n–**kite**. There is a p–morphism from $\mathfrak{G}_n$ onto the $\sharp N_n$–kite. Namely, collapse I_n and z into a single point. Hence Π is contained in the logic of all kites. The latter is not tabular. Hence Π is equal to it.

Theorem 8.2.9 (Maksimova, Esakia and Meskhi, Rautenberg). *Let Λ be a logic containing $\mathbf{S4}$. Λ is pretabular iff Λ is one of the following five logics: the logic of the chains, the logic of the clusters, the logics of the tacks, the logics of the fans, the logic of the kites.*

Notes on this section. The structure of finitely generated **K4**–algebras and of finitely generated **S4**–algebras — also called *interior algebras* — has been studied by a number of people, either directly or indirectly by means of the canonical frame; the first to mention is Krister Segerberg [**193**]. For interior algebras, see Wim Blok [**22**] and Fabio Bellissima [**4**]. The presentation here follows Marcus Kracht [**121**], who builds on Kit Fine [**66**]. It can be shown that $\mathcal{E}\,\mathbf{G}$ has $\aleph_0$ many pretabular logics. $\mathcal{E}\mathbf{K4}$ has $2^{\aleph_0}$ many pretabular logics, as was proved by Wim Blok.

Exercise 275. Show that all pretabular logics of **S4** are finitely axiomatizable.

Exercise 276. Let $\mathfrak{f}$ be a finite, rooted, reflexive frame. Show that $\mathsf{Th}(\mathfrak{f})$ is of codimension $\geq \sharp f$. *Hint.* Show that for each cardinality $< \sharp f$ there is a rooted frame $\mathfrak{g}$ with $\mathfrak{f} \twoheadrightarrow \mathfrak{g}$.

Exercise 277. As before, but with $\mathfrak{f}$ not necessarily reflexive.

Exercise 278. Let $\Lambda \supseteq \mathbf{K4}$ be pretabular. Show that $\mathcal{E}\Lambda \cong 1 + \omega^\star$. *Hint.* Show first that the sublattice of points of finite codimension must be linear.

Exercise 279. A variety of **K4**–algebras is locally finite iff it is of finite depth. *Hint.* Show that $dp(\leq n)$ is expressible with one variable. Now look at $\mathfrak{Fr}_\Lambda(1)$. It must be finite.

8.3. The Selection Procedure

This section will introduce one of the most important techniques in modal logic, that of the *selection procedure*, developed by Kit Fine ([**66**]) and Michael Zakharyaschev ([**245**]). The idea is that if we have a general frame refuting a formula φ, we can extract a finite countermodel for φ. We cannot expect that the new model will be based on a frame for the logic under consideration. However, as has been noted, many logics are actually closed under this new operation, so that for them this will be a proof of the finite model property. We begin by examining a special case. Suppose that we have a global model $\mathfrak{M} = \langle \mathfrak{F}, \beta \rangle$ and a formula φ. Let $X := sf(\varphi)$ and $x \in f$. Then there is a unique set $Y \subseteq X$ such that $\mathfrak{M}_x \vDash \chi$ iff $\chi \in Y$. Now define

$$mo_{\mathfrak{M}}(x) := \bigwedge_{\chi \in Y} \chi \wedge \bigwedge_{\chi \in X - Y} \neg \chi$$

If no confusion arises, we write $mo(x)$ rather than $mo_{\mathfrak{M}}(\varphi)$. $mo_{\mathfrak{M}}(x)$ is called the φ–**molecule** of y. We say that y is μ–**maximal** if it has molecule μ but no strict successor has molecule Y. Given φ, we say that y is **maximal** (for φ) if it is Y–maximal for some Y. If $\mathfrak{f}$ is noetherian it has no infinite strictly ascending chain and so every point y which is not itself maximal has a successor which is maximal for the molecule for y. Thus, given y there is a weak successor which is maximal for the molecule of y; it is denoted by y^{μ}. Let $\mathfrak{f}^{\mu}$ be the subframe of maximal points.

Lemma 8.3.1. *Assume $\langle \mathfrak{f}, \beta, x \rangle \vDash \varphi$. Let $\mathfrak{g} \sqsubseteq \mathfrak{f}$ be a subframe such that every point in f has a weak successor in g with identical molecule, and let $x \in g$. Then $\langle \mathfrak{g}, \beta, x \rangle \vDash \varphi$.*

Proof. By assumption there is a function $y \mapsto y^p$ such that y and y^p have the same molecule in $\mathfrak{f}$ and y^p is a weak successor of y contained in g. We may assume that $x^p = x$. By induction on $\chi \in sf(\varphi)$ we show that

$$\begin{array}{lcl} \langle \mathfrak{f}, \beta, y \rangle \vDash \chi & \Leftrightarrow & \langle \mathfrak{g}, \beta, y^p \rangle \vDash \chi \\ \langle \mathfrak{g}, \beta, z \rangle \vDash \chi & \Leftrightarrow & \langle \mathfrak{g}, \beta, z^p \rangle \vDash \chi \end{array}$$

Let $\chi = p$, a variable. Then, since y^p has the same molecule as y, they satisfy the same variables of φ, irrespective of the subframe they are in. So the base case is treated. The steps for $\wedge$ and $\neg$ are straightforward. Now let $\chi = \Diamond\psi$. If $\langle \mathfrak{f}, \beta, y \rangle \vDash \Diamond\psi$ then $\langle \mathfrak{f}, \beta, y^p \rangle \vDash \Diamond\psi$ by definition of y^p. There is a successor z such that $\langle \mathfrak{f}, \beta, z \rangle \vDash \psi$. By induction hypothesis, $\langle \mathfrak{g}, \beta, z^p \rangle \vDash \psi$. Then, as $y^p \lhd z$, z^p is a successor of y^p, so $\langle \mathfrak{g}, \beta, y^p \rangle \vDash \Diamond\psi$. Conversely, assume $\langle \mathfrak{g}, \beta, y^p \rangle \vDash \Diamond\psi$. Then for some successor z, $\langle \mathfrak{g}, \beta, z \rangle \vDash \psi$. We have that z^p is a weak successor of z, so it is a successor of y^p and of y. By induction hypothesis, $\langle \mathfrak{g}, \beta, z^p \rangle \vDash \psi$, and so $\langle \mathfrak{f}, \beta, y \rangle \vDash \Diamond\psi$. This shows the first claim. The second is similar. ∎

Let S be a subset of f. Put

$$\begin{array}{rclcrcl} \uparrow S & := & \{y : x \lhd y\} & \quad & \updownarrow S & := & \{y : x \unlhd y\} \\ \downarrow S & := & \{y : y \lhd x\} & \quad & \updownarrow S & := & \{y : y \unlhd x\} \end{array}$$

Definition 8.3.2. *Let $\langle \mathfrak{f}, \mathbb{F}\rangle$ be a $\boldsymbol{K4}$–frame. $a \in \mathbb{F}$ is called* **cofinal** *if $\downarrow a \supseteq \uparrow a$. $\mathfrak{G}$ is a* **cofinal subframe** *of $\mathfrak{F}$ if it is of the form $\mathfrak{F} \cap a$ for some cofinal $a \in \mathbb{F}$.*

Let φ be given and $\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi$. If $\mathfrak{f}$ is noetherian and $\mathfrak{g}$ contains the maximal points, then $\mathfrak{g}$ is cofinal in $\mathfrak{f}$. For if $y \in g$ and z a weak successor of y, then z^μ is weak successor, which is in g.

We return now to the general case. Recall the definition of a *molecule*. Given φ, we have $\leq 2^k$ molecules, where $k := \sharp sf(\varphi)$. We call a **molecule** also the set of points satisfying one and the same molecule. Moreover, $M(\varphi)$, or simply M, is the set of all molecules with respect to φ. M forms a partition of the original frame. Let $a \in \mathbb{F}$. Call x **critical** for a if for no weak successor y of x such that $y \in a$ we have $mo_{\mathfrak{M}}(y) = mo_{\mathfrak{M}}(x)$. This leads to the following definition

$$crit(\chi) := \bigvee_{\mu \in M} \mu \wedge \neg\chi \wedge \Box(\chi \to \neg\mu)$$

A point has **molecular span** (or **m–span**) N if it is in the set

$$N\text{–}span := \bigwedge_{\mu \in N} \Diamond\mu \bigwedge_{\mu \in M-N} \neg\Diamond\mu$$

The **molecular depth** (or **m–depth**) of a point x is the length of a maximal sequence $x = x_0 \lhd x_1 \lhd \ldots \lhd x_{n-1}$ such that x_{i+1} has lesser molecular span than x_i. (It follows then that $x_i \vec{\lhd} x_{i+1}$.) In particular, x is of m–depth 0 iff it is of minimal m–span. Notice that $x \in crit(\chi)$ iff every successor of x satisfying χ has lesser m–span (unless x is irreflexive). So, as can easily be seen, the operation of taking critical points can only be nontrivially iterated finitely many times, at most as many times as there are molecules.

Now let $\mathfrak{M} = \langle \mathfrak{F}, \beta\rangle$ be a global model. We define a frame $\mathfrak{S}(\varphi)$, a finite Kripke–frame $\mathfrak{z}(\varphi)$ and a p–morphism $\pi : \mathfrak{S}(\varphi) \twoheadrightarrow \mathfrak{z}(\varphi)$ such that the following holds.

1. $\mathfrak{s}(\varphi)$ is a cofinal subframe of $\mathfrak{f}$, and $\mathbb{S}(\varphi) \subseteq \mathbb{F}$.
2. $\pi : \mathfrak{S}(\varphi) \twoheadrightarrow \mathfrak{z}(\varphi)$ is the refinement map.
3. For every $x \in f$ there exists a weak successor $x^p \in \mathfrak{S}(\varphi)$ of x such that
   (1.) x^p has the same molecule as x;
   (2.) x^p has the same molecular depth as x.
4. For every $x \in s(\varphi)$, x is of molecular depth d iff $\pi(x)$ is of depth d in $\mathfrak{z}(\varphi)$.
5. For every $x \in z(\varphi)$, $\pi^{-1}(x)$ is definable by means of a formula of modal degree $2 + d \cdot dp(\varphi)$, d the molecular depth of x.

The frames $\mathfrak{S}(\varphi)$ and $\mathfrak{z}(\varphi)$ will be defined in stages. We define $\mathfrak{S}_d(\varphi)$, $\mathfrak{z}_d(\varphi)$ and maps $\pi_d : \mathfrak{S}_d(\varphi) \twoheadrightarrow \mathfrak{z}_d(\varphi)$. Here, $\mathfrak{S}_d(\varphi)$ is the frame of points of $\mathfrak{S}(\varphi)$ of molecular depth $< d$; π_d is the refinement map. $\mathfrak{z}_d(\varphi)$ is the set of points of depth $< d$ in $\mathfrak{z}(\varphi)$. We will have the following facts: $\mathfrak{S}_d(\varphi)$ is a generated subframe of $\mathfrak{S}_{d+1}(\varphi)$, consisting of the points x such that $\pi_{d+1}(x)$ is of depth $\leq d$. Furthermore, $\pi_{d+1} \upharpoonright s_d(\varphi) = \pi_d$. π_d is the refinement map. For $x \in z(\varphi)$ we denote by $\gamma(x)$ the formula defining the set $\pi^{-1}(x)$. It is clear from the previous remarks that $\gamma(x)$ is independent of d.

We start with $d = 1$. $\mathfrak{S}_1(\varphi)$ consists of all points of depth 0. These are the points critical for $\bot$. Moreover, the formulae $\gamma(x)$ are of degree 2. Now we define the points of $\mathfrak{S}_{d+1}(\varphi)$ on the basis of the points of $\mathfrak{S}_d(\varphi)$. Let Δ be the set of $\gamma(x)$, $x \in z_d(\varphi)$. Put $\delta^* := \bigvee \Delta$. A point y is of **molecular width** $\Gamma \subseteq \Delta$ if it satisfies the formula Γ–*width* defined by

$$\Gamma\text{–}width := \bigwedge_{\gamma\in\Gamma} \Diamond\gamma \wedge \bigwedge_{\gamma\in\Delta-\Gamma} \neg\Diamond\gamma$$

As before, the notion of *minimal molecular width* is defined by

$$mwidth := \neg\delta^* \wedge \Box\delta^* \vee \bigvee_{\Gamma\subseteq\Delta}(\Gamma\text{–}width \wedge \bigwedge_{\Sigma\subsetneq\Gamma} \Sigma\text{–}width \to \delta^*)$$

Among the points of minimal molecular width we select the points of minimal molecular span.

$$mspan := \neg\delta^* \wedge \begin{cases} & \bigvee_{\mu\in M(\varphi)} \mu \wedge \Box(\neg\mu \wedge \delta^*) \\ \vee & \bigvee_{N\subseteq M(\varphi)}(N\text{–}span \wedge \bigwedge_{O\subsetneq N}(O\text{–}span \to \delta^*)) \end{cases}$$

Finally

$$\zeta := mspan \wedge mwidth$$

Let now $x \in s_{d+1}(\varphi)$ iff $x \in s_d(\varphi)$ or x is of minimal molecular span and molecular width. For $x \in s_{d+1}(\varphi)$, let x have molecular span N and molecular width Γ; then put

$$\gamma(x) := N\text{–}span \wedge \Gamma\text{–}width \wedge \zeta$$

This is a formula of depth $k_{d+1} := k_d + dp(\varphi)$ if φ has depth at least 1. It is clear that $\mathfrak{S}_d(\varphi)$ is a generated subframe of $\mathfrak{S}_{d+1}(\varphi)$. Moreover, every point of $\mathfrak{S}_{d+1}(\varphi)$ has a strict successor in $\mathfrak{S}_d(\varphi)$. As internal sets we take the sets generated by $\beta(p_i)$. Let π_{d+1} be the refinement map of $\mathfrak{S}_{d+1}(\varphi)$. We show that the image of π_{d+1} is a finite frame of depth $d + 1$.

Lemma 8.3.3. *Suppose that x and y in $\mathfrak{S}_{d+1}(\varphi)$ have identical width, identical span and identical molecule in $\langle\mathfrak{F},\beta\rangle$. Then in $\mathfrak{S}_d(\varphi)$, $x \in a$ iff $y \in a$ for all internal sets. Hence $\pi_{d+1}(x) = \pi_{d+1}(y)$. Consequently, $\mathfrak{z}_{d+1}(\varphi)$ is finite.*

Proof. Every set is generated by $\beta(p_i)$, so we may reason by induction on the formula $\chi(\vec{p})$. For $\chi = p_i$ the case is clear, likewise the steps for $\neg$ and $\wedge$. Now let $\chi = \Diamond\eta$. Assume that $x \in \overline{\beta}(\Diamond\eta)$. Then there is a successor u of x satisfying η. Two cases arise. **Case 1.** $u \in \mathfrak{S}_d(\varphi)$. Then, as y has the same width as x, there is a point v such that $y \lhd v$ and v satisfies η. **Case 2.** $u \notin \mathfrak{S}_d(\varphi)$. Then u has the same width as x. It also has the same span as x. The atom of u is the span of x. u is critical, that is, it satisfies a molecule that no successor of layer $< d$ can satisfy. As y is of equal span as x and of equal m–span as x (by virtue of being of equal width) there exists a point v such that $y \lhd v$ and v satisfies the molecule of u. v cannot be of layer $< d$, since its molecule cannot be satisfied there. Hence v is of layer d. This means that v has the same span as u, the same width as u and the same atom as u. By induction hypothesis, $v \in \overline{\beta}(\eta)$. Hence $y \in \overline{\beta}(\Diamond\eta) = \overline{\beta}(\chi)$. And that had to be shown. □

Now we put for $x \in \mathfrak{z}(\varphi)$, $\gamma(x) := \gamma(y)$, where $\pi_{d+1}(y) = x$. By definition of $\mathfrak{z}(\varphi)$, this is independent of the representative. Now let x be a point of molecular depth d. By assumption, if $x \lhd y$ and y has lesser molecular span than x, y is of molecular depth $< d$. Moreover, x has a successor y of molecular depth $d - 1$ such that the molecule of x is not satisfied at any weak successor of y. By induction hypothesis, y has a weak successor y^p in $\mathfrak{S}_d(\varphi)$ which is of depth $d - 1$ in $\mathfrak{z}_d(\varphi)$. Clearly, $x \lhd y^p$. Therefore, $\gamma(y^p)$ is contained in the width of x. Consider a weak successor u of x which is of minimal width and minimal span. Then, by choice of x, u has the same span and same molecule as x, and $\gamma(y^p)$ is contained in its width. It follows that u has molecular depth d. This shows the claims.

Theorem 8.3.4 (Zakharyaschev, Fine). *Let φ be a formula and assume that $\langle \mathfrak{F}, \beta, x\rangle \vDash \varphi$. Let $k := \sharp sf(\varphi)$. Then there exist $\mathfrak{S}(\varphi)$ and $\mathfrak{z}(\varphi)$ such that the following holds.*

1. *$\mathfrak{S}(\varphi)$ is a cofinal subframe of $\langle \mathfrak{f}, \mathbb{G}\rangle$, where $\mathbb{G}$ is a subalgebra of $\mathfrak{F}$; the refinement of $\mathfrak{S}(\varphi)$ is $\mathfrak{z}(\varphi)$.*
2. *$\mathfrak{z}(\varphi)$ is finite and of depth $\leq 2^k$.*
3. *For every $x \in f$ there exists a $x^p \in s(\varphi)$ such that*
   (1.) *x^p is a weak successor of x;*
   (2.) *the molecule of x in $\langle \mathfrak{F}, \beta\rangle$ is the same as the molecule of x^p in $\langle \mathfrak{S}(\varphi), \beta\rangle$.*

*$\langle \mathfrak{S}(\varphi), \beta\rangle$ and $\mathfrak{S}(\varphi)$ is called the φ–**extract** of $\langle \mathfrak{F}, \beta\rangle$ and $\langle \mathfrak{z}(\varphi), \beta\rangle$ as well as $\mathfrak{z}(\varphi)$ the φ–**reduct**. x is called **quasi–maximal** if $x \in s(\varphi)$.*

Recall from Section 3.5 the definition of a subframe logic. It is a logic whose class of frames is closed under taking subframes. Now call Λ a **cofinal subframe logic** if for every $\mathfrak{F} \vDash \Lambda$ and every cofinal $\mathfrak{G} \subseteq \mathfrak{F}$ also $\mathfrak{G} \vDash \Lambda$.

Corollary 8.3.5 (Zakharyaschev). *Every cofinal subframe logic has the finite model property.*

Corollary 8.3.6 (Fine). *Every subframe logic has the finite model property.*

We remark here that there is a simple criterion for a logic to see whether it is a (cofinal) subframe logic.

Theorem 8.3.7. *Let Λ be a logic. If Λ has the finite model property and the set of finite models is closed under taking (cofinal) subframes, then Λ is a (cofinal) subframe logic.*

The proof is simple. Let Θ be the cofinal subframe logic whose set of finite frames equals the set of finite frames for Λ. Such a logic exists by assumption on

Λ. Since Λ has the finite model property, $\Theta = \Lambda$. Hence, **S4**, **S5**, **G** and **Grz** are subframe logics.

Exercise 280. Show that **K4.D** is a cofinal subframe logic but not a subframe logic.

Exercise 281. Show that the logics of width θ, the logics of tightness τ, the logics of fatness φ, the logics of depth δ, all are subframe logics, in all combinations.

Exercise 282. Let $\Lambda \supseteq$ **K4** be canonical and let the frames be determined by a set of universal sentences. Then Λ is a subframe logic.

8.4. Refutation Patterns

This section is devoted to the so-called *canonical formulae* by Michael Zakharyaschev. Basically, for every formula φ there exists a finite set of geometrical configurations, called *refutation patterns*, such that a frame refutes φ iff it realizes one of the refutation patters. The exclusion of a refutation pattern can be characterized by a modal axiom, and this axiom is called a canonical formula. The refutation patterns for φ can be constructed from the frames underlying minimal countermodels for φ, by observing how the selction procedure selects points and how the selected frame lies embedded in the whole frame. For concreteness, let us take a frame $\mathfrak{F}$. Let $\langle \mathfrak{F}, \beta, x\rangle \vDash \neg\varphi$. In $\mathfrak{F}$ lies cofinally embedded the φ–extract $\mathfrak{S}(\varphi)$. There is a contraction $\pi : \mathfrak{S}(\varphi) \twoheadrightarrow \mathfrak{z}(\varphi)$ onto the reduct. Now let us take a look at the other points of the frame and see how they lie with respect to $\mathfrak{S}(\varphi)$. More precisely, we are only interested in the way they lie with respect to $\mathfrak{z}(\varphi)$, because points that are being mapped onto the same element in the reduct have the same molecule. Call a subset V of points of $\mathfrak{z}(\varphi)$ a **view** if $V = \uparrow V$. Now take a point $x \in \mathfrak{f}$. The set $vw(x) = \pi[\uparrow x \cap z(\varphi)]$ is a view. It is called the ($\mathfrak{z}$–)**view** of x. A view V is **internal** if there is an $x \in s(\varphi)$ such that $V = vw(x)$, and **external** if there is an $x \in f - s(\varphi)$ such that $V = vw(x)$. Notice that a given view may be both external and internal; it may also be neiter internal nor external. We say that the frame **realizes an external** (**internal**) **view** V if V is the external (internal) view of some point of f. We will see that for a given frame two factors determine whether or not a model for φ can be based on it. One is to which frames it is subreducible and the other is which $\mathfrak{z}$–views it realizes. To make this precise, two more definitions are needed. We say that two points **0–agree** if they satisfy the same non–modal formulae, and that they **φ–agree** if they satisfy the same subformulae of φ (iff they have the same molecule).

Proposition 8.4.1 (Agreement). *Let $\langle \mathfrak{F}, \beta\rangle$ be a model, $\mathfrak{S}(\varphi)$ the φ–extract, $\mathfrak{z}(\varphi)$ the φ–reduct and $\pi : \mathfrak{S}(\varphi) \twoheadrightarrow \mathfrak{z}(\varphi)$ the refinement map. Let x and y have the same $\mathfrak{z}$–view. If x and y 0–agree, then they also φ–agree.*

Proof. By induction on the subformula χ it is shown that for two points x and y which 0–agree, $x \vDash \chi$ iff $y \vDash \chi$. The only nonobvious case is $\chi = \Diamond\psi$. Let $x \vDash \Diamond\psi$. Then there is a successor u of x such that $u \vDash \psi$. Furthermore, there exists a weak successor $u^p \in s(\varphi)$ of u such that $u^p \vDash \psi$. We then have $x \lhd u^p$ as well. By assumption, there exists w such that $y \lhd w$ and $\pi(w) = \pi(v)$. It follows that for all subformulae τ of φ, $u^p \vDash \tau$ iff $w \vDash \tau$. Choose $\tau = \psi$. This implies $w \vDash \psi$ and thus $y \vDash \chi$. ☐

Proposition 8.4.2 (Homogenization). *Let $\langle \mathfrak{F}, \beta \rangle$ be a model, x an external world, and Y a set of external worlds and let all points of $Y \cup \{x\}$ have identical view. Assume that Y is an internal set. Define γ as follows. $\gamma(p) := \beta(p) \cup Y$ if $x \in \beta(p)$, and $\gamma(p) := \beta(p) - Y$ otherwise. Let $y, y' \in Y$. Then y and y' φ–agree in $\langle \mathfrak{F}, \gamma \rangle$. Moreover, for $z \notin Y$ the molecule of z in $\langle \mathfrak{F}, \gamma \rangle$ is the same as the molecule of z in $\langle \mathfrak{F}, \beta \rangle$.*

Proof. Since Y is internal, γ is a valuation. Now, all internal points have the same valuation as before, likewise all external points $\notin Y$. Now we prove that if $z \notin Y$, then for all $\chi \in sf(\varphi)$ we have

$$\langle \mathfrak{F}, \gamma, z \rangle \vDash \chi \quad \Leftrightarrow \quad \langle \mathfrak{F}, \beta, z \rangle \vDash \chi$$

The only non–obvious case is $\chi = \Diamond\psi$. Let v be a successor of z such that $\langle \mathfrak{F}, \gamma, v \rangle \vDash \psi$. v has a weak successor v^p such that $\langle \mathfrak{F}, \beta, v^p \rangle \vDash \psi$. By induction hypothesis, $\langle \mathfrak{F}, \gamma, v^p \rangle \vDash \psi$. Hence, $\langle \mathfrak{F}, \beta, z \rangle \vDash \Diamond\psi$, since $z \lhd v^p$. Now let the latter be the case. Then, analogously, there is a v^p such that $\langle \mathfrak{F}, \beta, v^p \rangle \vDash \psi$. By induction hypothesis, $\langle \mathfrak{F}, \gamma, v^p \rangle \vDash \psi$ and so $\langle \mathfrak{F}, \gamma, z \rangle \vDash \Diamond\psi$, as required. ☐

The last lemma says that for external points we can make the valuation quite uniform, because in evaluating a formula we can 'skip' the external points. It is actually possible to derive the Homogenization Lemma from the first lemma. Finally, we are also interested in the possibility to insert points. Let $\langle \mathfrak{F}, \beta \rangle$ be a model. We then have a uniquely defined extract $\mathfrak{S}(\varphi)$ and the refinement map $\pi : \mathfrak{S} \twoheadrightarrow \mathfrak{z}$. We fix all these elements. We see that we can tinker with the external points quite drastically, taking them away if we want, and add new ones. However, what we need to know when we insert points is that its view (if external) must already be realized in the frame by a point x. Then we simply extend the valuation by giving the new point the valuation of x and we have again a model of φ. Thus, the structure of the external points is quite irrelevant as long as certain views are not realized. Notice also that if a view is realized then we always have a witness x giving us a valuation for the new point. Now go one step further. Take any frame $\mathfrak{F}$ with a cofinal subframe $\mathfrak{G}$ and a p–morphism $\pi : \mathfrak{G} \twoheadrightarrow \mathfrak{z}$. Then we can define $\mathfrak{z}$–views analogous to views. Namely, given a point x we put $vw(x) = \pi[\uparrow x]$ and call it the $\mathfrak{z}$–**view** of x. Notice that views depend on $\mathfrak{G}$ and π in addition to $\mathfrak{z}$. And we have the following theorem.

Proposition 8.4.3. *Let $\mathfrak{M} := \langle \mathfrak{F}, \beta \rangle$. Let $\mathfrak{S} = \mathfrak{S}(\varphi)$ be the φ–extract of $\langle \mathfrak{F}, \beta \rangle$ and $\pi : \mathfrak{S} \twoheadrightarrow \mathfrak{z}$ the refinement map. Consider a frame $\mathfrak{G}$ containing a cofinal subframe $\mathfrak{H}$ and $\rho : \mathfrak{H} \twoheadrightarrow \mathfrak{z}$, such that every external $\mathfrak{z}$–view of $\mathfrak{G}$ is realized in $\mathfrak{F}$. Moreover, let every point of g have a weak successor in h with identical view. Then there exists a valuation γ such that the φ–reduct of $\mathfrak{N} := \langle \mathfrak{G}, \gamma \rangle$ is isomorphic to $\mathfrak{z}$.*

Proof. By assumption there is a function $z \mapsto z^p : g \to h$ such that z^p is a weak successor of z with identical view. Furthermore, $z^p = z$ if $z \in h$. The valuation is as follows. (i) $x \in h$. Then $x \in \gamma(p)$ iff there is a $y \in s(\varphi)$ with $\pi(y) = \rho(x)$ and $y \in \beta(p)$. (ii) $x \notin h$. For each external view V take an external y_V in $\mathfrak{F}$ with identical $\mathfrak{z}$–view. y_V exists by assumption. Then let $x \in \gamma(p)$ iff $y_V \in \beta(p)$. This is a valuation on $\mathfrak{G}$. For $\gamma(p)$ is a union of sets of the form $\rho^{-1}(u)$, $u \in z(\varphi)$, and sets of points of indentical view. These are all internal sets. We claim the following.

$$\begin{array}{ll} (*) & \text{if } \pi(y) = \rho(x) \text{ then } mo_{\mathfrak{M}}(y) = mo_{\mathfrak{N}}(x) \\ (**) & mo_{\mathfrak{N}}(z) = mo_{\mathfrak{N}}(z^p) \end{array}$$

From this it immediately follows that the reducts are isomorphic. We show $(*)$ and $(**)$ by simultaneous induction on the subformulae of φ. For variables both hold by construction. The case of $\neg$ and $\wedge$ is clear. Now let $\chi = \Diamond\psi$. We show $(*)$ for χ. Let $\langle\mathfrak{F}, \beta, y\rangle \vDash \Diamond\psi$. Then there exists a z such that $y \lhd z$ and $\langle\mathfrak{F}, \beta, z\rangle \vDash \psi$. We may actually assume that $z \in s(\varphi)$. Then there exists a $v \in h$ such that $\pi(z) = \rho(v)$. Now, $\pi(y) = \rho(x)$ and $\pi(y) \lhd \pi(z)$ imply $\rho(x) \lhd \pi(z)$. Since ρ is a p–morphism, there is a v such that $x \lhd v$ and $\rho(v) = \pi(z)$. By $(*)$ for ψ, $\langle\mathfrak{G}, \gamma, v\rangle \vDash \psi$ and so $\langle\mathfrak{G}, \gamma, x\rangle \vDash \Diamond\psi$. Now let the latter be the case. Then there exists a v such that $x \lhd v$ and $\langle\mathfrak{G}, \gamma, v\rangle \vDash \psi$. By inductive hypothesis for $(**)$, $\langle\mathfrak{G}, \gamma, v^p\rangle \vDash \psi$. Now there exists a w such that $y \lhd w$ and $\pi(w) = \rho(v^p)$, and so $\langle\mathfrak{F}, \beta, w\rangle \vDash \psi$. It follows that $\langle\mathfrak{F}, \beta, y\rangle \vDash \Diamond\psi$. Now we show $(**)$ for $\chi = \Diamond\psi$. Assume that $\langle\mathfrak{G}, \gamma, z^p\rangle \vDash \Diamond\psi$. Then there exists a $y \rhd z^p$ such that $\langle\mathfrak{G}, \gamma, y\rangle \vDash \psi$. Hence $\langle\mathfrak{G}, \gamma, z\rangle \vDash \Diamond\psi$, since $z \unlhd z^p \lhd y$. Now assume that $\langle\mathfrak{G}, \gamma, z\rangle \vDash \Diamond\psi$. Then for some $y \rhd z$, $\langle\mathfrak{G}, \gamma, y\rangle \vDash \psi$. Since $y \unlhd y^p$, $z \lhd y^p$. Now z and z^p have the same $\mathfrak{z}$–view and so there is a u^p such that $z^p \lhd u^p$ and $\rho(u^p) = \rho(y^p)$. By $(*)$ for ψ, $\langle\mathfrak{G}, \gamma, u^p\rangle \vDash \psi$. Hence $\langle\mathfrak{F}, \beta, z^p\rangle \vDash \Diamond\psi$. ◻

The last theorem has in effect identified what the geometric condition for φ is. We need to know: (1.) the structure of the reducts and (2.) the admissible external views for each reduct. Since views are upward closed sets, that is, cones, the following definition emerges.

Definition 8.4.4. *Let $\mathfrak{z}$ be a finite Kripke–frame and $\mathfrak{V}$ be a set of cones of $\mathfrak{z}$. Then the pair $\mathbb{P} := \langle\rho, \mathfrak{V}\rangle$ is called a* ***refutation pattern****. We say that a frame $\mathfrak{F}$* ***satisfies*** *or* ***realizes*** *$\mathbb{P}$ if there is a subframe $\mathfrak{G}$ and a contraction $\mathfrak{G} \twoheadrightarrow \mathfrak{z}$ such that no external $\mathfrak{z}$–view is in $\mathfrak{V}$. $\mathfrak{v} \in \mathfrak{V}$ is called a* ***closed domain*** *of $\mathbb{P}$. If $\mathfrak{F}$ does not realize $\mathbb{P}$ we say that it* ***omits*** *$\mathbb{P}$.*

Given φ, there exist only finitely many reducts. On each reduct there exist finitely many closed domains, hence there are finitely many refutation patterns for φ. They can be calculated algorithmically. With φ given, enumerate all models of size $\leq 2^k$, $k := \sharp sf(\varphi)$. If necessary, reduce these models. This enumerates all reducts. Next try inserting a new point x somewhere and defining a valuation such that the reduct remains intact. If this is impossible, the view of x is a closed domain.

The nonsatisfaction of a refutation pattern can be characterized axiomatically. Given a refutation pattern $\langle \mathfrak{z}, \mathfrak{V} \rangle$, where $\mathfrak{z}$ is a frame based on the set $\{0, 1, \ldots, n-1\}$ and with root 0, we take for each i a distinct variable p_i and define

$$\begin{array}{rlll} real(\mathfrak{z}, \mathfrak{V}) := & & \bigwedge\langle p_i \to \neg p_j : i \neq j\rangle & (\neq) \\ & \wedge & \bigwedge\langle p_i \to \Diamond p_j : i \lhd j\rangle & (\lhd) \\ & \wedge & \bigwedge\langle p_i \to \neg\Diamond p_j : i \ntriangleleft j\rangle & (\ntriangleleft) \\ & \wedge & \bigwedge\langle \neg evr(\mathfrak{v}) : \mathfrak{v} \in \mathfrak{V}\rangle & (cd) \\ evr(\mathfrak{v}) := & & \bigwedge\langle \Diamond p_i : i \in \mathfrak{v}\rangle & \\ & \wedge & \bigwedge\langle \neg\Diamond p_i : i \notin \mathfrak{v}\rangle & \\ & \wedge & \bigwedge\langle \neg p_i : i < n\rangle & \end{array}$$

Consider a valuation $\beta : p_i \mapsto a_i$ into $\mathfrak{F}$ and a world x such that $\langle \mathfrak{F}, \beta, x\rangle \vDash p_0 \wedge \Box^+ real(\mathfrak{z}, \mathfrak{V})$. Then the set $\bigcup_i \beta(p_i)$ is a subframe which can be mapped onto $\mathfrak{z}$. This follows from $(\neq)$, $(\lhd)$ and $(\ntriangleleft)$. The formula $evr(\mathfrak{v})$ says that a point has ρ–view $\mathfrak{v}$ and is external. This is excluded by $real(\mathfrak{z}, \mathfrak{V})$. The cofinality requirement need not be stated separately, for we have the following fact.

Proposition 8.4.5. *Let $\langle \mathfrak{z}, \mathfrak{V}\rangle$ be an embedding pattern. If $\mathfrak{G}$ is a subframe in $\mathfrak{F}$ and $\mathfrak{G} \twoheadrightarrow \mathfrak{z}$ then $\mathfrak{G}$ is cofinal iff the empty $\mathfrak{z}$–view is not external.*

Now define

$$\gamma(\mathfrak{z}, \mathfrak{V}) := \Box^+ real(\mathfrak{z}, \mathfrak{V}) \to \neg p_0$$

Definition 8.4.6. *A formula φ is called a* **canonical formula** *if φ is of the form $\gamma(\mathbb{P})$ for some refutation pattern $\mathbb{P} = \langle \mathfrak{z}, \mathfrak{V}\rangle$.*

Proposition 8.4.7. *Let $\mathbb{P}$ be a refutation pattern. Then $\mathfrak{F} \nvDash \gamma(\mathbb{P})$ iff $\mathfrak{F}$ satisfies $\mathbb{P}$.*

Now consider φ again. We can specify all refutation patterns for φ. If φ is an axiom, the frames satisfying the refutation patterns are exactly those frames which must be excluded. For they allow to define a valuation refuting φ. We have seen that these patterns can be axiomatized by canonical formulae. Thus we have the following theorem.

Theorem 8.4.8 (Zakharyaschev). *There is an algorithm which for given formula φ returns a finite set of refutation patterns $\langle \mathfrak{z}_i, \mathfrak{V}_i\rangle$, $i < n$, such that $\boldsymbol{K4} \oplus \varphi = \boldsymbol{K4} \oplus \{\gamma(\rho_i, \mathfrak{V}_i) : i < n\}$.*

The consequences of this theorem are enormous and we will have to be content to list a few of them in the sequel. For the moment let us notice a few details. In [**245**], the closed domains are defined as *antichains* rather than cones. The effect is the same, because an antichain generates a cone, and each cone is generated by an antichain. However, there are more antichains than there are cones, so taking antichains introduces a redundancy here. The formulae $\gamma(\mathfrak{z}, \mathfrak{V})$ are therefore not identical with the α–formulae of [**245**]. We will switch freely between a definition of refutation pattern via cones and via antichains, taking whatever suits the purpose best.

Finally we have to discuss the case where external views are also internal, that is, when there is a point $x \in \mathfrak{z}$ such that $\uparrow x = \mathfrak{v}$ for $\mathfrak{v} \in \mathfrak{V}$. What happens if we exclude such an external view? There are two cases. First, assume that x is reflexive. Then $x \in \mathfrak{v}$, corresponding to an antichain $\{x\}$. Then nothing changes if in the frame we take all points with view $\mathfrak{v}$ as internal points, changing their valuation into that of x. Thus, forbidding such an external view has no effect. If, however, x is irreflexive, then forbidding the view of x to be external will have substantial effects.

Proposition 8.4.9. *Let $\langle \mathfrak{z}, \mathfrak{V} \rangle$ be a refutation pattern, and x a reflexive point. Then $\boldsymbol{K4} \oplus \gamma(\mathfrak{z}, \mathfrak{V}) = \boldsymbol{K4} \oplus \gamma(\mathfrak{z}, \mathfrak{V} \cup \uparrow x)$.*

It should be emphasized that even though the formulae $\gamma(\mathfrak{z}, \mathfrak{V} \cup \{\uparrow x\})$ and $\gamma(\mathfrak{z}, \mathfrak{V})$ are axiomatically equivalent over **K4**, they are not satisfied by the same models and so not deductively equivalent. A valuation refuting the first formula refutes the second; the converse does not necessarily hold. Namely, to satisfy the first formula, the view $\uparrow x$ may not be external, while for the second it is enough that $\uparrow x$ is internal, it may also be external. To prove Proposition 8.4.9, assume that $\mathfrak{M} = \langle \mathfrak{F}, \beta \rangle$ satisfies $\gamma(\mathfrak{z}, \mathfrak{V})$. Define a new valuation δ and a model $\mathfrak{N} := \langle \mathfrak{F}, \delta \rangle$ as follows. If $x \in \beta(p)$ let $\delta(p)$ be the union of $\beta(p)$ and all external points with view $\uparrow x$; if $x \notin \beta(p)$, let $\delta(p)$ be $\beta(p)$ minus the set of all external points with view $\uparrow x$ (check that this is an internal set, so δ is well–defined). In $\langle \mathfrak{F}, \delta \rangle$ the view $\uparrow x$ is now internal. To see that, note that the points of $\mathfrak{M}$ which were external and had view $\uparrow x$ have the same molecule as x in $\mathfrak{N}$, and it can be shown that they belong to the extract of $\mathfrak{N}$. So they are internal. Nothing else changed. Therefore, the φ–reduct of $\mathfrak{N}$ is the same as the φ–reduct of $\mathfrak{M}$, and they realize the same views except for $\uparrow x$. Playing with this distinction will be helpful sometimes.

In addition to $\gamma(\mathfrak{z}, \mathfrak{V})$ there is a formula $\gamma^\circ(\mathfrak{z}, \mathfrak{V})$, which results from $\gamma(\mathfrak{z}, \mathfrak{V})$ by dropping the cofinality requirement. We then have

Theorem 8.4.10. *A logic Λ is a cofinal subframe logic iff for every $\gamma(\mathfrak{z}, \mathfrak{V})$ we have that from $\gamma(\mathfrak{z}, \mathfrak{V}) \in \Lambda$ follows $\gamma(\mathfrak{z}, \varnothing) \in \Lambda$. A logic is a subframe logic iff $\gamma(\mathfrak{z}, \mathfrak{V}) \in \Lambda$ implies $\gamma^\circ(\mathfrak{z}, \varnothing) \in \Lambda$. Hence a cofinal subframe logic can be axiomatized by formulae of the form $\gamma(\mathfrak{z}, \varnothing)$, a subframe logic by axioms of the form $\gamma^\circ(\mathfrak{z}, \varnothing)$.*

Proof. Let $\gamma(\mathfrak{z}, \varnothing) \notin \Lambda$, that is, Λ admits the refutation pattern $\langle \mathfrak{z}, \mathfrak{V} \rangle$. Suppose there is a a frame $\mathfrak{F}$ with a cofinal subframe $\mathfrak{G}$ such that $\mathfrak{G} \twoheadrightarrow \mathfrak{z}$ respecting the external views. However, by assumption on Λ, $\mathfrak{G}$ is itself a frame for Λ and realizes the refutation pattern $\langle \mathfrak{z}, \varnothing \rangle$. Likewise for subframe logics. ◻

Theorem 8.4.11. *A logic $\boldsymbol{K4} \oplus \gamma(\mathfrak{z}, \mathfrak{V})$ is a splitting of $\boldsymbol{K4}$ by $\mathfrak{z}$ iff $\mathfrak{V} \cup \{\uparrow x : x \lhd x\}$ contains the set of all cones of $\mathfrak{z}$.*

Proof. Let $\mathfrak{F}$ be a frame, $\mathfrak{G} \subseteq \mathfrak{F}$ a cofinal subframe such that $\mathfrak{G} \twoheadrightarrow \mathfrak{z}$. If $\mathfrak{G}$ is a generated subframe, then no view is external. Hence no external view is realized. On the other hand, let no external view be realized. Then $\mathfrak{G}$ is a generated subframe. □

The canonical formulae do not provide an axiomatic basis for extensions of **K4**. This follows from the fact that there exists an extension of **K4** that does not possess an independent axiomatization. In the case of (cofinal) subframe logics the situation is different, though. Let Λ be a subframe logic and $\mathfrak{z}$ rooted and finite. Denote by $\Lambda/_{\mathrm{F}}\,\mathfrak{z}$ the smallest subframe logic containing Λ not having $\mathfrak{z}$ as a frame, and call $\Lambda/_{\mathrm{F}}\,\mathfrak{z}$ the **Fine–splitting** of Λ by $\mathfrak{z}$. It turns out that $\Lambda/_{\mathrm{F}}\,\mathfrak{z} = \Lambda \oplus \gamma^{\circ}(\mathfrak{z}, \varnothing)$. Any subframe logic Λ is a Fine–splitting $K4/_{\mathrm{F}}\,G$ with $G = \{\mathfrak{z} : \mathfrak{z} \notin \mathrm{Krp}(\Lambda), g \text{ rooted and finite}\}$. Analogously, let Λ be a cofinal subframe logic. The **Zakharyaschev–splitting** of Λ by $\mathfrak{z}$ is the least cofinal subframe logic containing Λ for which $\mathfrak{z}$ is not a frame. It is axiomatizable by $\Lambda \oplus \gamma(\mathfrak{z}, \varnothing)$. We write $\Lambda/_{\mathrm{Z}}\,\mathfrak{z}$. It follows from the fact that the cofinal subframe logics have the finite model property that we have a sublattice $\mathcal{S}\,\mathbf{K4}$ of subframe logics and a sublattice $\mathcal{CF}\,\mathbf{K4}$ of cofinal subframe logics, that all (and only) the rooted finite frames induce splittings and the splitting logics all have the finite model property. Under these circumstances we conclude the following. Put $\mathfrak{f} \prec_F \mathfrak{g}$ if $\mathfrak{g}$ is a contractum of a subframe of $\mathfrak{f}$, and $\mathfrak{f} \prec_Z \mathfrak{g}$ if $\mathfrak{g}$ is a contractum of a cofinal subframe of $\mathfrak{f}$. Then $\prec_Z \subseteq \prec_F$ and in both cases the lattices are isomorphic to the lattice of upper sets of finite rooted frames. Hence, on the finite frames the topology is the Alexandrov–topology and so the lattices are actually continuous. Moreover, there is no infinite strictly ascending chain of prime elements.

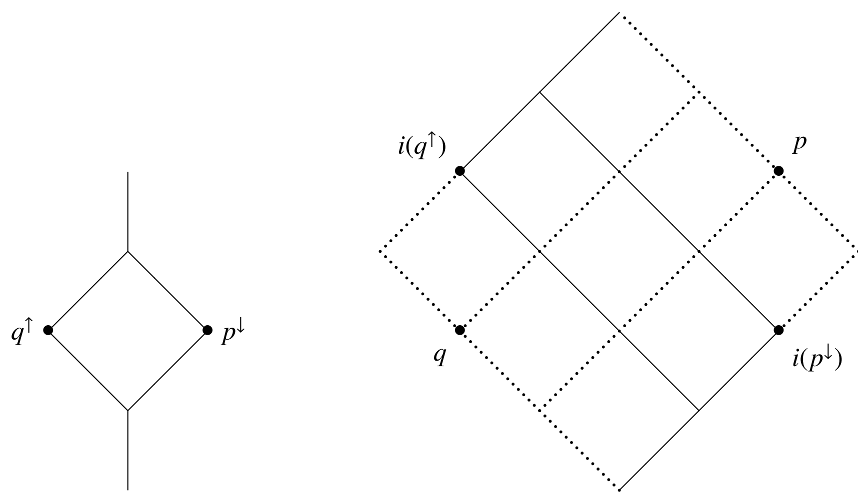
Corollary 8.4.12. *Both $\mathcal{S}\,\boldsymbol{K4}$ and $\mathcal{CF}\,\boldsymbol{K4}$ are continuous lattices and have a strong basis. The natural embeddings $\mathcal{S}\,\boldsymbol{K4} \subseteq \mathcal{CF}\,\boldsymbol{K4} \subseteq \mathcal{E}\,\boldsymbol{K4}$ are continuous maps.*

The continuity of the embeddings is a consequence of the fact that the upper limits and the lower limits of chains coincide with the respective limits in $\mathcal{E}\,\mathbf{K4}$. So any infinite intersection of (cofinal) subframe logics is again a (cofinal) subframe logic. In addition, we can study the splittings of the sublattices as induced splittings of the larger lattice. Notice namely, that if we have a locale $\mathfrak{L}$ and a point $p : \mathfrak{L} \twoheadrightarrow \mathbf{2}$ such that $p^{-1}(0)$ is a principal ideal and $p^{-1}(1)$ a principal filter, and we have a sublocale $i : \mathfrak{M} \rightarrowtail \mathfrak{L}$ such that the embedding respects all limits, then $i \circ p : \mathfrak{M} \twoheadrightarrow \mathbf{2}$ is a point with similar properties. To put this concretely, if $\langle p, q\rangle$ is a splitting of $\mathfrak{L}$ then $\langle p^{\downarrow}, q^{\uparrow}\rangle$ is a splitting of $\mathfrak{M}$ where

$$
\begin{aligned}
p^{\downarrow} &:= \bigsqcup \langle x : i(x) \leq p\rangle \\
q^{\uparrow} &:= \bigsqcap \langle x : i(x) \geq q\rangle
\end{aligned}
$$

$p^{\downarrow}$ is the largest $\mathfrak{M}$–logic below p, while $q^{\uparrow}$ is the smallest $\mathfrak{M}$–logic containing q. In the present context, $\Lambda^{\uparrow}$ is the (cofinal) subframe closure, the logic of all frames for which all (cofinal) subframes are frames for Λ, while $\Lambda^{\downarrow}$ is the (cofinal) subframe kernel, the logic of all (cofinal) subframes of frames for Λ. Figure 8.2 illustrates the situation. Take the distributive lattice to the left. It is a sublattice of the lattice to

FIGURE 8.2.

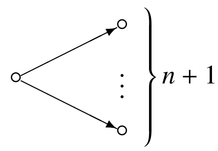


the right, the top and bottom elements are the same. The splitting $\langle p, q\rangle$ of the larger lattice induces the splitting $\langle p^{\downarrow}, q^{\uparrow}\rangle$ of the embedded lattice. The cofinal subframe logics provide a rich class of natural logics. We present them here without proof. In the exercises the reader is asked to supply some of the proofs, and to compare the Fine– and Zakharyaschev–splittings with standard splittings. We have $\mathbf{K4.1} := \mathbf{K4} \oplus \Box\Diamond p \to \Diamond\Box p$, $\mathbf{K4.2} := \mathbf{K4} \oplus \Diamond\Box p \to \Diamond\Box p$. $\mathbf{K4}_n$ is the logic of Kripke–frames in which there is no chain of points $\langle x_i : i < n+1\rangle$ such that $x_i \vec{\lhd} x_{i+1}$; $\mathbf{K4}.I_n$ is the logic of Kripke–frames in which there is no antichain of length $n+1$.

$$\begin{array}{lcl}
\mathbf{S4} & = & \mathbf{K4}/_{\mathrm{F}}\ \boxed{\bullet} \\
\mathbf{G} & = & \mathbf{K4}/_{\mathrm{F}}\ \boxed{\circ} \\
\mathbf{Grz} & = & \mathbf{K4}/_{\mathrm{F}}\ \{\boxed{\bullet}, \boxed{\infty}\} \\
\mathbf{K4.1} & = & \mathbf{K4}/_{\mathrm{Z}}\ \boxed{\infty} \\
\mathbf{K4.2} & = & \mathbf{K4}/_{\mathrm{Z}}\ F_1 \\
\mathbf{K4.3} & = & \mathbf{K4}/_{\mathrm{F}}\ F_1 \\
\mathbf{K4}_n & = & \mathbf{K4}/_{\mathrm{F}}\ L_n \\
\mathbf{K4}.I_n & = & \mathbf{K4}/_{\mathrm{F}}\ F_{n+1}
\end{array}$$

The notation is as follows. L_n is the set of frames whose reflexive closure is $\langle n, \geq\rangle$ and F_{n+1} is the set of frames whose reflexive closure is the frame $\mathfrak{d}_{n+1}$ of Figure 8.3. The exercises below shed some light on the splittings of $\mathcal{E}\,\mathbf{K4}$ as first outlined in

FIGURE 8.3. $\mathfrak{d}_{n+1}$



[**169**].

Notes on this section. The idea to study lattices of subframe logics rather than lattices of logics is a major theme of WOLTER [**244**]. This has turned out to be a more promising line of research than the investigation of entire lattices of logics, since the latter are too complex if the logics is not so strong (e. g. in the case of **S4.3**, **K5**).

Exercise 283. Show that $\mathbf{S4} = \mathbf{K4}/_{\mathrm{F}}\boxed{\bullet}$. Show also that **S4** is obtainable by splitting two frames from **K4**.

Exercise 284. Show that **Grz** is obtained by splitting two frames from **S4**, and that $\mathbf{Grz} = \mathbf{S4}/_{\mathrm{F}}\boxed{\infty}$. Verify that $\mathbf{S4.3} = \mathbf{S4}/_{\mathrm{F}}\mathfrak{d}_2$ and that **S4.3** is obtained by splitting two frames from **S4**.

Exercise 285. Show that $\mathbf{G} = \mathbf{K4}/_{\mathrm{F}}\boxed{\circ}$. Name a set U of frames such that $\mathbf{G} = \mathbf{K4}/_{\mathrm{Z}}U$. Show that there is no set N of frames such that $\mathbf{G} = \mathbf{K4}/N$.

Exercise 286. Show that **K4.1** and **K4.2** are cofinal subframe logics but not subframe logics.

Exercise 287. Show that there exist extensions of **K4** which are not cofinal subframe logics. Can you name one?

Exercise 288. (KRACHT [**126**].) Show that any splitting $\mathbf{G}/N$ for finite N has the finite model property.

Exercise 289. Let $\mathfrak{z}$ be a frame, $\mathfrak{V}$ and $\mathfrak{W}$ be sets of cones with $\mathfrak{V} \subseteq \mathfrak{W}$. Show that $\mathbf{K4} \oplus \gamma(\mathfrak{z}, \mathfrak{W}) \subseteq \mathbf{K4} \oplus \gamma(\mathfrak{z}, \mathfrak{V})$.

Exercise 290. Obviously, a subframe logic is a cofinal subframe logic. Given $\gamma^\circ(\mathfrak{z}, \mathfrak{V})$ compute explicitly a finite set C of cofinal subframe axioms such that $\mathbf{K4} \oplus \gamma^\circ(\mathfrak{z}, \mathfrak{V}) = \mathbf{K4} \oplus C$.

Exercise 291. (Robert Bull [**34**].) Show that every extension of **S4.3** has the finite model property. *Hint.* Show that every extension is a cofinal subframe logic.

8.5. Embeddability Patterns and the Elementarity of Logics

Refutation patterns are second–order conditions; some of them cannot be reduced to first–order conditions. An example is the logic **G** defined by the exclusion of $\boxed{\circ}$ as a subframe. Moreover, it is generally not decidable whether a logic is first–order or not (see [**44**]). However, as have noted Kit Fine and Michael Zakharyaschev, for cofinal subframe logics there is a close connection between the geometric conditions expressed by the refutation patterns and the elementarity of the conditions they impose on frames. (See [**66**] and [**246**].) We approach this circle of themes by defining a different condition associated with a refutation pattern $\mathbb{P} = \langle \mathfrak{z}, \mathfrak{V} \rangle$, called *embedding pattern* or *embeddability pattern*.

Definition 8.5.1. *Let $\mathfrak{z}$ be a finite frame and $\mathfrak{V}$ a set of cones. Then the Kripke–frame $\mathfrak{f}$ satisfies the* **(cofinal) embedding pattern** $\mathbb{P} := \langle \mathfrak{z}, \mathfrak{V} \rangle$ *if there is no (cofinal) embedding of $\mathfrak{z}$ into $\mathfrak{f}$ such that no view in $\mathfrak{V}$ is external in the frame generated by the image of $\mathfrak{z}$ in $\mathfrak{f}$. We write $\epsilon(\mathfrak{z}, \mathfrak{V})$ for the property of satisfying the cofinal embedding pattern and $\epsilon^\circ(\mathfrak{z}, \mathfrak{V})$ for the property of satisfying the embedding pattern $\mathbb{P}$.*

Clearly, $\epsilon(\mathfrak{z}, \mathfrak{V})$ and $\epsilon^\circ(\mathfrak{z}, \mathfrak{V})$ are first–order. Moreover, they are describable by a restricted $\forall\exists$–sentence. Namely, let $\mathfrak{z} = \langle z, \lhd \rangle$ be given. Let $z = \{0, 1, \dots, n-1\}$ and 0 be the root. Now $\neg\epsilon(\mathfrak{z}, \mathfrak{V})$ is the condition

$$(\exists x_0)(\exists x_1 \rhd x_0)\dots(\exists x_n \rhd x_0)\left(\begin{array}{ll} & \bigwedge_{i<j<n} x_i \not\doteq x_j \\ \wedge & \bigwedge_{i \lhd j} x_i \lhd x_j \\ \wedge & \bigwedge_{i \not\lhd j} x_i \not\lhd x_j \\ \wedge & (\forall y \rhd x_0)(\bigvee_{i<n} y \lhd x_i \vee y \doteq x_i) \\ \wedge & (\forall y \rhd x_0)(\bigvee_{i<n} y \doteq x_i \\ & \quad \vee(\bigwedge_{V \in \mathfrak{V}}(\bigvee_{i \in V} y \not\lhd x_i \\ & \qquad \wedge \bigvee_{i \notin V} y \lhd x_i))) \end{array}\right)$$

Here, the first three formulae describe the fact that we have an embedding, the fourth that this embedding is cofinal and the last that it satisfies the closed domain condition (no view from $\mathfrak{V}$ is external). As we can see, parts of the conditions are positive; it is not clear that these conditions define at all a modal class of frames. However, notice the following.

Lemma 8.5.2. *$\epsilon^\circ(\mathfrak{z}, \varnothing)$ is equivalent to a positive universal $\mathcal{R}^f$–sentence.*

Proof. Let $f = \{i : i < n\}$. Let $A(n)$ be the set of atomic fomulae of the form $x_i \doteq x_j$, $i, j < n$ or $x_i \lhd x_j$, $i, j < n$. Let $C \subseteq A(n)$; then let $r(C)$ be the frame $\langle f_C, \lhd_C \rangle$ where f_C is the factorization of f by the equivalence relation generated by

$\langle i, j\rangle$, where $x_i \doteq x_j \in C$. Then put $[i] \lhd_C [j]$ iff there is a $g \in [i]$ and an $h \in [j]$ such that $x_g \lhd x_h$. We call C a **barrier** if for no $D \supsetneq C$, $r(D)$ is isomorphic to $\mathfrak{f}$. Let B be the set of barriers. Now let

$$\alpha(\mathfrak{f}) := (\forall x_0)(\forall x_1 \rhd x_0) \dots (\forall x_{n-1} \rhd x_0)(\bigvee_{C \in B} \bigwedge C)$$

Then $\mathfrak{g} \vDash \alpha(\mathfrak{f})$ iff $\mathfrak{f}$ is not embeddable into $\mathfrak{g}$. $\alpha(\mathfrak{f})$ is positive, universal and restricted. ∎

Similarly $\neg\epsilon^\circ(\mathfrak{z}, \mathfrak{V})$ is $\exists\forall$. Now, just as $\gamma(\mathfrak{z}, \mathfrak{V})$ need not be elementary, $\epsilon(\mathfrak{z}, \mathfrak{V})$ need not be modal. We will investigate situations in which the embeddability conditions are modal and situations where the refutation patterns are elementary. In both cases, we may either consider special sets of patterns or special classes of frames. Let us first consider the general question of reducing the refutation patterns to first–order conditions in special classes, namely noetherian frames.

Theorem 8.5.3. *In noetherian frames of width θ every (cofinal) refutation pattern is a conjunction of finitely many (cofinal) embedding patterns. Moreover, a (cofinal) embedding pattern corresponds to a $\forall\exists$–sentence.*

Proof. Let $\mathfrak{f}$ be a noetherian frame, and $\langle \mathfrak{z}, \mathfrak{V}\rangle$ be a refutation pattern. We may assume that no closed domain is of the form $\uparrow x$, where x is reflexive. In that case, if there is a cofinal subframe $\mathfrak{g}$ of $\mathfrak{f}$ and a p–morphism $\pi : \mathfrak{g} \twoheadrightarrow \mathfrak{z}$ such that the closed domains are not external views, then let $\mathfrak{k}$ be the frame consisting of all points maximal in π^{-1}. $\pi \restriction k : \mathfrak{k} \twoheadrightarrow \mathfrak{z}$ and no closed domain is an external view. Moreover, no cluster of $\mathfrak{k}$ is larger than the largest cluster of $\mathfrak{k}$. Now let $\mathfrak{z}$ have ℓ many points. Let the largest cluster have size c. Then $\mathfrak{k}$ has at most $\ell \cdot c \cdot \theta$ many points. Take the embedding pattern based on $\langle \mathfrak{k}, \mathfrak{W}\rangle$ where $\mathfrak{W}$ consists of all closed domains V such that $\pi[V] \in \mathfrak{V}$. Then $\mathfrak{f}$ realizes that embedding pattern. So, let B be the set of all embedding patterns $\langle \mathfrak{k}, \mathfrak{W}\rangle$ based on frames with $\leq \ell \cdot c \cdot \theta$ points such that there exists a $\pi : \mathfrak{k} \twoheadrightarrow \mathfrak{z}$ and $\mathfrak{W}$ consists of all closed domains V with $\pi[V] \in \mathfrak{V}$. Then $\mathfrak{f}$ realizes the refutation pattern $\mathbb{P}$ iff it realizes some embedding pattern of B. ∎

Another important case is when we have no closed domains. Then the refutation pattern defines a cofinal subframe logic. Even though the following arguments extend to cofinal subframe logics we focus on subframe logics. Take a finite frame $\mathfrak{f}$ with root z. Let $Z := C(z)$. Call a **cluster sequence** of $\mathfrak{f}$ a sequence $\langle C_i : i < k\rangle$ of clusters of $\mathfrak{f}$ such that $C_0 = Z$ and C_{i+1} is an immediate successor cluster of C_i. It is clear that the length of a cluster sequence is bounded by the depth of $\mathfrak{f}$. Let $u(\mathfrak{f})$ be the set of pairs $\langle \Sigma, x\rangle$ such that Σ is a cluster sequence and x a member of the last cluster of Σ. Put $\langle \Sigma, x\rangle \blacktriangleleft \langle \Gamma, y\rangle$ if either (a) $\Sigma = \Gamma$ and $x \lhd y$ or (b) Σ is a proper prefix of Γ. Now put $\mathfrak{u}(\mathfrak{f}) := \langle u(\mathfrak{f}), \blacktriangleleft\rangle$ and call it the **cluster unravelling**. The map $d : \langle \Sigma, x\rangle \mapsto x$ is a p–morphism. Now let $\mathfrak{d}$ be any frame such that there exist p–morphisms $\pi : \mathfrak{u}(\mathfrak{f}) \twoheadrightarrow \mathfrak{d}$ and $\rho : \mathfrak{d} \twoheadrightarrow \mathfrak{f}$. Then $\mathfrak{d}$ is called a **disentangling** of $\mathfrak{f}$. (Disentanglings are called **descendants** in [**66**].)

Lemma 8.5.4. *Let $\mathfrak{f}$ be a finite frame and $\mathfrak{g}$ a noetherian frame. Suppose that $\mathfrak{g}$ is subreducible to $\mathfrak{f}$. Then there is a disentangling $\mathfrak{d}$ of $\mathfrak{f}$ which is embddable into $\mathfrak{g}$.*

Proof. Assume there is a subframe $\mathfrak{h}$ of $\mathfrak{g}$ and a p–morphism $\pi : \mathfrak{h} \twoheadrightarrow \mathfrak{f}$. $\mathfrak{h}$ is also noetherian. It suffices to show that there is a disentangling of $\mathfrak{f}$ which is embeddable into $\mathfrak{h}$. We select a subset of h in the following way. We define a map $\tau : u(\mathfrak{f}) \to g$ by induction on the length of the cluster sequence. Let $\Sigma = \langle Z \rangle$. Since $\mathfrak{h}$ is noetherian, the set $U := \pi^{-1}[Z]$ contains a cluster W isomorphic to Z. Namely, let $w \in U$ be such that $w \lhd v$ and $v \in U$ implies $v \lhd w$. Then two cases arise. **Case 1.** Z is degenerate. Then U is a disjoint union of degenerate clusters, and we may pick any $w \in U$. **Case 2.** Z is nondegenerate. Then a $\lhd$–maximal cluster contains at least as many points as Z. So, we pick a subset of size $\sharp Z$ from such cluster. We may in fact pick such a subset W on which π is injective. Finally, put $\tau(\langle \Sigma, x \rangle) := y$, where $y \in W$ and $\pi(y) = x$. Now let τ be defined on all pairs $\langle \Sigma, x \rangle$ where Σ is a cluster sequence of length $< d$. Let $\Gamma := \langle \Sigma, C \rangle$. Let x be an element of the last cluster of Σ and $y \in C$. Put $w := \tau(\langle \Sigma, x \rangle)$. Select a cluster D from the set $\pi^{-1}[C] \cap \uparrow w$. Such a cluster exists since π is a p–morphism and $\mathfrak{h}$ is noetherian. Moreover, we let D be such that π is injective on D. Now put $\tau(\langle \Gamma, y \rangle) := v$ where $v \in D$ and $\pi(v) = y$. This defines τ. Let $\mathfrak{d}$ be the image of $\mathfrak{u}(\mathfrak{f})$ under τ. Then τ induces a map from $\mathfrak{u}(\mathfrak{f})$ to $\mathfrak{d}$. This is easily seen to be a p–morphism. Finally, the restriction of π to $\mathfrak{d}$ is a p–morphism as well. So, $\mathfrak{d}$ is a disentangling of $\mathfrak{f}$ and a subframe of $\mathfrak{g}$. ◻

It is clear that if $\mathfrak{g}$ is cofinally subreducible to $\mathfrak{f}$ then some disentangling of $\mathfrak{f}$ is cofinally embeddable into $\mathfrak{g}$. It follows

Lemma 8.5.5. *Let $\mathbb{P} = \langle \mathfrak{z}, \varnothing \rangle$ be a refutation pattern without closed domains. Then there exists a finite set E of embedding patterns without closed domains such that a noetherian frame realizes $\mathbb{P}$ iff it realizes some member of E.*

Theorem 8.5.6. *Every (cofinal) subframe logic is Δ–elementary in the class of noetherian frames.*

This start is promising. So we have to study what happens if we lift the condition that the frames be noetherian. Here we face a problem. Now it is the case that a frame is subreducible to a finite frame $\mathfrak{f}$ without there being a finite subframe $\mathfrak{d}$ which can be mapped p–morphically to $\mathfrak{f}$. This situation however arises exclusively with clusters. A case in point is the chain $\omega^< = \langle \omega, < \rangle$. This chain can be mapped onto $\boxed{\circ}$, but $\boxed{\circ}$ cannot be embedded in it; also there does not exist any finite subframe of $\omega^<$ which can be mapped p–morphically onto $\boxed{\circ}$. Another case is the chain $\omega^\leq = \langle \omega, \leq \rangle$. It can can be mapped onto any cluster $\mathbf{n}$. However, if $n > 1$, there exists no finite subframe which can be mapped p–morphically onto $\mathbf{n}$; moreover, $\mathbf{n}$ is not embeddable into $\omega^\leq$. These two examples play a pivotal role here. Let $k^< := \langle k, < \rangle$ and $k^\leq := \langle k, \leq \rangle$ for every $k \in \omega$. Let $\mathfrak{f}$ be a frame, and C a cluster. Then the subframe based on C is totally local, and so the replacement of C by any other frame $\mathfrak{g}$ is well–defined. $\mathfrak{g}$ is an **immediate variant** of $\mathfrak{g}$ if there exists a cluster C in $\mathfrak{f}$ such that either (a) C

is improper and $\mathfrak{g} = \mathfrak{f}[k^{<}/C]$ for some $k \in \omega$, or (b) C is proper and $\mathfrak{g} = \mathfrak{f}[k^{\leq}/C]$ for some $k \in \omega$. To summarize, immediate variants are obtained from frames by replacing a nondegenerate cluster by a finite irreflexive chain or replacing a proper cluster by a finite reflexive chain. A **variant** of $\mathfrak{f}$ is obtained by iterating the process of forming immediate variants. Now say that a set S of frames is **quasi–closed under variants** if for each $\mathfrak{f} \in S$ and each cluster C of $\mathfrak{f}$, some variant $\mathfrak{f}[k^{<}/C]$ ($\mathfrak{f}[k^{\leq}/C]$) is subreducible to a member of S. A set T has the **finite embedding property** if a frame $\mathfrak{f}$ belongs to T exactly when each finite subframe belongs to T. Say that a logic has the **finite embedding property** if its class of frames has the finite embedding property.

Theorem 8.5.7 (Fine). *Let S be a set of finite frames. Then the set $\mathsf{Krp}(\boldsymbol{K4}/_F S)$ has the finite embedding property iff S is quasi–closed under variants. If either condition holds, the set of frames not subreducible to a member of S equals the set of frames into which no disentangling of a member of S is embeddable.*

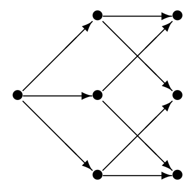
Proof. Suppose that S is not quasi–closed under variants. Then there exists a $\mathfrak{d} \in S$ such that no immediate variant is subreducible to a member of S. We may assume that $\mathfrak{f}$ is a frame of minimal size with this property. Take a cluster C of $\mathfrak{f}$ and consider the frame $\mathfrak{g} := \mathfrak{f}[\omega^{<}/C]$ if C is improper and $\mathfrak{g} := \mathfrak{f}[k^{\leq}/C]$ if C is proper. By assumption, no finite subframe is subreducible to any member of S. (Here we need the assumption that $\mathfrak{f}$ is of minimal size.) However, $\mathfrak{g}$ is subreducible to $\mathfrak{f}$. Hence, $\mathfrak{g} \notin \mathsf{Krp}(\mathsf{K4}/_F S)$ while every finite subframe $\mathfrak{h}$ of $\mathfrak{g}$ is contained in that class. So, $\mathsf{Krp}(\mathbf{K4}/_F S)$ does not have the finite embedding property. Now assume that S is quasi–closed under variants. Let $\mathfrak{g}$ be a frame which is subreducible to some $\mathfrak{f} \in S$. So, there exists a subframe $\mathfrak{h}$ and a p–morphism $\pi : \mathfrak{h} \twoheadrightarrow \mathfrak{f}$. Suppose there exists an improper cluster C of $\mathfrak{f}$ such that $\pi^{-1}[C]$ contains a cofinal subset of the form $\omega^{<}$. There is a variant $\mathfrak{z} := \mathfrak{f}[k^{<}/C]$ which is subreducible to a member of S. $\mathfrak{g}$ is subreducible to $\mathfrak{z}$, as can be shown. Hence $\mathfrak{g}$ fails the subframe condition for S. Similarly if C is proper. It follows that if $\mathfrak{g}$ is subreducible to $\mathfrak{f}$, some disentangling of $\mathfrak{f}$ is embeddable into $\mathfrak{g}$. ◻

Theorem 8.5.8 (Fine). *For a subframe logic $\Lambda \supseteq \boldsymbol{K4}$ the following are equivalent.*

1. $\mathsf{Krp}(\Lambda)$ *is definable by a set of universal, positive $\mathcal{R}^f$–sentences.*
2. Λ *is r–persistent.*
3. Λ *is Δ–elementary.*
4. Λ *is canonical.*
5. Λ *is compact.*
6. Λ *has the finite embedding property.*

Proof. The implication from (1.) to (2.) follows from Theorem 5.4.11. The implication from (2.) to (3.) is a consequence of Theorem 5.7.8. If (3.) holds, then (4.) follows from Theorem 5.7.11 and the fact that subframe logics are complete. A

Figure 8.4. The frame $\mathfrak{z}_3$



canonical logic is compact. This follows from Proposition 3.2.7. Now, suppose that (5.) holds. We establish (6.). Take an infinite frame $\mathfrak{f}$ such that every finite subframe satisfies Λ. Let $\Delta(\mathfrak{f})$ be the diagram. This diagram is finitely satisfiable, since each finite subset involves only a finite number of points. So, a finite subset describes a finite subframe, which is a frame for Λ. So the whole diagram is satisfiable. Thus, Λ has the finite embedding property. Now assume (6.). Then, let F be the class of rooted finite frames not being Λ–frames. Consider the set of sentences $\epsilon^\circ(\mathfrak{z}, \varnothing)$, $\mathfrak{z} \in F$. This is a set of universal $\mathcal{R}^f$–sentences. This set can be turned into a set of universal positive $\mathcal{R}^f$–sentences since F is inversely closed under contractions. ◻

Corollary 8.5.9 (Fine). *A subframe logic $\mathbf{K4}/_F S$ is canonical iff S is quasi–closed under variants.*

So, **S4** is canonical, since it is obtained by excluding $\boxed{\bullet}$, while **G** is not canonical, being obtained by excluding $\boxed{\circ}$. The latter set is not quasi–closed under variants; no variant of $\boxed{\circ}$ is subreducible to $\boxed{\circ}$. Likewise, **S5** is canonical while **Grz** is not. This concludes the discussion of subframe logics. Now in special circumstances, other refutation patterns yield elementary logics. A special case are logics of finite width. If we consider only noetherian frames, then elementarity is guaranteed. But we do not always need to assume that the frame is noetherian. For example if we consider splittings of irreflexive frames in logics of finite width. In this special case the assumption that the frame is noetherian can be dropped.

Theorem 8.5.10. *Each splitting condition by a finite irreflexive frame is elementary in the class of frames of width θ.*

In [**12**] it was claimed that this holds without assuming finite width. To see that the condition of finite width is essential look at the following frame. Let p be a prime number and $z_p := \{q\} \cup \{k^* : k \in p\} \cup \{k_* : k \in p\}$. We put

$$\lhd_p := \begin{cases} & \{\langle q, k^*\rangle : k < p\} \\ \cup & \{\langle q, k_*\rangle : k < p\} \\ \cup & \{\langle k^*, n_*\rangle : k, n < p, \text{ and } n = k \text{ or } n \equiv k+1 \pmod{p}\} \end{cases}$$

Finally, $\mathfrak{z}_p := \langle z_p, \lhd_p\rangle$. It is not hard to see that $\mathfrak{z}_p$ cannot be contracted onto $\mathfrak{z}_3$ unless $p = 3$. Let U be an ultrafilter over $\omega - \{0, 1, 2\}$ containing all cofinite sets.

Put $\mathfrak{z}_\infty := \prod_U \mathfrak{z}_{p_n}$, where p_n is the nth prime number. We show that $\mathfrak{z}_\infty$ can be mapped onto $\mathfrak{z}_3$ (in fact onto any $\mathfrak{z}_p$). Thus the condition 'does not contain a generated subframe which can be mapped onto $\mathfrak{z}_3$' is not elementary. $\mathfrak{z}_\infty$ is of depth 3, and every point of depth 1 sees exactly two points of depth 0, every point of depth 0 is seen by exactly two points of depth 1 (these properties are elementary, hence preserved by passing to an ultrapoduct). Let H be the set of points of depth 0 in $\mathfrak{z}_\infty$. Call a subset $C \subseteq H$ a **cycle** if it is closed under the operations

$$\overline{u} \mapsto \overline{u}_* + m := \langle (v_i)_* : v_i \equiv u_i + k \pmod{p_i} \rangle$$

for each $k \in \mathbb{Z}$. Pick from each cycle C a representative $r(C)$. Let Z be the set of all cycles. Now let

$$\begin{array}{rcl} H_0 & := & \bigcup_{C \in Z} \{r(C) + 3k : k \in \mathbb{Z}\} \\ H_1 & := & \bigcup_{C \in Z} \{r(C) + 3k + 1 : k \in \mathbb{Z}\} \\ H_2 & := & \bigcup_{C \in Z} \{r(C) + 3k + 2 : k \in \mathbb{Z}\} \end{array}$$

Let

$$\begin{array}{rcl} K_0 & := & \{\overline{u}^* : \overline{u}_* \in H_0\} \\ K_1 & := & \{\overline{u}^* : \overline{u}_* \in H_1\} \\ K_2 & := & \{\overline{u}^* : \overline{u}_* \in H_2\} \end{array}$$

Define $\pi : \mathfrak{z}_\infty \twoheadrightarrow \mathfrak{z}_3$ as follows. $\pi(\overline{q}) := q$, $\pi(x) = i^*$ iff $x \in K_i$ and $\pi(x) = i_*$ iff $x \in H_i$. We leave it to the reader to check that this is a p–morphism.

Notes on this section. Frank Wolter has shown in [**242**] that the minimal tense extension of a cofinal subframe logic has the finite model property iff it is Δ–elementary. Nevertheless, the minimal tense extension of any finitely axiomatizable cofinal subframe logic is decidable.

Exercise 292. Show that an extension of **G** is canonical iff it is of finite depth. Show that an extension of **Grz** is canonical iff it is of finite depth.

Exercise 293. $\mathbf{K4.1} = \mathbf{K4} \oplus \Box\Diamond p \to \Diamond\Box p$ is the cofinal subframe logic $\mathbf{K4}/_Z \boxed{\circ\circ}$. Show that it is not canonical but $\aleph_0$–canonical.

Exercise 294. (Continuing the previous exercise.) Say that a set of frames is **quasi–closed under nonfinal variants** if for each $\mathfrak{f} \in S$ and each nonfinal cluster C some variant $\mathfrak{f}[k^{<}/C]$ ($\mathfrak{f}[k^{\leq}/C]$) is cofinally subreducible to a member of S. Show that if S is quasi–closed under nonfinal variants, the cofinal subframe logic $\mathbf{K4}/_Z S$ is $\aleph_0$–canonical.

8.6. Logics of Finite Width I

Before we begin with logics of finite width proper, we will think somewhat more about how models can be made simple. A natural consequence will be that logics of finite width satisfy one of the central properties that allow to make models countable. Though we have in principle settled the question of characterizing the frames determined by an axiom, we will introduce here another line of thinking, which considers the question of how to make a model small without distorting it too much. As we have seen, there are ways to obtain very small models, based on frames whose cardinality depends only on φ. However, the price we have paid is that we do not know whether this operation ends in a model for the logic under consideration. We have seen that the logic must be a cofinal subframe logic if this is case generally. If not, we are however not without tools. Still, we can make the model well–behaved in certain respects. Some special cases are provided when we consider a local subframe $\mathfrak{g}$ of a Kripke-frame $\mathfrak{f}$. If $\mathfrak{g} \twoheadrightarrow \circ$, then we can contract $\mathfrak{g}$ in $\mathfrak{f}$, by locality, and so we can actually contract $\mathfrak{f}$ as soon as it contains no quasi–maximal points. For we know that if in a given model $\mathfrak{f}$ contains no quasi–maximal points, then all points realize the same external view; contracting them to a single point leaves the refutation pattern that is being realized intact. On the other hand, since it is a contraction, it preserves the fact that the frame is a frame for the logic.

PROPOSITION 8.6.1. *Let $\langle \mathfrak{F}, \beta, x\rangle \vDash \varphi$ and $\mathfrak{G}$ a local subframe containing no quasi–maximal points for φ, and $\sim$ a net on $\mathfrak{G}$. Let $\approx$ be the unique extension on $\mathfrak{F}$ and $\pi : \mathfrak{F} \to \mathfrak{K}$ the corresponding p–morphism. Then there exists a valuation γ and a y such that $\langle \mathfrak{K}, \gamma, y\rangle \vDash \varphi$.*

PROOF. Let $\langle \mathfrak{F}, \beta, x\rangle \vDash \varphi$ and $\sim$ a net on $\mathfrak{G}$, $\mathfrak{G}$ a local subframe of $\mathfrak{F}$ containing no quasi–maximal points. Denote by $\approx$ the extension of $\sim$. For $z \in g$ Let $[z] := \{y : z \approx y\}$. By assumption this is an internal set, and by the locality of $\mathfrak{G}$, all members of $[z]$ have the same view. For each $[z]$ let $r([z]) \in [z]$. Now put $\delta(y) := \beta(y)$ if $y \notin g$ and $\delta(y) := \beta(r([y]))$ if $y \in g$. By the Homogenization Theorem, the molecule of y in $\langle \mathfrak{F}, \beta\rangle$ is the same as the molecule of y in $\langle \mathfrak{F}, \delta\rangle$. It follows that $\langle \mathfrak{F}, \delta, x\rangle \vDash \varphi$. The p–morphism π corresponding to $\approx$ is admissible for δ. Denote by γ the valuation $\gamma(\pi(x)) := \beta(x)$. Then $\langle \mathfrak{K}, \gamma, \pi(x)\rangle \vDash \varphi$. ∎

For example, we may eliminate points x such that for some y, $x \vec{\lhd} y$ and $x \vec{\lhd} z$ implies $y \unlhd z$ for all z, on condition that x is not quasi–maximal, and we may eliminate all non quasi–maximal points in a proper cluster. Let $\mathfrak{F}$ be a frame, $N \subseteq f$ a subset (not necessarily internal). We say that $\mathfrak{G}$ results from **dropping** N if $\mathfrak{G} = \mathfrak{F} \cap (f - N)$. So, dropping points is another way of saying that we are passing to a subframe. In analogy to [**125**] we say that dropping N is **safe** if $\mathsf{Th}(\mathfrak{G}) \supseteq \mathsf{Th}(\mathfrak{F})$. We say that dropping N is **supersafe** if N can be safely dropped even when $\mathfrak{F}$ is a totally local subframe of another frame $\mathfrak{G}$. Evidently, if dropping a set is supersafe, it is also safe. The converse need not hold. A first case of (super)safe dropping is dropping points from a cluster. A more general case is dropping points which can be thought of as being collapsed by a p–morphism. For example, in the kite $\circ \oslash (\circ \oplus \circ) \oslash \circ$ we can drop one (or both) of the intermediate clusters.

Let us return to the problem of reducing a model of φ. Clearly, not all points are available for dropping; either because they are quasi–maximal or because they are required to keep the "structure" intact. But if a point can see another point of identical molecule, then it may be dropped. More generally, N is called **φ–covered** if every point in N has a successor outside of N with identical molecule.

LEMMA 8.6.2. *Let $\langle \mathfrak{F}, \beta\rangle$ be a model and N a φ–covered subset. Then every point in $f - N$ has the same molecule in $\langle \mathfrak{F} - N, \beta\rangle$ as it has in $\langle \mathfrak{F}, \beta\rangle$.*

PROOF. By Lemma 8.3.1. ∎

A note of caution is in order. In principle, the process of dropping can be iterated as many times as one likes. However, repeating it infinitely often requires careful argumentation. Just consider the frame $\langle \omega, <\rangle$. Dropping a single point is safe — the resulting frame is still isomorphic to $\langle \omega, <\rangle$. Hence, we might consider dropping 0, then 1, then 2, etc. In the limit, however, the whole structure is gone! There is an additional difference between dropping from a *frame* and dropping from a *model*. When we drop from a frame we are only interested in the fact that it is a frame for the logic. If on the other hand we want to drop from a model for φ we want to ensure in addition that after dropping we still have a model for φ. Now let in our previous case be $\varphi = p$ and $\beta(p)$. Each finite set is p–covered, but their union is not.

A second critical case is provided by $\mathfrak{b}_\omega := \circ \otimes \bigoplus_{i \in \aleph_0} \circ$. Consider the formula $\varphi = p \wedge \Diamond\neg p$, and let $\beta(p) := \{x_0\}$ where $\{x_0\}$ is the initial cluster. Then dropping a single point from the model, even a finite set, is permissible since every point is φ–covered. Again, if we drop all points, we no longer have a model for φ.

Evidently, if we want to avoid the dangers just described, we need to make sure that there are no infinite ascending chains and no infinite antichains.

DEFINITION 8.6.3. *A frame has the* **finite antichain property** *(**fap**) if every antichain is finite.*

DEFINITION 8.6.4 (Fine). *A frame is said to have the* **finite cover property** *(**fcp**) if for every set N there is a finite C such that $N \subseteq \downarrow C$.*

PROPOSITION 8.6.5. *If a frame is noetherian and has the finite antichain property it also has the finite cover property.*

PROOF. Now assume that $\mathfrak{f}$ is both noetherian and has the finite antichain property. Let $S \subseteq f$ be a set. Let $max\,S$ be the set of points without strict successors in S. We have $S \subseteq \downarrow max\,S$. For if $x = x_0$ is a point, and has a strict successor in S, then x_1 be such a successor. If x_1 has a strict successor in S, let x_2 be such a successor. By the ascending chain condition, this process must stop; and so we get a successor in S without a strict successor in S. $max\,S$ is a subframe which is of depth 1, thus a direct sum of clusters, $max\,S = \bigoplus_{i \in I} C_i$. Pick from each cluster C_j a representative x_j. Then $D := \{x_i : i \in I\}$ is an antichain, and so it must be finite. Moreover, $\downarrow D \supseteq maxS$ and so $\downarrow D \supseteq S$, by transitivity. ∎

The $\aleph_0$–kite $\circ \otimes (\bigoplus_{i \in \aleph_0} \circ) \otimes \circ$ is an example of a frame with the finite cover property but without the finite antichain property. It is quite important to have a geometrical intuition as to what points are eliminable from a frame. A useful case is this one.

Definition 8.6.6. *A cluster is* **slim** *if it has type $\varnothing$ or* $\mathbf{1}$*, that is, if it contains a single point. A frame is* **slim** *if every cluster is slim. A frame is* **almost slim** *if (1.) all clusters are finite and (2.) all but finitely many clusters are slim.*

Theorem 8.6.7. *Suppose that $\Lambda \supseteq$ **K4** is complete with respect to noetherian frames. Then it is complete with respect to almost slim noetherian frames.*

The proof is straightforward. Recall for that purpose the fact that in noetherian frames of bounded width all refutation patterns can be replaced by suitable embedding patterns. Consider a subset N of f such that every cofinal refutation pattern which can be realized on $\mathfrak{f}$ at all can be realized avoiding N. Then N can safely be dropped. For if φ fails at x in $\mathfrak{f}$, $x \notin N$, there is a cofinal embedding pattern (for φ) which can be realized in $\mathfrak{f}$. By assumption on N, the same pattern can be realized with no points in N. Then N may add some external views, but it is harmless to drop them. Now consider — for the converse – that φ fails at x in $\mathfrak{f} - N$. Then a cofinal embeddability pattern for φ can be realized in $\mathfrak{f} - N$. Now consider what happens when N is added. Then there is the possibility that some external views are actually added. If N is initial, that is, if $\mathfrak{f} - N$ is a generated subframe in $\mathfrak{f}$, then these external views are harmless. However, N need not be initial, and therefore a stronger notion is called for.

Definition 8.6.8. *Let $\mathfrak{f}$ be a frame. A set $N \subseteq f$* **avoids all configurations** *if for all cofinal subframes $\mathfrak{m}$ of $\mathfrak{f}$ there is an isomorphism $\iota : \mathfrak{m} \rightarrowtail \widehat{\mathfrak{m}}$, where $\widehat{\mathfrak{m}}$ is a cofinal subframe of $\mathfrak{f}$ such that $\widehat{\mathfrak{m}} \cap N = \varnothing$ and for any external $\mathfrak{m}$–view V which is realized in $\mathfrak{f}$ the $\widehat{\mathfrak{m}}$–view $\iota[V]$ is realized in $\mathfrak{f} - N$.*

Proposition 8.6.9. *Let $\mathfrak{f}$ be noetherian and let N avoid all configurations in $\mathfrak{f}$. Then dropping N is supersafe.*

Proof. Obviously, dropping a set which avoids all configurations is safe. So it is enough to show that it is safe when $\mathfrak{f}$ is embedded as a totally local subframe. So, let $\mathfrak{f}$ be a local subframe in $\mathfrak{g}$ and let $\mathfrak{m}$ be a cofinal subframe of $\mathfrak{g}$. Then $\mathfrak{n} := \mathfrak{m} \cap f$ is a subframe of $\mathfrak{f}$. Then there exists an isomorphic copy $\widehat{\mathfrak{n}}$ in $\mathfrak{f}$ such that all external $\mathfrak{n}$–views (for $\mathfrak{f}$) are realized for $\widehat{\mathfrak{n}}$ in $\mathfrak{f}$. Let $\widehat{\mathfrak{m}} := (\mathfrak{m} - \mathfrak{n}) \cup \widehat{\mathfrak{n}}$. We must ensure that $\widehat{\mathfrak{m}}$ is cofinal. This is evidently so if $\mathfrak{f}$ is not itself cofinal. For then every point in $\mathfrak{f}$ has a strict successor outside of $\mathfrak{f}$, so $\mathfrak{m} - \mathfrak{n}$ is cofinal, and thus $\widehat{\mathfrak{m}}$ is as well. However, if $\mathfrak{f}$ is cofinal, $\mathfrak{n}$ is cofinal in $\mathfrak{f}$ and so is $\widehat{\mathfrak{n}}$ by construction, and so $\widehat{\mathfrak{m}}$ is cofinal in $\mathfrak{g}$. Fix a bijection $\mathfrak{m} \mapsto \widehat{\mathfrak{m}}$ and extend it to a bijection $\mathfrak{n} \mapsto \widehat{\mathfrak{n}}$. Modulo this bijection, $\mathfrak{n}$–views are translated into $\widehat{\mathfrak{n}}$–views and back. Now let a view be realized in $\mathfrak{g}$, say by x. Two cases arise. First, $x \notin f$. Then the $\mathfrak{n}$–view of x is $\mathfrak{n}$ or $\varnothing$, depending on where x is situated with respect to $\mathfrak{f}$. If it is $= \mathfrak{n}$, then the $\widehat{\mathfrak{n}}$–view is $\widehat{\mathfrak{n}}$, and if it is $\varnothing$ then the

$\widehat{\mathfrak{n}}$–view is $\varnothing$ as well. Thus, the $\widehat{\mathfrak{m}}$–view and the $\mathfrak{m}$–view of x correspond. (Recall that $\mathfrak{f}$ is assumed to be totally local.) Next, assume that $x \in f$. Then by assumption on N, there is an $\widehat{x}$ realizing the $\widehat{\mathfrak{n}}$–view in $\mathfrak{f} - N$ corresponding to the $\mathfrak{n}$–view of x. Now, the $\mathfrak{m} - \mathfrak{n}$–view of x and the $\mathfrak{m} - \mathfrak{n}$–view of $\widehat{x}$ concide. For if $x \lhd y$ and $y \notin \mathfrak{f}$ then also $\widehat{x} \lhd y$, and conversely, by the fact that $\mathfrak{f}$ is totally local in $\mathfrak{g}$. So, the $\widehat{\mathfrak{m}}$–view of $\widehat{x}$ corresponds to the $\mathfrak{m}$–view of x. And that had to be shown. ◻

After all these theoretical results we will finally prove some useful, concrete results.

Theorem 8.6.10. *(1.) Let* $\circ \otimes \mathfrak{f}$ *be a* $\boldsymbol{S4}$*–frame of depth* $\omega + 1$*. If* $\mathfrak{f}$ *is of finite tightness then the initial cluster can be supersafely dropped. (2.) Let* $\bullet \otimes \mathfrak{f}$ *be a* $\boldsymbol{G}$*–frame of depth* $\omega + 1$*. If* $\mathfrak{f}$ *is of finite tightness then the initial cluster can supersafely be dropped.*

Proof. Show that in the cases given the cluster avoids all geometrical configurations. ◻

These are generic cases; one should not be mislead into easy generalizations, see the exercises. However, as a guideline, if we have a frame $\mathfrak{f}$ of depth ω which looks completely regular in shape, then in $\mathfrak{f} \otimes \mathfrak{f}$ the first occurrence of $\mathfrak{f}$ can be dropped. We will meet situations like this later on.

Theorem 8.6.11 (Fine). *Let* Λ *be a logic of finite width. Then the weak canonical frames are noetherian and have the finite cover property.*

Proof. Let Λ be of width θ. Recall the structure theory for finitely generated **K4** algebras. We have seen there how the points of layer α can be constructed on the basis that we have constructed the points of depth $< \alpha$. However, in that construction, α was finite. Here we will show that the procedure extends to all ordinals α. Let us assume then that we have constructed all points of depth $< \alpha$. We show that each point sees a point of minimal width in the set of points of depth $< \alpha$. For let $x = x_0$ be a point not of depth $< \alpha$. Suppose that there is a chain $\langle x_i : i \in \omega \rangle$ such that for each $i < \omega$ there is a w_i such that $x_i \lhd w_i$ but $x_{i+1} \ntriangleleft w_i$. We cannot have $w_{i+1} \lhd w_i$. Otherwise $x_{i+1} \lhd w_{i+1} \lhd w_i$ which was excluded. So the sequence $\langle w_i : i \in \omega \rangle$ is non–descending. Since each w_i is incomparable with at most θ points in the chain, there must be an ascending subchain. Contradiction. Consequently, there exists no infinite ascending chain $\langle x_i : i \in \omega \rangle$ with decreasing width, and we must have a successor for x of minimal width. Within the set of points of minimal width we choose again the set of points of minimal atomic span, and this is the desired set of points of depth α. Since the frame can be enumerated with the ordinals, every point will eventually be assigned a depth. So there are no ascending chains, and every point has a depth.

Now consider an arbitrary set N of points, not necessarily internal. For each point x in N there is a weak successor of minimal depth in N. The set of these points will be called *maxN*. *maxN* is a sum of clusters and there are at most θ clusters. Pick a representative from each cluster. This is a cover for N. Hence N has a cover of size $\leq \theta$. ◻

Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$ be a **K4**–frame. Call a point x **eliminable** if for all a such that $x \in a$ there exists a strong successor y such that $y \in a$. Call $\mathfrak{F}$ **separable** if there are no eliminable points.

Theorem 8.6.12 (Fine). *Every logic containing* ***K4*** *is complete with respect to separable frames.*

Proof. Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F} \rangle$ be a descriptive frame. Now drop from f all eliminable points. Denote the resulting set by f° and the resulting Kripke–frame by $\mathfrak{f}^\circ$. (f° is not an internal set.) We show that the algebra induced by $\mathbb{F}$ on $\mathfrak{f}^\circ$ is isomorphic to the original one, and so their logics coincide. First, the map $a \mapsto a \cap f^\circ$ is a boolean homomorphism. Next, we have to show that it respects $\blacklozenge$. So let $x \in f^\circ$ and $x \in \blacklozenge d$. Then x has a successor $y \in d$; y has a noneliminable weak successor $y^\gamma \in d$. For the set d contains an ascending chain $\langle x_i : i \in \alpha \rangle$ such that $x_{i+1} \ntriangleleft x_i$ which is cofinal in d. Suppose this chain is strictly ascending, that is, $x_i \vec{\triangleleft} x_{i+1}$ for all $i < \omega$. Consider the set U of all sets containing almost all points from this chain. Then U is contained in an ultrafilter $U^\star$. Then $U^\star \in a$, and moreover, $U^\star$ is final in d. Contradiction. Then $x \triangleleft y^\gamma$ and so $x \in \blacklozenge(d \cap f^\circ)$. The converse is easy. ◻

Corollary 8.6.13. *Let* Λ *be of finite width and* $\mathfrak{F}$ *a finitely generated, refined, separable* Λ*–frame. Then* $\mathfrak{F}$ *is atomic.*

Proof. Let x be a point. By separability there is a set d such that no strong successor of x is in d. **Case 1.** $x \ntriangleleft x$. Then put $e := d \cap \blacksquare - d$. This set contains x and is an antichain. Consequently, it is finite. By refinedness, $\{x\}$ is an internal set. **Case 2.** $x \triangleleft x$. Let $N := \{y : x \ntriangleleft y\}$ and let $C := maxN$, the set of points y such that $y \in N$ and if $y \vec{\triangleleft} z$ then $z \notin N$. For each $z \in C$ let c_z be a set such that $x \in \blacksquare - c_z$ but $x \in c_z$ (this exists by refinement and the fact that $x \notin C$) and b_z a set such that $z \in b_z$ but for no strong successor y, $y \in b_z$. Put

$$e := d \cap \bigcap \langle -b_z \cap \blacksquare - c_z : z \in C \rangle$$

Let $y \in e$. We claim that $x \triangleleft y \triangleleft x$. First, if $y \in N$ then there is a $z \in C$ such that $y \trianglelefteq z$. If $y = z$ then $y \in C$. Hence $y \notin -b_y$, so $y \notin e$; if $y \notin C$ then $y \triangleleft z$ for some z and so $y \notin \blacksquare - c_z$, from which $y \notin e$. Hence, e is disjoint with N. It follows that for $y \in e$, $x \ntriangleleft y$ cannot hold. So assume now that $x \triangleleft y$. If $y \ntriangleleft x$ then $y \notin d$ and so $y \notin e$. Hence also $y \triangleleft x$. It follows that e is a subset of the cluster of x. Since the algebra of sets is finitely generated, e is finite. The frame is refined, and so the set $\{x\}$ is easily shown to be internal. ◻

Theorem 8.6.14 (Fine). *Every logic of finite width is* $\aleph_0$*–canonical. Every logic of finite width is complete with respect to countable frames with the finite cover property, without ascending chains and of depth* $< \varepsilon$*, where* ε *is the smallest uncountable ordinal.*

Proof. Let $\mathfrak{F}$ be a weak canonical frame for Λ. Then $\mathfrak{F}$ is noetherian and has the finite cover property. It is also atomic. Take a refutation pattern $\langle \mathfrak{r}, \mathfrak{V} \rangle$. We can assume that there are no sets of the form $\uparrow x$ in $\mathfrak{B}$ for a reflexive x. Suppose that $\langle \mathfrak{r}, \mathfrak{B} \rangle$ can be realized in $\mathfrak{F}$. Then there exists a cofinal subframe $\mathfrak{m}$ and a p–morphism $\pi : \mathfrak{m} \twoheadrightarrow \mathfrak{r}$ satisfying the closed domain condition. Let $\mathfrak{n}$ consist of the sets $max\pi^{-1}(x)$ for all $x \in r$. $\pi \restriction n : \mathfrak{n} \twoheadrightarrow \mathfrak{r}$. The subframe $\mathfrak{n}$ is cofinal and satisfies the closed domain conditions for $\langle \mathfrak{r}, \mathfrak{B} \rangle$. Then there are only finitely many points in $\pi^{-1}(x)$. We have just seen that these sets are internal in a separable frame. Hence adding new sets does not change the logic. This shows that the logics are weakly canonical. Since there are only countably many formulae in the language (the restriction to finitely many generators plays no role here), there can be only countably many internal sets of the form $\{x\}$, hence only countably many points. The smallest uncountable ordinal is ε, so if a frame is countable, it is of depth $< \varepsilon$. ◻

So all logics of finite width have Kuznetsov–index $\leq \aleph_0$. To get a more fine–grained view of the matter, let us define the **ordinal Kuznetsov–index** $OKz(\Lambda)$ of a logic complete with respect to noetherian frames as follows. $OKz(\Lambda)$ is the supremum of all ordinal numbers α such that there is a formula $\varphi \notin \Lambda$ and a Λ–frame of depth α refuting φ but no Λ–frame of depth $< \alpha$ exists refuting φ. It is shown in [**125**] that for every $\alpha < \omega^2$ there is a logic Λ with $OKz(\Lambda) = \alpha$. This can be extended up to ω^ω. Whether the ordinal Kuznetsov–index can be greater than ω^ω is unknown. We conjecture that this is false for logics of finite width.

We may cash out a useful result from the previous proof. Call a set a an **interval** if it is of the form $[x, y] := \{z : x \trianglelefteq z \trianglelefteq y\}$, $(x] := \{y : y \trianglelefteq x\}$, $]x) := \{y : x \vec{\triangleleft} y\}$ or of the form $\{x\}$. Call a set **simple** if it is a finite union of intervals. It is a matter of direct verification that in a noetherian frame with the finite antichain property the simple sets are closed under all operations and are the least such set containing all singleton sets.

Theorem 8.6.15. *Let Λ be a logic of finite width. Then Λ is complete with respect to countable frames $\mathfrak{F}$ of finite width which are atomic, separable and such that the internal sets are the simple sets.*

Theorem 8.6.14 deserves special attention. We have shown that *all logics of finite width are complete*. Moreover, they are complete with respect to countable frames. We will state this explicitly once again, but we will be more specific about the structure of the frames for these logics. Take a subset of the form

$$s_\kappa = \{y : dp(y) = \omega\kappa + \beta_0 \text{ for some } \beta_0 < \omega\}$$

s_κ is the disjoint sum of at most θ many sets Γ_i which are $\oplus$–indecomposable. Such sets are called **galaxies**. We say a galaxy in s_κ has **depth** κ, and a point in that galaxy has **galactic** depth κ. The **local** depth of a point is just the unique $\beta_0 < \omega$ such that $dp(x) = \omega \cdot \kappa + \beta_0$. So, a point is determined both by its galactic depth and its local depth. A logic has **galactic finite model property** if it is complete with respect to frames of finite galactic depth. In case of finite width this means that there are only finitely many galaxies. Many notions are now generalized to galaxies, such as *being* *linear*. Moreover, we can in some sense consider the frame as a *frame of galaxies*. Namely, let $gal(\mathfrak{F})$ be the set of galaxies of $\mathfrak{f}$. Put $\Gamma \lhd_{\mho} \Delta$ if for every $x \in \Gamma$ there is a $y \in \Delta$ such that $x \lhd y$. Finally, let

$$\mho(\mathfrak{f}) := \langle gal(\mathfrak{F}), \lhd_{\mho} \rangle$$

$\mho(\mathfrak{F})$ is called the **frame of galaxies** of $\mathfrak{f}$. It is not necessarily a contractum of $\mathfrak{f}$.

Proposition 8.6.16. *Let $\mathfrak{f}$ be a noetherian frame. Then $\mho(\mathfrak{f})$ has fatness 1.*

Proof. Let $\Gamma \lhd_{\mho} \Delta \lhd_{\mho} \Gamma$. Let $x \in \Gamma$. Then there exists a $y \in \Delta$ such that $x \lhd y$. There exist a κ_0, κ_1 and β_0, β_1 such that the depth of x is $\omega\kappa_0 + \beta_0$ and the depth of y is $\omega\kappa_1 + \beta_1$. Furthermore, $\kappa_1 < \kappa_0$ or $\beta_1 \leq \beta_0$. From $\Delta \lhd_{\mho} \Gamma$ it follows that $\kappa_0 = \kappa_1$. Hence, Γ and Δ are not disconnected subsets of s_{κ_0}, and so they are identical. ◻

We put $\mho^2(\mathfrak{f}) := \mho(\mho(\mathfrak{f}))$. There is a different construction which runs as follows. Let a **hypergalaxy** of depth κ be a maximal $\oplus$–indecomposable subset of the set of points of depth $\omega^2\kappa + \omega\beta_1 + \beta_0$, for some β_0, β_1. For two hypergalaxies Γ_2 and Δ_2 put $\Gamma_2 \lhd_{\mho_2} \Delta_2$ if for all $x \in \Gamma_2$ and all $y \in \Delta_2$ we have $x \lhd y$. Now let $gal_2(\mathfrak{f})$ be the set of hypergalaxies of $\mathfrak{f}$ and

$$\mho_2(\mathfrak{F}) := \langle gal_2(\mathfrak{f}), \lhd_{\mho_2} \rangle$$

It seems plausible that $\mho^2(\mathfrak{F}) \cong \mho_2(\mathfrak{f})$. However, this is false. We leave a proof of that as an exercise and indicate only why this is in fact not to be expected. For note that for two hypergalaxies $\mathbb{G}$, $\mathbb{D}$ we have $\mathbb{G} \lhd_{\mho_2} \mathbb{D}$ iff for all $x \in \mathbb{G}$ and all $y \in \mathbb{D}$, $x \lhd y$. On the other hand, let $[\mathbb{G}]$ be the set of galaxies contained in $\mathbb{G}$, and $[\mathbb{D}]$ the set of galaxies contained in $\mathbb{D}$. In $\mho(\mho(\mathfrak{F}))$, $[\mathbb{G}] \lhd [\mathbb{D}]$ iff for all $\Gamma \in [\mathbb{G}]$ there exists a $\Delta \in [\mathbb{D}]$ such that for all $x \in \Gamma$ some $y \in \Delta$ exists such that $x \lhd y$. The latter is clearly a stronger condition than the previous one. However, there is an important case where the two coincide.

Proposition 8.6.17. *Let $\mathfrak{f}$ be a noetherian frame of finite tightness.*

1. *If Γ is not initial in $\mho(\mathfrak{f})$, it is infinite.*
2. *$\mho(\mathfrak{f})$ is galactically linear.*
3. *$\mho_2(\mathfrak{f}) \cong \mho^2(\mathfrak{f})$.*

Proof. Let Γ be noninitial. Then there exists an immediate predecessor Δ of Γ. Suppose that there does not exist an infinite Γ' such that $\Gamma \ntriangleleft_{\mho} \Gamma'$ and $\Gamma' \ntriangleleft_{\mho} \Gamma$. Then Δ, Γ are part of the same galaxy. Contradiction. Hence there exists an infinite immediate successor Γ' of Δ. Suppose Γ' is incomparable with Γ. Then the frame is not of finite tightness. So, Γ is the only immediate successor of Δ. Hence $\mho(\mathfrak{f})$ is galactically linear and of fatness 1, by Proposition 8.6.16. Moreover, take hypergalaxies $\mathbb{G}$ and $\mathbb{D}$. Let $\mathbb{G} \lhd_{\mho_2} \mathbb{D}$. Let Γ and Δ be galaxies and $\Gamma \subseteq \mathbb{G}$, $\Delta \subseteq \mathbb{D}$. Then $\Gamma \lhd_{\mho} \Delta$ or $\Delta \lhd_{\mho} \Gamma$, by galactic linearity. Suppose now that $\mathbb{G} \neq \mathbb{D}$. Then $\Gamma \vec{\lhd} \Delta$. Suppose $\mathbb{G} = \mathbb{D}$. Then for every $x \in \mathbb{G}$ there is a $y \in \mathbb{G}$ such that $x \lhd x$. Let Δ the galaxy of least depth in $\mathbb{G}$. Then $\Delta \lhd_{\mho} \Delta$. Furthermore, for every $\Gamma \subseteq \mathbb{G}$, $\Gamma \lhd_{\mho} \Delta$, by galactic linearity. This shows that $\mathbb{G} \lhd_{\mho^2} \mathbb{D}$. ◻

This approach can be generalized as follows. Let ξ be an ordinal number. For every ordinal κ put

$$s(\xi, \kappa) := \{x : dp(x) = \omega^{\xi} \cdot \kappa + \beta, \beta < \omega^{\xi}\}$$

Call a $\oplus$–indecomposable subset Γ of $s(\xi, \kappa)$ a ξ–**hypergalaxy** of depth κ. Thus, **galaxies** are **1–hypergalaxies** and hypergalaxies are **2–hypergalaxies**. We let $Gal_{\xi}(\mathfrak{f})$ be the set of ξ–hypergalaxies, and $\Gamma \lhd_{\xi} \Delta$ iff for all $x \in \Gamma$ and $y \in \Delta$ we have $x \lhd y$. Finally, $\mho_{\xi}(\mathfrak{f}) := \langle Gal_{\xi}(\mathfrak{f}), \lhd_{\xi}\rangle$. We mention also an interesting particular case, $\xi = 0$. It turns out that the 0–hypergalaxies are exactly the clusters, and $C \lhd_0 D$ iff $C \lhd D$. We call $\mho_0(\mathfrak{f})$ the **skeleton** of $\mathfrak{f}$.

Lemma 8.6.18. *Let $\mathfrak{f}$ be a noetherian frame. Then there is a p–morphism $\mathfrak{f} \twoheadrightarrow \mho_0(\mathfrak{f})$.*

The proof is very simple and is omitted. Notice that for $\xi > 0$ this need not hold.

There is no restriction on ξ. In principle, it is possible to define the notion of an ω–**hypergalaxy**. We believe, however, that logics of finite width are complete with respect to frames of depth $< \omega^{\omega}$. That is to say, we conjecture that the ordinal Kuznetsov–index of logics of finite width is $\leq \omega^{\omega}$. If that is so, the need of considering ω–hypergalaxies does not arise. At present, however, this is only speculation. We will show that the conjecture holds for extensions of **S4** of finite width and finite tightness, and we will see later that it also holds for all logics of finite width and finite tightness.

Theorem 8.6.19. *Every extension of **S4** of finite width and finite tightness has the galactic finite model property. Equivalently, it is complete with respect to frames of hypergalactic depth 1.*

Proof. Let $\mathfrak{f}$ be a frame for Λ, and let $\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi$. Consider the φ–extract. It is a cofinal subframe of finite depth hence finite. Hence only finitely many galaxies contain (quasi–)maximal points. The p–morphism contracting such a galaxy to a single reflexive point is locally admissible. So, we can reduce such galaxies. If $\mathfrak{f}$ had depth $\omega \cdot \beta_1 + \beta_0$, it now has depth β_1. If after that step we still have infinitely many galaxies left, i. e. if $k - 1 > 1$, we iterate the procedure, infinitely often if necessary. After completion we have reached a frame of finite galactic depth on which a model for φ is based. ◻

Exercise 295. Give an example of a frame which has the finite cover property but is not noetherian. *Hint.* There is a linear frame with this property.

Exercise 296. Show that all logics containing **S4** of tightness 1 have the finite model property.

Exercise 297. Show that all extensions of $\mathbf{S4}/k_2$ have the finite model property. *Hint.* Only the 0–slice may contain a proper antichain.

Exercise 298. Let $\mathfrak{f}$ be of depth ω. Show that when $\circ \otimes \mathfrak{f}$ is local but not totally local, $\circ$ may not be safely dropped.

Exercise 299. Show that in general for a frame $\mathfrak{f}$ of depth ω, $\circ$ in $\circ \otimes \mathfrak{f}$ cannot be safely dropped. Likewise for $\bullet \otimes \mathfrak{f}$.

Exercise 300. Show that in general, if $\mathfrak{f}$ is an **S4**–frame of depth ω, the initial cluster of $\mathbf{n} \otimes \mathfrak{f}$ may not be safely dropped if $n > 1$.

Exercise 301. Show that no nonempty set in a finite frame avoids all configurations. Show that, nevertheless, there is safe dropping from finite models.

8.7. Logics of Finite Width II

Consider a logic containing **K4.3**. We know that it is complete with respect to linear noetherian frames. Moreover, we know that we can assume these frames to be almost slim. Let us see how we can simplify the structure of frames even more. Let $\langle \mathfrak{f}, \beta \rangle$ be a Λ–model for φ, $\mathfrak{f}$ noetherian and linear. Consider a segment Γ without maximal points. Then $\mathfrak{f} \cong \mathfrak{f}_1 \otimes \Gamma \otimes \mathfrak{f}_2$. **Case 1.** $\Gamma \cong \Gamma' \otimes \mathbf{n} \otimes \Gamma''$. Then there is a p–morphism $\Gamma \twoheadrightarrow \circ \otimes \Gamma''$. By Proposition 8.6.1, there is a model for φ on the frame

$$\mathfrak{f}_1 \otimes \circ \otimes \Gamma'' \otimes \mathfrak{f}_2$$

Case 2. Γ contains no reflexive point. Then $\mathfrak{g} \cong \alpha^{op}$, α an ordinal. However, if $\alpha > \omega$ it can be shown that all points of depth $\geq \omega$ in α^{op} can be supersafely dropped. It follows from this consideration that logics containing **K4.3** are complete with respect to frames of the following form

$$\mathfrak{f}_k \otimes \omega^{op} \otimes \mathfrak{f}_{k-1} \otimes \omega^{op} \otimes \ldots \otimes \omega^{op} \otimes \mathfrak{f}_0$$

where each $\mathfrak{f}_i$ is finite and linear. It follows that every extension of **S4.3** has the finite model property.

Theorem 8.7.1 (Bull, Fine). *Every extension of $S4.3$ has the finite model property and is finitely axiomatizable. Hence there are only countably many such extensions, and all are decidable.*

Proof. We know already that all extensions have the finite model property. An extension is characterized by a set of canonical axioms $\gamma(\mathfrak{m}, \mathfrak{V})$, where $\mathfrak{m}$ is linear. Since all closed domains of the form $\uparrow x$, we can dispense with the closed domains entirely. For we have $\mathbf{S4.3} \oplus \gamma(\mathfrak{m}, \mathfrak{V}) = \mathbf{S4.3} \oplus \gamma(\mathfrak{m}, \varnothing)$. (Hence all extensions of **S4.3** are cofinal subframe logics, from which it follows once again that they have the finite model property.) We have to show that every extension is finitely axiomatizable. We $\mathfrak{m} \rightarrowtail \mathfrak{n}$ iff $\mathfrak{n} \twoheadrightarrow \mathfrak{m}$ iff $\mathfrak{m}$ is subreducible to $\mathfrak{n}$. Consider the partial order $\prec$ on chains of clusters defined by $\mathfrak{n} \prec \mathfrak{m}$ iff $\mathfrak{m}$ is a subframe of $\mathfrak{n}$. This ordering can be construed as chains–over–$\langle \omega, \geq \rangle$. The latter is a well–partial order (see [**135**]) and so is then the order $\lhd$. Now consider $\mathfrak{f} \otimes \mathbf{m} \sqsubseteq \mathfrak{g} \otimes \mathbf{n}$ iff $\mathfrak{g} \lhd \mathfrak{f}$ and $m \geq n$. Then this actually defines the order of *being–a–cofinal–subframe–of*, and thus the converse of *being–reducible–to*, the one we are interested in. Now this order is a product of two well partial orders, and hence itself a well partial order. So there are no infinite antichains, and hence every logic is finitely axiomatizable. ◻

This more or less finishes the case of **S4.3**. If we drop reflexivity, things get a bit more awkward. First of all, the frames $\circ \otimes \mathfrak{ch}(n)$ where $\mathfrak{ch}(n)$ is an irreflexive chain of length n, form an infinite antichain.

Proposition 8.7.2. *There are $2^{\aleph_0}$ many logics in $\mathcal{E}$**K4.3**.*

Call a frame **almost irreflexive** if all but finitely many clusters consist of a single irreflexive point.

Theorem 8.7.3. *Every extension of **K4.3** has the galactic finite model property. Moreover, it is complete with respect to finite chains of almost irreflexive galaxies.*

Corollary 8.7.4. ***G.3** has the finite model property. It is weakly canonical but not canonical. Every proper extension of **G.3** is tabular.*

Proof. The first statement actually follows from the fact that **G.3** is a subframe logic, but can be shown also by showing that every finite **G.3**–configuration can be realized on a finite chain. A proper extension must therefore have one of the finite chains not among its models. But then almost all of them are not among the models. The logic is then the logic of a finite chain, hence tabular. **G.3** is weakly canonical, being of finite width. But it is not canonical. Consider namely the following set

$$\{\Box(p_i. \to .\Diamond p_{i+1} \wedge (\neg p_i \wedge \neg\Diamond p_i)) : i \in \omega\}$$

Each finite subset is satisfiable on a finite frame, so the set is consistent. A Kripke–frame underlying a model for this set must have a strictly ascending chain, so cannot be a **G.3**–frame. ◻

Proposition 8.7.5. *The logic of the frame $\circ \otimes \omega^{op}$ is not finitely axiomatizable and fails to have the finite model property. $\mathsf{Th}(\circ \otimes \omega^{op})$ is the lower cover of **G.3** in $\mathcal{E}$**K4.3**. Hence every proper extension is finitely axiomatizable and has the finite model property.*

Proof. Add to **K4.3** the axiom $ft(\leq 1)$ and the subframe axiom saying that a reflexive cluster must be strictly initial, that is, not properly preceded by any other point. Call this logic **Ref**. Now add all splitting axioms for $\circ \otimes \mathfrak{ch}(n)$, $n \in \omega$. The resulting logic is not finitely axiomatizable, because this axiomatization is independent. Moreover, the only infinite frames for this logic are frames of the form α^{op} or $\circ \otimes \alpha^{op}$ with α an ordinal. Now, similar reasoning as in the Corollary 8.7.4 shows that for an infinite ordinal α we have $\mathsf{Th}(\circ \otimes \alpha^{op}) = \mathsf{Th}(\circ \otimes \omega^{op})$. Thus we have axiomatized the logic of $\mathsf{Th}(\circ \otimes \omega^{op})$. Any proper extension does not have this frame among its models, hence must be a logic of irreflexive chains. Consequently, every proper extension contains **G.3**. ◻

Finally, we turn to the question of decidability. Recall from Section 2.6 that a logic which is finitely axiomatizable is decidable if it is complete with respect to an enumerable set of effective algebras. So, we have to show that for finitely axiomatizable logics there exists such a set of algebras. We know that logics of finite width are complete with respect to frames $\langle \mathfrak{f}, \mathbb{F}\rangle$ where $\mathfrak{f}$ is a noetherian frame of finite width and finite fatness and $\mathbb{F}$ is the algebra of simple sets. For the purpose of the next definition, a computable function from $Z \to \wp(\omega)$, where Z is a possibly infinite set, is a function that yields a finite set for each $z \in Z$ and which can be computed.

Definition 8.7.6. *Let $\mathfrak{F} = \langle \mathfrak{f}, \mathbb{F}\rangle$ be a frame. Call $\mathfrak{F}$ **simple** if it is of finite width and finite fatness and $\mathbb{F}$ is the set of simple sets. Call $\mathfrak{F}$ **effective** if $f = \omega$, (A) '$x \lhd y$' is decidable for all $x, y \in \omega$, (B) '$]y) = \varnothing$' is decidable for all $y \in \omega$, and (C) there are computable functions $a, c, \ell : \omega \to \wp(\omega)$, and $d, p, q : \omega^2 \to \wp(\omega)$ such that (1.) $a(x)$ is an antichain with $x \ntrianglelefteq z$ for all $z \in a(x)$, such that $x \ntrianglelefteq y$ implies that there is a $z \in a(x)$ with $y \trianglelefteq z$ (2.) $\ell(x)$ is the cluster containing x, (3.) $c(x)$ is an antichain such that $y \vec{\lhd} x$ iff y has a weak successor in $c(x)$, (4.) $d(x, y)$ is an antichain such that $z \lhd x$ and $z \ntrianglelefteq y$ iff $u \trianglelefteq z$ for some $u \in d(x, y)$, (5.) $p(x, y)$ is an atichain such that for all z: $z \trianglelefteq w$ for some $w \in p(x, x')$ iff $z \trianglelefteq x$ and $z \trianglelefteq y$, and (6.) $q(x, y)$ is an atichain such that for all z: $w \trianglelefteq z$ for some $w \in q(x, x')$ iff $x \trianglelefteq z$ and $y \trianglelefteq z$.*

We remark here that this definition of effectiveness is only useful for noetherian frames of finite width and finite tightness. For other classes of frames, other definitions have to be found.

Theorem 8.7.7. *Let $\mathfrak{F}$ be simple. If $\mathfrak{F}$ is effective, the algebra $\mathfrak{F}_+$ is effective.*

Proof. First of all, $\mathbb{F}$ is countable and there is a computable bijection between the set of finite sets of intervals and ω. (Basically, an interval is a set $[x, y]$ hence a pair of natural numbers, a set $(x]$, a set $[x)$ or a set $\{x\}$. Hence a set in $\mathbb{F}$ can be represented by a sequence $\langle W, X, Y, Z\rangle$ where W is a finite sequence of pairs of natural numbers and X, Y and Z finite sequences of natural numbers. It is a standard fact that there is a computable bijection between ω and such quadruples.) We have to show how the operations can be computed. Union is clear. Next intersection. It is easy to see that since $[x, y] = [x) \cap (y]$, it is enough to show that we can compute the intersections of two open intervals. By definition of p and q we have

$$[x) \cap [y) = \bigcup_{w \in q(x,y)} [w)$$

and

$$(x] \cap (y] \ = \bigcup_{w \in p(x,y)} (w]$$

Now $\blacklozenge$. If $x \lhd x$ then $\blacklozenge\{x\} = (x]$. If $x \ntriangleleft x$ then $\blacklozenge\{x\} = \bigcup_{y \in c(x)} (y]$. $\blacklozenge(x] = \blacklozenge\{x\}$ and $\blacklozenge[y, x] = \blacklozenge\{x\}$. Therefore the case $\blacklozenge\,]x)$ remains. Suppose first that $]x)$ is empty (this is decidable). Then $\blacklozenge\,]x) = \varnothing$. So, suppose that $]x) \neq \varnothing$. Let $Y := \{y \in a(x) :]y) \neq \varnothing\}$. By assumption (A) and the computability of a, Y is computable. We have

$$\blacklozenge\,]x) = (x] \cup \,]x) \;\cup \bigcup_{y \in Y} (y]$$

For let $z \in \blacklozenge]x)$. Then there is a y such that $x \lhd y$ and $z \lhd y$. If $y = x$, $y \in (x]$. If $x \lhd z$ then $z \in]x)$. Assume therefore $x \ntriangleleft z$. Then there exists a $y \in a(x)$ such that $z \trianglelefteq a(x)$. Moreover, y must have a successor (and this must be a point in $]x)$, by definition of $a(x)$). Hence $y \in Y$. The converse inclusion is established similarly. Finally, we treat the complement. This is by far the most involved case. First the singletons. Put $L := \ell(x) - \{x\}$.

$$-\{x\} = \,]x) \cup \bigcup_{y \in L} \{y\} \cup \bigcup_{y \in a(x)} (y]$$

For if $y \neq x$ then either (a) $x \lhd y \lhd x$, or (b) $x \lhd y \ntriangleleft x$ or (c) $x \ntriangleleft y$. In case (a), $y \in L$, in case (b) $y \in]x)$. In case (c) there is a $z \in a(x)$ such that $y \trianglelefteq z$. So $y \in (z]$. Since x is not contained in the right hand side, equality is shown. Next the sets $(x]$. Put $D := \{d(y, x) : y \in a(x)\}$ and $A := a(x) - \{x\}$. It is checked that

$$-(x] = \bigcup_{z \in D, y \in A} [z, y] \cup \,]x) \cup \bigcup_{y \in A}]y)$$

The sets $-]x)$ are computed as follows.

$$-]x) = \ell(x) \cup \bigcup_{y \in a(x)} (z]$$

For if it is not the case that $x \vec{\lhd} y$ then either (a) $x \lhd y \lhd x$ or (b) $x \ntriangleleft y$. In case (a) $y \in \ell(x)$ in case (b) $y \in (z]$ for some $z \in a(x)$. Finally the sets $-[x, y]$. $u \in -[x, y]$ iff either $x \ntriangleleft u$ or $u \ntriangleleft y$ iff $u \in -(y]$ or $u \in -[x)$. We can compute $-(y]$, so we need to know how to compute $-[x)$. It is easily seen that

$$-[x) = \bigcup_{y \in a(x)} (z]$$

Finally, we want to show that for two unions of intervals b, c, it is decidable whether '$b = c$'. This is equivalent to the decidability of the emptiness of a given union of intervals. This in turn is equivalent to the problem to decide whether an interval is empty. (a) $(x]$ is never empty. (b) $\{x\}$ is never empty. (c) $[x, y]$ is empty iff $x \neq y$ and $x \ntriangleleft y$. This is decidable by (A). (d) Whether $]x)$ is empty is decidable by (B). ◻

Definition 8.7.8. *Let $\mathfrak{F} = \langle \omega, \lhd, \mathbb{F} \rangle$ be simple. $\mathfrak{F}$ is called **supereffective** if (a) it is effective and (b) there is a computable function j from the set of embedding patterns into $\wp(\omega)$ such that for every embedding pattern $\epsilon = \epsilon(\mathfrak{r}, \mathfrak{B})$ the set $j(\epsilon)$ is a finite set of points such that whenever ϵ is realizable in $\mathfrak{F}$ there is an embedding $p : r \to j(\epsilon)$ realizing ϵ in $\mathfrak{F}$.*

Theorem 8.7.9. *Suppose that $\mathfrak{F}$ is supereffective. Then $\mathsf{Th}\,\mathfrak{F}$ is decidable.*

Proof. It is enough to show that for every embedding pattern it is decidable whether or not it is realizable. Take $\epsilon = \epsilon(\mathfrak{r}, \mathfrak{V})$. Then whether or not ϵ is realizable in $\mathfrak{F}$ can be checked by trying all embeddings $p : r \to j(\epsilon)$. There are only finitely many of them. Hence it is enough to show that for every embedding p it is decidable whether p satisfies all closed domain conditions. So take a closed domain $\mathfrak{v} \in \mathfrak{V}$. We need to decide whether $\mathfrak{v}$ is an external view, that is, whether there is a point x not in $p[r]$ such that $p[\mathfrak{v}]$ is the set of points $p(y)$ seen by x. For that we need to check whether the set b is empty, where

$$b := -p[r] \cap \bigcap_{y \in \mathfrak{v}} \blacklozenge\{p(y)\} \cap \bigcap_{y \notin \mathfrak{v}} -\blacklozenge\{p(y)\}$$

b can be computed, since the frame is effective. Whether b is empty is decidable. ◻

Theorem 8.7.10. *Let Λ be a logic of finite width. Suppose that Λ is finitely axiomatizable and complete with respect to a recursively enumerable class of frames which are supereffective. Then Λ is decidable.*

Proof. Let $\mathfrak{C}$ be a recursively enumerable class of supereffective frames such that Λ is $\mathfrak{C}$–complete. Since Λ is finitely axiomatizable it is recursively enumerable. Let $\Lambda = \mathbf{K4} \oplus \varphi$. It is decidable for $\mathfrak{F} \in \mathfrak{C}$ whether $\mathfrak{F} \vDash \varphi$. Hence the class of $\mathfrak{C}$–frames for Λ is recursively enumerable. Let it be $\mathfrak{D}$. Hence Λ is complete with respect to a recursively enumerable class of effective algebras, and so co–recursively enumerable. ◻

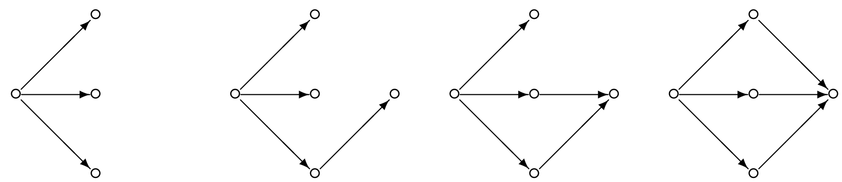
K4.3 is complete with respect to finite chains of the form

$$\mathfrak{g}_k \oslash \omega^{op} \oslash \mathfrak{g}_{k-1} \oslash \omega^{op} \oslash \ldots \oslash \omega^{op} \oslash \mathfrak{g}_0$$

where each $\mathfrak{g}_i$ is finite. It is not hard to see that these frames are supereffective. However, this is not all we have to show. For we need to be able to determine whether or not a given Kripke–frame $\mathfrak{f}$ is a frame for a logic. It remains to show that we can decide for such frames and a given refutation pattern whether it is satisfiable in the frame. This is not hard to see. For let $\gamma(\mathfrak{r}, \mathfrak{V})$ be given; let $\sharp\mathfrak{r} = k$. For cofinal embeddability, we need to check only final segements of ω of depth $\leq k$. So, $\langle \mathfrak{r}, \mathfrak{V} \rangle$ is satisfiable in this frame iff it is satisfiable in a finite subframe which can be computed from the original frame and the refutation pattern. The latter problem is now clearly decidable because we have to check only finitely many cases. Now given an extension of **K4.3** by means of finitely many canonical formulae, we can enumerate the frames of the above form, because we can decide whether the refutation patterns are satisfiable. Now let a (canonical) formula φ be given. At the same time that we are checking these frames for whether they satisfy the canonical axioms for the logic we can also check whether or not φ can be refuted. Again, this is decidable.

Theorem 8.7.11 (Alekseev & Zakharyaschev). *Every finitely axiomatizable extension of* ***K4.3*** *is decidable.*

FIGURE 8.5.



This theorem is actually the first general theorem on decidability of modal logics without assuming the finite model property.

Logics of width 2 play a significant role in the lattice of extensions of **S4**. We will concentrate on such extensions here. We begin by showing that $\mathbf{S4}.I_2 = \mathbf{S4}.wd(\leq 2)$ has an exceptional status with respect to logics of finite width in that it has a splitting representation.

THEOREM 8.7.12. *$\mathbf{S4}.wd(\leq 2)$ is a finite splitting of $\mathbf{S4}$ and hence of $\mathbf{K4}$.*

PROOF. Consider the frames in Figure 8.5. We will show that there is a finite set R of finite frames of width 3 such that for all refined frames $\mathfrak{F}$ for **S4** of width 2 there exists an $\mathfrak{r} \in R$ such that some generated subframe $\mathfrak{g}$ of $\mathfrak{f}$ is contractible to $\mathfrak{r}$. Suppose that $\mathfrak{b}_3 = \circ \oslash (\circ \oplus \circ \oplus \circ)$ is a subframe of a refined **S4**–frame $\mathfrak{F}$. Then we have four points x, y_0, y_1 and y_3 such that x sees y_i for all $i < 3$ and $y_i \lhd y_j$ iff $i = j$. There are sets b_i, $i < 3$, such that $y_i \in b_j$ iff $j = i$. Put $Z := -\blacklozenge b_0 \cap -\blacklozenge b_1 \cap -\blacklozenge b_2$. If Z is empty, the first frame is cofinally embeddable. Now suppose that Z is not empty. Z is a definable subset and the map contracting Z to a single point is a p–morphism; for Z is successor closed. Thus without loss of generality we may assume $Z = \{z\}$. Let $B_0 := \blacklozenge b_0 \cap -\blacklozenge b_1 \cap -\blacklozenge b_2$, $B_1 := -\blacklozenge b_0 \cap \blacklozenge b_1 \cap -\blacklozenge b_2$ and $B_2 := -\blacklozenge b_0 \cap -\blacklozenge b_1 \cap \blacklozenge b_2$. Then $y_i \in B_i$. The map sending B_i to a single point for each $i < 3$ is a p–morphism. Hence we may also assume that $B_i = \{y_i\}$. Now four cases arise. **Case 1.** z is incomparable with all y_i. Collapse z and y_2 into a single point. This is a p–morphism and we now have an antichain of size 3 in which each point is of depth. (First picture in Figure 8.5.) **Case 2.** z is comparable with exactly one of the y_i. Then without loss of generality $y_2 \vec{\lhd} z$. (Second picture in Figure 8.5.) **Case 3.** z is comparable with exactly two of the y_i. Then without loss of generality $y_1 \vec{\lhd} z$ and $y_2 \vec{\lhd} z$. (Third picture in Figure 8.5.) **Case 4.** All three y_i are comparable with z. Then $y_i \vec{\lhd} z$ for all $i < 3$. (Last picture of Figure 8.5.)

The set R is produced as follows. Take a frame $\mathfrak{m}$ from Figure 8.5. It consists of an antichain $Y := \{y_0, y_1, y_2\}$, a root x and (with one exception) also a point z of depth 0. If $i < j < 3$ then let q_{ij} be a new point. Let $\mathfrak{m}_0 := m \cup \{q_{01}, q_{02}, q_{12}\}$ and let $x \vec{\lhd} q_{ij}$ and $q_{ij} \lhd y_k$ iff $k = i$ or $k = j$. Then $\mathfrak{m}$ is a subframe of $\mathfrak{m}_0$. There exist 5

FIGURE 8.6.

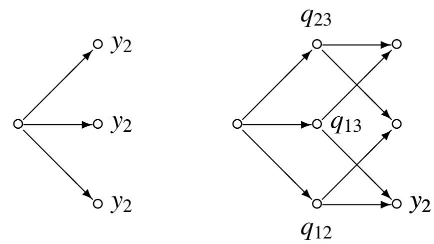
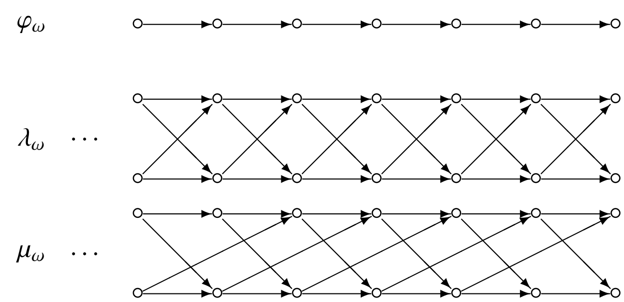


FIGURE 8.7.



subframes which contain $\mathfrak{m}$ and which are contained in $\mathfrak{m}_0$. R is the set of all such frames. (See Figure 8.6.)

We show that $\mathfrak{F}$ contains a generated subframe which can be mapped onto some member of $\mathfrak{F}$. We may assume that x generates $\mathfrak{F}$ (otherwise take the subframe generated by x). Let $i < j < 3$. Choose k such that $\{1, 2, 3\} = \{i, j, k\}$. Define sets $C_{ij} := \blacklozenge b_k \cap -\blacklozenge b_j \cap -\blacklozenge b_i$ and $X := \blacklozenge b_0 \cap \blacklozenge b_1 \cap \blacklozenge b_2$. The sets C_{ij} (if nonempty) can be mapped onto a single point. After that X can be mapped onto a single point. After this process we have a frame of R. ◻

There are a number of important frames which will play a fundamental role in the theory of finite width. They exhibit a certain regular pattern connected with their tightness. The simplest of them are shown in Figure 8.7.

The first of the three galaxies is called *photonic*, the second *leptonic* and the third *mesonic*. Their indecomposable generated subframes are the elementary particles from which the galaxies are composed. So, there is only one *photon*, namely $\circ$, and there are two leptons, $\circ$ and $\circ \oplus \circ$. There is an infinite series of mesons. So the mesonic galaxy is itself a meson. It is immediately checked that φ_ω is of tightness 0, λ_ω of tightness 1 and μ_ω of tightness 2. In a similar fashion we can create 'heavier' particles. It is perhaps intuitively clear how we will proceed, but let us introduce some notions that will help us with more complex structures as well. Let $\mathfrak{f}$ be a noetherian Kripke–frame of width 2 and depth β. Then there exist sequences $\mathfrak{c} = \langle c_\alpha : \alpha < \beta\rangle$ and $\mathfrak{d} = \langle d_\alpha : \alpha < \beta\rangle$ such that (1.) c_α and d_α are clusters of $\mathfrak{f}$ of depth α, (2.) $c_\alpha \lhd c_{\alpha'}$ and $d_\alpha \lhd d_{\alpha'}$ for all $\alpha > \alpha'$, and (3.) every point is contained in some c_α or d_α. Notice that it is not required that c_α and d_α are distinct. We call $\mathfrak{c}$ and $\mathfrak{d}$ the two **spines** of $\mathfrak{f}$. The division of $\mathfrak{f}$ into spines is arbitrary but the results are independent of it.

Let us restrict our attention to a single galaxy for the moment. Such a galaxy is fixed in its structure by two things: (1.) the cardinality of the clusters c_α and d_α, (2a.) the maximal local depth of a cluster in the $\mathfrak{d}$–chain seen by c_α, (2b.) the maximal local depth of a cluster in the $\mathfrak{c}$–chain seen by d_α. Let us call the **index** of $x \in c_\alpha$ (or of c_α itself for that matter) the order type of $\uparrow d_\alpha - \uparrow x$ and the index of $y \in d_\alpha$ the inverse order type of $\uparrow c_\alpha - \uparrow y$, which is the same as its depth as a frame. This is therefore an ordinal number. This cumbersome definition takes care of the case where there is no finite index. (This can happen! Think of two parallel chains of galaxies.) The proof of the next proposition is left as an exercise.

Proposition 8.7.13. *The following holds.*

1. $ind(x) \geq 0$.
2. $ind(x) \leq dp(x)$.
3. *If y immediately precedes x then* $ind(x) \leq ind(y) \leq ind(x) + 1$.
4. *If y immediately precedes x and* $ind(y) = ind(x)+1$ *then y is nonbranching.*

Lemma 8.7.14. *Extensions of $\mathbf{S4}$ of width 2 are complete with respect to noetherian frames of bounded index.*

Proof. Consider a formula φ rejected on a (noetherian) frame. We can safely drop a nonbranching point if it is nonmaximal. Thus, as there are at most $\sharp sf(\varphi)$ maximal points in a single spine, the index of a point is bounded by $\sharp sf(\varphi)$. ◻

To have a bounded index is in this case the same as being of bounded tightness. It is now easily deduced that extensions of **S4** of width 2 are galactically linear. Moreover, we know that frames can be assumed to be almost thin. Call a frame $\mathfrak{h}$ (k, ℓ)–**hadronic** if there is a partition into two spines such that every point of the first spine has index k, and every point of the second spine has index ℓ. Call a frame $\mathfrak{h}$ almost (k, ℓ)–**hadronic** if there is a p–morphism $\pi : \mathfrak{h} \twoheadrightarrow \mathfrak{k} \circledcirc \circ$ such that $card(\pi^{-1}(x)) = 1$ for almost all x, and $\mathfrak{k}$ is (k, ℓ)–hadronic.

Theorem 8.7.15. *The set $\mathfrak{H}$ of reflexive frames of width 2 and finite galactic depth in which all galaxies are almost hadronic is recursively enumerable. Moreover, all members of $\mathfrak{H}$ are supereffective.*

Proof. We show that the set of almost hadronic frames is enumerable and that these frames are supereffective. The general case is rather cumbersome and not revealing. A slice is uniquely characterized by a quadruple of natural numbers $\tau = \langle i, a, j, b\rangle$, where i is the index of a point of the first spine, a the size of the cluster of that point, j the index of a point of the second spine, and b the size of its cluster. Hence a reflexive frame of width 2 is uniquely characterized by a sequence $\langle \tau_n : n \in \alpha\rangle$, $\tau_n = \langle i_n, a_n, j_n, b_n\rangle$, where $\alpha < \omega + 1$. A frame is almost hadronic if there is a $n_0 < \alpha$, such that for all $n > n_0$, $\tau_{n_0} = \tau_n = \langle k, 1, \ell, 1\rangle$. (Or, by switching the spines, almost all types are $= \langle \ell, 1, k, 1\rangle$.) Thus we need three things: k, ℓ, n_0 and the sequence $T = \langle \tau_i : i < n_0\rangle$. This is a finite set. What needs to be shown is that given these four parameters we can compute the relations and operations of the frame. (This shows that the frame is effective.) For that, points can be represented as triples $\langle i, n, p\rangle$ where $i = 1$ or $i = 2$; $n < \omega$; $p < a_n$ if $i = 1$ and $p < b_n$ otherwise. For any triple it can be decided whether it is a member of the frame $\langle k, \ell, n_0, T\rangle$. Next it can be decided whether $x := \langle i, n, p\rangle \lhd y := \langle i', n', p'\rangle$ as follows. If $i = i'$, $x \lhd y$ iff $n \geq n'$. If $i = 1$ and $i' = 2$ then $x \lhd y$ iff $n' \leq n - i_n$, where i_n is given by τ_n for $n < n_0$ and by k for $n \geq n_0$. If $i = 2$ and $i' = 1$ then $x \lhd y$ iff $n' \leq n - j_n$. Further, whether $]x)$ is empty is equivalent to $n = 0$ and so decidable. If $x = \langle 1, n, p\rangle$ then $c(x) = \{\langle 1, n, q\rangle : q \leq a_n\}$; if $x = \langle 2, n, p\rangle$ then $c(x) = \{\langle 2, n, q\rangle : q \leq b_n\}$. $a(x)$ can be computed as follows. Let $x = \langle 1, n, p\rangle$. If $i_n = 0$ then $a(x) = \{x\}$ otherwise $a(x) = \{x, \langle 2, n - j_n + 1, 0\rangle\}$ (if such a point exists) and $a(x) = \{x\}$ otherwise. Similarly for $x = \langle 2, n, p\rangle$. Similarly the existence of the other functions is proved. This shows that the frames are all effective. Now we show that they are supereffective. Let $\epsilon(\mathfrak{r}, \mathfrak{V})$ be an embedding pattern and let it be realizable in $\mathfrak{F}$. We show that it is realizable in the set of all points of depth $\leq n_0 + k \cdot \sharp r$ if $k \geq \ell$ and $\leq n_0 + \ell \cdot \sharp r$ otherwise. Without loss of generality we assume that $k \geq \ell$. For let $q : r \to f$ be an embedding. Take W the set of all $x \in r$ such that $q(r)$ is of depth $> n_0$ and for all y such that $x \vec{\lhd} y$, the depth of $q(y)$ is $\leq q(x) - k$. Let x be a member of W. Then the cluster of x is degenerate. Then let q' be defined by $q'(y) := q(y)$ for $y \notin W$ and $q'(x) := \langle i, n-1, 0\rangle$ for $x = \langle i, n, p\rangle \in W$. We shall show that q' is an embedding realizing $\epsilon(\mathfrak{r}, \mathfrak{V})$. Suppose that $x \lhd y$ and $x, y \in W$. Then $q(x) \lhd q(y)$ depends only on the difference between the depths of $q(x)$ and $q(y)$ (and the spine of them) and so $q'(x) \lhd q'(y)$ iff $q(x) \lhd q(y)$. Or $x, y \notin W$ and $q(x) \lhd q(y)$ iff $q'(x) \lhd q'(y)$, since $q'(x) = q(x)$ and $q'(y) = q(y)$. Or $x \in W$ and $y \notin W$. Then $q(x) \lhd q(y)$, since the depth of $q(x)$ exceeds that of $q(y)$ by more than k. Then the difference between $q'(x)$ and $q'(y)$ is $\geq k$ and so $q'(x) \lhd q'(y)$. Similarly it is shown that an external view is realized by q iff it is realized by q'. Consequently, as long as $W \neq \varnothing$ this operation is applicable and reduces the maximum depths of points $q(x)$, $x \in r$, by 1. The procedure ends when $W = \varnothing$. Then if x immediately precedes y, the depth of $q(x)$ minus the depth of $q(y)$ is at most k. ◻

Theorem 8.7.16. *Extensions of $\mathbf{S4}$ of width 2 are complete with respect to frames with almost hadronic galaxies. Moreover, each finitely axiomatizable extension of $\mathbf{S4}$ of width 2 is decidable.*

Proof. Let Λ be of width 2. Since Λ is galactically linear, the index of a point is always finite. We know moreover that it can be bounded. The sequences $\langle ind(x_\alpha) : x \in c_\alpha\rangle$ and $\langle ind(x_\alpha) : x_\alpha \in d_\alpha\rangle$ are bounded by a number k. Moreover, in a galaxy there are for any model only finitely many maximal points, and so by Proposition 8.7.13 we can assume that these sequences are almost non–increasing. Hence, they must be stationary from some depth $\alpha^\star$ onwards. Put $k^\star := ind(c_{\alpha^\star})$ and $\ell^\star := ind(d_{\alpha^\star})$. Then the galaxy is almost $(k^\star, \ell^\star)$–hadronic. **S4** has the galactic finite model property by Theorem 8.6.19. Thus we can assume models to be finite sequences of almost hadronic frames. The theorem now follows from the previous theorem. ◻

Extensions of **S4** of finite width in general enjoy a rather nice finiteness property, namely that they are complete with respect to frames with finitely many segments. Namely, take any model for φ based on a frame. Each segment which does not contain a maximal point for φ can be reduced p–morphically to a single point. Thus, in analogy to the case of fatness, we can already assume that almost all segments are one–membered. Now, take such a one–point segment $\{x\}$. If it is of depth $\omega \cdot \lambda$, then it can be dropped. Otherwise it precedes directly another segment. In almost all cases this segment is of the form $\{y\}$, one–membered. In that case, we can collapse x into y, reaching a further reduction. It is not hard to show that this leaves us with at most twice as many segments as there are maximal points.

Theorem 8.7.17. *All extensions of S4 of finite width are complete with respect to frames with finitely many indecomposable segments.*

Corollary 8.7.18. *All extensions of S4 of tightness 2 have the finite model property.*

Proof. Indecomposable segments consist of points of same depth. So, the logics are complete with respect to frames of finite depth. Hence they have the finite model property. ◻

We can push this result a little bit further. Consider a frame of width 2 and tightness 3. Then the points of infinite depth are all eliminable because any embedding pattern using these points can be avoided.

Theorem 8.7.19. *All logics containing S4 of width 2 and tightness 3 have the finite model property.*

Exercise 302. A logic is called **dense** if it contains the axiom $\Diamond p \to \Diamond\Diamond p$. Show that this corresponds with the refutation pattern $\gamma^\circ(\varnothing_1 \otimes \varnothing_0, \{0\})$. Show that all logics of dense linear frames have the finite model property.

Exercise 303. Show that there are countably many logics of dense linear orders, each finitely axiomatizable.

Exercise 304. (Segerberg [**194**].) Show that all quasi–normal extensions of **S4.3** are normal.

Exercise 305. Show that all extensions of **S4.3** of the form $\mathsf{Th}\,\mathfrak{f}$ for a single frame are Halldén–complete.

Exercise 306. (Spaan [**202**].) Show that for every consistent logic containing **S4.3** satisfiability of a formula is NP–complete. *Hint.* NP–hardness is clear. Show that given φ, there is a model of size at most $|\varphi| + 1$.

Exercise 307. Show Proposition 8.7.13

Exercise 308. Show that all dense logics of width 2 have the finite model property.

Exercise 309. Give a detailed proof of Theorem 8.7.19.

8.8. Bounded Properties and Precomplete Logics above S4

In George Schumm [**190**] the following definition is given. A logic Θ **bounds** a property $\mathfrak{P}$ of logics if Θ fails to have $\mathfrak{P}$ while all proper extensions of Θ have $\mathfrak{P}$. Often such a logic is said to have **pre–**$\mathfrak{P}$. (For example: pretabular, precomplete.) Let us call a property **essentially bounded** if every logic without $\mathfrak{P}$ is contained in a logic that bounds $\mathfrak{P}$. If the inconsistent logic has $\mathfrak{P}$ then no consistent logic can bound $\mathfrak{P}$. So the concept of boundedness is only interesting for properties which the inconsistent logic has.

Theorem 8.8.1. *Let Θ be finitely axiomatizable. Then the property '**contains** Θ' is bounded.*

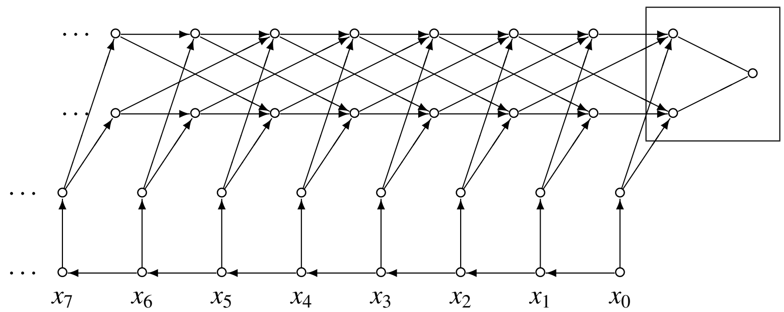
The proof of this theorem is left as an exercise. It uses only the fact that the finitely axiomatizable logics are the compact logics, and that the lattice of logics is algebraic. Notice that a lower cover of a finitely axiomatizable logic need not be finitely axiomatizable again. For tabular transitive logics this is correct, though. This is the deeper reason for correctness of the following theorem.

Theorem 8.8.2. *The property '**is of codimension less than n**' is a bounded property in the lattice $\mathcal{E}$ **K4**.*

Proof. There exist only finitely many logics of codimension $\leq n$ in $\mathcal{E}$ **K4**. They are tabular and finitely axiomatizable. Hence for each logic Θ of codimension n there is a finite set $L(\Theta)$ of lower covers, such that every logic properly contained in Θ is contained in a member of $L(\Theta)$. Let L be the union of $L(\Theta)$, Θ of codimension n. Then any logic of codimension $\not\leq n$ is contained in a member of L. ◻

In this section we will examine completeness and incompleteness of logics containing **S4**. In [**62**], Kit Fine has constructed a logic containing **S4** which is not

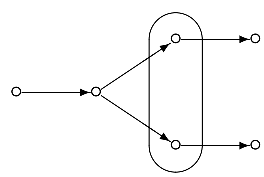
FIGURE 8.8. $\mathfrak{C}$



complete. We have met this logic in Section 7.6. We will show that it is also **pre–complete**. The underlying frame of $\mathfrak{C}$ is shown again in Figure 8.8. The sets are the sets which are finite or cofinite subsets of the three upper layers and finite or cofinite on the lower layer.

THEOREM 8.8.3. *The logic of* $\mathfrak{C}$ *is precomplete. Moreover, each proper extension has the finite model property.*

PROOF. We have $\mathsf{Th}\,\mathfrak{F} \supseteq$ **Grz**. The following embedding pattern is satisfiable but cannot be satisfied on a frame. (The oval encloses the only closed domain of this embedding pattern.)



The reason is simply that the root point must always be a part of the head. Hence it has a successor which is also a head part etc. There exists therefore an ascending chain of points satisfying this embedding pattern. Hence the underlying frame does not satisfy **Grz**.

Now we show that the logic of $\mathfrak{C}$ is pre–complete, and that every proper extension has the finite model property. Consider the subframe generated by x_1. If the points of depth < 2 are contracted onto a single point (they are enclosed in a box in Figure 8.8) then we get a p–morphism onto $\mathfrak{C}$. Hence if a formula is satisfiable at x_1 it is satisfiable at x_0. So, by induction, φ is satisfiable at x_i, $i \in \omega$, iff it is satisfiable at x_0. Let φ be such that $\mathsf{Th}\,\mathfrak{C} \oplus \varphi$ is different from $\mathsf{Th}\,\mathfrak{C}$. Then φ is not satisfiable at any point x_i. Then $\mathsf{Th}\,\mathfrak{C} \oplus \varphi \supseteq \mathsf{Th}\,\mathfrak{C}^{<\omega}$. This logic has the finite model property. □

The situation changes drastically when we shift to logics of finite width. By Theorem 8.6.14, there are no pre–complete extensions of finite width. What we will show is that there are 13 logics of finite width which have pre–finite model property, and that finite model property is essentially bounded in the lattice of logics of finite width. Let Θ be a logic of finite width that does not have the finite model property. We will show that Θ is contained in one of 13 logics that will be described below.

LEMMA 8.8.4. *Let Θ be a logic containing* $\mathbf{S4}$ *without the finite model property. Then* $\mu_\omega \circledcirc \circ$ *is a frame for* Θ.

PROOF. We may assume that $\Theta = \mathsf{Th}\,\mathfrak{F}$ for some $\mathfrak{F}$. Consider $\mathfrak{G} := \mathfrak{F}^{<\omega}$. Suppose that $\mathfrak{G}$ is decomposable into infinitely many segments. Then

$$\mathfrak{G} = \ldots \circledcirc \mathfrak{K}_3 \circledcirc \mathfrak{K}_2 \circledcirc \mathfrak{K}_1 \circledcirc \mathfrak{K}_0$$

for some $\mathfrak{K}_i$, $i \in \omega$. Then all $\mathfrak{K}_i$ are finite. They are totally local, and so the contraction of $\mathfrak{K}_i \twoheadrightarrow \circ$ induces a p–morphism of $\mathfrak{F}$ reducing that component. Let $k \in \omega$. Denote by $\mathfrak{F}_k$ the result of contracting the frame $\ldots \mathfrak{K}_{k+2} \circledcirc \mathfrak{K}_{k+1} \circledcirc \mathfrak{K}_k$ onto a single point. Then $\mathsf{Th}\,\mathfrak{F} = \bigcap_k \mathsf{Th}\,\mathfrak{F}_k$. Hence one of the $\mathsf{Th}\,\mathfrak{F}_k$ fails to have the finite model property. We may assume therefore that $\mathfrak{G}$ is not decomposable into infinitely many segments. Then it contains a single infinite segment. Contract all other segments into a single point. Then we have reduced $\mathfrak{G}$ to a single indecomposable segment. Moreover, we can assume that there is a single galaxy of depth 0. Otherwise, we contract all but one galaxy into a single point of depth 0. Then $ti(> 2)$ is satisfiable in $\mathfrak{G}$. Therefore there exist points x_0, y_0, y_1 and u such that $u \vec{\lhd} x_0$, $u \vec{\lhd} y_1 \vec{\lhd} y_0$, and $x_0 \parallel y_1$, $x \parallel y_0$. Contract all points which do not see x_0 or y_0 into a single point. Now we construct the points of depth 2. There exists a point $x_1 \vec{\lhd} x_0$ of depth 2. We may assume that x_1 sees a point y_i. Since the segment is indecomposable, $i = 0$. Now for the points os depth 3. There exist x_2 and y_2 such that $x_2 \vec{\lhd} x_1$ and $y_2 \vec{\lhd} y_1$. x_2 sees a y_i. By indecomposability, $i = 1$. Likewise, $y_2 \lhd x_j$ for some $j < 2$. Again $j = 1$. And so on. Hence, the segment is isomorphic to μ_ω. Now, $\mu_\omega \twoheadrightarrow \mu_\omega \circledcirc \circ$. ◻

The following is proved using Proposition 8.6.9.

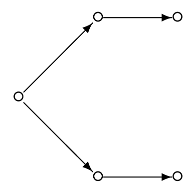
LEMMA 8.8.5. $\mathsf{Th}\,\mu_\omega \circledcirc \mu_\omega = \mathsf{Th}\,\mu_\omega$.

It follows that if Θ is without the finite model property, there is a Θ–frame $\mathfrak{g} \circledcirc \mu_\omega \circledcirc \circ$ such that $\mathsf{Th}\,\mathfrak{g} \circledcirc \mu_\omega \circledcirc \circ$ fails to have the finite model property. Suppose that (A), (B), (C) and (D) hold.

(A) $\mathfrak{g}$ is of fatness 1.
(B) $\mathfrak{g}$ is of tightness 3.
(C) $\mathfrak{g}$ does not contain $\mathfrak{p}$.
(D) $\mathfrak{g}$ is of width 2.

Then the logic of $\mathfrak{g} \circledcirc \mu_\omega \circledcirc \circ$ has the finite model property. For an indecomposable segment of $\mathfrak{g}$ is either of tightness 1 and so of depth 1, or it is of tightness 2 and then isomorphic to μ_ω. We leave the verfication of this fact to the reader. So, one of (A), (B), (C) and (D) must fail. Suppose (A) fails. Then $\mathfrak{g}$ contains a cluster of fatness

FIGURE 8.9. $\mathfrak{p}$



> 1. Let it be C. Let $\mathfrak{h}$ be the subframe of $\mathfrak{g}$ generated by C. Then $\mathfrak{h} \twoheadrightarrow C \twoheadrightarrow \mathbf{2}$ or $\mathfrak{h} \twoheadrightarrow C \otimes \circ \twoheadrightarrow \mathbf{2} \otimes \circ$. In the first case, $\mathsf{Th}\, \mathfrak{g} \otimes \mu_\omega \otimes \circ \subseteq \mathsf{Th}\, \mathbf{2} \otimes \mu_\omega \otimes \circ$. In the second case $\mathsf{Th}, \mathfrak{g} \otimes \mu_\omega \otimes \circ = \mathsf{Th}\, \mathbf{2} \otimes \circ \otimes \mu_\omega \otimes \circ = \mathsf{Th}\, \mathbf{2} \otimes \mu_\omega \otimes \circ$. The last equation follows from Theorem 8.6.10. Now suppose that (B) fails; we assume for simplicity that (D) does not fail. Call $\mathfrak{u}$ the frame consisting of a chain of length 2 parallel to a single point. This frame is embeddable into $\mathfrak{g}$. $\mathfrak{g}$ is noetherian. There exists chain $x_4 \vec{\lhd} x_3 \vec{\lhd} x_2 \vec{\lhd} x_1 \vec{\lhd} x_0$ and a point y such that $x_4 \vec{\lhd} y$ and $y \parallel x_0$ (and so $y \parallel x_1$, $y \parallel x_2$ and $y \parallel x_3$ as well). It follows that we may choose the points in such a way that x_i immediately succeeds x_{i+1} and also y immediately succeeds x_4. Take the subframe $\mathfrak{h}$ of $\mathfrak{g}$ generated by x_4. We may assume that all clusters have size 1. Take a point u different from y or x_i; then either $y \lhd u$ or $x_0 \lhd u$. Therefore, we may contract all points different from y and the x_i onto a single point v. **Case 1.** $x_0 \ntriangleleft v$. Then contract y and v to a single point. **Case 2.** $y \ntriangleleft v$. Then contract x_0 and v to a single point. **Case 3.** Both $y \lhd v$ and $x_0 \lhd v$. Hence, we may assume that $\mathfrak{g}$ is of the form $\mathfrak{u} \otimes \mu_\omega$ or $\mathfrak{u} \otimes \circ \otimes \mu_\omega$. In the latter case we may use Theorem 8.6.10 again and drop the $\circ$. Hence we have shown that the logic of the frame is contained in $\mathfrak{u} \otimes \mu_\omega$. Now suppose that (C) fails. Then as before we may assume that we have points $x_2 \vec{\lhd} x_1 \vec{\lhd} x_0$ and $x_2 \vec{\lhd} y_1 \vec{\lhd} y_0$. For simplicity we assume again that the frame is of width 2. In the same way we reduce $\mathfrak{g}$ to the form $\mathfrak{p} \otimes \mu_\omega$ or $\mathfrak{p} \otimes \circ \otimes \mu_\omega$. Aagain, by Theorem 8.6.10 the logic of the latter is identical to the logic of the former. This leaves us with (D) to discuss. Suppose we have an antichain $X = \{x_0, x_1, x_2\}$. By dropping intermediate galaxies (after contraction to a point and then dropping that point supersafely) we can reduce $\mathfrak{f}$ to a decomposition of two galaxies $\mathfrak{g} \otimes \mu_\omega$, where $\mathfrak{g}$ contains an antichain of size 3. We contract all points seen by either member of X into a single point z, if such points exist. This is a p–morphism. Furthermore, consider a point $y \parallel X$. The map contracting the interval $[y, z]$ onto z is a p–morphism. Therefore we may assume from now on that X is an antichain of maximal size. We have now the following structure at depth 0 and 1 of $\mathfrak{g}$ shown in Figure 8.10. The first situation can be reduced to the last by Theorem 8.6.10. The third can be reduced to the last as well by contracting x_1 and z. This leaves us with two frames. Now we consider the points immediately preceding members of X. Given an antichain Z, call u a Z–**unifier** if for all $z \in Z$ $u \vec{\lhd} z$ and if $u \vec{\lhd} v$ then $z \unlhd v$ for some $z \in Z$. If $\sharp Z = k$,

FIGURE 8.10.

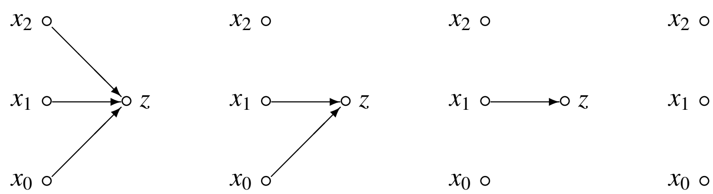
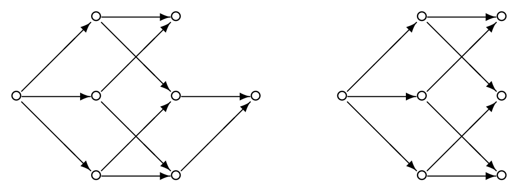


FIGURE 8.11.



call u a k–unifier. If $Z = \{z\}$ and u a Z–unifier, then the map collapsing u into z is a p–morphism. We may therefore assume that we have no 1–unifier. For a subset $Y \subseteq X$ of cardinality at least 2 denote by y_Y the Y–unifier. For certain Y there may not exist such a y_Y. **Case 1.** There is an X–unifier, y_X. Then we take the subframe generated by y_X. (This gives two possibilities, depending on whether X is of depth ω or whether x_0 and x_1 are of depth $\omega + 1$, see Figure 8.10. For the first possibility can be eliminated through dropping z and the third by a contraction.) **Case 2.** y_X does not exist. **Case 2a.** Assume that all 2–unifiers y_Y exist and that there is a $\{y_{01}, y_{02}, y_{12}\}$–unifier w. Then we take the subframe generated by w. This gives 2 distinct frames, shown in Figure 8.11. **Case 2b.** For two sets $Y, Z \subseteq X$ of cardinality 2 there exist y_Y and y_Z and a $\{y_Y, y_Z\}$–unifier w. We take the subframe generated by w. (This gives 3 possibilities, depending on whether X is entirely of depth ω or whether two points, say x_0 and x_1 are of depth $\omega + 1$. In the latter case we have to distinguish whether or not $\{x_0, x_1\} \in \{Y, Z\}$.) **Case 2c.** For only one set Y of cardinality 2 there exists y_Y. Let $x_i \in X - Y$. Then $\{x_i, y_Y\}$ is an antichain and has a unifier, w. We take the frame generated by w. (Again 3 possibilities.)

THEOREM 8.8.6. *Finite model property is a bounded property in the lattice of logics of finite width containing* $\mathbf{S4}$*. Moreover, there are 3 logics of width 2 bounding finite model property, and 13 logics of width 3.*

We close this section with an overview over bounded properties. For most

TABLE 1. Bounded Properties

property	bounded
finite codimension	yes
tabularity	yes
finite model property	yes
completeness	yes
compactness	yes
finite axiomatizability	yes
elementarity	yes
decidability	yes
interpolation	yes
Halldén–completeness	yes

claims the proof is rather easy. **G.3** bounds the following properties: finite codimension, tabularity, compactness and elementarity. Above we have shown that completeness is a bounded property. We have shown in Theorem 7.5.15 that there is a logic bounding finite axiomatizability. Now we turn to decidability. Let $\mathfrak{a}(0) := \bullet$ and $\mathfrak{a}(1) := \bullet \oplus \circ$. Take an infinite sequence $\alpha : \omega \to \{0, 1\}$. Then let $\mathfrak{f}_\alpha$ be the following Kripke–frame

$$\mathfrak{f}_\alpha := \ldots \oslash \mathfrak{a}(\alpha(3)) \oslash \mathfrak{a}(\alpha(2)) \oslash \mathfrak{a}(\alpha(1)) \oslash \mathfrak{a}(\alpha(0))$$

It is easily seen that for any two different sequences α and β that $\mathsf{Th}\,\mathfrak{f}_\alpha \neq \mathsf{Th}\,\mathfrak{f}_\beta$. Hence we obtain the following theorem.

THEOREM 8.8.7 (Blok). *There are $2^{\aleph_0}$ many pretabular logics in $\mathcal{E}\,\boldsymbol{K4}$.*

Now it is obvious that there must be also pretabular logics which are undecidable. These logics bound decidability since all proper extensions are tabular and so decidable, by Theorem 1.6.1. The logic of the two one–point frames (axiomatized by $\Diamond p \to p$) is not Halldén–complete. For $\Diamond\top \vee \Box\bot$ is a theorem, but neither $\Diamond\top$ nor $\Box\bot$ is a theorem. However, all proper extensions are Halldén–complete. In order to show that interpolation is bounded it is in fact enough to show that there is a logic of **S4** of finite codimension which fails to have interpolation. This is left as an exercise.

Exercise 310. (WIM BLOK [**21**].) Let $\mathfrak{b}_3 := \circ \oslash (\circ \oplus \circ \oplus \circ)$ and $\mathfrak{h} := \mathfrak{b}_3 \oslash \mu_\omega$. Show that the logic of $\mathfrak{h}$ fails to have finite model property. Show that it is the union of two logics which have finite model property. *Hint.* Define the subframe logic which consists of frames of fatness 1, width 2, tightness 4, such that there exists at most one chain of length three parallel to a point. Show that the theory of $\mathfrak{h}$ can be obtained by splitting two frames from this logic, and that splitting either frame results in a logic with finite model property.

Exercise 311. (Continuing the previous exercise.) Show that $\mathsf{Th}\,\mathfrak{h}$ is a splitting of **K4** by finitely many frames. Hence, splitting does not always preserve the finite model property.

Exercise 312. Let $\mathfrak{D}$ be the irreflexive counterpart of $\mathfrak{C}$ of Figure 8.8. Show that the logics $\mathsf{Th}\,\mathfrak{D} \otimes \alpha^{op}$ for $\alpha < \omega$ form an ascending chain of logics. The limit of this chain is $\mathsf{Th}\,\mathfrak{D} \otimes \omega^{op}$.

Exercise 313. Prove Theorem 8.8.1.

Exercise 314. Let Θ be the logic of the frame $\mu \otimes \lambda_\omega$ where $\mu = \circ \otimes (\circ \oplus (\circ \otimes \circ))$. Show that Θ cannot be finitely axiomatized. Show that any proper extension can be finitely axiomatized, however. (See [**125**].)

*__Exercise 315.__ Show that there are $2^{\aleph_0}$ many pre–complete logics in the lattice $\mathcal{E}$ **Grz**.

Exercise 316. Show that if $\alpha \neq \beta$ are infinite sequences of 0 and 1 then $\mathsf{Th}\,\mathfrak{f}_\alpha \neq \mathsf{Th}\,\mathfrak{f}_\beta$.

Exercise 317. Let Λ bound decidability. Show that (a) Λ is $\sqcap$–irreducible, (b) if Λ is recursively enumerable then it is not the lower limit of a chain of logics.

Exercise 318. Show that there exists a tabular logic in $\mathcal{E}$ **S4** which fails to have interpolation. *Hint.* Show that there exists a logic which fails to have superamalgamation.

Exercise 319. Show that the logic of $\circ \otimes \bullet$ bounds 0–axiomatizability. *Hint.* $\circ \otimes (\bullet \oplus \bullet)$ satisfies the same constant formulae as $\circ \oplus \bullet$.

8.9. Logics of Finite Tightness

We advise the reader to study Section 10.4 before entering this section. We will show here that logics of finite width and finite tightness are decidable if finitely axiomatizable. The proof uses methods from the theory of finite automata, which are provided in Section 10.4.

Lemma 8.9.1. *Let Λ be a logic extending $\boldsymbol{K4}$ and let φ be a formula. Suppose that $\sharp var(\varphi) \leq k$. Then $\varphi \in \Lambda$ iff $\varphi \in \Lambda.ft(\leq 2^k)$.*

Proof. Clearly, if $\varphi \in \Lambda$ then $\varphi \in \Lambda.ft(\leq 2^k)$. Now suppose that $\varphi \notin \Lambda$. Then $\mathfrak{Can}_\Lambda(k) \nvDash \varphi$. Now, $\mathfrak{Can}_\Lambda(k)$ is of fatness $\leq 2^k$; so it is a frame for $\Lambda.ft(\leq 2^k)$. Hence $\varphi \notin \Lambda.ft(\leq 2^k)$. ∎

Let θ be fixed. A θ–**accessibility matrix** is a $\theta \times \theta$–matrix over ω such that $a_{ii} = 0$ for all $i < \theta$. A θ–**fatness vector** is a θ–vector of elements in ω. A θ–**type** is a pair $\langle A, v\rangle$, where A is a θ–accessibility matrix and v a θ–fatness vector such that if $a_{ij} = 0$ for some $i \neq j$ then $v_i = v_j$ and for all $k < \theta$ $a_{ik} = a_{jk}$ and $a_{ki} = a_{kj}$. (It follows that $a_{ij} = a_{ii} = a_{ji} = a_{jj} = 0$.) A (θ, τ, π)–**type** is a type such that (i) $a_{ij} \leq \tau$ for all $i, j < \theta$, (ii) $v_i \leq \pi$ for all $i < \theta$. The set of (θ, τ, π)–types is finite and denoted by $\mathsf{Tp}(\theta, \tau, \pi)$ or simply by Tp. Types are intended to code the structure of frames. A sequence of (θ, τ, π)–types $\langle t^\alpha : \alpha < \gamma\rangle$ is called a **frame sequence** if the following holds. If $t^\alpha = \langle A^\alpha, v^\alpha\rangle$ then $a^\alpha_{ij} \leq a^{\alpha+1}_{ij} \leq a^\alpha_{ij} + 1$ for all $i, j < \theta$ such that $i \neq j$. A sequence is called **rooted** if $\gamma = \beta + 1$ and $a^\beta_{ij} = 0$ for all $i, j < \theta$.

For a rooted noetherian Kripke–frame $\mathfrak{f}$ of width θ, tightness τ and fatness φ a rooted frame sequence of (θ, τ, φ)–types is constructed as follows. Let γ be the depth of $\mathfrak{F}$. Then there exist θ many sequences $\Sigma(i) = \langle C^i_\alpha : \alpha < \gamma\rangle$ for $i < \theta$ and a map κ from the set of clusters of $\mathfrak{F}$ into the set θ such that (i) the restriction of κ to the clusters of depth α is injective, (ii) for a cluster C of $\mathfrak{F}$ of depth α, $C = C^{\kappa(C)}_\alpha$, (iii) C^i_α immediately succeeds $C^i_{\alpha+1}$, $\alpha + 1 < \gamma$. Now a sequence $\langle t^\alpha : \alpha < \gamma\rangle$ of types is defined as follows. $t^\alpha = \langle A^\alpha, v^\alpha\rangle$, where v^α_i is the type of C^i_α, and a^α_{ij} is the least number k such that $\alpha = \beta + k$ and $C^i_\alpha \trianglelefteq C^j_\beta$. (Since the frames are of tightness τ, k exists, is finite and $\leq \tau$. Clearly, $a_{ii} = 0$ by choice of the clusters. Each t^α is a type. For let $a^\alpha_{ij} = 0$. Then $C^i_\alpha \trianglelefteq C^j_\alpha$. Hence the clusters, being of identical depth, are equal. So, $a_{ji} = 0$. Furthermore, $a_{ik} = a_{jk}$ and $a_{ki} = a_{kj}$ for all k. Finally, $a^{\alpha+1}_{ij} = a^\alpha_{ij}$ or $a^{\alpha+1}_{ij} = a^\alpha_{ij} + 1$. Now suppose that $\mathfrak{F}$ is rooted with root cluster C. Let C be of depth β. Then $\gamma = \beta + 1$. Moreover, all clusters of depth β are identical; so, $a^\beta_{ij} = 0$ for all $i, j < \theta$. So, the frame sequences is rooted. It should be noted that the type sequence of $\mathfrak{f}$ depends on the division into sequences of clusters. The results are independent of this, however. Denote by $\mathsf{Seq}(\mathfrak{f})$ the set of all possible type sequences for $\mathfrak{f}$. Given a type sequence, let $\mathsf{Fr}(\Sigma)$ be the frame corresponding to Σ. This is unique.

Proposition 8.9.2. *For every rooted frame sequence Σ of (θ, τ, π)–types there exists a rooted noetherian frame $\mathfrak{f}$ of width θ, tightness τ and fatness π such that $\Sigma \in \mathsf{Seq}(\mathfrak{f})$.*

Proof. Now let $\Sigma = \langle t^\alpha : \alpha < \gamma\rangle$ be a rooted frame sequence. Let Γ be the set of pairs $\langle \alpha, i\rangle$ such that $\alpha < \gamma$ and $a_{ki} \neq 0$ for all $k < i$. Γ represents the set of clusters of $\mathfrak{F}$. For $C \in \Gamma$ has type v^α_i. Now put $\langle \alpha, i\rangle \lhd \langle \beta, j\rangle$ if (1) $i = j$ and (1a) $\alpha > \beta$ or (1b) $\alpha = \beta$ and $v^\alpha_i \neq \varnothing$, (2) $i \neq j$ and $\alpha \geq \beta + a_{ij}$. (Note that in case (2), $a_{ij} \neq 0$ by choice of the set of clusters.) We will leave it to the reader to show that $\lhd$ as defined is transitive, and that $\Sigma \in \mathsf{Seq}(\mathfrak{f})$. ◻

Now let $\Sigma = \langle t^\alpha : \alpha < \gamma\rangle$. Let κ be an ordinal such that $\omega\kappa \leq \gamma$. Put

$$\Sigma(\kappa) := \langle t^\alpha : \omega\kappa \leq \alpha < \omega(\kappa + 1)\rangle$$

$\Sigma(\kappa)$ corresponds to the galaxy of depth κ in $\mho(\mathfrak{F})$. Hence we call $\Sigma(\kappa)$ a **galaxy sequence**. Put

$$\mho(\Sigma) := \langle \Sigma(\kappa) : \omega\kappa \leq \gamma \rangle$$

Just as with frames, we may study Σ be means of the sequence of galaxy sequences.

Definition 8.9.3. *A class $\mathfrak{X}$ of (noetherian) frames is* ***type regular*** *if there exists a regular language R such that $\mathfrak{f} \in \mathfrak{X}$ iff for all κ such that $\omega\kappa \leq dp(\mathfrak{f})$, $\Sigma(\kappa) \in R$.*

Lemma 8.9.4. *The class of all noetherian frames of width θ, tightness τ and fatness π is regular. For regular classes $\mathfrak{X}$ and $\mathfrak{Y}$, $\mathfrak{X} \cup \mathfrak{Y}$, $\mathfrak{X} \cap \mathfrak{Y}$ and $\mathfrak{X} - \mathfrak{Y}$ are regular.*

The proof of this theorem is left as an exercise. We note that if $\tau \leq \tau'$ and $\pi \leq \pi'$ then $\mathsf{Tp}(\theta, \tau, \pi) \subseteq \mathsf{Tp}(\theta, \tau', \varphi')$. A regular class $\mathfrak{X}$ of (θ, τ, π)–frames is therefore also a regular class of (θ, τ', π')–frames. Moreover, if R is a regular expression defining $\mathfrak{X}$ as a class of (θ, τ, π)–frames, then R is a regular expression defining $\mathfrak{X}$ as a class of (θ, τ', π')–frames. This is useful to know. Now let us see why we can limit ourselves to type regular frames. Suppose we want to know which frames can refute φ. Then, alternatively, we can ask ourselves which *refutation patterns* can be realized. Finally, since we have bounded width and no ascending chains, we can limit ourselves to *embedding patterns*. Take an embedding pattern $\epsilon(\mathfrak{d}, \mathfrak{V})$. Let us for the moment ignore the closed domains. Let $\mathfrak{d} = \mathfrak{d}_{n-1} \otimes \ldots \otimes \mathfrak{d}_0$ be a decomposition into indecomposable parts. Then $\iota : \mathfrak{d} \to \mathfrak{f}$ is a cofinal embbedding into $\mathfrak{f}$ if $\iota \restriction \mathfrak{d}_0$ is a cofinal embedding, $\iota \restriction \mathfrak{d}_i$ is an embedding for $0 < i < n$ and ι preserves the decomposition, that is,

$$\iota[\mathfrak{d}] = \iota[\mathfrak{d}_{n-1}] \otimes \ldots \otimes \iota[\mathfrak{d}_0]$$

Now, let us consider the simplest possible case, that of the embeddability of an indecomposable part, cofinal or not. Recall the definition of the *antiframe*. We have defined $x \shortparallel y$ by $x \ntriangleleft y \ntriangleleft y$, and defined the antiframe of $\mathfrak{f} = \langle f, \triangleleft \rangle$ to be the frame $\langle f, \shortparallel \rangle$. Now, if $\mathfrak{f}$ is indecomposable then its antiframe is connected. Moreover, if $\mathfrak{f}$ is a subframe of $\mathfrak{g}$, then the antiframe of $\mathfrak{f}$ is embedded as a subframe in the antiframe of $\mathfrak{g}$. Consequently, in our case $\iota[\mathfrak{d}_i]$ must be connected in $\langle f, \shortparallel \rangle$. However, by the fact that the target frame $\mathfrak{f}$ is of tightness τ, the points of $\mathfrak{d}_i$ cannot be too far apart. Namely, if $x \shortparallel y$ then $\iota(x) \shortparallel \iota(y)$; but the latter can only be if $\iota(x)$ and $\iota(y)$ are in the same galaxy (recall that $\mathfrak{f}$ is galactically linear) and their local depths differ by at most τ. If $\mathfrak{d}$ is indecomposable, the antiframe is connected, and for each pair x, y of points there is a path of length at most $\sharp\mathfrak{d}$ connecting x to y in the antiframe. This shows the following lemma.

Lemma 8.9.5. *Let $\mathfrak{d}$ be an indecomposable frame of cardinality ℓ and let $\iota : \mathfrak{d} \to \mathfrak{f}$ be an embedding into a noetherian frame of tightness τ. Then for any two points x and y of $\mathfrak{d}$ the points $\iota(x)$ and $\iota(y)$ are in the same galaxy, and the local depths differ by at most $\tau \cdot \ell$.*

Lemma 8.9.6. *Let $\mathfrak{d}$ be indecomposable, and $\epsilon^{(\circ)}(\mathfrak{d}, \mathfrak{V})$ a (cofinal) embedding pattern. The class of frames that realize $\epsilon(\mathfrak{d}, \mathfrak{V})$ is type regular.*

Proof. We show the criterion of embeddability first. Let $\Sigma \in \mathsf{Seq}(\mathfrak{f})$. Clearly, $\epsilon^{(\circ)}(\mathfrak{d}, \mathfrak{V})$ is realizable iff $\epsilon^{(\circ)}(\mathfrak{d}, \mathfrak{V})$ is realizable in a galaxy of $\mathfrak{f}$ iff $\epsilon^{(\circ)}(\mathfrak{d}, \mathfrak{V})$ is realizable in a subframe of points consisting of the slices $\alpha, \alpha + 1, \ldots, \alpha + \tau \cdot \ell$, for some α. Hence there exists a finite set T of sequences of length $\leq \tau \cdot \ell$ such that $\epsilon^{(\circ)}(\mathfrak{d}, \mathfrak{V})$ is realizable in $\mathfrak{f}$ iff no member of T is a subword of some $\Sigma(\kappa)$. For each κ, the language of types not containing a member of T as a subword, is a regular language. Namely, for each $t \in T$ we can write a term $\vec{t}$ defining just the language $\{t\}$. Then the term $\mathsf{Tp}^* \cdot \vec{t} \cdot \mathsf{Tp}^*$ defines the language of sequences containing t. Now take the union of all these terms (this is well–defined since T is finite). The intersection of the just defined language and the language of frame sequences is again regular and can be described by a regular term. For cofinal embeddability there exists a finite set S of sequences of length $\leq \tau \cdot \ell$ such that $\epsilon(\mathfrak{d}, \mathfrak{V})$ is realizable iff $\Sigma(0)$ does not start with a member of S. This is a regular language, as is shown similarly. ◻

So, the next step is to consider embedding patterns consisting of a decomposable frame. Here, we meet a subtle problem. If $\mathfrak{d} = \mathfrak{m} \otimes \mathfrak{n}$ we may not necessarily conclude that an embedding pattern for $\mathfrak{d}$ can be realized by embedding $\mathfrak{m}$ before $\mathfrak{n}$. Whether this division can be made depends on the closed domains of that embedding pattern. Recall that a closed domain can be viewed as a cone, i. e. an upper set, or alternatively as an antichain. It is the latter that suits our purpose best here. Let

$$\mathfrak{d} = \mathfrak{d}_{n-1} \otimes \mathfrak{d}_{n-2} \otimes \ldots \otimes \mathfrak{d}_1 \otimes \mathfrak{d}_0$$

Each antichain $\mathfrak{v}$ is contained in one and only one segment of $\mathfrak{d}$. An antichain is called an **outer antichain** of $\mathfrak{d}_i$ if $\mathfrak{d}_i = \mathop{\uparrow}\mathfrak{v}$, else it is called an **inner antichain**. An antichain is an outer (inner) antichain of $\mathfrak{d}$ if it is an outer (inner) antichain of some segment (which is unique given the antichain). The outer chains glue the segments to each other in the following way.

Lemma 8.9.7. *Let $\iota : \mathfrak{d}_1 \otimes \mathfrak{d}_0 \to \mathfrak{f}$ be an embedding and $\mathfrak{v}$ an outer antichain of $\mathfrak{d}_0$. If ι satisfies the closed domains condition for $\mathfrak{v}$ then $\mathfrak{d}_1$ and $\mathfrak{d}_0$ are in the same galaxy and the depth of a maximal point of $\iota[\mathfrak{d}_1]$ is at most τ larger than the maximal depths in $\iota[\mathfrak{d}_0]$.*

Proof. Let $z \lhd \iota[\mathfrak{v}]$. Then z may not be an external point; hence $z \in \iota[\mathfrak{d}_1]$. On the other hand, there is such a z at depth $\leq max\ dp[\iota(\mathfrak{v})] + \tau$. ◻

By a **compartment** of $\mathfrak{d}$ with respect to $\mathfrak{V}$ we understand a maximal chain of the form $\mathfrak{d}_{i+j-1} \otimes \ldots \otimes \mathfrak{d}_i$ of segments of $\mathfrak{d}$ such that for every $k < j-1$ there is an outer antichain for $\mathfrak{d}_{i+k}$ in $\mathfrak{V}$.

Lemma 8.9.8. *Let $\mathfrak{d} = \mathfrak{d}_k \otimes \ldots \otimes \mathfrak{d}_0$ be a decomposition of $\mathfrak{d}$ in compartments with respect to $\mathfrak{V}$. Then the embedding pattern $\epsilon(\mathfrak{d}, \mathfrak{V})$ can be satisfied iff there is an embedding $\iota : \mathfrak{d} \to \mathfrak{f}$ which satisfies all embedding patterns $\langle \mathfrak{d}_i, \mathfrak{V} \restriction \mathfrak{d}_i \rangle$ considering points of depth $\leq max\{dp(x) : x \in \mathfrak{v}\} + \tau$.*

Proof. If ι satisfies the embedding pattern for $\mathfrak{d}$ then it satisfies the embedding patterns for the compartments. For the converse, let ι respect the decomposition and satisfy the embdeding patterns for the compartments. Then ι is an embedding for $\mathfrak{d}$. Furthermore, let $\mathfrak{v} \in \mathfrak{V}$. Then $\mathfrak{v} \subseteq \mathfrak{d}_i$ for some i. Let $z \lhd \iota[\mathfrak{v}]$. We have to show that z is in $\iota[\mathfrak{d}]$. Suppose that $\mathfrak{v}$ is an outer antichain; then, by definition of a compartment, it is not an outer antichain of the compartment, so all antichains are inner antichains of the compartments. This means that there is no antichain $\mathfrak{v}$ in a compartment $\mathfrak{d}_i$ such that $\uparrow z \cap \hat{\uparrow} \mathfrak{d}_i = \hat{\uparrow} \mathfrak{v}$ for an internal point not belonging to $\mathfrak{d}_i$ itself. Or, to put it another way, internal points outside of $\mathfrak{d}_i$ cannot realize views forbidden by the closed domain conditions for $\mathfrak{d}_i$. So if such a view is realized, it is externally realized. But we know that we need not look very far for such points, namely only up to depth $maxdp(\iota[\mathfrak{v}]) + \tau$. ◻

Notice that it is not enough to satisfy all local patterns for compartments, since they are formulated such that we need only consider points of the subframe generated by these compartments. For the overall embedding, intermediate points have to be checked as well. If a decomposition of the original frame is given, for example into galaxies, then the condition is satisfied since it does not have to be enforced that the decomposition is respected. The proof of the next proposition is now obvious.

Lemma 8.9.9. *Let $\epsilon(\mathfrak{d}, \mathfrak{V})$ be a cofinal embedding pattern. Let $\mathfrak{f} = \bigotimes_{\alpha<\beta} \mathfrak{g}_\alpha$ be an arbitrary frame. The embedding pattern can be satisfied in $\mathfrak{f}$ iff there is a decomposition of $\mathfrak{d} = \mathfrak{d}_{k-1} \otimes \ldots \otimes \mathfrak{d}_0$ respecting the compartments and a monotonic sequence of numbers $\langle \alpha_i : i < k \rangle$ such that $\alpha_0 = 0$ and the cofinal embedding pattern $\langle \mathfrak{d}_0, \mathfrak{V} \restriction \mathfrak{d}_0 \rangle$ is satisfiable in $\mathfrak{g}_0$, the embedding patterns $\epsilon(\mathfrak{d}_i, \mathfrak{V} \restriction \mathfrak{d}_i)$ are satisfiable in $\mathfrak{g}_{\alpha_i}$, $0 < i < k$.*

The consequence is that we have a preservation theorem concerning decompositions. Namely, if $\epsilon(\mathfrak{d}, \mathfrak{V})$ is satisfiable in $\mathfrak{f} \otimes \mathfrak{g}$ then it is also satisfiable in $\mathfrak{f} \otimes \mathfrak{h} \otimes \mathfrak{g}$, simply by checking that the same embedding modulo identification of the parts does the job. This implies among other the following.

Theorem 8.9.10. *Let Λ be of finite width and finite tightness. Then dropping nonfinal galaxies is supersafe. Hence Λ has the galactic finite model property.*

Having progressed this far we now need a notational system for frames which consist of a single hypergalaxy. The essential idea is the following. Each regular language is defined by a regular expression. Let us take a set $\mathbb{R}$ of regular expressions over the set of types. $\mathbb{R}$ is called a **block** if (0) $\mathbb{R}$ is finite, (i) every frame sequence belongs to some member of $\mathbb{R}$, and (ii) no frame sequence belongs to different members of $\mathbb{R}$. In other words, the members of $\mathbb{R}$ partition the set of frame sequences. A **galactic** $\mathbb{R}$**–expression** is a regular expression $\mathfrak{E}$ over $\mathbb{R}$. A sequence $\langle \Sigma(k) : k < \alpha \rangle$ $(\alpha \leq \omega)$ is **accepted** by $\mathfrak{E}$ if there is a sequence $\langle R_i : i < \alpha \rangle$ of terms in $\mathbb{R}$ which is accepted by $\mathfrak{E}$ and R_i accepts $\Sigma(i)$ for all $i < \alpha$. Likewise, a frame $\mathfrak{f}$ of hypergalactic depth 1 is accepted if $\mho(\Sigma)$ is accepted by $\mathfrak{E}$ for some $\Sigma \in \mathsf{Seq}(\mathfrak{f})$. (We note that in principle, even if $\Sigma \in \mathsf{Seq}(\mathfrak{f})$ is accepted, some other $\Sigma' \in \mathsf{Seq}(\mathfrak{f})$ might not be accepted. This is however harmless.) A **hypergalactically regular class** of frames is a class $\mathfrak{X}$ of frames such that there exists a block $\mathbb{R}$ and a galactically regular expression $\mathfrak{E}$ over $\mathbb{R}$ such that $\mathfrak{f} \in \mathfrak{X}$ iff for each hypergalaxy Γ_2 of $\mathfrak{f}$, $\mathfrak{E}$ accepts Γ_2. We show that the class of frames of frames of hypergalactic depth 1 satisfying a given embedding pattern is hypergalactically regular. From this everything follows. We already have completeness with respect to noetherian frames of hypergalactic depth 1.

Proposition 8.9.11. *Let $\mathfrak{X}$ and $\mathfrak{Y}$ be hypergalactically regular classes of noetherian (θ, τ, π)–frames. Then $\mathfrak{X} \cap \mathfrak{Y}$, $\mathfrak{X} \cup \mathfrak{Y}$ and $\mathfrak{X} - \mathfrak{Y}$ are also hypergalactically regular. The class of all frames is hypergalactically regular. Moreover, $\mathfrak{X}$ is also hypergalactically regular as a class of (θ, τ', π')–frames for $\tau \leq \tau'$ and $\pi \leq \pi'$.*

Proof. First, let $\mathfrak{E}$ define $\mathfrak{X}$ and $\mathfrak{F}$ define $\mathfrak{Y}$. If $\mathfrak{E}$ and $\mathfrak{F}$ are regular expressions over the same block $\mathbb{R}$, the claim is clear. So, let $\mathfrak{E}$ be defined over $\mathbb{R}$ and $\mathfrak{F}$ defined over $\mathbb{S}$. Let $\mathbb{T}$ the set of all expressions defining the languages $L(R) \cap L(S)$, $R \in \mathbb{R}$ and $S \in \mathbb{S}$, if nonempty. These expressions exist, since regular languages are closed under intersection. $\mathbb{T}$ as just defined is a block. Let $R \in \mathbb{R}$. Define $R^\dagger$ by $R^\dagger := \bigcup\langle T : T \in \mathbb{S}, L(T) \subseteq L(R)\rangle$. The latter is well–defined, since the union is finite. By the definition of $\mathbb{T}$, $L(R) = L(R^\dagger)$. Call $\mathfrak{E}^\dagger$ the result of replacing each occurrence of R by $R^\dagger$, for each $R \in \mathbb{R}$. $\mathfrak{E}^\dagger$ is a regular expression over $\mathbb{T}$. The class of frames accepted by $\mathfrak{E}^\dagger$ is exactly $\mathfrak{X}$. Similarly we can produce a term $\mathfrak{F}^\ddagger$ over $\mathbb{T}$ defining $\mathfrak{Y}$. Now the first claim is clear. For the second claim, note that the class of all frames corresponds to the term $(\bigcup_{R \in \mathbb{R}} R)^*$. For the last claim, note that $\mathfrak{E}$ accepts $\mathfrak{X}$ also as a class of (θ, τ', π')–frames for $\tau \leq \tau'$ and $\pi \leq \pi'$. ◻

Theorem 8.9.12. *The set of noetherian frames of width θ, tightness τ and fatness π and depth $\leq \omega^2$ realizing a given embedding pattern is hypergalactically regular.*

Proof. Let $\mathfrak{d} = \mathfrak{d}_{k-1} \oslash \mathfrak{d}_{k-2} \oslash \ldots \oslash \mathfrak{d}_1 \oslash \mathfrak{d}_0$ be a division of $\mathfrak{d}$ into compartments. Then define $\mathbb{R}$ as follows. For each pair $i, j < k$ such that $i \leq j$ let $\mathfrak{d}[i, j] := \mathfrak{d}_j \oslash \mathfrak{d}_{j-1} \oslash \ldots \oslash \mathfrak{d}_i$ and $E[i, j]$ be the regular expression defining the class of galactic frames in which $\epsilon^\circ(\mathfrak{d}[j, i], \mathfrak{V} \upharpoonright \mathfrak{d}[i, j])$ is realizable, for $i > 0$, and $\epsilon(\mathfrak{d}[j, i], \mathfrak{V} \upharpoonright \mathfrak{d}[j, i])$ if $i = 0$. Let X define the class of frames into which no $\mathfrak{d}[i, j]$ is embeddable (confinally, if $i = 0$). Take all possible intersections of those $E[i, j]$ which are nonempty. The set of the terms thus defined together with X forms a block. Call it $\mathbb{R}$. Let the union of all terms of $\mathbb{R}$ be R. Let $H := \langle i_x : x < r\rangle$ be a strictly ascending sequence of natural numbers $< k - 1$. Put

$$\mathfrak{E}[H] := E[0, i_0] \cdot R^* \cdot E[i_0 + 1, i_1] \cdot \ldots \cdot E[i_{r-2} + 1, i_{r-1}] \cdot R^* \cdot E[i_{r-1} + 1, k - 1] \cdot R^*$$

Now let $\mathfrak{E}$ be the union of all $\mathfrak{E}[H]$. Since there are only finitely many sequences H, this is well–defined. It is not hard to check that $\mathfrak{E}$ admits a frame $\mathfrak{f}$ consisting of a single hypergalaxy iff $\epsilon(\mathfrak{d}, \mathfrak{V})$ is realizable. ◻

Corollary 8.9.13. *Let Λ be a finitely axiomatized logic of width θ, tightness τ and fatness π. Then the class of rooted frames for Λ of hypergalactic depth 1 is hypergalactically regular.*

Corollary 8.9.14. *Let Λ be of finite width and finite tightness. If Λ is finitely axiomatizable, it is decidable.*

Proof. Let φ be given. We want to decide whether $\varphi \in \Lambda$. For π large enough, this is equivalent to the problem '$\varphi \in \Lambda.ft(\leq \pi)$'. Λ is complete with respect to noetherian frames of hypergalactic depth 1. It determines a hyperregular set in that class. Likewise, the set of frames of hypergalactic frames in which φ is satisfiable is a hyperregular set. Now their intersection is also hyperregular. Moreover, terms representing these sets can be algorithmically computed. Hence we are done if we can decide whether for a given $\mathfrak{E}$ we can decide whether or not there exists a frame accepted by $\mathfrak{E}$. To that end, we decide whether there are nonempty $\mathbb{R}$–sequences falling under $\mathfrak{E}$. This is decidable on the basis of $\mathfrak{E}$. If there exists such a sequence, there exists a frame accepted by $\mathfrak{E}$. ◻

Exercise 320. Show Lemma 8.9.4.

CHAPTER 9

Logics of Bounded Alternativity

9.1. The Logics Containing $\mathbf{K.alt}_1$

Throughout this chapter we will deal with logics with one or several operators each of which satisfies an axiom of the form $\mathbf{alt}_n$ for some n. These are the logics of *bounded alternative* or logics of *bounded alternativity*. The study of such logics has been initiated by several insights and questions. First, any relation on a set can be seen as the union of sufficiently many partial functions. Hence, any monomodal Kripke–frame can be viewed as a frame for polymodal $\mathbf{K.alt}_1$ where the distinction between the operators has been lost. We will study this interpretation under the name of colouring and decolouring of polymodal frames. The second motivation comes from studying attribute–value formalisms, as used in computer science and linguistics. It has been realized that modal logics containing $\mathbf{K.alt}_1$ provide an ideal basis for studying such formalisms. It quickly turned out that the quasi–functional logics bear close resemblance to Thue–systems and this connection has been exploited to derive numerous undecidability results in modal logic. Many of these results were known already to Albert A. Muchnik, but only few got published (see [**55**] and [**56**]). Krister Segerberg in [**197**] has investigated the lattice $\mathcal{E}\,\mathbf{K.alt}_1$. Most of his results were reproved by Fabio Bellissima in [**5**], who added some facts about the lattice $\mathcal{E}\,\mathbf{K.alt}_2$. The logics of bounded alternative form a neglected area of modal logic, with most results generally being unknown. However, as we will see, they provide an ideal source of many powerful negative results on modal logic, as well as an ideal tool for modelling certain mathematical structures.

We start the investigation by classifying the extensions of $\mathbf{K.alt}_1$. We have seen in Section 3.2 that extensions of $\mathbf{K.alt}_1$ are canonical, and it is not hard to see that they all are df–persistent. We derive from this the fact that all extensions are complete, and so we can classify the extensions just by looking at the frames for $\mathbf{K.alt}_1$. Furthermore, it is enough to look at the rooted frames. We can view them as ordinary frames $\langle f, \lhd\rangle$ or as structures $\langle f, s\rangle$, where s is a partial function such that $s(x)$ is the $\lhd$–successor of x, if such a successor exists, and undefined else. A map $\pi : f \to g$ is a p–morphism from $\langle f, s\rangle$ to $\langle g, s'\rangle$ iff for all $x \in f$, $\pi(s(x))$ exists iff $s'(\pi(x))$ exists, and then both are equal. To say that $\mathfrak{f}$ is rooted is to say that there is a point $x \in f$ such that $f = \{s^n(x) : n \in \omega\}$. There are three possibilities. **Case 1.** There is an $n \in \omega$ such that $s^n(x)$ is not defined. Assume that $s^{n-1}(x)$ exists. In that case we have a chain of length n. We abbreviate this frame by $\mathfrak{ch}_n$. **Case 2.** $s^n(x)$ exists for all n. In that case, there are two further possibilities. **Case 2a.** All $s^n(x)$ are different. Then our frame is isomorphic to the natural numbers with the successor function. We denote this frame by $\mathfrak{ch}_\omega$. **Case 2b.** There are m, n such that $s^m(x) = s^n(x)$. Let m, n be minimal such that $m < n$. Then put $p := n - m$ and $y := s^m(x)$. We have $s^p(y) = y$, and consequently, for all natural numbers, $s^{i+p}(y) = s^i(y)$. Hence, y is in a cycle of length p (since n was chosen minimally). The elements $x, s(x), s^2(x), \dots, s^{m-1}(x)$ form a chain leading into the cycle. We denote this frame by $\mathfrak{cyc}_{m,p}$.

Proposition 9.1.1. *(i.)* $\mathfrak{ch}_m$ *is a generated subframe of* $\mathfrak{ch}_n$ *iff* $m \leq n$*, but not a generated subframe of either* $\mathfrak{ch}_\omega$ *or* $\mathfrak{cyc}_{n,p}$*. No frame for* $\boldsymbol{K.alt}_1$ *maps p–morphically onto* $\mathfrak{ch}_m$*, except for* $\mathfrak{ch}_m$*. (ii.)* $\mathfrak{cyc}_{m,p}$ *is a generated subframe of* $\mathfrak{cyc}_{n,q}$ *iff* $p = q$ *and* $m \leq n$*, and a p–morphic image of* $\mathfrak{cyc}_{n,q}$ *iff* $m \leq n$ *and* p *divides* q*.* $\mathfrak{cyc}_{m,p}$ *is a p–morphic image of* $\mathfrak{ch}_\omega$*, but no image of* $\mathfrak{ch}_n$ *for any* n*.*

The proof of these facts is easy in all cases except for the p–morphic images of $\mathfrak{cyc}_{m,p}$. In that case we have that if $\pi(x) = u$ then $\pi(s(x)) = s(u)$, from which above claims can be derived. We now have a full overview of the relation between these frames. Furthermore, for finite $\mathfrak{f}$ and $\mathfrak{g}$ we have $\mathsf{Th}\,\mathfrak{f} \subseteq \mathsf{Th}\,\mathfrak{g}$ iff $\mathfrak{g}$ is a p–morphic image of a generated subframe of $\mathfrak{f}$. Finally, let us observe that all frames except for $\mathfrak{ch}_\omega$ are finite. However, $\mathsf{Th}\,\mathfrak{ch}_\omega$ has the finite model property. For consider a formula φ and $\mathfrak{ch}_\omega \nvDash \varphi$. Without loss of generality we can assume that φ does not hold at the root. Then for $n + p \geq dp(\varphi)$ we get $\mathfrak{cyc}_{n,p} \nvDash \varphi$. This is so, because the $dp(\varphi)$–transits of the roots of $\mathfrak{ch}_\omega$ and $\mathfrak{cyc}_{n,p}$ are isomorphic. Likewise it is shown that $\mathfrak{ch}_\omega = \bigcap\langle \mathsf{Th}\,\mathfrak{ch}_n : n \in \omega\rangle$.

Proposition 9.1.2.

$$\begin{array}{lcl} \mathsf{Th}\,\mathfrak{ch}_n & = & \boldsymbol{K.alt}_1 \oplus \neg\Diamond^n\top \\ \mathsf{Th}\,\mathfrak{ch}_\omega & = & \boldsymbol{K.alt}_1.\boldsymbol{D} \\ \mathsf{Th}\,\mathfrak{cyc}_{n,q} & = & \boldsymbol{K.alt}_1.\boldsymbol{D} \oplus \Box^n(p \leftrightarrow \Diamond^q p) \end{array}$$

Proof. By correspondence theory. Notice that we can describe the structure of these frames by first–order restricted sentences, namely in the case of $\mathfrak{ch}_n$ by $(\forall y \rhd^{n+1} x)\mathsf{f}$, in the case of $\mathfrak{ch}_\omega$ by $(\exists y \rhd x)\mathsf{t}$, and in the case of $\mathfrak{cyc}_{n,q}$ by $(\forall y \rhd^n x)(\forall z \rhd^q y)(z \doteq y)$. All these are clear instances of Sahlqvist–formulae. The axioms above are the modal correspondents. ∎

Notice also that the conditions expressed by these formulae are derivable in $\mathfrak{Df}$, the class of differentiated frames, another proof that these logics are df–persistent.

Theorem 9.1.3 (Segerberg). $\boldsymbol{K.alt}_1$ *is the logic of the chains* $\mathfrak{ch}_n$*,* $n \in \omega$*. Every extension of* $\boldsymbol{K.alt}_1$ *is df–persistent and has the finite model property. Moreover,* $\boldsymbol{K.alt}_1.\boldsymbol{D}$ *is pretabular. Each proper extension of* $\boldsymbol{K.alt}_1$ *not containing* $\boldsymbol{K.alt}_1.\boldsymbol{D}$ *is of the form* $\Lambda \sqcap \mathsf{Th}\,\mathfrak{ch}_n$ *for some* $n \in \omega$*, where* $\Lambda \in \mathcal{E}\,\boldsymbol{K.alt}_1.\boldsymbol{D}$*. Hence in* $\mathcal{E}\,\boldsymbol{K}.\mathbf{alt}_1$ *there are only countably many elements.*

Proof. For the first claim let φ be a formula of modal depth d and let it have a model on a rooted frame $\mathfrak{f}$. If $\mathfrak{f}$ is not a finite chain then it is a p–morphic image of $\mathfrak{ch}_\omega$, so a model for φ can be based on $\mathfrak{ch}_\omega$. However, the d–transit of the root of $\mathfrak{ch}_d$ is isomorphic to the d–transit of any point in $\mathfrak{ch}_\omega$, so a model can be based on $\mathfrak{ch}_d$ as well. Consider a proper extension Λ of **K.alt**$_1$.**D**. Suppose there are infinitely many $\mathfrak{cyc}_{n,p}$ which are frames for Λ. Then the set $\{n + p : \mathfrak{cyc}_{n,p} \vDash \Lambda\}$ is unbounded. Hence every formula satisfiable in $\mathfrak{ch}_\omega$ is satisfiable in a frame for Λ, so that $\Lambda \subseteq \mathsf{Th}\,\mathfrak{ch}_\omega$. Hence all $\mathfrak{cyc}_{n,p}$ are frames for Λ, a contradiction, since then $\Lambda =$ **K.alt**$_1$.**D**. Every logic containing **K.alt**$_1$ is determined by its frames of the form $\mathfrak{ch}_n$, $n \in \omega$, and the frames of the form $\mathfrak{cyc}_{n,p}$. Of the first there are finitely many, since the extension is proper. Then there is a largest chain $\mathfrak{ch}_n$ contained in the set and the set of chains for Λ is the set of chains of the form $\mathfrak{ch}_m$, $m \leq n$. The logic of the cycles for Λ is an extension of **K.alt**$_1$.**D** and this proves the claim. Finally, to see that the lattice of extensions is countable, observe that **K.alt**$_1$.**D** has at most as many proper extensions as there are finite subsets of $\omega \times \omega$. $\omega \times \omega$ is countable, and the set of finite sets of a finite set is countable again. So $\mathcal{E}\,$**K.alt**$_1$.**D** is countable. Now, an extension Λ of **K.alt**$_1$ is characterized by the number of the largest chain admitted by it and the logic $\Lambda \sqcup$ **K.alt**$_1$.**D**. So there are at most $\omega \times \omega$ logics, again a countable number. □

From this theorem we can get a complete overview of the lattice $\mathcal{E}\,$**K.alt**$_1$ by describing the corresponding T_D–space. Recall that if $\mathfrak{S}$ is a poset then $\Phi(\mathfrak{S})$ is the space of the topology having as closed sets all unions of sets of the form $\uparrow x$. Now, given two topological spaces $\mathfrak{X}$ and $\mathfrak{Y}$, define $\mathfrak{X} \rtimes \mathfrak{Y}$ as follows. A set is closed iff it is a finite union of sets $A \times B$ where (i) $A \neq X$ or (ii) $A = X$ and $B = Y$. It is immediately verified that this is a topological space.

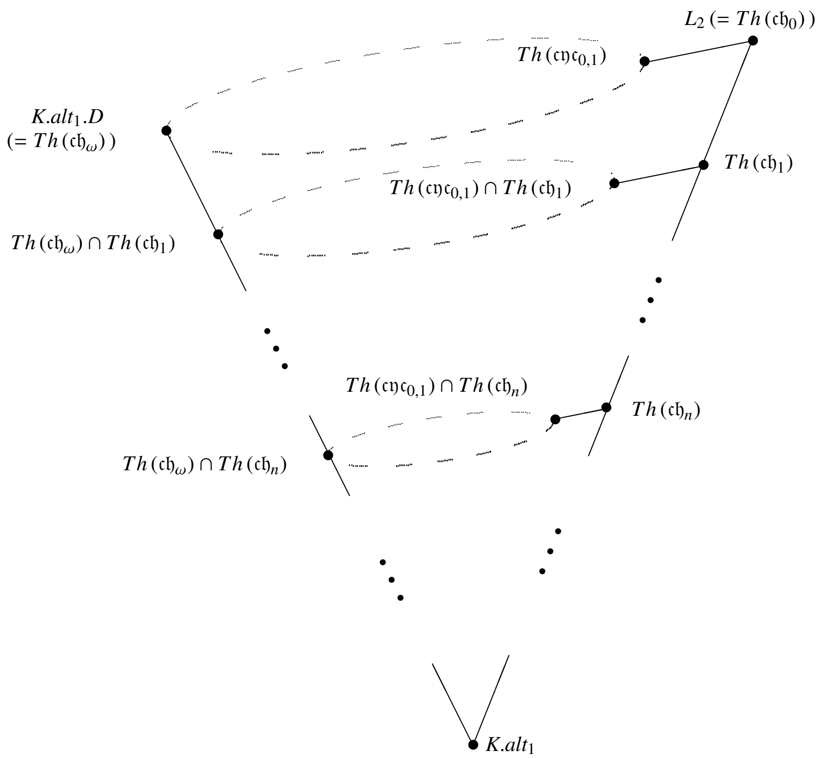
Theorem 9.1.4. *Let* $\alpha := \langle \omega, \leq \rangle^{op}$, *and* $\mu := \langle \omega - \{0\}, | \rangle^{op}$.

$$\begin{array}{lcl} \mathfrak{ISpc}(\mathcal{E}\,\boldsymbol{K.alt}_1.\boldsymbol{D}) & \cong & \Phi(\alpha \times \mu) \\ \mathfrak{ISpc}(\mathcal{E}\,\boldsymbol{K.alt}_1) & \cong & \Phi(\mu) \rtimes \Phi(\alpha \times \mu) \end{array}$$

α derives from the additive structure over ω ($n \leq m$ iff there exists a k such that $n + k = m$), and μ from the multiplicative structure over $\omega - \{0\}$ (with $n \mid m$ iff there exists a $k > 0$ such that $nk = m$). Note that $\alpha = \langle \omega, \geq \rangle$, but for stating the theorem it is better to display the similarity between α and μ. For a proof of the first fact notice that every proper extension is tabular, and so its corresponding set in the space is finite. Hence the topology is indeed the weak topology. Now, an extension of **K.alt**$_1$ is an intersection of an extension of **K.alt**$_1$.**D** with a logic of the form $\mathsf{Th}\,\mathfrak{ch}_n$. This representation is unique. Hence, any logic extending **K.alt**$_1$ corresponds to a closed set in $\mathfrak{ISpc}$(**K.alt**$_1$.**D**) and a closed set in $\Phi(\alpha)$. However, if the latter is the entire space, the extension is not proper, and conversely. In that case, the first set is the entire space as well. This shows the correctness of the representation.

Theorem 9.1.5. *All logics in* $\mathcal{E}\,\boldsymbol{K.alt}_1$ *are finitely axiomatizable, have the global finite model property and are globally decidable.*

FIGURE 9.1. $\mathcal{E}\,\boldsymbol{K.alt}_1$



PROOF. If $\Lambda = \mathbf{K.alt}_1.\mathbf{D}$, the claim is clearly true, so let Λ be a proper extension. Every extension of $\mathbf{K.alt}_1.\mathbf{D}$ is either this logic itself or tabular. In that case, all three claims follow. The logics $\mathsf{Th}\,\mathfrak{ch}_n$ also have all these properties. Any finite intersection of logics which have the global finite model property also has the global finite model property, so we need to show that the intersection is finitely axiomatizable. This is unproblematic if $\Lambda \supsetneq \mathbf{K.alt}_1.\mathbf{D}$, since in that case Λ is tabular. Thus, we have to study logics of the form $\mathbf{K.alt}_1.\mathbf{D} \sqcap \mathsf{Th}\,\mathfrak{ch}_n$. In that case $\Lambda = \mathbf{K.alt}_1 \oplus \neg\Diamond^n\top \vee \Diamond^{n+1}\top$. ◻

PROPOSITION 9.1.6. *Every quasi–normal extension of $\boldsymbol{K.alt}_1.\boldsymbol{D}$ is normal.*

PROOF. A quasi–normal extension is the theory of some set of pointed frames $\langle \mathfrak{f}, w_0\rangle$, where w_0 is the root of $\mathfrak{f}$. We will show for every $x \in f$ we have $\mathsf{Th}\,\langle \mathfrak{f}, w_0\rangle \subseteq \mathsf{Th}\,\langle \mathfrak{f}, x\rangle$. It follows that $\mathsf{Th}\,\langle \mathfrak{f}, w_0\rangle = \mathsf{Th}\,\mathfrak{f}$. So, take an $\mathfrak{f}$. **Case 1.** $\mathfrak{f} = \mathfrak{ch}_\omega$. Then the claim is immediate. **Case 2.** $\mathfrak{f} = \mathfrak{cyc}_{m,p}$. Then the transit of x in $\mathfrak{f}$ is of the form $\mathfrak{cyc}_{n,p}$

for $n \leq m$. By Proposition 9.1.1 there exists a map $\mathfrak{cyc}_{m,p} \twoheadrightarrow \mathfrak{cyc}_{n,p}$. It maps the root of the former frame onto the root of the latter frame. Hence $\mathsf{Th}\,\langle \mathfrak{f}, w_0 \rangle \subseteq \mathsf{Th}\,\langle \mathfrak{f}, x \rangle$. ◻

Proposition 9.1.7. *An extension of* ***K.alt***$_1$ *is Halldén–complete iff it is the logic of* $\mathfrak{ch}_1$, $\mathfrak{ch}_\omega$ *or* $\mathfrak{cyc}_{n,p}$ *for some* $n, p \in \omega$.

Proof. A logic is Halldén–complete only if it is the logic of the one–element chain or contains the axiom **D**. Let the latter be the case. In case Λ is the theory of a single frame $\mathfrak{cyc}_{u,v}$, it is the quasi–normal theory of $\langle \mathfrak{cyc}_{u,v}, x \rangle$, where x is the generating point. Hence, using Theorem 1.6.5 we see that the logic is Halldén–complete. If it is not the logic of a single frame, then there exist frames $\mathfrak{cyc}_{m,p}$ and $\mathfrak{cyc}_{n,q}$ such that there is no frame for Λ having both of them as its p–morphic image. Call these frames **minimal**. There exists φ such that φ is refutable on $\mathfrak{cyc}_{m,p}$ but not on $\mathfrak{cyc}_{n,q}$, and a ψ which is refutable on $\mathfrak{cyc}_{m,q}$ but not on $\mathfrak{cyc}_{n,p}$. We can assume that φ and ψ are disjoint in variables. Then $\varphi \vee \psi$ is not refutable on either frame, or of a p–morphic image of the two. So, if Λ is characterized by these two minimal frames, $\varphi \vee \psi$ cannot be refuted on any frame for Λ, though φ and ψ both are refutable. Likewise we can proceed if Λ is characterized by n minimal frames, picking a formula φ_n for each minimal frame. (Choose, for example, the diagrams of each of these frames, prefixed by a sufficiently large $\Box^{\leq n}$.) ◻

Exercise 321. Show that **K.alt**$_1$ has $2^{\aleph_0}$ many quasi–normal extensions.

Exercise 322. Show that **K.alt**$_2$ has $2^{\aleph_0}$ many normal extensions. *Hint.* These extensions can even be axiomatized by constant axioms.

Exercise 323. Show that **K.alt**$_1$.**D** is the only pretabular logic in $\mathcal{E}$ **K.alt**$_1$.

Exercise 324. Show that Λ splits $\mathcal{E}$ **K.alt**$_1$.**D** iff $\Lambda = \mathsf{Th}\,\mathfrak{ch}_n$ for some $n \in \omega$. The logics **K.alt**$_1 / \mathfrak{ch}_n$ are called *Chellas–Hughes–Logics* in [**197**]. Show that they can be axiomatized by constant formulae.

Exercise 325. Show that $\mathcal{E}$ **K.alt**$_1$.**D** has exactly one splitting, namely the reflexive, one–point frame.

Exercise 326. Show that every element in $\mathcal{E}$ **K.alt**$_1$.**D** has $\aleph_0$ many lower covers.

Exercise 327. Show that $\mathcal{E}$ **K.alt**$_1$.**D** has $2^{\aleph_0}$ automorphisms.

Exercise 328. (Spaan [**202**].) Show that **K.alt**$_1$ is NP–complete. *Hint.* Estimate the size of models.

9.2. Polymodal Logics with Quasi–Functional Operators

In this section we will turn to the logics containing $\bigotimes_{i<\kappa} \mathbf{K.alt}_1$. The principal aim is to define a set of formulae with which any extension can be axiomatized and which is geometrically perspicuous.

Definition 9.2.1. *Let Θ be a modal logic, and φ, ψ formulae. We say that φ and ψ are **axiomatically equivalent over** Θ if $\Theta \oplus \varphi = \Theta \oplus \psi$.*

Obviously, if φ and ψ are deductively equivalent in Θ then they are also axiomatically equivalent. The converse does not hold, however. For example, if $\psi = p$ and $\varphi = q$, $p \neq q$, then $\Theta \oplus \varphi = \Theta \oplus \bot = \Theta \oplus \psi$. But if Θ is consistent, φ and ψ are not deductively equivalent. For then $p \leftrightarrow q \in \Theta$, from which $p \leftrightarrow \neg p \in \Theta$, hence $\bot \in \Theta$.

Recall that a model for polymodal $\mathbf{K.alt}_1$ is thought of as a frame or alternatively as a partial algebra with κ many partial unary functions. An extension of polymodal $\mathbf{K.alt}_1.\mathbf{D}$ is complete with respect to *semigroup* models. To be precise, take a frame $\langle f, \langle \lhd_j : j < \kappa \rangle\rangle$. Let $x \in f$ and $j < \kappa$. Then if x has a j–successor y we put $j(x) := y$ and let $j(x)$ be undefined otherwise. For a sequence σ we define inductively $\sigma(x)$ to be that element which can be reached from x 'following' the path defined by σ, if that path exists. In extensions of polymodal $\mathbf{K.alt}_1.\mathbf{D}$ all paths exist, and they form a semigroup under concatenation. Now, our first aim is to derive a special normal form for formulae. Notice, namely that we have for each operator the theorems

$$\begin{array}{ll} \Diamond_j p. \leftrightarrow .\Diamond_j \top \wedge \Box_j p\ , & \Diamond_j (p \wedge q). \leftrightarrow .\Diamond_j p \wedge \Diamond_j q\ , \\ \Box_j p. \leftrightarrow .\Box_j \bot \vee \Diamond_j p\ , & \Box_j (p \vee q). \leftrightarrow .\Box_j p \vee \Box_j q\ . \end{array}$$

If we apply the standard algorithm for obtaining normal forms of Section 2.7 we observe that a normal form of degree $n+1$ is a disjunction of formulae ψ of some ψ_i, $i < n$, where each ψ_i is a conjunction of (i) a normal form μ of degree 0 and (ii) for each j, either $\Box_j \bot$ or $\Box_j \chi \wedge \Diamond_j \chi$, where χ is some normal form form of degree n. The following definition was proposed in [**91**].

Definition 9.2.2. *A formula φ is a **strict canonical** $\boldsymbol{alt}_1$**–formula** if it is of the following form*

$$\bigwedge_{\sigma \in W^\top} \Diamond^\sigma \top \wedge \bigwedge_{\sigma \notin W^\bot} \neg \Diamond^\sigma \top \wedge \bigwedge_{i<k} \Diamond^{\sigma_i} p \leftrightarrow \Diamond^{\tau_i} p$$

*Here, σ_i, τ_i are finite sequences of indices, and $W^\top, W^\bot$ finite sets of such sequences. A formula φ is a **canonical** $\boldsymbol{alt}_1$**–formula** if it is a disjunction of strict canonical formulae φ_i, $i < n$, such that $var(\varphi_i) \cap var(\varphi_j) = \varnothing$ if $i \neq j$.*

Proposition 9.2.3 (Grefe). *There is an effective algorithm reducing each formula φ into a canonical $\bigotimes \boldsymbol{K.alt}_1$–formula ψ which is axiomatically equivalent with φ over $\bigotimes \boldsymbol{K.alt}_1$.*

Proof. **Step 1.** Let φ be given. Eliminate all symbols different from variables, $\top$, $\neg$, $\wedge$, $\vee$ and $\Diamond_j$ by their respective definitions. Next move unary operators into the scope of binary ones, i. e. negation into the scope of conjunction and disjunction, and $\Diamond_j$ into the scope of conjunction and disjunction as well. The latter is admissible, by the laws of distribution given above. Next, replace subformulae of the form $\Diamond_j\neg\psi$ by $\neg\Diamond_j\psi \wedge \Diamond_j\top$. Finally, move conjunction out of the scope of disjunction. If none of these operations are applicable, we have obtained a formula $\varphi^{(1)}$ which is deductively equivalent to φ and of the form $\bigwedge_{i<m}\varphi_i^{(1)}$, where for each i there exist finite sets of finite sequences over κ, W^+ and W^-, such that

$$\varphi_i^{(1)} := \bigvee_{\sigma\in W^\top} \Diamond^\sigma\top \vee \bigvee_{\sigma\in W^\perp} \neg\Diamond^\sigma\top \vee \bigvee_{\sigma\in W^+} \Diamond^\sigma p_{\sigma,i} \vee \bigvee_{\sigma\in W^-} \neg\Diamond^\sigma p_{\sigma,i}$$

The variables $p_{\sigma,i}$ need not be distinct for distinct σ or i. **Step 2.** There exists a renaming of the variables which makes the formulae $\varphi_i^{(1)}$ disjoint in variables. In general, the axiom $\alpha \wedge \beta$ is equal in force with the axiom $\alpha \dot{\wedge} \beta$, which results from making α and β disjoint in variables. Now assume that $p_{\sigma,i}$ occurs only positively in $\varphi_i^{(1)}$. Then we can replace $\Diamond^\sigma p_{\sigma,i}$ by $\perp$. Likewise, if $p_{\sigma,i}$ occurs only negatively. (That this results in an axiomatically equivalent formula has been the content of an exercise of Section 5.5.) Now take Σ^+ to be the set of paths in W^+ with associated variable $p = p_{\sigma,i}$ and Σ^- the set of paths in W^- with associated variable p. By construction, Σ^+ and Σ^- are not empty. Let χ be the disjunction

$$\chi := \bigvee_{\sigma\in\Sigma^+} \Diamond^\sigma p \vee \bigvee_{\sigma\in\Sigma^-} \neg\Diamond^\sigma p$$

Replace χ by χ' which is defined by

$$\chi' := \bigvee_{\sigma\in\Sigma^+,\tau\in\Sigma^-} \Diamond^\sigma p_{\sigma,\tau} \vee \neg\Diamond^\tau p_{\sigma,\tau}$$

We show that χ and χ' are axiomatically equivalent. Since χ is a substitution instance of χ', we must have $\bigotimes \mathbf{K.alt}_1 \oplus \chi \subseteq \bigotimes \mathbf{K.alt}_1 \oplus \chi'$. The converse inclusion remains to be shown. Now let $\mathfrak{g} \nvDash \chi'$, say $\langle \mathfrak{g}, x\rangle \nvDash \chi'$. Then all paths τ from Σ^- must exist. Furthermore, assume a path $\sigma \in \Sigma^+$ exists starting from x. Then its endpoint is distinct from all endpoints of paths τ starting from x. Let $\beta(p) := \{\sigma(x) : \sigma \in \Sigma^-\}$. Then $\langle \mathfrak{g}, \beta, x\rangle \nvDash \chi$, as required. Let $\varphi_i^{(2)}$ result from $\varphi_i^{(1)}$ by performing the substitution just mentioned. A slight modification of the previous argument shows that $\varphi_i^{(2)}$ is axiomatically equivalent to $\varphi_i^{(1)}$. This concludes the second step. Put $\varphi^{(2)} := \bigwedge_{i<m}\varphi_i^{(2)}$. $\varphi^{(2)}$ is a conjunction of formulae of the type $\varphi_i^{(2)}$ with the distinctive property that each variable occurs in at most one $\varphi_i^{(2)}$ and there exactly once positively and once negatively. **Step 3.** For each variable p of $\varphi^{(2)}$, there exists an $i < m$ and two subformulae of $\varphi_i^{(2)}$ of the form $\Diamond^\sigma p$ and $\neg\Diamond^\tau p$. $\varphi_i^{(2)}$ is therefore deductively equivalent to a formula $\psi \vee \Diamond^\sigma p \vee \neg\Diamond^\tau p$ for some ψ not containing p. Replace $\varphi_i^{(2)}$ by $\psi \vee \neg\Diamond^\tau\top \vee (\Diamond^\sigma p \leftrightarrow \Diamond^\tau p)$. This results in a deductively equivalent formula. Perform this operation for every occurring variable. The resulting formulae are denoted by $\varphi_i^{(3)}$. Rewrite $\varphi_i^{(3)}$ by moving disjunction out of the scope of conjunction. After that $\varphi_i^{(3)}$ is a disjunction of formulae $\psi_{i,j}$ which are conjunctions of formulae which are either constant and of the form $\Diamond^\sigma\top$ or $\neg\Diamond^\tau\top$ or they are of the form $\Diamond^\sigma p \leftrightarrow \Diamond^\tau p$. Moreover, $\varphi^{(3)}$ is deductively equivalent to the disjunction of the $\psi_{i,j}$ for some suitable set E of pairs $\langle i, j\rangle$. In each $\psi_{i,j}$, replace all variables by a single one among them. This yields the formula $\tau_{i,j}$. By choice of the variables, $var(\tau_{i,j}) \cap var(\tau_{m,n}) = \varnothing$ for $\langle i, j\rangle \neq \langle m, n\rangle$. It is easy to check that $\tau_{i,j}$ and $\psi_{i,j}$ are axiomatically equivalent and so is their disjunction. Finally, let $\varphi^{(4)}$ be the disjunction of the $\tau_{i,j}$, $\langle i, j\rangle \in E$. $\varphi^{(4)}$ is axiomatically equivalent to $\varphi^{(3)}$. $\varphi^{(4)}$ is in alt$_1$–canonical form. ◻

Proposition 9.2.4. *Every extension of polymodal $\boldsymbol{K.alt}_1$ can be axiomatized by formulae of the form*

$$\chi \vee \bigvee_{i<d}\left(\bigwedge_{\sigma\in W_i^\top}\Diamond^\sigma\top \wedge \bigwedge_{\sigma\in W_i^\perp}\neg\Diamond^\sigma\top \wedge \bigwedge_{j<k}(\Diamond^{\sigma_{ij}}\top \wedge \Diamond^{\tau_{ij}}\top \wedge \Diamond^{\sigma_{ij}}p_i \leftrightarrow \Diamond^{\tau_{ij}}p_i)\right)$$

where $var(\chi) = \varnothing$ and $\sigma_{ij} \neq \tau_{ij}$.

Corollary 9.2.5. *Every extension of $\bigotimes \boldsymbol{K.alt}_1.\boldsymbol{D}$ is axiomatizable by formulae of the form $\bigvee_{i<m}\Diamond^{\sigma_i}p_i \leftrightarrow \Diamond^{\tau_i}p_i$ for some finite sequences σ_i, τ_i. Moreover, these axioms correspond in the class of differentiated frames to*

$$(\forall x)[\bigvee_{i<m}(\forall y \rhd^{\sigma_i} x)(\forall z \rhd^{\tau_i} x)(y \doteq z)]\ .$$

Corollary 9.2.6. *Each extension of polymodal $\boldsymbol{K.alt}_1$ is of special Sahlqvist rank 0. Every extension of polymodal $\boldsymbol{K.alt}_1.\boldsymbol{D}$ is of pure, special Sahlqvist rank 0.*

We can cash out on this characterization as follows. Call a statement of the form

$$\Diamond^\sigma\top \wedge \Diamond^\tau\top \wedge \Diamond^\sigma p \leftrightarrow \Diamond^\tau p$$

a **path–equation**, and denote it by $\sigma \approx \tau$. A path equation states that both paths exist and that they are equal:

$$(\exists y \rhd^\sigma x)\mathrm{t} \wedge (\exists y \rhd^\tau x)\mathrm{t} \wedge (\forall y \rhd^\sigma x)(\forall z \rhd^\tau x)(y \doteq z)$$

An extension of polymodal $\mathbf{K.alt}_1$ is then characterized by a disjunction of conjunctions of path equations and constant formulae. The constant formulae are either existence or nonexistence statements for certain paths. An extension of $\bigotimes \mathbf{K.alt}_1\mathbf{D}$ is axiomatizable by a set of disjunctions of path equations.

Finally, recall that an elementary class of frames is modally definable iff it is closed under contractions, generated subframes, disjoint unions and ultrafilter extensions, while also the complement is closed under ultrafilter extensions. In the present context the last restriction can be dropped.

Proposition 9.2.7. ($\kappa < \aleph_0$.) *A class of frames for logics of bounded operator alternative is modally definable iff it is closed under contractions, generated subframes, disjoint unions and ultrafilter extensions.*

Proof. We have to show that the class is inversely closed under ultrafilter extensions. We show that $\mathfrak{g}$ is isomorphic to a generated subframe of $\mathsf{ue}(\mathfrak{g})$, which will establish the theorem. The elements of $\mathfrak{g}$ correspond to the principal ultrafilters of $\mathsf{ue}(\mathfrak{g})$. Thus we have to show that no principal ultrafilter can see a non–principal ultrafilter. So let U_x be the ultrafilter of sets containing x and let T be non–principal. The set of points seen in one step by x via j is finite, consisting, say, of $y_0, \dots, y_{n-1}$, where $n \leq \kappa$. For each $i < n$ there exists a set a_i such that $y_i \in a_i$ and $a_i \notin T$. Put $b := \bigcup_{i<n} a_i$. Then $b \notin T$, since T is an ultrafilter. On the other hand, $x \in \blacksquare_j b$. Thus $U_x \not\lhd_j T$. □

Corollary 9.2.8. ($\kappa < \aleph_0$.) *A class of frames of bounded operator alternativity is modally definable iff it is closed under contractions, generated subframes, disjoint unions and* ***ultraproducts****.*

Proof. Let $\mathcal{K}$ be modally definable. Then it is elementary, since its logic is r–persistent. It follows that $\mathcal{K}$ is closed under ultraproducts. Conversely, assume that $\mathcal{K}$ is closed under ultraproducts. By Theorem 5.7.18, the ultrafilter extension of a frame $\mathfrak{f}$ is a contractum of an ultrapower of $\mathfrak{g}$. Therefore $\mathcal{K}$ is closed under ultrafilter extensions as well. □

The lattice of extensions of $\mathcal{E} \bigotimes_{j<\kappa} \mathbf{K.alt}_1$ is quite complex despite the fact that the extensions are axiomatizable by such simple axioms. We will restrict our attention to $\kappa = 2$, but it is clear that the results can be generalized.

Proposition 9.2.9. *$\mathcal{E}(\boldsymbol{K.alt}_1 \otimes \boldsymbol{K.alt}_1)$ contains countably many tabular logics of codimension 1. $\mathcal{E}(\boldsymbol{K.alt}_1.\boldsymbol{D} \otimes \boldsymbol{K.alt}_1.\boldsymbol{D})$ contains countably many tabular logics of codimension 2.*

Proof. Consider the frames $\mathfrak{f}\text{–}\mathfrak{cyc}_n$ defined on the set $2n + 1$ where $i \lhd j$ iff $i + 1 = j$ and i even, or $i = 2n - 1$ and $j = 2n$, and $i \blacktriangleleft j$ iff $j = i + 1$ and i is odd or $i = 2n$ and $j = 0$. The frames $\mathfrak{f}\text{–}\mathfrak{cyc}_n$ are almost like a cycle where the relations between neighbouring elements alternate. There is one position where this regularity is disturbed and this accounts for the fact that $\mathfrak{f}\text{–}\mathfrak{cyc}_n$ has no proper contracta. Since it has no proper generated subframes either, its logic is of codimension 1. In bimodal $\mathbf{K.alt}_1.\mathbf{D}$ we have exactly one logic of codimension 1, which is the logic of the one–point frame, which is reflexive with respect to all relations. Now for this case take a prime number $q \neq 2$ and define on q the relations $\lhd$ and $\blacktriangleleft$ as follows. We let $0 \lhd 0$ and $q - 1 \blacktriangleleft q - 1$. For even numbers i we put $i \blacktriangleleft i + 1 \blacktriangleleft i$ and for odd numbers i we put $i \lhd i + 1 \lhd i$. No other relations hold. This defines the frame $\mathfrak{d}\text{–}\mathfrak{cyc}_q$. Now define a function $f : q \to q$ by taking $f(i)$ to be the $\lhd$–successor of the $\blacktriangleleft$–successor of i. It is verified in the case of $q = 5$ that $f(0) = 2$, $f(2) = 4$, $f(4) = 5$, $f(5) = 3$, $f(3) = 1$ and $f(1) = 0$. In general, we have $f^q(x) = x$ but $f^r(x) \neq x$ for $r < q$. The powers of f form a cyclic group of order q under composition. Now take a p–morphism π of $\mathfrak{d}$–$\mathfrak{cyc}_q$ onto a frame $\mathfrak{e}$. Then the function f induces a function g on $\mathfrak{e}$ defined as follows; $g(x) := \pi(f(y))$ where $y \in \pi^{-1}(x)$. (By the conditions on p–morphisms this is well–defined.) Moreover, the powers of g under composition must form a group, too. For g is invertible, since $g^q(x) = x$. However, the cyclic group of order q has only two images, itself and the group of order 1. In the first case p must be injective, in the second case $\mathfrak{e}$ has one element. For $\mathfrak{d}$–$\mathfrak{cyc}_q$ is connected under f, that is, for any two elements i, j there exists an r such that $j = f^r(i)$. $\mathfrak{e}$ must have the same property with respect to g. So $\mathfrak{e}$ has one element, and thus the theory of $\mathfrak{d}$–$\mathfrak{cyc}_q$ has codimension 2. ◻

Since the finite frames are countable for finite κ, this is the best possible result. Since all extensions of polymodal $\mathbf{K.alt}_1$ are complete this leaves us with the possibility of maximal consistent logics which are complete but not tabular, that is, which are determined by a single, infinite frame. The key to such logics lies in the existence of infinite sequences over a given alphabet which are nonrepeating. These are sequences in which no subword is eventually repeated over and over. By a simple cardinality argument there are $2^{\aleph_0}$ sequences which are nonrepeating. However, the result that we are going to prove is much stronger; it shows that there are $2^{\aleph_0}$ many infinite sequences over the alphabet $\{0, 1\}$ such that no subword is repeated five times consecutively.

Definition 9.2.10. *For a natural number k and a finite sequence σ, let $k \times \sigma$ denote the sequence consisting of the k–fold repetition of σ. Let $\alpha : \omega \to \{0, 1\}$ be an infinite sequence and $\sigma = s_0 s_1 \dots s_{n-1}$ a finite string over $\{0, 1\}$. The* ***index*** *of σ in α, $H_\alpha(\sigma)$ is the maximal ordinal number $j < \omega + 1$ such that α contains a subword of the form $j \times \sigma$. That is to say, $H_\alpha(\sigma) = j$ iff there is an r such that for all $i < j$ and $e < n$, $\alpha(r + n \cdot i + e) = s_e$.*

For an infinite sequence α let $\mathfrak{g}_\alpha$ be the frame based on ω and $i \lhd j$ iff $j = i + 1$ and $\alpha(i) = 0$ and $i \blacktriangleleft j$ iff $j = i + 1$ and $\alpha(i) = 1$. If $H_\alpha(s) = \omega$, then there is a point from which the sequence α repeats the sequence σ periodically. Put $E := \{a : H_\alpha(\sigma) \in \omega, \textit{for all } \sigma \in \{0, 1\}^*\}$. Thus E is the set of sequences in which every subword has finite index. It is clear that E has the cardinality $2^{\aleph_0}$. It is the logics of $\mathfrak{g}_\alpha$ for $\alpha \in E$ that we want to use as examples of logics of codimension 1. However, there are two obstacles. The difficulty lies in showing that (i) they are of codimension 1 and (ii) that they all give rise to distinct logics. We will solve these problems as follows. We show first that even if the theory of $\mathfrak{g}_\alpha$ is not maximal, it has no finite frames. The second problem is solved by constructing special sequences, in which even $H_\alpha(\sigma) \leq 4$ for all σ.

Lemma 9.2.11. *Let σ be a non–empty finite string with index m. Then*

$$\mathfrak{g}_\alpha \vDash \neg\Diamond^{(m+1)\times\sigma}\top.$$

LEMMA 9.2.12. *Let* $\mathfrak{g}_\alpha$ *be a frame such that* $\alpha \in E$ *and let* $\mathfrak{f}$ *be a finite frame. Then* $\mathsf{Th}\,\mathfrak{g}_\alpha \not\subseteq \mathsf{Th}\,\mathfrak{f}$.

PROOF. Let $\mathfrak{F} = \langle f, \lhd, \blacktriangleleft\rangle$. It is enough to show this for such finite $\mathfrak{f}$ for which the theory is of codimension 1. In that case $\mathfrak{f}$ has no proper generated subframes. **Case 1.** $\mathfrak{f}$ is the one–element frame with empty relations. Then $\Diamond\top \vee \blacklozenge\top \in \mathsf{Th}\,\mathfrak{g}_\alpha - \mathsf{Th}\,\mathfrak{f}$. **Case 2.** Not both $\lhd$ and $\blacktriangleleft$ are empty. Let $\mathfrak{f}$ have n elements. Then $\mathfrak{f}$ is a full cycle, that is, there is a sequence $\sigma = s_0 s_1 \dots s_{n-1}$ over $\{0, 1\}$ and a world w_0 such that $w_{i+1} = s_i(w_i)$, $i < n-1$ and $w_0 = \sigma_{n-1}(w_{n-1})$. Put $m := H_\alpha(\sigma)$. Then $\neg\Diamond^{(m+1)\times\sigma}\top \in \mathsf{Th}\,\mathfrak{g}_\alpha - \mathsf{Th}\,\mathfrak{f}$, as required. ◻

DEFINITION 9.2.13. *Let* $X \subseteq \omega$. *We define a sequence* a_X *to be the concatenation of the words* z^X_n, *which are defined as follows. We put* $z^X_n := l^X_n r^X_n$. l^X_n *and* r^X_n *— or simply* l_n *and* r_n *from now on — are defined inductively by*

$$\begin{array}{llllllll} l_0 & := & 0, & r_0 & := & 1, & \text{if} & 0 \notin X, \\ l_0 & := & 00, & r_0 & := & 11, & \text{if} & 0 \in X, \\ l_{n+1} & := & l_n r_n, & r_{n+1} & := & r_n l_n, & \text{if } n+1 \notin X, & \\ l_{n+1} & := & l_n r_n l_n r_n, & r_{n+1} & := & r_n l_n r_n l_n, & \text{if } n+1 \in X. & \end{array}$$

In order to understand this definition, let us note some properties of this sequence. First, it can be divided from left to right into blocks of equal length of the form $l_0 r_0$ or $r_0 l_0$, and starting from the first occurrence of l_n, it can be divided into blocks of the form $l_n r_n$ and $r_n l_n$. These blocks have length 2^k for some k. Let us define a new sequence from a_X by forgetting z^X_0 and dividing the remaining sequence into parts of the form $l_0 r_0$ and $r_0 l_0$, as indicated. The sequence $l_0 r_0$ is replaced by 0 (its first symbol) and $r_0 l_0$ is replaced by 1 (again its first symbol). The resulting sequence is a sequence of the form a_Y where $Y = \{n : n+1 \in X\}$. The proof is by induction. So, if the sequence taken from a certain point onwards and is thinned by taking only the 2^k–next symbol, we get a sequence of type a_Y. In particular, if $0 \in X$ then we may replace in the entire sequence blocks of the form $l_0 r_0 = 0011$ by 0 and blocks of the form $r_0 l_0 = 1100$ by 1. Then we get a sequence of the form a_Y with $Y = \{n : n+1 \in X\}$. The point where the sequence $l_n r_n$ appears for the first time, will be denoted by o_n. Furthermore, if $0 \notin X$ then the subsequence $\langle a_X(2n), a_X(2n+1)\rangle$ contains a zero and a one, and if $0 \in X$ then $\langle a_X(4n), a_X(4n+1), a_X(4n+2), a_X(4n+3)\rangle$ is either 0011 or 1100. Thus, in any subsequence the number of zeros may differ from the number of ones by at most four. (For each subsequence of length can be divided into a sequence $\vec{p} \cdot \vec{w}_0 \cdot \vec{w}_1 \cdot \ldots \cdot \vec{w}_{n-1} \cdot \vec{s}$ where the $\vec{w}_i$ are equal to 0011 or 1100, $\vec{p}$ a postfix of 1100 or 0011 of length ≤ 3, and $\vec{s}$ a prefix of length ≤ 3 of 0011 or 1100. Each $\vec{w}_i$ contains two ones and two zeros, so it is balanced. In $\vec{p}$ and $\vec{s}$ the number of ones and zeros can differ at most by 2.)

LEMMA 9.2.14 (Grefe). *For all* $X \subseteq \omega$ *and all finite strings* σ $H_{a_X}(\sigma) \leq 4$. *In particular* $a_X \in E$.

Proof. Let $lng(\sigma)$ denote the length of a string. Let σ have the length $2^k \cdot u$, u an odd number. Suppose, a_X contains the subsequence $5 \times \sigma$. **Case 1.** There is an n such that $lng(l_n r_n) = 2^k$. Let $5 \times \sigma$ begin at the number r and let $t \in \omega$ the unique number such that $o_n + (t-1) \cdot 2^k < r \leq o_n + t \cdot 2^k$. (This number exists since $o_n < lng(l_n r_n) = 2^k$.) Then $a_X(o_n + (t+i) \cdot 2^k) = a_X(o_n + i + j \cdot u) \cdot 2^k)$ for all $i < lng(s)$ and $j < 5$. We may then form the sequence $b(i) := a(o_n + 2^k \cdot i)$ and get $b = a_Y$ for some Y. Then $\langle b(t), b(t+1), \dots, b(t+u-1)\rangle$ has uneven length. So it has one more zeros than ones ore one more ones than zeros. Since this sequence repeats five times consecutively, we have found a subsequence in which the number of ones and zeros differs by at least 5, in contradiction to the fact that b is of the form a_Y. **Case 2.** There is no n such that $lng(l_n r_n) = 2^k$. Then there is an n such that $lng(l_n r_n l_n r_n) = 2^k$. Now proceed as in Case 1. □

Theorem 9.2.15 (Grefe). *The lattice $\mathcal{E}\,\mathsf{Th}\,\mathfrak{ch}_4 \otimes \mathsf{Th}\,\mathfrak{ch}_4$ has $2^{\aleph_0}$ many logics of codimension 1.*

Proof. The monomodal fragments of $\mathsf{Th}\,\mathfrak{g}_{a_X}$ contain the theory of the four–element chain, since neither 0 nor 1 may be iterated five times. $\mathsf{Th}\,\mathfrak{g}_{a_X}$ is contained in a maximal consistent logic. It is enough to show that for $X \neq Y$, $\mathsf{Th}\,\mathfrak{g}_{a_X} \sqcup \mathsf{Th}\,\mathfrak{g}_{a_Y}$ is inconsistent. Let n be the smallest number on which X and Y disagree, say, $n \in X - Y$. Now, assume that $n = 0$. Then $\mathfrak{g}_{a_Y} \vDash \Box\Box\Box\bot$. However, r_2^X contains four consecutive zeros. Moreover, l_2^X may not be repeated more than four times, so that wherever we are, after at most $k = 5 \cdot lng(l_2^X)$ steps r_2^X must occur, and hence a sequence with three consecutive zeros. Thus $\mathfrak{g}_{a_Y} \vDash \neg \boxtimes^{\leq k} \Diamond\Diamond\Diamond\top$. For $n = m + 1$ the proof is performed with the subword l_m in place of 0 and r_{m+2}^X in place of r_2^X. □

Notes on this section. In Edith Spaan [**202**] is is shown that $\mathbf{K.alt}_1$ is locally NP–complete and globally PSPACE–complete, while polymodal $\mathbf{K.alt}_1$ is locally NP–complete and globally EXPTIME–complete.

Exercise 329. With this set of exercises we will shed some more light on the results by Wim Blok of the preceding chapter. First, show that any tabular κ–modal logic is co–covered by $2^{\aleph_0}$ complete logics, $\kappa > 1$. *Hint.* Use the fact that the inconsistent logic has this property.

Exercise 330. Show that any tabular monomodal logic contained in $\mathsf{Th}\,\boxed{\bullet}$ is co–covered by $2^{\aleph_0}$ complete logics. *Hint.* Use the simulation theorem.

Exercise 331. Let the degree of **nonfinite model property** of a logic Λ be defined to be the cardinality of the spectrum of logics with the same class of finite models as Λ. Show that a tabular κ–modal logic has degree of nonfinite model property $2^{\aleph_0}$ unless $\kappa = 1$ and $\Lambda = \mathbf{K}_1 \oplus \bot$.

Exercise 332. (Spaan [**202**].) Show that polymodal $\mathbf{K.alt}_1$ is NP–complete.

9.3. Colourings and Decolourings

A bimodal frame in which each operator satisfies $\mathbf{alt}_1$ can be seen as a kind of monomodal $\mathbf{alt}_2$–frame in which the successors are discriminated by the modalities. We can define two operations on these frames going in opposite directions; we call them *colouring* and *decolouring*. Decolouring is the process of forgetting the distinction between the relations, and colouring is the (re)introduction of the distinction. Any monomodal Kripke–frame can in principle be treated in this way, given enough additional relations. If each point of the frame has at most κ many successors then given κ many operators the successors can be distinguished by the new accessibility relations. Each of the new accessibility relations satisfies $\mathbf{K.alt}_1$ and may thus be interpreted as a partial function on the frame. In this way, modal logic seems generally reducible to the study of (polymodal) $\mathbf{K.alt}_1$. Moreover, alternative semantics in terms of functions on frames rather than relations — as proposed e. g. in van Benthem [**13**] — can be studied as well. Therefore, in this section we will show how to perform such a reduction and where its limits are. Basically, the limitation is to logics which are of bounded alternativity. The reason is that otherwise the logics are not necessarily complete and the operation of colouring may not be definable on enough frames to guarantee that the logic is recoverable from its coloured counterpart.

There are two approaches to colouring and decolouring. One is to view this as a relation between bimodal and monomodal frames. The other is to view it as a relation between bimodal and trimodal frames. In the first approach the monomodal frame contains the relation that is 'split' in its bimodal coloured version. In the second approach the three modalities are merged. So, the three–modal frame is obtained by a fusion of the bimodal coloured frame and its uncoloured variant. (This has been brought to my attention by Kit Fine.) Both approaches have their merits. Prima facie colouring seems to be an operation that is always defined. Unfortunately, we have to introduce two restrictions. Namely, we must restrict ourselves to Kripke–frames, and moreover to Kripke–frames in which a point has a $\lhd$–successor iff it has a $\blacktriangleleft$–successor.

Definition 9.3.1. *Let* $\mathfrak{f} = \langle f, \lhd, \blacktriangleleft \rangle$ *be a Kripke–frame for* $\boldsymbol{K.alt}_1 \otimes \boldsymbol{K.alt}_1(\Diamond\top \leftrightarrow \blacklozenge\top)$*. Then* $\mathsf{D}(\mathfrak{f}) := \langle f, \lhd \cup \blacktriangleleft \rangle$ *is called the* ***decolouring*** *of* $\mathfrak{f}$*. Conversely, if* $\mathfrak{g} = \langle g, \leqslant \rangle$ *is a monomodal Kripke–frame then any* $\mathfrak{f}$ *such that* $\mathsf{D}(\mathfrak{f}) = \mathfrak{g}$ *is called a* ***colouring*** *of* $\mathfrak{g}$*. We denote by* $\mathsf{C}(\mathfrak{g})$ *the set of colourings of* $\mathfrak{g}$*, and for a class* $\mathcal{K}$ *of monomodal Kripke–frames,* $\mathsf{C}\,\mathcal{K} := \mathsf{D}^{-1}[\mathcal{K}]$.

Let us first present an example of an uncolourable frame. Take the set $f = \omega \times 4$. Let $\mathfrak{f} = \langle f, \leqslant \rangle$ where

$$\begin{array}{rcl} \leqslant & := & \{\langle\langle n,0\rangle, \langle n,0\rangle\rangle : n \in \omega\} \\ & \cup & \{\langle\langle n,0\rangle, \langle n,1\rangle\rangle : n \in \omega\} \\ & \cup & \{\langle\langle n,2\rangle, \langle n,3\rangle\rangle : n \in \omega\} \\ & \cup & \{\langle\langle n,2\rangle, \langle n+1,3\rangle\rangle : n \in \omega\} \end{array}$$

Put $Y := (\omega \times \{0\}) \cup (\omega \times \{2\})$ and $X := (\omega \times \{1\}) \times (\omega \times \{3\})$. The sets X and Y are definable by constant formulae, for example $\boxminus\bot$ and $\Diamond \boxminus \bot$. Finally, let $\mathbb{F}$ be the algebra of sets generated by the finite sets. This is exactly the algebra $\mathbb{A}$ of sets a such that $a \cap X$ is finite or cofinite in X and $a \cap Y$ is finite or cofinite in Y. To show that $\mathbb{A} = \mathbb{F}$, let us first note that $\mathbb{A}$ contains all finite sets. Moreover, $\mathbb{A} \subseteq \mathbb{F}$ because if we have a finite set a, we have $a \cap X \in \mathbb{F}$ and $a \cap Y \in \mathbb{F}$, as well as $X \cap (-a)$ and $Y \cap (-a)$. So we can compose any set which is finite or cofinite in X and finite or cofinite in Y. $\mathbb{F} \subseteq \mathbb{A}$ is clear. It is also routine to check that $\mathbb{A}$ is closed under the natural operations.

Proposition 9.3.2. $\mathfrak{F} := \langle \mathfrak{f}, \mathbb{A}\rangle$ *is not colourable.*

Proof. Let $\langle f, \lhd, \blacktriangleleft, \mathbb{A}\rangle$ be a colouring of $\langle f, \leqslant, \mathbb{A}\rangle$. Then $\leqslant \,=\, \lhd \cup \blacktriangleleft$. Any point in $\omega\times\{2\}$ sees exactly one point in $\omega\times\{3\}$ in each direction. Hence the sets $\Diamond(\omega\times\{3\})$ and $\blacklozenge(\omega\times\{3\})$ are both cofinite in $\omega\times\{2\}$. But $\Diamond(\omega\times\{1\})$ and $\blacklozenge(\omega\times\{1\})$ cannot both be cofinite in $\omega\times\{0\}$. Without loss of generality we assume that $\omega\times\{0\} - \Diamond(\omega\times\{1\})$ is infinite. Then $\Diamond X = \Diamond(\omega\times\{1\}) \cup \Diamond(\omega\times\{3\})$ is neither finite nor cofinite in Y, and so not in $\mathbb{A}$. Thus no colouring exists on $\mathfrak{F}$. ∎

The next question is the restriction to frames which validate the axiom $\blacklozenge\top \leftrightarrow \Diamond\top$. If we are interested in the mere operations of colouring and decolouring such a restriction is strictly speaking unnecessary. However, consider the monomodal frame $\mathfrak{m} := \langle\{0,1\}, \leqslant\rangle$ where $\leqslant := \{\langle 0,1\rangle, \langle 1,1\rangle\}$. Consider the bimodal frame $\mathfrak{b} := \langle\{0,1\}, \lhd, \blacktriangleleft\rangle$ where $\lhd = \{\langle 0,1\rangle\}$ and $\blacktriangleleft = \{\langle 1,1\rangle\}$. Then $\mathfrak{m}$ has a one–point contractum, but $\mathfrak{b}$ does not. This is not a good situation; we want that a colouring of $\mathfrak{m}$ must be a frame $\mathfrak{c}$ such that each p–morphic image $\mathfrak{n}$ of $\mathfrak{m}$ can be coloured in such a way that it is a contractum of $\mathfrak{c}$. This gives rise to the restriction that a point must have a $\lhd$–successor iff it has a $\blacktriangleleft$–successor.

Definition 9.3.3. *Let* Λ *be an extension of* $\boldsymbol{K.alt}_2$. *Then* Λ*, defined by* $\Lambda^c := \mathsf{Th}\,(\mathsf{C}\,[\mathsf{Krp}\,\Lambda])$ *is called the* ***colouring of*** Λ. *Let* Θ *be an extension of the logic*

$$\bigotimes_2 \boldsymbol{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top\ .$$

Then $\Theta^d := \mathsf{Th}\,(\mathsf{D}\,[\mathsf{Krp}\,\Theta])$ *is called the* ***decolouring kernel of*** Θ.

In this section we will be concerned with the algebraic properties of these maps, with a syntactic description of the operations of colouring and decolouring as well as the properties of the corresponding semantic maps. First, as it is easy to see, colouring and decolouring are isotone maps between the poset reducts of the lattices of logics.

Theorem 9.3.4. *The colouring map (as a map between posets) is left adjoined to the decolouring kernel map. Moreover, for every extension* Θ *of* $\boldsymbol{K.alt}_2$, $\Theta = \Theta^{cd}$.

Proof. We have $\mathsf{DD}^{-1}[\mathcal{K}] = \mathcal{K}$ for all classes of monomodal Kripke–frames. Putting $\mathcal{K} := \mathsf{Krp}\,\Theta$ shows $\Theta = \Theta^{cd}$. For by Kripke–completeness, $\Theta = \mathsf{Th}\,\mathsf{Krp}\,\Theta$.

Let Λ be a bimodal logic, Θ a monomodal logic. Then $\Lambda \supseteq \Theta^c$ iff $\mathsf{Krp}\,\Lambda \subseteq \mathsf{Krp}\,\Theta^c = \mathsf{C}\,[\mathsf{Krp}\,\Theta]$. The latter implies $\mathsf{D}[\mathsf{Krp}\,\Lambda] \subseteq \mathsf{DD}^{-1}\,[\mathsf{Krp}\,\Theta]$ and so $\mathsf{D}\,\mathsf{Krp}\,\Lambda \subseteq \mathsf{Krp}\,\Theta$, since $\mathsf{DD}^{-1}\,[\mathcal{K}] = \mathcal{K}$ for all classes of Kripke–frames. The latter is equivalent to $\Lambda^d \supseteq \Theta$. It is checked that the converse holds as well. ◻

The connection between frames and colourings is introduced by the following translation.

Definition 9.3.5. *The map* $^c : \mathcal{P}_1 \to \mathcal{P}_2$ *is defined by*

$$\begin{array}{lcl} p^c & := & p \\ (\varphi \wedge \psi)^c & := & \varphi^c \wedge \psi^c \\ (\neg\varphi)^c & := & \neg\varphi^c \\ (\Diamond\varphi)^c & := & \Diamond\varphi^c \vee \blacklozenge\varphi^c \end{array}$$

Moreover, $\Delta^c := \{\varphi^c : \varphi \in X\}$ *for a set* $\Delta \subseteq \mathcal{P}_1$.

Proposition 9.3.6. *Let* $\mathfrak{b}$ *be a Kripke–frame for* $\boldsymbol{K.alt}_1 \otimes \boldsymbol{K.alt}_1(\Diamond\top \leftrightarrow \blacklozenge\top)$*,* $x \in b$ *and* β *a valuation on* $\mathfrak{b}$*. Then*

$$\langle \mathfrak{b}, \beta, x\rangle \vDash \varphi^c \qquad \Leftrightarrow \qquad \langle \mathsf{D}(\mathfrak{b}), \beta, x\rangle \vDash \varphi$$

Moreover, $\mathfrak{b} \vDash \varphi^c$ *iff* $\mathsf{D}(\mathfrak{b}) \vDash \varphi$.

Proof. By induction on φ. Only the step for $\varphi = \Diamond\psi$ is nonobvious.

$$\begin{array}{ll} & \langle \mathfrak{b}, \beta, x\rangle \vDash (\Diamond\psi)^c \\ \text{iff} & \langle \mathfrak{b}, \beta, x\rangle \vDash \Diamond\psi^c \text{ or } \langle \mathfrak{b}, \beta, x\rangle \vDash \blacklozenge\psi^c \\ \text{iff} & \text{for some } y \rhd x \quad \langle \mathfrak{b}, \beta, y\rangle \vDash \psi^c \text{ or for some } y \blacktriangleright x \quad \langle \mathfrak{b}, \beta, y\rangle \vDash \psi^c \\ \text{iff} & \text{for some } y(\rhd \cup \blacktriangleright)x \quad \langle \mathfrak{b}, \beta, y\rangle \vDash \psi^c \\ \text{iff} & \text{for some } y \geqslant x \quad \langle \mathsf{D}(\mathfrak{b}), \beta, y\rangle \vDash \psi \\ \text{iff} & \langle \mathsf{D}(\mathfrak{b}), \beta, x\rangle \vDash \varphi. \end{array}$$

The second claim follows immediately. If $\mathfrak{b} \nvDash \varphi^c$ then $\langle \mathfrak{b}, \beta, x\rangle \vDash (\neg\varphi)^c$ for some β, x. Hence $\langle \mathsf{D}(\mathfrak{b}), \beta, x\rangle \vDash \neg\varphi$ and so $\mathsf{D}(\mathfrak{b}) \nvDash \varphi$. And conversely. ◻

The syntactic correlate of colouring of a bimodal logic is also rather easily defined.

Proposition 9.3.7. *Suppose* $\Lambda = \mathbf{K.alt}_2 \oplus \Delta$*. Then* $\Lambda^c = \bigotimes_2 \mathbf{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \Delta^c$.

Proof. Let $\mathfrak{m}$ be a monomodal frame for $\mathbf{K.alt}_2$. From Proposition 9.3.6 we deduce that $\mathfrak{m} \vDash \varphi$ iff $\mathsf{C}(\mathfrak{m}) \vDash \varphi^c$. Alternatively, $\mathfrak{m} \in \mathsf{Krp}\,\mathbf{K.alt}_2 \oplus \varphi$ iff $\mathsf{C}(\mathfrak{m}) \subseteq \mathsf{Krp} \bigotimes_2 \mathbf{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \varphi^c$. From this the theorem follows. ◻

Theorem 9.3.8. *The colouring map is a homomorphic embedding of the locale* $\mathcal{E}\,\boldsymbol{K.alt}_2$ *into the locale* $\mathcal{E}\,(\bigotimes_2 \mathbf{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top)$.

Proof. Using the results of Section 2.9 it is not hard to see that colouring commutes with (infinite) joins. To see that it commutes with meets, suppose that $\Lambda_1 = \mathbf{K}.\mathbf{alt}_1 \oplus \Delta_1$ and $\Lambda_2 = \mathbf{K}.\mathbf{alt}_2 \oplus \Delta_2$. Then

$$\begin{array}{ll} & \Lambda_1^c \sqcap \Lambda_2^c \\ = & (\bigotimes_2 \mathbf{K}.\mathbf{alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \Delta_1^c) \sqcap (\bigotimes_2 \mathbf{K}.\mathbf{alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \Delta_2^c) \\ = & \bigotimes_2 \mathbf{K}.\mathbf{alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \{\boxtimes^{\leq m}\varphi \dot{\vee} \boxtimes^{\leq m}\psi : \varphi \in \Delta_1^c, \psi \in \Delta_2^c, m \in \omega\} \\ = & \bigotimes_2 \mathbf{K}.\mathbf{alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top \oplus \{(\boxtimes^{\leq m}\varphi \dot{\vee} \boxtimes^{\leq m}\psi)^c : \varphi \in \Delta_1, \psi \in \Delta_2, m \in \omega\} \\ = & (\Lambda_1 \sqcap \Lambda_2)^c \end{array}$$

(Note that the definition of $\boxtimes$ depends on the language. Hence the step from the second to the third line is correct.) Finally, we have to show that the colouring is injective. So, let $\Lambda := \mathbf{K}.\mathbf{alt}_2 \oplus \Delta$ and $\Theta := \mathbf{K}.\mathbf{alt}_2 \oplus \Delta'$ and $\Lambda \neq \Theta$. Without loss of generality there exists a Kripke–frame $\mathfrak{m}$ such that $\mathfrak{m} \vDash \Delta$, but $\mathfrak{m} \nvDash \psi$ for some $\psi \in \Delta'$. Then there exists a colouring $\mathfrak{b}$ of $\mathfrak{m}$ such that $\mathfrak{b} \nvDash \psi^c$. $\mathfrak{b} \vDash \Delta^c$. This shows $\Lambda^c \neq \Theta^c$. ∎

Proposition 9.3.9. *Let Θ be an extension of $\bigotimes_2 \boldsymbol{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top$. Then* $\Theta^d = \{\varphi : \varphi^c \in \Theta\}$.

Proof. $\varphi \in \Theta^d$ iff $\varphi \in \mathsf{Th}\,\mathsf{D}[\mathsf{Krp}\,\Theta]$ iff for all $\mathfrak{b} \in \mathsf{Krp}\,\Theta$ we have $\mathsf{D}(\mathfrak{b}) \vDash \varphi$ iff for all $\mathfrak{b} \in \mathsf{Krp}\,\Theta$ we have $\mathfrak{b} \vDash \varphi^c$ iff $\varphi^c \in \Theta$. ∎

Theorem 9.3.10. *The decolouring operation* D *commutes with the class operators of taking generated subframes, contraction images, disjoint unions and ultraproducts.*

The proof is routine and left as an exercise. We note that from Corollary 9.2.8 we derive

Corollary 9.3.11. *For any class $\mathcal{K}$ of bimodal Kripke–frames for the logic $\bigotimes_2 \boldsymbol{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top$ it holds that* $\mathsf{D}\,\mathsf{M}\,\mathcal{K} = \mathsf{M}\,\mathsf{D}\,\mathcal{K}$. *Here,* $\mathsf{M}\,\mathcal{K} := \mathsf{Krp}\,\mathsf{Th}\,\mathcal{K}$.

Theorem 9.3.12 (Grefe). *The colouring map map leaves invariant **finite axiomatizability** and **local finite model property**.*

Proof. Finite model property is clear from the definition and the construction of decolouring. Moreover, if Λ is finitely axiomatizable, so is Λ^c, by Proposition 9.3.7. For the converse, let Λ^c be finitely axiomatizable. Then $\mathsf{Krp}\,\Lambda^c$ is an elementary class. Therefore $\mathsf{Krp}\,\Lambda = \mathsf{D}\,\mathsf{Krp}\,\Lambda^c$ is an elementary class, since D commutes with Up, by Theorem 9.3.10. ∎

Theorem 9.3.13. C *commutes with taking generated subframes and disjoint unions of Kripke–frames. Moreover, for contraction images* Cn, $\mathsf{Cn}\,\mathsf{C}\,\mathcal{K} \subseteq \mathsf{C}\,\mathsf{Cn}\,\mathcal{K}$.

Proof. **Generated Suframes.** Let $\mathfrak{g}$ be a Kripke–frame, and $\mathfrak{f} \rightarrowtail \mathfrak{g}$. Then any colouring of $\mathfrak{g}$ induces a colouring of $\mathfrak{f}$. Any colouring on $\mathfrak{f}$ can be extended to a colouring of $\mathfrak{g}$. **Disjoint Unions.** Let $\mathfrak{g}_i$ be Kripke–frames, $i \in I$. Pick a colouring

$\mathfrak{h}_i$ on each $\mathfrak{g}_i$. It is easy to check that $\bigoplus_{i\in I}\mathfrak{h}_i$ is a colouring of $\bigoplus_{i\in I}\mathfrak{g}_i$. Also, every colouring on $\bigoplus_{i\in I}\mathfrak{g}_i$ gives rise to a colouring of each individual $\mathfrak{g}_i$. **Contraction Images**. Assume that $\mathfrak{g}$ and $\mathfrak{h}$ are monomodal Kripke–frames, and $\pi : \mathfrak{g} \twoheadrightarrow \mathfrak{h}$. Let $\mathfrak{h}^c$ be a colouring of $\mathfrak{h}$. We have to show that there is a colouring $\mathfrak{g}^c$ of $\mathfrak{g}$ such that $\pi : \mathfrak{g}^c \twoheadrightarrow \mathfrak{h}^c$. So let $x, y \in g$ and $x \leqslant y$. we have to decide whether $x \lhd y$ or whether $x \blacktriangleleft y$. Three cases arise. (i.) x has a single successor. This successor is y. Then we put $x \lhd y$ and $x \blacktriangleleft y$. (ii.) x has two successors y and z and $\pi(y) \neq \pi(z)$. Then $x \lhd y$ iff $\pi(x) \lhd \pi(y)$ and $x \blacktriangleleft y$ iff $\pi(x) \blacktriangleleft \pi(y)$; and similarly for z. (iii.) x has two successors y and z but $\pi(y) = \pi(z)$. In this case we choose as to whether $x \lhd y$ and so $x \blacktriangleleft z$ or $x \blacktriangleleft y$ and so $x \lhd z$. With this definition made, let us show that $\pi : \mathfrak{g}^c \twoheadrightarrow \mathfrak{h}^c$. So, let $x \lhd y$. Then $x \leqslant y$ and so $\pi(x) \leqslant \pi(y)$. It remains to be seen that $\pi(x) \lhd \pi(y)$. If $\pi(y)$ is the only $\leqslant$–successor of $\pi(x)$, then $\pi(x) \lhd \pi(y)$ by virtue of definition of a colouring. So assume that $\pi(x)$ has two successors, u_1 and u_2. By construction, if $x \lhd u_1$ then $u_1 = \pi(y)$, and so $\pi(x) \lhd \pi(y)$. Similarly the case $x \blacktriangleleft y$ is dealt with. This shows the first condition on p–morphisms. For the second condition, assume $\pi(x) \lhd u$. Then $\pi(x) \leqslant u$ and for some y we have $x \leqslant y$ and $\pi(y) = u$. Then $x \blacktriangleleft y$ or $x \lhd y$. Clearly, x has a $\leqslant$–successor. **Case 1.** x has exactly one successor. Then both $x \blacktriangleleft y$ and $x \lhd y$ and we are done. **Case 2.** x has two successors and their images are distinct. Then $x \lhd y$ by definition of $\lhd$. **Case 3.** x has two successors y_1, y_2, and both are being mapped onto u. Then we have $x \lhd y_1$ or $x \lhd y_2$, by construction. This concludes the proof of the fact that π is a p–morphism. ◻

Lemma 9.3.14. *Let $\mathcal{K}$ be a class of Kripke–frames for* $\mathbf{K.alt}_2$. *Then* $\mathsf{Up}\,\mathsf{C}\,\mathcal{K} \subseteq \mathsf{C}\,\mathsf{Up}\,\mathcal{K}$.

Proof. Let $\mathfrak{g}_i$, $i \in I$, be Kripke–frames in $\mathcal{K}$, and choose a colouring $\mathfrak{g}_i^c$ of each frame. Let U be an ultrafilter on I. We want to show that $\prod_U \mathfrak{g}_i^c$ is isomorphic to a colouring of an ultraproduct of the $\mathfrak{g}_i^c$. We take $\mathfrak{h} := \prod_U \mathfrak{g}_i$. We let $\overline{x}_U \lhd \overline{y}_U$ iff $\{i : x_i \lhd y_i\} \in U$. This definition does not depend on the choice of the representatives. Moreover, it shows that the identity map is an isomorphism from $\prod_U \mathfrak{g}_i^c$ onto $\mathfrak{h}^c$. ◻

Theorem 9.3.15 (Grefe). *Let $\mathcal{K}$ be a modally definable class of Kripke–frames for $\boldsymbol{K.alt}_2$. Then $\mathsf{C}\,\mathcal{K}$ is a modally definable class of Kripke–frames for $\boldsymbol{K.alt}_1 \otimes \boldsymbol{K.alt}_1(\Diamond\top \leftrightarrow \blacklozenge\top)$.*

Proof. By the previous theorems, $\mathsf{C}\,\mathcal{K}$ is closed under generated subframes, contractions, disjoint unions and ultraproducts. By Corollary 9.2.8 we get that $\mathsf{M}\,\mathsf{C}\,\mathcal{K} \subseteq \mathsf{C}\,\mathsf{M}\,\mathcal{K}$, since M consists in taking generated subframes of contractions of ultraproducts. If $\mathcal{K} = \mathsf{M}\,\mathcal{K}$ then $\mathsf{C}\,\mathcal{K} \subseteq \mathsf{M}\,\mathsf{C}\,\mathcal{K} \subseteq \mathsf{C}\,\mathsf{M}\,, \mathcal{K} = \mathsf{C}\,\mathcal{K}$. And so equality holds. ◻

Theorem 9.3.16. $\mathsf{Th}\,\mathsf{Up}\,\mathsf{C}\,\mathcal{K} = \mathsf{Th}\,\mathsf{C}\,\mathsf{Up}\,\mathcal{K}$.

Proof. The inclusion '$\supseteq$' follows from Lemma 9.3.14. For the other inclusion assume $\varphi \notin \mathsf{Th}\ \mathsf{C}\ \mathsf{Up}\,\mathcal{K}$. Then for some $\mathfrak{g}_i$ there exists a colouring $\mathfrak{h}$ of $\prod_U \mathfrak{g}_i$ such that $\mathfrak{h} \nvDash \varphi$, say $\langle \mathfrak{h}, \beta, \overline{x}_U\rangle \vDash \neg\varphi$. Let δ be the modal depth of φ. Now look at the

δ–transit of $\overline{x}_U$ in $\mathfrak{h}$. It is finite, and hence

$$A := \{i : \mathfrak{Tr}^{\delta}_{\mathfrak{g}_i}(x_i) \cong \mathfrak{Tr}^{\delta}_{\mathfrak{h}^d}(\overline{x}_U)\} \in U$$

For each $i \in A$ fix an isomorphism ι_i from the δ–transit of x_i in $\mathfrak{g}_i$ onto the δ–transit of $\overline{x}_U$ in $\mathfrak{h}$. We colour the δ–transit of x_i in $\mathfrak{g}_i$ in the way prescribed by $\mathfrak{h}$. That is, we put $x_i \lhd y_i$ iff $\iota_i(x_i) \lhd \iota_i(y_i)$. This defines for each $i \in I$ a partial colouring (if $i \notin A$, then nothing is prescribed so far). Choose any colouring on the $\mathfrak{g}_i$ that extends the partial colouring. This defines $\mathfrak{g}_i^c$. Then let $\mathfrak{k} := \prod_U \mathfrak{g}_i^c$. It is routine to show that the δ–transit of $\overline{x}_U$ in $\mathfrak{k}$ is isomorphic to the δ–transit of $\overline{x}_U$ in $\mathfrak{h}$. Hence $\mathfrak{k} \nvDash \varphi$. And $\mathfrak{k} \in \mathsf{Up}\,\mathsf{C}\,\mathcal{K}$. ◻

THEOREM 9.3.17. *For any class $\mathcal{K}$ of $\boldsymbol{K.alt}_2$–Kripke–frames*

$$\mathsf{M}\,\mathsf{C}\,\mathcal{K} = \mathsf{C}\,\mathsf{M}\,\mathcal{K}$$

PROOF. Both are modal classes. And their theory is identical. Hence they are identical. ◻

Exercise 333. Suppose Λ^c is defined on the basis of an axiomatization as in Proposition 9.3.7. Verify syntactically that this definition is independent of the chosen axiomatization of Λ.

Exercise 334. Let Θ be an extension of $\bigotimes_2 \mathbf{K.alt}_1 \oplus \Diamond\top \leftrightarrow \blacklozenge\top$. Show syntactically that $\{\varphi : \varphi^c \in \Theta\}$ is a logic.

Exercise 335. Show Theorem 9.3.10.

Exercise 336. Show with a particular counterexample that $\mathsf{C}\,\mathsf{Cn}\,\mathcal{K} \subseteq \mathsf{Cn}\,\mathsf{C}\,\mathcal{K}$ does not hold in general.

***Exercise 337.** Generalize the results of this section to polymodal logics. Show that extensions of n–modal $\mathbf{K.alt}_1$ can be interpreted as extensions of $\mathbf{K.alt}_n$.

9.4. Decidability of Logics

Consider a finite set A of symbols, called **alphabet**. By A^* we denote the set of finite strings over A, including the empty string, denoted by ϵ. Strings will be denoted by a vector arrow, e. g. $\vec{x}$, $\vec{y}$ etc. An **equation** is a pair $\langle \vec{v}, \vec{w}\rangle \in A^* \times A^*$, written $\vec{v} \approx \vec{w}$. A **Thue–process** over A is a finite set of equations. Given a Thue–process $\mathbb{P}$ we write $\vec{y} \approx^1_{\mathbb{P}} \vec{z}$ iff there exist $\vec{c}, \vec{d} \in A^*$ and $\vec{v} \approx \vec{w} \in \mathbb{P}$ such that $\vec{y} = \vec{c} \cdot \vec{v} \cdot \vec{d}$ and $\vec{z} = \vec{c} \cdot \vec{w} \cdot \vec{d}$ or $\vec{y} = \vec{c} \cdot \vec{w} \cdot \vec{d}$ and $\vec{z} = \vec{c} \cdot \vec{v} \cdot \vec{d}$. We say also that $\vec{z}$ is **one–step derivable** from $\vec{y}$. Thus, $\vec{z}$ is one–step derivable from $\vec{y}$ iff it can be produced from $\vec{y}$ by replacing a substring matching one side of an equation in $\mathbb{P}$ by the other side of that equation. We define $\vec{y} \approx^n_{\mathbb{P}} \vec{z}$ inductively by $\vec{y} \approx^0_{\mathbb{P}} \vec{z}$ iff $\vec{y} = \vec{z}$ and $\vec{y} \approx^{n+1}_{\mathbb{P}} \vec{z}$ iff there exists $\vec{x}$ such that $\vec{y} \approx^1_{\mathbb{P}} \vec{x} \approx^n_{\mathbb{P}} \vec{z}$. Finally, $\vec{y} \approx_{\mathbb{P}} \vec{z}$ iff there exists an $n \in \omega$ such that

$\vec{y} \approx^n_{\mathbb{P}} \vec{z}$. We say that $\vec{z}$ is **n–step derivable from** $\vec{y}$ if $\vec{y} \approx^n_{\mathbb{P}} \vec{z}$ and that it is **derivable** if $\vec{y} \approx_{\mathbb{P}} \vec{z}$. A Thue–process is **decidable** if for given $\vec{y}, \vec{z}$ it is decidable whether or not $\vec{z} \approx \vec{y}$ is derivable in it. A Thue–process is called **trivial** if $\vec{x} \approx \epsilon$ is derivable for all $\vec{x} \in A^*$ iff $a \approx \epsilon$ is derivable for all $a \in A$. The following facts will be used in sequel.

Theorem 9.4.1 (Post, Markov, Rabin). *Let A contain at least two symbols.*

1. *There exists undecidable Thue–processes.*
2. *The set of decidable Thue–processes is undecidable.*
3. *The set of trivial Thue–processes is undecidable.*

These statements were shown in [**163**], [**156**] and [**167**], respectively. Proofs of these fact can also be found in many textbooks. A Thue–process can be seen as a presentation of a *semigroup*. Recall from Section 1.3 the notion of a presentation. Let SG the theory of semigroups in the language $1, \cdot$. For example, $SG = \{x \approx x \cdot 1, x \approx 1 \cdot x, (x \cdot y) \cdot z \approx x \cdot (y \cdot z)\}$. The free semigroup over A, $\mathfrak{Fr}_{SG}(A)$ is isomorphic to $\langle A^*, \epsilon, \cdot\rangle$. Given a Thue–Process $\mathbb{P}$, $\langle A, \mathbb{P}\rangle$ is a presentation of a semigroup denoted by $\mathfrak{Fr}_{SG}(A)/\mathbb{P}$, and we have $\vec{x} \approx_{\mathbb{P}} \vec{y}$ iff $\vec{x}$ and $\vec{y}$ are identical elements of $\mathfrak{Fr}_{SG}(A)/\mathbb{P}$. This is the most standard way of thinking about Thue–processes. We can derive two rather useful results. First, take the first–order theory of two unary functions, f_0 and f_1. These functions are interpreted on a semigroup generated by two elements, say 0 and 1, as concatenating a word with one of the two generators. Namely, we put $f_0(x) := x \cdot 0$ and $f_1(x) := x \cdot 1$. Since 0 and 1 generate this group, for any word $\vec{u}$ the function $f_{\vec{u}} : x \mapsto x \cdot \vec{u}$ can be composed from these two functions. Hence an equation $\vec{u} \approx \vec{v}$ can be rewritten into the statement

$$(\ddagger) \qquad (\forall x)(f_{\vec{u}}(x) \doteq f_{\vec{v}}(x))$$

Now, a semigroup satisfies the equation $\vec{u} \approx \vec{v}$ iff it satisfies the first–order condition $(\forall x)(f_{\vec{u}}(x) \doteq f_{\vec{v}}(x))$. Thus, the first–order theory of the semigroup generated by a Thue–process $\mathbb{P}$ is exactly axiomatized by the the axioms ($\ddagger$) for $\langle \vec{u}, \vec{v}\rangle \in \mathbb{P}$. So, we can encode the derivability of the Thue–process into first–order logic over two unary functions. The unary function f_0, on the other hand, can be viewed as a binary relation $\lhd_0$ which satisfies two postulates. (i) $(\forall xyz)(x \lhd_0 y \wedge x \lhd_0 z. \rightarrow .y \doteq z)$ and (ii) $(\forall x)(\exists y)(x \lhd_0 y)$. Similarly with f_1. Hence we conclude that the $\forall\exists$–theory (i. e. the set of sentences of that theory which are of complexity at most $\forall\exists$) of two binary relations is undecidable. Given a Thue–process we build a canonical process–frame $\mathfrak{c}_{\mathbb{P}}$ as follows. $\approx_{\mathbb{P}}$ is a congruence on A^*. The equivalence class of $\vec{x}$ in this congruence is denoted by $[\vec{x}]$. So, $[\vec{x}] = \{\vec{y} : \vec{y} \approx_{\mathbb{P}} \vec{x}\}$. Then $c_{\mathbb{P}} := \{[\vec{x}] : \vec{x} \in A^*\}$. We put $A := \{w, b\}$ (w stands for *white* and b for *black*.)

$$\begin{array}{lcl} A^*/\mathbb{P} & := & \{[\vec{x}] : \vec{x} \in A^*\} \\ \lhd & := & \{\langle [\vec{z}], [\vec{z} \cdot w]\rangle : \vec{z} \in A^*\} \\ \blacktriangleleft & := & \{\langle [\vec{z}], [\vec{z} \cdot b]\rangle : \vec{z} \in A^*\} \\ \mathfrak{c}_{\mathbb{P}} & := & \langle A^*/\mathbb{P}, \lhd, \blacktriangleleft\rangle \end{array}$$

In what is to follow we will not always distinguish between sequences and the equivalence classes of these sequences modulo $\mathbb{P}$. Fix a sequence $\vec{x}$. The map $\pi : [\vec{y}] \mapsto [\vec{x}] \cdot [\vec{y}]$ is a p–morphism of $\mathfrak{c}_{\mathbb{P}}$ onto the transit of $[\vec{x}]$. For if $[\vec{y}] \lhd [\vec{z}]$ then $[\vec{z}] = [\vec{y} \cdot w]$, and then also $[\vec{x}] \cdot [\vec{z}] = [\vec{x}] \cdot [\vec{y} \cdot w]$, from which $\pi([\vec{y}]) = [\vec{x}] \cdot [\vec{y}] \lhd [\vec{x}] \cdot [\vec{z}] = \pi([\vec{z}])$. Next, if $\pi([\vec{y}]) = [\vec{x}] \cdot [\vec{y}] \lhd [\vec{u}]$, then $[\vec{u}] = [\vec{x}] \cdot [\vec{y} \cdot w]$. So, let $[\vec{v}] := [\vec{y} \cdot w]$. Then $\pi([\vec{v}]) = [\vec{x}] \cdot [\vec{y} \cdot w] = [\vec{u}]$. And $[\vec{y}] \lhd [\vec{v}]$. Similarly for $\blacktriangleleft$.

With a Thue process we associate two logics, namely

$$\begin{aligned} \Sigma_{\mathbb{P}} &= \mathbf{K.alt}_1 \otimes \mathbf{K.alt}_1 \oplus \{\Diamond^{\vec{v}} p \leftrightarrow \Box^{\vec{w}} p : \vec{v} \approx \vec{w} \in \mathbb{P}\} \\ \Lambda_{\mathbb{P}} &= \mathbf{K.alt}_1.\mathbf{D} \otimes \mathbf{K.alt}_1.\mathbf{D} \oplus \{\Diamond^{\vec{v}} p \leftrightarrow \Box^{\vec{w}} p : \vec{v} \approx \vec{w} \in \mathbb{P}\} \\ &= \Sigma_{\mathbb{P}} \oplus \Diamond\top \oplus \blacklozenge\top \end{aligned}$$

The following theorem is easy to prove with the help of Theorem 3.5.3.

Proposition 9.4.2. *Let $\mathbb{P}$ be a Thue–process. $\Sigma_{\mathbb{P}}$ is a subframe logic. It is of (pure) Sahlqvist rank 0 and its frames are characterized by the properties*

$$(\forall x)(\forall y \rhd^{\vec{v}} x)(\forall z \rhd^{\vec{w}} x)(y \doteq z), \qquad \langle \vec{v}, \vec{w} \rangle \in \mathbb{P}$$

$\Sigma_{\mathbb{P}}$ has the finite model property and is locally decidable.

Proposition 9.4.3. *Let $\mathfrak{g}$ be a rooted Kripke–frame for $\Lambda_{\mathbb{P}}$. Then there exists a contraction $\mathfrak{c}_{\mathbb{P}} \twoheadrightarrow \mathfrak{g}$.*

Proof. Let w_0 be the root of $\mathfrak{g}$. Then let $s : A^* \to g$ be defined by $s(\epsilon) := w_0$. Further, $s(\vec{x} \cdot w)$ is the unique element y such that $s(\vec{x}) \lhd y$. Similarly, $s(\vec{x} \cdot b)$ is the unique element y' such that $s(\vec{x}) \blacktriangleleft y'$. It is checked that if $\vec{x} \approx_{\mathbb{P}} \vec{y}$ then $s(\vec{x}) = s(\vec{y})$. Hence, the function defines a function $t([\vec{x}]) := s[([\vec{x}])]$. It is readily verified that t is a p–morphism. It is onto by the fact that $\mathfrak{g}$ is rooted at w_0. ∎

Proposition 9.4.4. $\mathsf{Th}\,\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle = \mathsf{Th}\,\mathfrak{c}_{\mathbb{P}} = \Lambda_{\mathbb{P}}$.

Proof. We know that $\mathsf{Th}\,\langle \mathfrak{c}_{\mathbb{P}}, [\vec{x}]\rangle \supseteq \mathsf{Th}\,\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle$ from the fact that the transit of $[\vec{x}]$ is the p–morphic image of $\mathfrak{c}_{\mathbb{P}}$. Hence $\mathsf{Th}\,\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle$ is normal and identical to $\mathsf{Th}\,\mathfrak{c}_{\mathbb{P}}$. Now, $\Lambda_{\mathbb{P}}$ and $\mathsf{Th}\,\mathfrak{c}_{\mathbb{P}}$ are extensions of $\mathbf{K.alt_1.D} \otimes \mathbf{K.alt_1.D}$. First of all, $\mathfrak{c}_{\mathbb{P}}$ is a frame for $\Lambda_{\mathbb{P}}$. Hence $\mathsf{Th}\,\mathfrak{c}_{\mathbb{P}} \supseteq \Lambda_{\mathbb{P}}$. Also, let $\mathfrak{g}$ be a rooted Kripke–frame for $\Lambda_{\mathbb{P}}$. Then $\mathfrak{g}$ is a contractum of $\mathfrak{c}_{\mathbb{P}}$. So, $\mathsf{Th}\,\mathfrak{c}_{\mathbb{P}}$ and $\Lambda_{\mathbb{P}}$ have the same rooted Kripke–frames. Hence — being complete — they are identical. ∎

Theorem 9.4.5. *$\Lambda_{\mathbb{P}}$ is decidable iff $\mathbb{P}$ is decidable.*

Proof. Suppose that $\mathbb{P}$ is undecidable. Then the problem '$\vec{x} \approx_{\mathbb{P}} \vec{y}$' is undecidable. Hence the problem '$\Diamond^{\vec{x}} p \leftrightarrow \Diamond^{\vec{y}} p \in \Lambda_{\mathbb{P}}$' is undecidable (since it is equivalent to '$\vec{x} \approx_{\mathbb{P}} \vec{y}$'). Now suppose that $\mathbb{P}$ is decidable. Given a formula φ, we want to be able to decide whether it is satisfiable in a $\Lambda_{\mathbb{P}}$–frame. By Proposition 9.4.4, it suffices to be able to decide whether φ is satisfiable in $\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle$. It is possible to convert φ algorithmically into a disjunction of formulae χ which have the form

$$\bigwedge_{i<n} \Diamond^{\vec{x}_i} \mu_i$$

where $\vec{x}_i \in A^*$ and μ_i a conjunction of variables or negated variables. (Moreover, no variable occurs both simply and negated.) It is enough if we are able to decide whether χ is satisfiable in $\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle$. Now consider a valuation γ such that $\langle \mathfrak{c}_{\mathbb{P}}, \gamma, [\epsilon]\rangle \vDash \chi$. Then $\langle \mathfrak{c}_{\mathbb{P}}, \gamma, [\vec{x}_i]\rangle \vDash \mu_i$, for all $i < n$. And conversely. γ exists iff there do not exist i and j and a variable p such that $[\vec{x}_i] = [\vec{x}_j]$ and p is a conjunct of μ_i, $\neg p$ a conjunct of μ_j. Since $\mathbb{P}$ is decidable, the problem '$[\vec{x}_i] = [\vec{x}_j]$' is decidable. Hence we can decide whether γ exists. ◻

We can extract the following characterization of Halldén–completeness.

Theorem 9.4.6 (Grefe). *An extension Λ of bimodal $\boldsymbol{K.alt}_1$ is Halldén–complete iff*

1. $\Lambda = \boldsymbol{K}(\Box\bot) \otimes \boldsymbol{K}(\Box\bot)$ *or*
2. *there are $s, t \in \omega$ such that* $\Lambda = \boldsymbol{K}(\Box\bot) \otimes \mathsf{Th}\,\mathfrak{c}\mathfrak{y}\mathfrak{c}_{s,t}$ *or* $\Lambda = \mathsf{Th}\,\mathfrak{c}\mathfrak{y}\mathfrak{c}_{s,t} \otimes \boldsymbol{K}(\Box\bot)$ *or*
3. $\Lambda = \Lambda_{\mathbb{P}}$ *for some (possibly infinite) Thue–process.*

Proof. Since Halldén–completeness transfers under fusion the logics of the first two types are Halldén–complete. The logics of the third kind, however, are also Halldén–complete since they are determined by a single matrix by Theorem 1.6.5. This matrix corresponds to $\langle \mathfrak{c}_{\mathbb{P}}, [\epsilon]\rangle$. Now let Λ be a logic containing bimodal $\mathbf{K.alt}_1$. If it contains the axiom $\Box\bot$, then it is of the form $\mathbf{K.alt}_1(\Box\bot) \otimes \Theta$ with Θ Halldén–complete. Hence, it is of the first or the second type. Likewise if it contains the axiom $\blacksquare\bot$. If neither is the case, Λ is an extension of bimodal $\mathbf{K.alt}_1.\mathbf{D}$. Let φ be an axiom for Λ. We can put it into canonical form; it is then a disjunction of formulae disjoint in variables. Hence, by Halldén–completeness the disjunction is trivial, and so φ is actually a conjunction of path–equations. And so Λ is of the form $\Lambda_{\mathbb{P}}$ for some possibly infinite $\mathbb{P}$. ◻

$\Lambda_{\mathbb{P}}$ extends $\Sigma_{\mathbb{P}}$ by two constant axioms, $\Sigma_{\mathbb{P}}$ does not have the global finite model property, by constructive reduction. Similarly for decidability.

Proposition 9.4.7. *Let $\mathbb{P}$ be a Thue–process. (i) If $\mathbb{P}$ presents an infinite semigroup then $\Sigma_{\mathbb{P}}$ fails to have the global finite model property. (ii) If $\mathbb{P}$ is undecidable then $\Sigma_{\mathbb{P}}$ is globally undecidable.*

Exercise 338. Show that it is undecidable whether or not given a Thue–process $\mathbb{P}$ the semigroup presented by $\mathbb{P}$ is finite.

Exercise 339. Show that it is undecidable for every n whether the semigroup presented by $\mathbb{P}$ has $\leq n$ (exactly n) elements.

Exercise 340. Generalize Theorem 9.4.6 to arbitrarily many operators.

Exercise 341. Show that the first–order theory of a single binary relation is undecidable. *Hint.* Use the modal simulation theorem.

Exercise 342. Let $\mathbb{P}$ be a Thue–process. Show that $\Sigma_{\mathbb{P}}$ is in NEXPTIME.

9.5. Decidability of Properties of Logics I

In the remaining sections of this chapter we will discuss questions of decidability of properties of logics. Recall that we study these questions in the following general setting.

Definition 9.5.1. *Let P be a subset of $\mathcal{E}\,\mathbf{K}_\kappa$. P is* ***decidable*** *if for every finite set Δ of formulae in $\mathcal{P}_\kappa$ the problem '$\boldsymbol{K}_\kappa \oplus \Delta \in P$' is decidable.*

In general sets P will be determined by certain properties. So, we will say that a property of logics is **undecidable** if the corresponding subset of the lattices have that property. It will turn out that properties of logics are undecidable in the overwhelming number of cases. There might be a general reason for this, but right now we will just walk through a number of properties and discuss their decidability. The first property we will discuss is *identity to a given logic*, and related to it *inclusion in a given logic*. Recall from Section 7.1 the following theorem.

Proposition 9.5.2. *Λ is decidable iff the problem '$\boldsymbol{K}_\kappa \oplus \varphi \subseteq \Lambda$' is decidable. In other words, Λ is decidable iff the set $\downarrow\Lambda$ is a decidable subset of $\mathcal{E}\,\boldsymbol{K}_\kappa$.*

Theorem 9.5.3. *For monomodal logics,* ***consistency*** *is decidable.*

Proof. By Makinson's Theorem, $\Lambda = \mathbf{K} \oplus \bot$ iff $\Lambda \not\subseteq \mathsf{Th}\boxed{\bullet}$ and $\Lambda \not\subseteq \mathsf{Th}\boxed{\circ}$. The latter is decidable. ∎

This is an exceptional fact of monomodal logic, just in the same way as Makinson's Theorem is unique for monomodal logics. Consistency is undecidable as soon as we have two modal operators. The way to see this is as follows. First, by the results of the previous section it is undecidable whether a bimodal logic is equal to $\mathsf{Th}\boxed{\circ|\circ}$, where $\boxed{\circ|\circ}$ is the one–point bimodal frame which is reflexive in both relations. For example, take the logics $\Sigma_{\mathbb{P}}$ corresponding to Thue–systems. The logics $\Sigma_{\mathbb{P}}$ can have the following one–point frames as models, $\boxed{\bullet|\bullet}$, $\boxed{\bullet|\circ}$, $\boxed{\circ|\bullet}$, $\boxed{\circ|\circ}$, standing for the one point frame with the two relations being either reflexive or irreflexive. Now, $\mathbb{P} \vDash w \approx \epsilon$ iff neither $\boxed{\bullet|\bullet}$ nor $\boxed{\bullet|\circ}$ is a frame for $\Sigma_{\mathbb{P}}$. Define

$$\Theta := \mathsf{Th}\boxed{\bullet|\bullet} \sqcap \mathsf{Th}\boxed{\bullet|\circ}.$$

Θ is a subframe logic. Then $\mathbb{P} \vDash w \approx \epsilon$ iff $\Theta \sqcup \Sigma_{\mathbb{P}}$ is inconsistent. It is not decidable whether $\mathbb{P} \vDash w \approx \epsilon$, otherwise it is decidable whether a Thue–process is trivial. Notice that $\Theta \sqcup \Sigma_{\mathbb{P}}$ is a subframe logic.

Theorem 9.5.4. ($\kappa \geq 2$.) *Decidability is undecidable. Moreover, consistency of elementary subframe logics is undecidable.* ($\kappa = 1$.) $\{\mathsf{Th}\boxed{\bullet}\}$ *is undecidable.*

This theorem has a number of consequences. The first concerns the property of being a subframe logic. The argument is shown for $\kappa = 1$, but can easily be lifted to a logics with several operators.

Corollary 9.5.5. *It is undecidable whether a given logic is a subframe logic.*

Proof. *(Version 1.)* Let Θ be a bimodal logic. Consider Θ^s. We show that it is closed under subframes iff Θ is inconsistent. This is undecidable. Suppose that Θ^s is a subframe logic. Let $\mathfrak{F}$ be a Θ–frame. Then let $\mathfrak{G}$ be the subframe over the points satisfying $\alpha \vee \beta$. This is not a frame for **Sim** unless $\mathfrak{G}$ is empty. So, $\mathfrak{F}$ is either empty or $\mathfrak{F} \cong \boxed{\bullet}$. Hence Θ is inconsistent. Now suppose that Θ is inconsistent. Then the frames of Θ^s are the empty frame and $\boxed{\bullet}$. So, Θ^s is a subframe logic. ∎

Proof. *(Version 2. ($\kappa > 1$.))* Consider the logics $\Lambda_{\mathbb{P}}$. They are complete and elementary. Now, $\Lambda_{\mathbb{P}}$ is a subframe logic iff $\mathfrak{c}_{\mathbb{P}}$ contains only a single point. Namely, consider the subframe $\mathfrak{e}$ based on $[\epsilon]$. Either for all $\lhd_j$, $[\epsilon] \lhd_j [\epsilon]$ and then $\mathfrak{c}_{\mathbb{P}}$ has only one point, or $[\epsilon] \not\lhd_j [\epsilon]$ for some j, and then $\mathfrak{e} \nvDash \Diamond_j \top$ while $\mathfrak{c}_{\mathbb{P}} \vDash \Diamond_j \top$. ∎

Furthermore, *tabularity* and *weak transitivity* are undecidable. For notice that the two are equivalent in logics of bounded alternativity. So suppose that *weak transitivity is decidable*. First we decide whether or not $\Lambda_{\mathbb{P}}$ is weakly transitive. If not, it $\mathbb{P}$ unequal to the trivial process. But if $\Sigma_{\mathbb{P}}$ is weakly transitive, the canonical Thue–frame $\mathfrak{c}_{\mathbb{P}}$ is finite, and $\Lambda_{\mathbb{P}}$ is decidable, and we can check whether it contains the equations $\epsilon \approx a$, $a \in A$. So we can decide of $\mathbb{P}$ whether it is identical to the trivial process. Contradiction.

Theorem 9.5.6. ***Tabularity*** *and* ***weak transitivity*** *are undecidable.*

Theorem 9.5.7. *It is undecidable whether or not a monomodal logic contains* ***K4***.

Proof. Consider the logic $(\Sigma_{\mathbb{P}} \sqcup \Theta)^s$, the simulation of $\Sigma_{\mathbb{P}} \sqcup \Theta$. It is transitive iff $\mathbb{P}$ is trivial. ∎

Now let us deal with *decidability*. If we have more than one operator, then local decidability is undecidable, because $\Lambda_{\mathbb{P}}$ is decidable iff $\mathbb{P}$ is (Theorem 9.4.5). Now, $\Sigma_{\mathbb{P}}$ is locally decidable. Let us assume we can decide whether or not $\Sigma_{\mathbb{P}}$ is globally decidable. Then we can actually decide whether $\mathbb{P}$ is trivial. This goes as follows. Take $\mathbb{P}$ and check first whether $\Sigma_{\mathbb{P}}$ is globally decidable. If not, $\mathbb{P}$ is not trivial. If $\Sigma_{\mathbb{P}}$ is globally decidable, then decide whether or not $\Diamond\top; \blacklozenge\top \Vdash_{\Sigma_{\mathbb{P}}} p \leftrightarrow \Diamond p; p \leftrightarrow \blacklozenge p$. $\mathbb{P}$ is trivial iff both formulae are derivable. This is decidable. Similarly we can show that it is undecidable whether $\Sigma_{\mathbb{P}}$ has the global finite model property. Namely, assume it is decidable whether or not $\Sigma_{\mathbb{P}}$ has the global finite model property. Take a Thue–process $\mathbb{P}$ and check whether $\Sigma_{\mathbb{P}}$ has the global finite model property. If not, $\mathbb{P}$ cannot be trivial. If it does, however, $\Lambda_{\mathbb{P}}$ is decidable, and we can then decide whether $\mathbb{P}$ is trivial. We leave it to the reader to supply the argument that it is undecidable whether $\Lambda_{\mathbb{P}}$ has the local finite model property.

Theorem 9.5.8. ($\kappa \geq 2$.) *It is undecidable whether or not a logic is locally decidable. Moreover, it is undecidable whether or not a logic is globally decidable even when it is known that it is locally decidable.*

Theorem 9.5.9. ($\kappa \geq 2$.) *It is undecidable whether or not a logic has the local finite model property. Moreover, it is undecidable whether or not a logic has the global finite model property even when it is known that it has the local finite model property.*

The next result has first been obtained by Lilia Cagrova in [**44**]. The present proof appeared first in Carsten Grefe [**91**].

Theorem 9.5.10. ($\kappa \geq 2$.) *It is undecidable whether a first–order condition is modally definable.*

Proof. Let T_1 be the elementary theory of two binary relations. T_1 is universal and undecidable. The formula

$$\alpha_0 := (\forall x)[(\forall y \rhd x)(y \not\doteq x) \wedge (\forall y \blacktriangleright x)(y \not\doteq x)]$$

expresses irreflexivity and is not modally definable, because it is not positive. Now consider the formula $\beta := \alpha_0 \vee \gamma$, where γ is arbitrary. Then if for a Kripke frame $\mathfrak{f}$ we have $\mathfrak{f} \vDash \alpha_0$, then also $\mathfrak{f} \vDash \beta$. Now suppose that β is modally definable. Then it must hold in all Kripke frames for $\mathbf{K.alt}_1 \otimes \mathbf{K.alt}_1$ by the fact that every frame is the p–morphic image of an irreflexive frame (using unravellings). Thus $T_1 \vdash \alpha_0 \vee \gamma$, whence $T_1; \neg\alpha_0 \vdash \gamma$. Suppose now that β is not modally definable. Then $T_1 \nvdash \beta$, that is, $T_1; \neg\alpha_0 \nvdash \gamma$, for otherwise β holds in all frames and is therefore modally definable (for example by the true constant). Hence if we are able to show that $T_2 := T_1 \cup \{\neg\alpha_0\}$ is undecidable, we have succeeded in showing that modal definability (of β) is undecidable.

Now consider the theory $T_3 := T_1 \cup \{\alpha_1\}$ with

$$\begin{array}{rll} \alpha_1 := & (\exists x) & [(\forall y \rhd x)(y \doteq x) \wedge (\forall y \blacktriangleright x)(y \doteq x) \\ & \wedge & (\forall z)\{(\forall y \rhd z)(y \not\doteq x) \wedge (\forall y \blacktriangleright z)(y \not\doteq x)\}] \end{array}$$

Since $\vdash \alpha_1 \to \neg\alpha_0$ we have that if T_2 is decidable, so is $\{\zeta : T_2 \vdash \alpha_1 \to \zeta\} = \{\zeta : T_1; \neg\alpha_0; \alpha_1 \vdash \zeta\} = \{\zeta : T_1; \alpha_1 \vdash \zeta\} = \{\zeta : T_3 \vdash \zeta\}$. So we are done if we have shown that T_3 is undecidable. Now, consider a frame $\mathfrak{f}$ for T_1. If we add an inaccessible, reflexive point, that is, if we form the disjoint union $\mathfrak{f} \oplus \mathfrak{r}$, where $\mathfrak{r}$ is the one–point, reflexive frame, then we have a T_3–frame. And a T_3–frame is a T_1–frame. Denote by $T^{\forall}$ the set of sentences of T which have the complexity $\forall$. By standard model theory, $T_3^{\forall} = T_1^{\forall} = T_1$. Hence, since T_1 is undecidable, so is $T_3^{\forall}$ and a fortiori T_3. ☐

Corollary 9.5.11. ($\kappa \geq 2$.) *It is undecidable for Sahlqvist logics of rank 0 whether they are of pure rank 0.*

Theorem 9.5.12. ($\kappa \geq 2$.) *It is undecidable whether a logic is a fusion of monomodal logics.*

Proof. Suppose it is decidable whether a κ–modal logic is fusion of monomodal logics. Let $\mathbb{P}$ be a Thue–process over κ. Then it is decidable whether $\Lambda_{\mathbb{P}}$ is a fusion of monomodal logics. We show that it is decidable whether $\mathbb{P}$ is trivial. For assume that $\Lambda_{\mathbb{P}}$ is in fact a fusion of monomodal logics. Then by Theorem 9.1.3 it has the finite model property and is decidable. So, it is decidable whether or not $\Lambda_{\mathbb{P}}$ is the logic of a one–point frame. Therefore it is decidable whether $\mathbb{P}$ is trivial. Now assume that $\Lambda_{\mathbb{P}}$ is not the fusion of monomodal logics. Then $\mathbb{P}$ is not trivial. ◻

Finally, by invoking the Simulation Theorem, the following is proved without any condition on the number of operators.

Theorem 9.5.13. *The following properties of finitely axiomatizable logics are generally undecidable on the basis of a finite axiomatization*

- *identity to and inclusion in a given tabular logic,*
- *independent axiomatizability (for $\kappa > 1$),*
- *tabularity, weak transitivity,*
- *local decidability,*
- *global decidability, with local decidability given,*
- *local finite model property,*
- *global finite model property, with local finite model property given,*
- *modal definability for elementary conditions,*
- *being a subframe logic.*

The first claim should be read as follows. There is no general algorithm that decides, given a tabular logic Λ and a formula φ, whether or not $\mathbf{K} \oplus \varphi \subseteq \Lambda$, and whether or not $\mathbf{K}\oplus\varphi = \Lambda$. In particular, there is no algorithm deciding these problems for $\Lambda = \mathsf{Th}\boxed{\bullet}$ in case $\kappa = 1$ and $\Lambda = \mathbf{K}_\kappa \oplus \bot$ for $\kappa > 1$. The application of the Simulation Theorem is in each cases straightforward. Notice that the simulation of a subframe logic is not a subframe logic, so in this case nothing follows for $\kappa = 1$. Likewise, Corollary 9.5.11 has no analogue, even though a somewhat less interesting variant could be formulated. The reader may find it useful to describe the fact about monomodal logics that this theorem gives rise to.

Notes on this section. Many results of this section have been obtained by Alexander Chagrov, Lilia Chagrova, and Michael Zakharyaschev, see for example [**44**], [**41**]. The method in these papers is mainly simulating the action of a Minsky machine in a logical frame. This construction is rather complicated compared to the methods of this section. However, it achieves stronger results (whenever applicable) namely for extensions of **K4** and for intermediate logics.

Exercise 343. Show that *being n–transitive* is undecidable for any $n \geq 1$.

Exercise 344. Show that an extension of **G** is canonical iff it contains an axiom of the form $\Box^n\bot$. Show that the latter is equivalent to the logic being contained in **G.3**.

Exercise 345. Show that *being of codimension n* is undecidable for any given n, with the exception of $\kappa = 1$ and $n = 0$.

9.6. Decidability of Properties of Logics II

The previous results have been more or less straightforward consequences of the classical theorems on Thue–processes. For many properties, however, there is a rather effective tool for establishing their undecidability. It is due to S. Thomason [**212**]. Suppose we are interested in the decidability of a property $\mathfrak{P}$. Suppose further that the inconsistent polymodal logics have $\mathfrak{P}$ and that we have found a finitely axiomatizable logic Λ which lacks $\mathfrak{P}$. If $\mathfrak{P}$ transfers under fusion, then undecidability of $\mathfrak{P}$ follows simply from the fact that consistency is undecidable for bimodal logics. For consider logics of the form $\Lambda \otimes \Theta$, Θ finitely axiomatized. This logic has $\mathfrak{P}$ iff Θ is inconsistent. For if Θ is consistent, so is $\Lambda \otimes \Theta$, and since Λ fails to have $\mathfrak{P}$, $\Lambda \otimes \Theta$ fails to have $\mathfrak{P}$, too. If, however, Θ is inconsistent, so is $\Lambda \otimes \Theta$ and has $\mathfrak{P}$ by assumption. We refer to this argument as *Thomason's Trick*. Clearly, if $\mathfrak{P}$ also transfers under simulation, then undecidability of $\mathfrak{P}$ can be shown for monomodal logics. An impressive list of properties can be treated in this way. The following list is by no means exhaustive.

Theorem 9.6.1. *The following properties of logics are undecidable.*

1. $\mathfrak{Krp}$*–elementarity,*
2. α*–compactness,*
3. α*–canonicity,*
4. *global completeness given local finite model property,*
5. *(local/global) interpolation,*
6. *(local/global) Halldén–completeness.*

Proof. Using Thomason's Trick. The inconsistent logic is $\mathfrak{Krp}$–elementary, α–compact, α–canonical, has the global finite model property, is Halldén–complete and has interpolation. Hence, single negative examples must be found for all of these properties. **G** is a logic that is not $\mathfrak{Krp}$–elementary (1.), **G** $\otimes$ **K** is not 1–compact (see Section 6.5) (2.) and therefore not 1–canonical (3.). (4.) follows from the conjunction of Theorem 9.6.3 and 9.6.4. For interpolation and Halldén–completeness we need to ensure that a Λ exists not having these properties which is also complete with respect to atomic frames. For (local/global) Halldén–completeness take **K**. For interpolation take **S4.3**. ◻

This method can be adapted to bounded properties as well. Notice that bounded properties generally fail to transfer under fusion. For example, if Λ is pre–tabular, then $\Lambda \otimes \Theta$ need not be pre–tabular even if Θ is tabular. However, if $\Theta = \mathsf{Th}\boxed{\bullet}$, or $\Theta = \mathsf{Th}\boxed{\circ}$, pre–tabularity is preserved and reflected. Moreover, for any other logic $\Lambda \otimes \Theta$ is not pre–tabular, by Makinson's Theorem. To show that pre–tabularity is undecidable it just remains to show that the set $\{\mathsf{Th}\boxed{\bullet}, \mathsf{Th}\boxed{\circ}\}$ is an undecidable subset of $\mathcal{E}\,\mathbf{K}$. This is relatively straightforward given the results of the previous section and is therefore given as an exercise.

We still have to show that there exist logics which have the local finite model property but are globally incomplete. Here is such an example, the logic Θ. It is a 3–modal logic based on $\Box_i$, $i < 3$. We write $\Box$ for $\Box_0$, $\blacksquare$ for $\Box_1$ and $\boxminus$ for $\Box_2$.

$$\begin{array}{rll}
\Theta_o := & \mathbf{K4.3} \otimes \mathbf{K.alt}_1 \otimes \mathbf{K.alt}_1 & \\
& \oplus & p \to \blacksquare\Diamond p \\
& \oplus & \Diamond\blacklozenge q \vee \blacklozenge\Diamond q \to q \vee \Diamond q \\
& \oplus & \Diamond\!\!\!\!\!-\, p \to \Diamond p \\
& \oplus & \boxminus\boxminus\bot \\
& \oplus & \boxminus\Box\boxminus\bot \\
& \oplus & \Diamond\Diamond\!\!\!\!\!-\, p \wedge \Diamond\!\!\!\!\!-\, q \to \Diamond\Diamond\!\!\!\!\!-\,(p \wedge \Diamond q) \\
\Theta := & \Theta_o \oplus \Diamond p \to \Diamond(p \wedge \neg\Diamond p) &
\end{array}$$

Θ_o is Sahlqvist and therefore $\mathfrak{Krp} \cup \mathfrak{D}$–elementary. Θ contains in addition the axiom **G** for $\Box$. A Kripke–frame $\langle f, \lhd, \blacktriangleleft, \leqslant\rangle$ satisfies the axioms of Θ iff the following conditions are met.

(a) $\lhd$ is a linear irreflexive order such that there are no infinite ascending chains.
(b) If $x \blacktriangleleft y_1, y_2$ then $y_1 = y_2$.
(c) $\lhd$ contains the converse of $\blacktriangleleft$.
(d) If $x \lhd y \blacktriangleleft z$ or if $x \blacktriangleleft y \lhd z$ then either $x = z$ or $x \lhd z$.
(e) If $x \leqslant y_1$ and $x \leqslant y_2$ then $y_1 = y_2$.
(f) If $x \leqslant y$ then $x \lhd y$.
(g) If $x \leqslant z$ and $z = y$ or $z \lhd y$ then y has no $\leqslant$–successor.
(h) If $x \lhd y$, $x \leqslant x^+$ and $y \leqslant y^+$, then $y^+ \leqslant x^+$.

The following is easily checked by inspection of all axioms.

Proposition 9.6.2. *Θ_o and Θ are subframe logics.*

Now suppose we have a formula φ and a global model $\langle\mathfrak{F}, \beta\rangle$. Put $X := sf(\varphi)$. The φ–**span** of a point $x \in f$ is the set of all subformulae $\Diamond\psi$ of φ such that $\langle\mathfrak{F}, \beta, x\rangle \vDash \Diamond\psi$. If $\langle\mathfrak{F}, \beta, x\rangle \vDash \tau$ for some $\tau \in X$ then there exists a point y such that $x = y$ or $x \lhd y$ and y has least φ–span among all points satisfying τ. This point is necessarily irreflexive. (For if it is reflexive, then it also satisfies $\Diamond\tau$, and so it satisfies $\Diamond(\tau \wedge \neg\Diamond\tau)$. Thus, it has a successor z satisfying τ and τ is not in the φ–span of z. Since τ is in the φ–span of y, y has not been chosen minimal. Contradiction.)

We now assume that φ is in normal form and that $\langle \mathfrak{F}, \beta, w_0\rangle \vDash \varphi$. We can actually assume that φ is not in the φ–span of w_0. Moreover, we assume $\mathfrak{F}$ to be descriptive; hence its underlying Kripke–frame is a Θ_o–frame. We now define a selection procedure on points as follows. In each step a point is selected by a so–called **request**. A request is a pair $\langle y, \psi\rangle$ where $\langle \mathfrak{F}, \beta, y\rangle \vDash \psi$ and $\psi \in X$. In the selection step we answer the request by selecting for $\langle y, \psi\rangle$ a point z. Depending on ψ, this may give rise to new requests. After the selection, $\langle y, \psi\rangle$ is removed from the list of requests. The starting set of requests is $\{\langle w_0, \varphi\rangle\}$. The answers on the request $\langle y, \psi\rangle$ are as follows. If $\psi = \tau_1 \wedge \tau_2$, then the request is answered by adding the requests $\langle y, \tau_1\rangle$ and $\langle y, \tau_2\rangle$. (y is then said to be selected for τ_1 and τ_2 from y.) If $\psi = \tau_1 \vee \tau_2$, then if τ_1 holds at y the answer is $\langle y, \tau_1\rangle$ and y is selected for τ_1 from y. Otherwise τ_2 holds at y, and we answer with $\langle y, \tau_2\rangle$. If $\psi = \Diamond\tau_1$, then a z is chosen such that $y \lhd z$ and $\langle \mathfrak{F}, \beta, z\rangle \vDash \tau_1$, and the request is answered with $\langle z, \tau_1\rangle$. z is said to be selected for τ_1 from y. Analogously with $\psi = \blacklozenge\tau_1$ and $\psi = \diamondvert\tau_1$. If $\psi = \Box\tau_1$ or $= \blacksquare\tau_1$ or $= \boxminus\tau_1$, then the request is answered by dropping it, i. e. no new request is being made. Similarly if $\psi = p$, a variable. We take g the set of all points that have been selected. It is easy to see that g is finite. g is not necessarily an internal set. However, Θ_o is a subframe logic and $\mathfrak{F}$ is a descriptive Θ frame. Hence its underlying Kripke–frame, $\mathfrak{f}$, is a Θ_o–frame, and so the subframe $\mathfrak{g}$ defined by g is actually a Θ_o–frame. Moreover, if we let $\gamma(p) := \beta(p) \cap g$ it is established by easy induction that each point satisfies the formulae it has been selected for. In particular, $\langle \mathfrak{g}, \gamma, x\rangle \vDash \varphi$.

We now massage $\mathfrak{g}$ into a Θ–frame. To do this, we first analyse what could have gone wrong in case $\mathfrak{g} \nvDash \Theta$. In that case $\mathfrak{g}$ may contain $\lhd$–self–accessible points (that is a point x such that $x \lhd x$). Let us call them **improper points**. Improper points can be avoided as an answer to a request of the form $\langle y, \Diamond\tau\rangle$ as we have seen above. Moreover, if $x \ntriangleleft x$ and $x \blacktriangleleft y$ then $y \ntriangleleft y$, so improper points are not selected from proper points by answering a request of the form $\blacklozenge\tau$. (To see this, notice that by (d), if x is proper, and $x \blacktriangleleft y$ then y is the immediate $\lhd$–predecessor of x. It is not difficult to verify that in that case y is also proper.) Therefore, improper points can only arise through a request of the form $\diamondvert\tau$. Thus, let y be selected to answer a request $\langle x, \diamondvert\tau\rangle$. Then we claim that in the frame generated by y in $\mathfrak{F}$, no point has a $\leqslant$–successor. This holds for all $\lhd$–successors of y, by (f). Now let $y \blacktriangleleft z$. Then $z = y$ or $z \lhd y$, by (d). Hence, z is already in the transit of y. Moreover, $z \lhd z$, and the same argument can be repeated with z. (To see this, notice that $z \lhd y$. since $y \blacktriangleleft z$, by (c). Second, $y \lhd y \blacktriangleleft z$, and so $y = z$ or $y \lhd z$. In the first case $z \lhd z$ since $y \lhd y$. In the second case, $z \lhd z$ follows from $z \lhd y$ and $y \lhd z$, by transitivity of $\lhd$.)

Now take an improper point $x \in g$. Let $C(x) := \{y \in g : x \lhd y \lhd x\}$. We can assume that (i) $C(x) = \{x_i : i < n\}$ and $x_i \blacktriangleleft x_j$ iff $j \equiv i \pmod n$ and that (ii) $y \leqslant x$ for some $y \in g$. We let $\Omega_n := \langle \mathbb{Z}, \lhd_\Omega, \blacktriangleleft_\Omega, \leqslant_\Omega\rangle$, where $\lhd_\Omega := >$, $\blacktriangleleft_\Omega := \{\langle i, i+1\rangle : i \in \mathbb{Z}\}$ and $\leqslant_\Omega := \varnothing$. This is a Θ_o–frame. We assume its set of points to be disjoint from that of g. Now define the cluster substitution for $C(x)$ with respect to y by

$$\begin{array}{rcl} g^+ & := & (g - C(x)) \cup \mathbb{Z} \\ \lhd^+ & := & \left\{ \begin{array}{ll} & \lhd \cap (g - C(x))^2 \cup \lhd_\Omega \\ \cup & \{\langle u, v\rangle : u \in g, v \in \mathbb{Z}, u \lhd x\} \\ \cup & \{\langle v, u\rangle : u \in g, v \in \mathbb{Z}, x \lhd u\} \end{array} \right. \\ \blacktriangleleft^+ & := & \blacktriangleleft \cap (g - C(x))^2 \cup \blacktriangleleft_\Omega \\ \leqslant^+ & := & \leqslant \cap (g - C(x))^2 \ \cup \{\langle y, 0\rangle\} \end{array}$$

This defines the frame $\mathfrak{g}^+$. As is easily checked, it is a Θ_o–frame. Furthermore, the map $\pi : x \mapsto x$ for $x \notin \mathbb{Z}$, $\pi : k \cdot n + i \mapsto x_i$ is a p–morphism $\pi : \mathfrak{g}^+ \twoheadrightarrow \mathfrak{g}$. (This is clear for the relation $\lhd$, and easy to check for the relation $\leqslant$. In the case of $\blacktriangleleft$, one only has to observe that the cluster $C(x)$ is a connected, isolated component of the graph $\langle g, \blacktriangleleft\rangle$. This, however, follows from (c).) Put $\gamma^+(p) := \pi^{-1}[\gamma(p)]$. Then $\langle \mathfrak{g}^+, \gamma^+, w_0\rangle \vDash \varphi$. We perform this substitution successively for all clusters of $\mathfrak{g}$ containing an improper point. This yields a frame $\mathfrak{h}$ and a valuation δ such that $\langle \mathfrak{h}, \delta, w_0\rangle \vDash \varphi$. It is a Kripke–frame for Θ_o. On the basis of our original selection procedure we define a new selection procedure on $\langle \mathfrak{h}, \delta\rangle$ that will yield a finite model on a frame for Θ.

Assume that the selection procedure began with $\langle w_0, \varphi\rangle$. w_0 is proper. Let y be the first improper point selected. Then y is in a cluster which gets replaced by Ω_n. We select 0 in the new model, instead of y. Now we consider the effect this can have on our future selection. For each point we define the **selection history** to be a sequence of subformulae of φ by induction on the selection process as follows. x is assigned $\langle \varphi\rangle$. If y is assigned $\langle \psi_i : i < n\rangle$ and z is selected by $\langle y, \psi_{n-1}\rangle$ with answer $\langle z, \tau_1\rangle$ ($\{\langle z, \tau_1\rangle, \langle z, \tau_2\rangle\}$), then z is assigned $\langle \psi_0, \dots, \psi_{n-1}, \tau_1\rangle$ ($\langle \psi_0, \dots, \psi_{n-1}, \tau_1\rangle$ and $\langle \psi_0, \dots, \psi_{n-1}, \tau_2\rangle$). z may of course be identical to y, so that points may have several histories; but each history σ is assigned to only one point, denoted by $a(\sigma)$. Now we define a function b from selection histories into h as follows. If $a(\sigma)$ is proper, then $b(\sigma) := a(\sigma)$. Otherwise, consider the shortest history σ leading to an improper point z. Then $b(\sigma) := 0$, where 0 is the distinguished in the frame replacing the cluster of $a(\sigma)$. We claim that any history such that $a(\sigma)$ is improper extends a shortest history of this kind only by $\wedge$, $\vee$ and $\blacklozenge$ moves. To show this claim, assume that u_0 is the first improper point in a history σ. We have to show that all improper points of σ are obtained by a series of $\wedge$, $\vee$ and $\blacklozenge$–moves from u_0. Let us observe the following. The last move in a history that leads to an improper point is not a $\Diamond$–move. For they select (by our choice) proper points. Moreover, if x is proper and $x \blacktriangleleft y$ then y is also proper. Hence, an improper point arises either by making a $\leqslant$–move (A), or from an improper point by making a $\blacktriangleleft$–move (B). Consider (B) first. We have the situation that x is improper and that $x \blacktriangleleft y$. Then by (d), since $x \lhd x$ we deduce that $x \lhd y$ or $x = y$ (and then also $x \lhd y$). Moreover, by (c), $y \lhd x$. Hence, x and y are contained in the same $\lhd$–cluster. So y is improper, too. This finishes the case (B). Now let (A) be the case. We show that once (A) has occurred in a selection sequence it will not happen again. To that end let us take u_0 and a sequence $\langle u_i : i < n+1\rangle$ such that u_{i+1} is a successor via one of the relations. Assume that no $\lhd$–successor is improper. First of all, note that $u_0 \lhd u_i$ for all $i < n+1$. For u_0 is improper, and so $u_0 \blacktriangleleft u_1$ implies $u_0 \lhd u_1$, as we have seen. Furthermore, $u_0 \leqslant u_1$ implies $u_0 \lhd u_1$, by (f). Now consider the a point u_{k+1}. Assume that $u_0 \lhd u_k$. Then $u_k \leqslant u_{k+1}$ cannot occur, by (g). If $u_k \lhd u_{k+1}$ then $u_0 \lhd u_{k+1}$ by transitivity of $\lhd$. Finally, let $u_k \blacktriangleleft u_{k+1}$. Then $u_0 \lhd u_k \blacktriangleleft u_{k+1}$ yield $u_0 = u_{k+1}$ or $u_0 \lhd u_{k+1}$. Since u_0 is improper, $u_0 \lhd u_{k+1}$ in both cases. This shows, by (g), that no u_k has a $\leqslant$–successor. It follows that if u_{k+1} is improper, $u_k \blacktriangleleft u_{k+1}$ and u_k is improper too. And this proves our claim that an improper point is in the $\blacktriangleleft$–transit of u_0, the first improper point in a history.

Thus, let τ be improper and σ the shortest subhistory such that $a(\tau)$ is improper. We have defined $b(\tau)$. $b(\sigma)$ is now defined by induction on its length. (In each of the steps the choice of value under b is unique.) Finally, let k be the set of points selected by the new procedure. k is finite, and defines a subframe of $\mathfrak{h}$, $\mathfrak{k}$. Let $\epsilon(p) := k \cap \delta(p)$. Then as before we conclude that $\langle \mathfrak{k}, \epsilon, w_0\rangle \vDash \varphi$. Moreover, $\mathfrak{k}$ is a frame for Θ. We have shown the following.

Theorem 9.6.3. *Θ has the local finite model property.*

Now we show

Theorem 9.6.4. *Θ is not globally complete.*

Proof. We claim that (i) $\lozenge\top \to \blacklozenge\lozenge\top \nVdash_\Theta \neg\lozenge\top$ and (ii) no model in which $\lozenge\top \to \blacklozenge\lozenge\top$ holds globally but $\lozenge\top$ is locally satisfied is based on a Kripke–frame for Θ. That (ii) is the case is rather easy to see. Suppose we have a Θ–frame $\mathfrak{F}$ such that $\mathfrak{F} \vDash \lozenge\top \to \blacklozenge\lozenge\top$ and that there exists a point x such that $\langle \mathfrak{F}, x\rangle \vDash \lozenge\top$. Then $x \vDash \blacklozenge\lozenge\top$ and so there exists a x_1 such that $x \blacktriangleleft x_1$ and $x_1 \vDash \lozenge\top$. Inductively one can show that there exists a chain $\langle x_i : i \in \omega\rangle$ such that $x_0 = x$ and $x_i \blacktriangleleft x_{i+1}$ for all i. Moreover, there exist points y_i such that $x_i \leqslant y_i$, by (h). By the postulates of Θ we have $y_i \lhd y_{i+1}$ for all i. This means that $\mathfrak{F}$ contains an ascending chain with respect to $\lhd$. Hence, the underlying Kripke–frame violates the axiom **G** for $\lhd$, thus is not a Θ–frame. Now we still have to show (i). Take a copy of the natural numbers $\mathbb{N}$, members of which we denote by underlining, e. g. $\underline{5}$; and a (disjoint) copy of $\mathbb{Z}$, members of which are denoted by overstriking, e. g. $\overline{-5}$. We put $\overline{k} \lhd \underline{\ell}$ for all $k \in \mathbb{Z}$ and $\ell \in \mathbb{N}$; $\overline{k} \lhd \overline{\ell}$ iff $k < \ell$; and $\underline{k} \lhd \underline{\ell}$ iff $k > \ell$. So, with respect to $\lhd$ we have placed $\mathbb{Z}$ 'before' $\mathbb{N}^{op}$. We put $\underline{k} \blacktriangleleft \underline{\ell}$ iff $\ell = k+1$; $\overline{k} \blacktriangleleft \overline{\ell}$ iff $\ell = k-1$ and no other points are related with respect to $\blacktriangleleft$. Finally, for $k > 0$ we put for $\overline{-k} \leqslant \overline{k}$, and no other relations shall hold. We take as the algebra $\mathbb{O}$ of sets the 0–generated algebra of sets. This defines the frame Ω. Now observe that $\Omega_\sharp$ satisfies all postulates except for **G**. Hence everything depends on finding a set of internal sets such that **G** is also valid. To show that Ω fulfills **G**, let us prove that $\mathbb{O}$ is nothing but finite unions of intervals with respect to $\lhd$, where an interval is a set of the form

$[x, y] := \{z : x \lhd z \lhd y\} \cup \{x, y\}$, $x, y \in f$, or of the form $[\infty, x] := \{y : y \lhd x\} \cup \{x\}$, where ∞ is just an artificial symbol. We claim that these sets are 0–definable. (A set c is 0–definable in $\mathfrak{F}$ if there exists a variable free formula C such that $\langle \mathfrak{F}, x\rangle \vDash C$ iff $x \in c$. It is easy to see that c is 0–definable iff it is contained in the 0–generated subalgebra of $\mathfrak{F}_+$.) To that end, observe that $[\infty, \underline{k}]$ is 0–definable, by the fact that $[\infty, \underline{0}]$ is defined by $\top$ and $\Diamond[\infty, \underline{k}] = [\infty, \underline{k+1}]$. Furthermore, we have $[\infty, \overline{-1}] = \lozengeminus[\infty, \underline{0}]$, $[\infty, \overline{k+1}] = \blacklozenge[\infty, \overline{k}]$, and $[\infty, \overline{k-1}] = \Diamond[\infty, \overline{k}]$. Thus all sets of the form $[\infty, x]$ are 0–definable, and from that follows that all finite unions of intervals are 0–definable. Secondly, we show that this algebra is already closed under all operations. The set operations are clear. Now $\Diamond[x, y] = [\infty, y^+]$, where $y \blacktriangleleft y^+$, and $\blacklozenge[x, y] = [x^-, y^-]$, where $x^- \blacktriangleleft x$, $y^- \blacktriangleleft y$, if these points exist. (The other cases are also straightforward.) Next $\blacklozenge$. If $x = y = \underline{0}$ then $\blacklozenge[x, y] = \varnothing$. Otherwise, $\blacklozenge[\infty, y] = [\infty, y^-]$. Now finally $\lozengeminus$. Consider $[x, y]$. If $x = \underline{k}$ then $\lozengeminus[x, y] = \varnothing$. Hence let $x = \overline{k}$. We may assume $k > 0$, otherwise $\lozengeminus[x, y] = \lozengeminus[\overline{1}, y]$. If $y = \underline{\ell}$, then $\lozengeminus[x, y] = \lozengeminus[\overline{k}, \underline{\ell}] = [+\infty, \overline{-k}]$. If $y = \overline{\ell}$ then we can assume $y > x$ and so $\lozengeminus[x, y] = \lozengeminus[\overline{k}, \overline{\ell}] = [\overline{-\ell}, \overline{-k}]$. Indeed, the set of sets is closed under all operations. We have to show now that $\Omega \vDash \Diamond p \to \Diamond(p \wedge \neg\Diamond p)$. Take a valuation assigning an internal set c to p. Then this set has a largest element with respect to $\lhd$. For it is a union of intervals $[x_i, y_i]$. Among the y_i there is a largest with respect to $\lhd$, say y_0. Then if $\langle \Omega, \beta, x\rangle \vDash \Diamond p$ we have $x \lhd y_0$ and so $\langle \Omega, \beta, x\rangle \vDash \Diamond(p \wedge \neg\Diamond p)$, since $\langle \Omega, \beta, y_0\rangle \vDash p \wedge \neg\Diamond p$ by choice of y_0. To end the proof, we notice that $\Omega \vDash \lozengeminus\top \to \blacklozenge\lozengeminus\top$ but $\Omega \nvDash \neg\lozengeminus\top$. For the first note that $\langle \Omega, x\rangle \vDash \lozengeminus\top$ iff $x = \overline{k}$ for some $k < 0$. Then $\overline{k} \blacktriangleleft \overline{k-1}$, and $\langle \Omega, \overline{k-1}\rangle \vDash \lozengeminus\top$. For the second notice that $\langle \Omega, \overline{-1}\rangle \nvDash \neg\lozengeminus\top$. ∎

Corollary 9.6.5. *Θ is locally complete. Its extension by a universal modality, $\Theta^{\blacksquare}$, is not locally complete.*

This follows immediately with Theorem 3.1.13.

Exercise 346. This is another set of exercises that deal with Blok's Alternative. This time we deal with incompleteness phenomena. Show that any of the logics of Section 9.2 based on sequences of bounded index could have been taken. Show based on these considerations that the lattice of 3–modal logics has $2^{\aleph_0}$ co–atoms, which are incomplete.

Exercise 347. Show that in the lattice of 3–modal logics any logic of finite codimension has degree of incompleteness $2^{\aleph_0}$.

Exercise 348. Show that in the lattices of κ–modal logics, $\kappa \neq 0$, any consistent logic of finite codimension has degree of incompleteness $2^{\aleph_0}$. What if the logic is inconsistent?

Exercise 349. Show based on the preceding considerations that no consistent modal logic of finite codimension can be obtained by iterated splittings of $\mathbf{K}_\kappa$. *Hint.* Otherwise, show that it can be co–covered by at most countably many logics.

Exercise 350. Let $\mathcal{C}$ be any of the following complexity classes: NP, PSPACE, EXPTIME. Show that given a monomodal formula φ it is undecidable whether or not $\mathbf{K}_1 \oplus \varphi$ is $\mathcal{C}$–computable. Likewise, show that it is undecidable whether or not $\mathbf{K}_1 \oplus \varphi$ is $\mathcal{C}$–hard ($\mathcal{C}$–complete).

Exercise 351. Show that $\{\mathsf{Th}\boxed{\bullet}, \mathsf{Th}\boxed{\circ}\}$ is an undecidable subset of $\mathcal{E}\,\mathbf{K}$.

Exercise 352. (Continuing the previous exercise.) Show that pre–tabularity, pre–finite model property and pre–completeness are undecidable.

Exercise 353. Show that the set of pre–tabular logics of $\mathcal{E}\,\mathbf{S4}$ is decidable.

CHAPTER 10

Dynamic Logic

10.1. PDL — A Calculus of Compound Modalities

Propositional Dynamic Logic, **PDL** for short, can be seen as a special kind of polymodal logic and this is — at least implicitly — the way we will handle it here. This has the advantage that we can use the results of the previous chapters to a large extent. Nevertheless, **PDL** has a different syntax. **PDL** concentrates on compound modalities; the idea is that it is worthwile to investigate the structure of compound modalities separately, because many compound modalities arise naturally from the relational interpretation. For example, if we have an operator $\Box$ based on the relation $\lhd$ and an operator $\blacksquare$ based on the relation $\blacktriangleleft$, then the compound modality $\Box p \wedge \blacksquare p$ is based on the union $\lhd \cup \blacktriangleleft$ of the two, and $\Box\blacksquare p$ is based on $\blacktriangleleft \circ \lhd$. The perspective from which **PDL** views these things is from the perspective of relations. To be able to define an operator based on the union or composition of two relations is important. For the main interest for doing **PDL** is in reasoning about computer programs, or more generally, about *actions*. A computer program can be seen simply as a relation between memory states of a computer. For example, if our computer has three memory cells, x, y and z, storing integer numbers, the program $z := x + y$ is a relation between triples of integer numbers. In order not to confuse a program with the relation we call the latter the *extension* of a program on a given computer. Thus, $\langle\langle 3, -4, 7\rangle, \langle 3, -4, -1\rangle\rangle$ is in the extension of the program $z := x + y$, but $\langle\langle 1, 1, 0\rangle, \langle 0, 1, 2\rangle\rangle$ is not. The particular power of programming languages comes from the fact that programs can be combined. We can namely

Compose Programs: From two programs α and β we form the composition $\alpha;\beta$ which is the program defined by executing first α and then β.

Test: We can ask whether certain facts hold such as '$x = y$?', by means of which we can ask whether the number assigned to x is identical to the number assigned to y.

Combine Tests: We can use standard boolean connectives such as true, not and and to build more complex tests.

Combine Programs Logically: That is, we can use logical gates such as

$$\text{if } \varphi \text{ then } \alpha \text{ else } \beta \text{ fi}$$[1]

where φ is a statement and α and β programs. This means that we execute first a test for φ. If φ holds, α is executed; and if not, β is executed.

Iterate Programs: For a statement φ and a program α, the program

$$\text{while } \varphi \text{ do } \alpha \text{ od}$$

will execute α and continue to do so as long as the condition φ is momentarily satisfied, and the program

$$\text{until } \varphi \text{ do } \alpha \text{ od}$$

will execute α as long as φ is momentarily not satisfied.

This is the basic inventory of standard programming languages like Algol, Pascal and their derivatives, with those parts stripped away which belong to specific interpretations of the symbols (such as real or integer numbers, or sets, or lists) and the usual input/output routines. They can be reintroduced as a set of basic (or elementary) programs, from which the complicated routines are built up. To be realistic, we would need to talk at this point about memory cells and assignments. So in fact, a real model of a computer would have to use some first–order logic. But the purely propositional part is not only interesting (and decidable), but already very powerful.

Now a program is in some sense a relation of states in a computer, and the list of constructions we have just given can be produced from a small list of basic constructions. **PDL** starts with two different sorts of symbols, propositions and programs. There are at the basic level propositional variables, propositional constants and program constants. The set of basic program constants is denoted by Π_0 and consists of $\alpha_0, \alpha_1, \ldots$. The number of these constants varies. There are no program variables. Propositions can be composed as before with the boolean connectives. Programs are combined using the connectives

$\alpha \cup \beta$	set theoretic union
$\alpha;\beta$	relational composition
α^*	reflexive transitive closure

Any proposition φ can be converted into a program, by using the question mark '?', called *test*. If we test for φ, we write 'φ?' and read that φ *test*. The program φ? operates as follows. If φ is the case — we say then that the program φ? succeeds — the next clause is carried out. If φ is not the case — we say the program φ? fails — then no subsequent programs will be carried out. For example, the program $(\varphi?;\alpha) \cup (\neg\varphi?;\beta)$ performs α if φ is the case and β if φ is not the case. Thus it is identical to if φ then α else β fi. Notice that because we use the union, we save the entire

[1]To write an if–then–else–clause with the help of fi is standard practice, though not always explained. The word fi has actually no meaning. It is used simply to mark the end of the clause. This saves brackets while writing and makes a program optically perspicuous. The same applies to the pair do ... od below.

program from failing, even though some of its parts may fail individually. Thus, **PDL** makes crucial use of the fact that programs are allowed to be nondeterministic. This sounds absurd if we think of numerical calculations, where a correct program should yield a definite result, but is really rather useful in connection with inbuilt choices, as we have in fact just seen. Finally, any program α defines a modality $[\alpha]$ and its dual $\langle\alpha\rangle$. Basically, the extension of the program is the relation on which the modality is based. The statement $[\alpha]\varphi$ can be read as *at the end of all possible computations of α, φ holds*. Dually, the statement $\langle\alpha\rangle\varphi$ means *there is a computation for α at the end of which φ holds*.

One can think of various fragments, extensions and refinements of **PDL**. The fragment of **PDL** that does not use the star, is called **EPDL**, which is short for **elementary PDL**. It turns out to be a notational variant of polymodal logic. Second, there is *test free* **PDL**, of which we will make certain use. A particularly interesting extension is the addition of the *converse operator*. Given a program α, $\alpha^{\smile}$ will denote the converse program, i. e. the backward execution of α. Although for the standard interpretation this makes little sense, because standard computer languages do not need such an operation, it makes sense in reasoning about the behaviour of programs and actions. Also, dynamic logic is increasingly used in the semantics of natural language (see for example Jeroen Groenendijk and Martin Stokhof [**92**], Jan van Eijck and Fer–Jan de Vries [**218**]), and in reasoning about actions (see Vaughan Pratt [**165**] and Brigitte Penther [**160**]), to name just a few. Many connectives in natural languages make reference to the converse, such as *until*, *since* etc., though mostly in connection with tense only.

10.2. Axiomatizing PDL

In this section we are going to axiomatize **PDL** and prove the correctness of this axiomatization. As it turns out, the newly added program constructors are pretty harmless, with the exception of the Kleene Star. The latter, however, is a relatively difficult operator. We will see that there is no axiom system that can guarantee that α^* is in all cases the reflexive transitive closure of α. However, it is possible to give an axiomatization such that at least in Kripke–frames α^* has this property. Thus let us start with the latter. Let Π_0 be given; a member of Π_0 is denoted by ζ or η. A **dynamic Kripke–frame** over Π_0 is a pair $\mathfrak{f} = \langle f, \sigma\rangle$, where $\sigma : \Pi_0 \to 2^{f\times f}$ assigns to each basic program a binary relation on f. A **valuation** is a function

$\gamma : var \cup cons \to 2^f$ and the satisfaction clauses are as follows.

$$
\begin{array}{lcl}
\langle \mathfrak{f}, \gamma, x\rangle \vDash p_i & \Leftrightarrow & x \in \beta(p_i) \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash \neg\varphi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \nvDash \varphi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash \varphi \wedge \psi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \vDash \varphi, \psi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash [\varphi?]\psi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \vDash \varphi \to \psi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash [\zeta]\varphi & \Leftrightarrow & \text{for all } y \text{ such that } \langle x, y\rangle \in \sigma(\zeta) \qquad \langle \mathfrak{f}, \gamma, y\rangle \vDash \varphi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash [\alpha;\beta]\varphi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \vDash [\alpha][\beta]\varphi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash [\alpha \cup \beta]\varphi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \vDash [\alpha]\varphi; [\beta]\varphi \\
\langle \mathfrak{f}, \gamma, x\rangle \vDash [\alpha^*]\varphi & \Leftrightarrow & \langle \mathfrak{f}, \gamma, x\rangle \vDash \varphi; [\alpha]\varphi; [\alpha^2]\varphi; [\alpha^3]\varphi; \ldots
\end{array}
$$

Alternatively, given β and σ, we can extend σ to a map $\overline{\sigma}$ from all programs to binary relations over f and β to a map $\overline{\beta}$ from all propositions to subsets of f.

$$
\begin{array}{lcl}
\overline{\beta}(p_i) & := & \beta(p_i) \\
\overline{\beta}(\neg\varphi) & := & f - \overline{\beta}(\varphi) \\
\overline{\beta}(\varphi \wedge \psi) & := & \overline{\beta}(\varphi) \cap \overline{\beta}(\psi) \\
\overline{\beta}([\alpha]\varphi) & := & \{x : (\forall y)(\langle x, y\rangle \in \overline{\sigma}(\alpha) \Rightarrow y \in \overline{\beta}(\varphi))\} \\
 & & \\
\overline{\sigma}(\zeta) & := & \sigma(\zeta) \\
\overline{\sigma}(\varphi?) & := & \{\langle x, x\rangle : x \in \overline{\beta}(\varphi)\} \\
\overline{\sigma}(\alpha \cup \beta) & := & \overline{\sigma}(\alpha) \cup \overline{\sigma}(\beta) \\
\overline{\sigma}(\alpha;\beta) & := & \overline{\sigma}(\alpha) \circ \overline{\sigma}(\beta) \\
\overline{\sigma}(\alpha^*) & := & \overline{\sigma}(\alpha)^*
\end{array}
$$

In order to make the notation perspicuous we often write $x \xrightarrow{\alpha} y$ if $\langle x, y\rangle \in \overline{\sigma}(\alpha)$; in other words, we write $x \xrightarrow{\alpha} y$ to state that from x there exists an α–transition to y. This is a popular notation in computer science. The reader may verify that the two definitions of acceptance of formulae are one and the same, that is, we have $\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi$ iff $x \in \overline{\beta}(\varphi)$ for all propositions. Now we proceed to the promised axiomatization. As usual, the only rule of inference is modus ponens (in the local case), and the rules of substitution and necessitation are admissible. The axioms are in addition to those of boolean logic the following.

(**bd**.)	$\vdash [\alpha](p \to q). \to .[\alpha]p \to [\alpha]q$
(**df**?.)	$\vdash [p?]q. \leftrightarrow .p \to q$
(**df** $\cup$.)	$\vdash [\alpha \cup \beta]p. \leftrightarrow .[\alpha]p \wedge [\beta]p$
(**df**; .)	$\vdash [\alpha;\beta]p. \leftrightarrow .[\alpha][\beta]p$
(**cls**.)	$\vdash [\alpha^*]p. \leftrightarrow .p \wedge [\alpha][\alpha^*]p$
(**ind**.)	$\vdash [\alpha^*](p \to [\alpha]p). \to .p \to [\alpha^*]p$

This logic is standardly known as **PDL**. These axioms are due to KRISTER SEGERBERG, see for example [**195**]. Notice that α and β are not variables of the logic, but meta–variabes for programs. Hence the above postulates are not axioms but schemes of axioms. **PDL** is in general not finitely axiomatizable; the axiom system is recursive (i. e. decidable). Our first theorem concerns the correctness of this axiomatization. This means in informal terms that if we view the programs of **PDL** as separate and consider the above axioms as restrictions on the definition of these programs, then it will turn out that any Kripke–frame for the above axioms reduces to a dynamic Kripke–frame. To state this precisely, let us introduce the notion of a **PDL–Kripke–frame**. This is a pair $\langle f, \tau\rangle$ which satisfies the axioms of **PDL** not involving test. A **PDL–Kripke–model** is a triple $\langle f, \tau, \beta\rangle$ such that $\langle f, \tau\rangle$ is a **PDL**–Kripke–frame and the axioms (df?.) are satisfied. (In this case we say that τ and β are **compatible**.)

THEOREM 10.2.1. *Let $\langle f, \tau\rangle$ be a* $\boldsymbol{PDL}$*–Kripke–frame. Then*

1. $\tau(\alpha \cup \beta) = \tau(\alpha) \cup \tau(\beta)$.
2. $\tau(\alpha;\beta) = \tau(\alpha) \circ \tau(\beta)$.
3. $\tau(\alpha^*) = \tau(\alpha)^*$.

PROOF. Suppose, $\langle f, \tau\rangle \vDash [\alpha \cup \beta]p \leftrightarrow [\alpha]p \wedge [\beta]p$. We know that this axiom is Sahlqvist, and it is easily checked that it corresponds to the following property.

$$(\forall x)(\forall y)(x \overset{\alpha\cup\beta}{\to} y \leftrightarrow x \overset{\alpha}{\to} y \vee x \overset{\beta}{\to} y)$$

Thus, (df$\cup$.) holds iff $\tau(\alpha \cup \beta) = \tau(\alpha) \cup \tau(\beta)$. Similarly it is shown that (df;.) holds iff $\tau(\alpha;\beta) = \tau(\alpha) \circ \tau(\beta)$. The axiom (cls.)

$$[\alpha^*]p \to p \wedge [\alpha][\alpha^*]p$$

has as its dual

$$p \vee \langle\alpha; \alpha^*\rangle p. \to .\langle\alpha^*\rangle p$$

which is a conjunction of $p \to \langle\alpha^*\rangle p$, corresponding to the reflexivity of α^*, and $\langle\alpha; \alpha^*\rangle p. \to .\langle\alpha^*\rangle p$. The latter is also Sahlqvist and corresponds to the condition

$$(\forall x)(\forall y \overset{\alpha}{\leftarrow} x)(\forall z \overset{\alpha^*}{\leftarrow} y)(x \overset{\alpha}{\to} z)$$

In other words, the relation $\tau(\alpha^*)$ is reflexive and successor closed with respect to $\tau(\alpha)$. It therefore contains the reflexive transitive closure of $\tau(\alpha)$. That it is *exactly* the closure is the effect of the induction axiom (ind.). Namely, let x be given. Assume

$$\langle \mathfrak{f}, x\rangle \vDash [\alpha^*](p \to [\alpha]p). \to .p \to [\alpha^*]p$$

Put $\beta(p) := \{y : (\exists n \in \omega)(x \overset{\alpha^n}{\to} y)\}$. Then

$$\langle \mathfrak{f}, \beta, x\rangle \vDash p; [\alpha^*](p \to [\alpha]p)$$

and so $\langle \mathfrak{f}, \beta, x\rangle \vDash [\alpha^*]p$. Hence $x \overset{\alpha^*}{\to} y$ implies $x \overset{\alpha^n}{\to} y$ for some $n \in \omega$. ◻

The last theorem established that $\langle f, \tau\rangle$ is effectively a dynamic frame if we concentrate on test–free propositions. The test presents a problem, not so much for the axiomatization as for a proper formulation of the result. For suppose that in the last theorem we had stated that $\langle f, \tau\rangle$ satisfies (df?.). Then since τ is given, β is actually fixed. For we have

$$\langle f, \tau\rangle \vDash [p?]\bot. \leftrightarrow .p \to \bot$$

which is the same as

$$\langle f, \tau\rangle \vDash \langle p?\rangle\top. \leftrightarrow .p$$

Thus $\beta(p?) = \{x : (\exists y)(\langle x, y\rangle \in \tau(p?))\}$. However, intuitively we want the assignment of relations to tests be regulated by β and not by τ. This problem can be solved as follows. We abandon β and *define* it as above. Therefore, the whole information is already in the map τ. Then the pair $\langle f, \tau\rangle$ can serve both as a frame and as a model. The variability of β as opposed to that of σ will be accounted for by the notion of *variants*. Given $\langle f, \tau\rangle$ we call $\langle f, \widehat{\tau}\rangle$ a **variant** if $\widehat{\tau} \restriction \Pi_0 = \tau \restriction \Pi_0$. We now say that the **frame** $\langle \mathfrak{f}, \tau\rangle$ satisfies φ iff for all variants $\widehat{\tau}$ of $\langle \mathfrak{f}, \tau\rangle$ we have $\langle \mathfrak{f}, \widehat{\tau}\rangle \vDash \varphi$. Now what happens if the frame $\langle f, \tau\rangle$ satisfies (df?.)? Take any point x, and assume that there is a y such that $\langle x, y\rangle \in \widehat{\tau}(p?)$ for some variant $\widehat{\tau}$. Then $\beta(p)$ contains x, so p is true at x. Assume furthermore $\beta(q) = \{x\}$ (that is to say, $\widehat{\tau}(q?) = \{\langle x, x\rangle\}$). Then $\langle f, \widehat{\tau}, x\rangle \vDash p \to q$ and so $\langle f, \widehat{\tau}, x\rangle \vDash [p?]q$, so that $\langle f, \widehat{\tau}, y\rangle \vDash q$, which is to say, $y = x$. Hence if the frame $\langle f, \tau\rangle$ satisfies (df?.) then $\tau(p?) \subseteq \Delta_f$, the diagonal on f. Moreover, given the definition of β we get that $\langle x, x\rangle \in \tau(p?)$ iff $x \in \beta(p)$, since x has a $p?$–successor iff $x \in \beta(p)$. This finally justifies to restrict the attention to dynamic frames. To recapitulate, the problem lies in the fact that the basic programs are actually constants and so are all test–free programs. The test $p?$, however, behaves like a variable, though not induced by the assignment $\sigma : \Pi_0 \to 2^{f\times f}$ but by the basic assignment β.

Let us turn now to generalized frames. In principle, they are defined in full analogy to the modal case, if we just take dynamic logic to be a particular case of modal logic, where we have an infinite stock of operators with the postulates as above. This is not so straightforward for reasons mentioned earlier. Nevertheless, we can form the canonical structure $\mathfrak{Can}_{\mathbf{PDL}}(var)$ in the following way. The worlds are as usual maximally consistent sets, and the accessibility relation is as in Section 2.8, namely $U \xrightarrow{\alpha} V$ iff for all $[\alpha]\varphi \in U$ we have $\varphi \in V$.

Proposition 10.2.2. *In the canonical structure,* $\mathfrak{Can}_{\mathbf{PDL}}(var)$*,* $\tau(\alpha \cup \beta) = \tau(\alpha) \cup \tau(\beta)$*,* $\tau(\alpha;\beta) = \tau(\alpha) \circ \tau(\beta)$*, and* $\tau(\varphi?) = \{\langle U, U\rangle : \varphi \in U\}$*. Moreover,* $\tau(\alpha^*) \supseteq \tau(\alpha)^*$.

Proof. The first two and the last are straightforward from the fact that the postulates (**df**$\cup$.) and (**df**;.) are elementary and correspond to the fact that the relation for $\alpha \cup \beta$ $(\alpha;\beta)$ is the union (composition) of the relations for α and β. Similarly, the last claim follows from (**cls**.). For $\varphi?$ observe that $U \xrightarrow{\varphi?} V$ iff for all ζ if $[\varphi?]\zeta \in U$ we have $\zeta \in V$ iff for all ζ if $\varphi \to \zeta \in U$ we have $\zeta \in V$. Now two cases arise. First, if $\varphi \in U$, then $\varphi \to \zeta \in U$ iff $\zeta \in U$, so that we get $U \xrightarrow{\varphi?} V$ iff $V \supseteq U$ iff $U = V$, since both are maximally consistent. The second case is $\varphi \notin U$. Then $[\varphi?]\bot \in U$, and so no world is φ?–accessible. □

The problem with the canonical structure is the bad behaviour of the star. We only know that the relation for α^* is at least the transitive closure of the relation for α, and is successor closed, but we really cannot say more. This is no accident, for **PDL** fails to be frame–compact. Hence, a completeness proof for **PDL** is not straightforward. In fact, many authors have overlooked this rather subtle fact, and it took some time until a correct proof appeared that did not make the assumption that the canonical model has an underlying Kripke–frame which also satisfies **PDL**. We add here that **PDL** is also not compact. It is in fact not even 1–compact, which means that there is a set of constant formulae that have no model based on a **PDL**–Kripke–frame. This is left to the reader as an exercise.

Standardly, **PDL** has infinitely many variables and infinitely many programs. However, as in the modal case we can restrict ourselves as well as to finite variable fragments as well as to finitely many basic operators. Let us fix the set of variables p_i, $i \in \omega$, and ζ_i, $i \in \omega$. Let us denote by $\mathbf{PDL}(m,n)$ the restriction of **PDL** to the subset of variables p_i, $i < m$, and programs ζ_i, $i < n$. Similarly, the notation $\mathbf{PDL}(\omega,n)$ and $\mathbf{PDL}(m,\omega)$ is used. Then $\mathbf{PDL} = \mathbf{PDL}(\omega,\omega)$ and

$$\mathbf{PDL} = \bigcup_{m,n\in\omega} \mathbf{PDL}(m,n)$$

It turns out that $\mathbf{PDL}(\omega,n)$ is weakly transitive for $n < \omega$, while **PDL** is not. The master modality of $\mathbf{PDL}(\omega,n)$ is γ^* with $\gamma := \bigcup_{i<n} \zeta_i$.

Let us now turn to $\mathbf{PDL}^{\smile}$, the extension of **PDL** by the converse. There is an axiomatization of the converse as follows.

$$(\mathbf{df}^{\smile}_{+}.) \qquad \vdash p \wedge \langle\alpha\rangle q. \to .\langle\alpha\rangle(q \wedge \langle\alpha^{\smile}\rangle p)$$
$$(\mathbf{df}^{\smile}_{-}.) \qquad \vdash p \wedge \langle\alpha^{\smile}\rangle q. \to .\langle\alpha^{\smile}\rangle(q \wedge \langle\alpha\rangle p)$$

This is not exactly the way in which the converse is standardly axiomatized. In tense logic, the postulates $p \to [\alpha]\langle\alpha^{\smile}\rangle p$ and $p \to [\alpha^{\smile}]\langle\alpha\rangle p$ are normally used. This is a matter of convenience. It is also possible to replace the second postulate by $\langle\alpha^{\smile\smile}\rangle p \leftrightarrow \langle\alpha\rangle p$. All these boil down to the same. For notice that the postulates $(\mathbf{df}^{\smile}_{\pm}.)$ are r–persistent, so that in the canonical frame we do have $\tau(\alpha^{\smile}) = \tau(\alpha)^{\smile}$. The other postulates have the same effect on the canonical structure.

Finally, there are important program constants that we can define, namely skip and fail, defined by $\top$? and $\bot$?, respectively. skip is interpreted by the diagonal, and fail by the empty relation. There are many more operators that get studied with respect to computer application, such as since and until, both binary connectives for formulas, *nominals* (see Solomon Passy and Tinko Tinchev [**159**]) and many more. An overview of some of these different logics is given in Robert Goldblatt [**81**]. We will study some of these in detail. A very important logic from a practical point of view is the logic **DPDL**, which in addition to **PDL** contains the axiom $\langle\zeta\rangle p \to [\zeta]p$ for all basic programs ζ. This codifies the fact that an execution of a program yields at most one outcome. We say that the program is *deterministic*, and this is why **DPDL** contains the additional letter 'D'. Notice that it is only the *basic* programs which are required to be deterministic; for even if α is deterministic, α^* as well as $\alpha \cup (\alpha; \alpha)$ need not be.

Exercise 354. Formulate **PDL** instead of * with the operator $^+$ of transitive closure.

Exercise 355. Show that the boolean connectives can all be dropped from **PDL** without loss of expressivity, retaining only the propositional variables and constants, including $\top$ and $\bot$.

Exercise 356. Show that the postulates

$$\begin{aligned}
&p \to \langle\alpha^*\rangle p,\\
&\langle\alpha; \alpha^*\rangle p \to \langle\alpha^*\rangle p,\\
&[\alpha^*](p \to [\alpha]p) \to (p \to [\alpha^*]p)
\end{aligned}$$

are all independent. That means, show that no logic axiomatized by two together contains the third.

Exercise 357. Show that **PDL** is not 1–compact. *Hint.* The only source for failure of 1–compactness can be the induction axiom.

Exercise 358. Recall that **EPDL** is the fragment of **PDL** without the star. Show that in **EPDL** every formula is deductively equivalent to a formula without program connectives, that is, without $\cup$, ;, *, and ?.

Exercise 359. Show that $\mathbf{PDL} = \bigcup_{m,n\in\omega} \mathbf{PDL}(m, n)$. *Hint.* Use the compactness of the derivability in **PDL**.

10.3. The Finite Model Property

One of the earliest results on **PDL** was the proof by Fischer and Ladner [**69**] that **PDL** has the finite model property. The proof is interesting in two respects; first of course for the result itself, and second for the notion of downward closure that it uses. Notice that one of the first problems that one has to solve is that of defining a good notion of *subformula*. Of course, there is a syntactic notion of a subformula, but it is of no help in proofs by induction on subformulae. For what we need is a notion that allows us to assess the truth of formula in a model by induction on its subformulas. Consider, for example, the formula $\langle\alpha;\beta\rangle p$. The only proper subformula, syntactically speaking, is p. So, all we can say is that $\langle\alpha;\beta\rangle p$ is true at a point iff there is a $\alpha;\beta$–successor at which p holds — period. But $\alpha;\beta$ is a complex program and we would like to do an induction using only the basic programs to start with. Hence, let us take $\langle\alpha\rangle\langle\beta\rangle p$ also as a subformula. Then we can break down the truth definition of $\langle\alpha;\beta\rangle p$ into two parts. $\langle\alpha;\beta\rangle p$ is true at a node x if there is a α–successor y at which $\langle\beta\rangle p$ holds. The latter is the case if there is a β–successor z of y such that p holds at z. It is these considerations that lead to the following definition, taken from [**69**].

Definition 10.3.1. *Let X be a set of **PDL**–formulas. The **Fischer–Ladner closure** of X, denoted by $FL(X)$, is the smallest set containing X such that*

(fld.) *$FL(X)$ is closed under subformulae.*
(fl?.) *If $[\varphi?]\psi \in FL(X)$ then also $\varphi, \psi \in FL(X)$.*
(fl;.) *If $[\alpha;\beta]\varphi \in FL(X)$ then also $[\alpha][\beta]\varphi \in FL(X)$.*
(fl∪.) *If $[\alpha \cup \beta]\varphi \in FL(X)$ then also $[\alpha]\varphi, [\beta]\varphi \in FL(X)$.*
(fl*.) *If $[\alpha^*]\varphi \in FL(X)$ then also $[\alpha][\alpha^*]\varphi \in FL(X)$.*

Lemma 10.3.2. *If X is finite, so is $FL(X)$.*

The proof of this theorem is left to the reader. An effective bound on $FL(X)$ a priori can also be given (namely, the sum of the lengths of the formulae contained in X). The following proof of the finite model property of **PDL** is due to Rohit Parikh and Dexter Kozen. The original proof of M. J. Fischer and R. E. Ladner procedes from the tacit assumption that **PDL** is complete (and therefore that the interpretation of α^* is the reflexive transitive closure of the interpretation of α).

Theorem 10.3.3 (Fischer & Ladner, Kozen & Parikh). ***PDL** has the local finite model property.*

Proof. Take a formula φ and let $S(\varphi)$ be the collection of all those atoms in the boolean algebra generated by $FL(\varphi)$ which are **PDL**–consistent. These can be represented as subsets of $FL^{\neg}(\varphi)$, which is defined by $FL(\varphi) \cup \{\neg\chi : \chi \in FL(\varphi)\}$. We write W both for the set and the conjunction over its members. A dynamic Kripke–frame is based on $S(\varphi)$ via

$$V \xrightarrow{\alpha} W \qquad \Leftrightarrow \qquad Con_{\mathbf{PDL}} V \wedge \langle\alpha\rangle W$$

Moreover, we put $\beta(p) := \{W : p \in W\}$. If φ is consistent there is a $W \in S(\varphi)$ such that $\varphi \in W$. This concludes the definition of the dynamic frame, which we denote by $\mathfrak{S}(\varphi)$, and the valuation β. That $\langle\mathfrak{S}(\varphi), \beta, W\rangle \vDash \varphi$ is a consequence of the next three lemmas. ∎

Lemma 10.3.4. *For any program α, if $Con_{\mathbf{PDL}} V \wedge \langle\alpha\rangle W$ then $V \xrightarrow{\alpha} W$ in $\langle\mathfrak{S}(\varphi), \beta\rangle$.*

Proof. We do an induction on α. The basic case is covered by the definition. Also, if $\alpha = \psi?$ and $V \wedge \langle\psi?\rangle W$ is consistent, then by the axiom (**df?**.) we have that $V \wedge W$ is consistent. V and W, being atoms, must be equal, and $\psi \in V$, and so $V \xrightarrow{\psi?} W$, as required. Next, let $\alpha = \beta \cup \gamma$. Let $V \wedge \langle\beta \cup \gamma\rangle W$ be consistent. Then by (**df**$\cup$.) $V \wedge (\langle\beta\rangle W \vee \langle\gamma\rangle W)$ is consistent. By induction hypothesis therefore $V \xrightarrow{\beta} W$ oder $V \xrightarrow{\gamma} W$. Third, assume $\alpha = \beta;\gamma$ and $V \wedge \langle\beta;\gamma\rangle W$ is consistent. Then by (**df;**.), $V \wedge \langle\beta\rangle\langle\gamma\rangle W$ is consistent, which is to say, $V \wedge \langle\beta\rangle(\top \wedge \langle\gamma\rangle W$. Now, $\top \equiv \bigvee_{Z \in S(\varphi)} Z$, so that we can rewrite the latter into the disjunction of all $V \wedge \langle\beta\rangle(Z \wedge \langle\gamma\rangle W)$, $Z \in S(\varphi)$. Thus, by induction hypothesis there exists a $Z \in S(\varphi)$ such that $V \xrightarrow{\beta} Z$ and $Z \xrightarrow{\gamma} W$. Finally, let $\alpha = \beta^*$. Assume that $V \wedge \langle\beta^*\rangle W$ is consistent. Let $\mathfrak{B}$ be the closure of $\{V\}$ under β–successors. We wish to show that $W \in \mathfrak{B}$. To see that put $\chi := \bigvee_{Z \in \mathfrak{B}} Z$. Then $\chi \wedge \langle\beta\rangle\neg\chi$ is equivalent to the disjunction of all $T \wedge \langle\beta\rangle U$, where $T \in \mathfrak{B}$ but $U \notin \mathfrak{B}$. Each disjunct is individually inconsistent by induction hypothesis, for $T \xrightarrow{\beta} U$ does not hold. Thus, $\chi \wedge \langle\beta\rangle\neg\chi$ is inconsistent, from which $\vdash_{\mathbf{PDL}} \chi \to [\beta]\chi$. By necessitation, $\vdash_{\mathbf{PDL}} [\beta^*](\chi \to [\beta]\chi)$, and so

$$\vdash_{\mathbf{PDL}} V \to [\beta^*](\chi \to [\beta]\chi).$$

By $V \in \mathfrak{B}$ we also have $\vdash_{\mathbf{PDL}} V \to \chi$, and thus finally $\vdash_{\mathbf{PDL}} V \to [\beta^*]\chi$. This shows that $V \wedge \langle\beta^*\rangle\neg\chi$ is inconsistent, thus $W \in \mathfrak{B}$, since we have assumed that $V \wedge \langle\beta^*\rangle W$ is consistent. ◻

Lemma 10.3.5. *For every $\langle\alpha\rangle\chi \in FL(\varphi)$ and $V \in S(\varphi)$, $\langle\alpha\rangle\chi \in V$ iff there exists a $W \in S(\varphi)$ such that $V \xrightarrow{\alpha} W$ and $\chi \in W$.*

Proof. From left to right follows from the previous theorem. Now for the other direction. Again, we do induction on α. If α is basic, and $V \xrightarrow{\alpha} W$ with $\chi \in W$, then $\langle\alpha\rangle W \vdash \langle\alpha\rangle\chi$. Since $V \wedge \langle\alpha\rangle W$ is consistent by definition of α, so is $V \wedge \langle\alpha\rangle\chi$. By definition of $S(\varphi)$, $\langle\alpha\rangle\chi \in V$. The case of test is easy. Now let $\alpha = \beta \cup \gamma$. If $V \xrightarrow{\beta\cup\gamma} W$ then either $V \xrightarrow{\beta} W$ or $V \xrightarrow{\gamma} W$, from which by induction hypothesis and the fact that $\langle\beta\rangle\chi \in FL(\varphi)$ and $\langle\gamma\rangle\chi \in FL(\varphi)$ we have $\langle\beta\rangle\chi \in W$ or $\langle\gamma\rangle\chi \in W$. By (**df**$\cup$.) we have in either case $\langle\beta \cup \gamma\rangle\chi \in W$. Next let $\alpha = \beta;\gamma$. If $V \xrightarrow{\beta;\gamma} W$ and $\langle\beta;\gamma\rangle\chi \in FL(\varphi)$ then $\langle\gamma\rangle\chi \in FL(\varphi)$ as well as $\chi \in FL(\varphi)$. Assume $\chi \in W$. By assumption on the model there is a Z such that $V \xrightarrow{\beta} Z \xrightarrow{\gamma} W$. Then by induction hypothesis $\langle\gamma\rangle\chi \in Z$ and $\langle\beta\rangle\langle\gamma\rangle\chi \in V$ from which by use of (**df**;.) we get $\langle\beta;\gamma\rangle\chi \in V$. Finally, let $\alpha = \beta^*$. If $V \xrightarrow{\beta^*} W$ there exists a chain

$$V = V_0 \xrightarrow{\beta} V_1 \xrightarrow{\beta} V_2 \xrightarrow{\beta} \dots \xrightarrow{\beta} V_n = W.$$

Now, assuming $\chi \in V_n$ we get $\langle\beta^*\rangle\chi \in V_n$ by (**cls**.) and so $\langle\beta\rangle\langle\beta^*\rangle\chi \in V_{n-1}$, since the latter formula is in $FL^{\neg}(\varphi)$. By (**cls**.), $\langle\beta^*\rangle\chi \in V_{n-1}$. Inductively, we get $\langle\beta^*\rangle\chi \in V_i$ for all i, and so $\langle\beta^*\rangle\chi \in V$. ◻

Lemma 10.3.6. *For all $\chi \in FL^{\neg}(\varphi)$ and all $V \in S(\varphi)$, $\langle\mathfrak{S}(\varphi), \beta, V\rangle \vDash \chi$ iff $\chi \in V$.*

Proof. By induction on χ. If it is a variable, then this is true by definition of β. The induction steps for the boolean connectives are straightforward. There remains the case $\chi = \langle\alpha\rangle\psi$. $\langle\mathfrak{S}(\psi), \beta, V\rangle \vDash \langle\alpha\rangle\psi$ iff $V \xrightarrow{\alpha} W$ for some W such that

$\langle \mathfrak{S}(\varphi), \beta, W\rangle \vDash \psi$. By induction hypothesis, this is the case iff $\psi \in W$, and by the previous lemma the latter holds iff $\langle\alpha\rangle\psi \in V$. □

Corollary 10.3.7. *$\boldsymbol{PDL}$ has the global finite model property.*

Proof. Assume $\varphi \nVdash_{\mathbf{PDL}} \psi$. Take Π' to be the set of basic programs occurring in φ or ψ and put $\gamma := \bigcup_{\alpha\in\Pi'} \alpha$. Then $[\gamma^*]\varphi \nvdash_{\mathbf{PDL}} \psi$. By Theorem 10.3.3 there is a finite model $\langle \mathfrak{f}, \beta, x\rangle \vDash [\gamma^*]\psi; \neg\psi$ based on a dynamic frame $\mathfrak{f}$ for $\mathbf{PDL}(\Pi')$. We may assume that $\mathfrak{f}$ is rooted at x and that every node is a γ^*–successor of x. Now interpret the programs of $\Pi_0 - \Pi'$ arbitrarily. This gives a dynamic frame $\mathfrak{f}^+$ for $\mathbf{PDL}(\Pi_0)$ such that $\langle \mathfrak{f}^+, \beta, x\rangle \vDash [\gamma^*]\varphi; \neg\psi$. However, by construction this means $\langle \mathfrak{f}^+, \beta\rangle \vDash \varphi$ and so we have found a countermodel. □

Now let us turn to $\mathbf{PDL}^{\smile}$. Start with the same definition of $V \xrightarrow{\alpha} W$. We have to show only that $V \xrightarrow{\alpha^{\smile}} W$ iff $W \xrightarrow{\alpha} V$. Assume the first. Then $V \wedge \langle\alpha^{\smile}\rangle W$ is consistent. From this follows with ($\mathbf{df}^{\smile}_{-}$.)

$$V \wedge \langle\alpha^{\smile}\rangle(W \wedge \langle\alpha\rangle V)$$

This last formula must be consistent too, and so must be $W \wedge \langle\alpha\rangle V$. Hence $W \xrightarrow{\alpha} V$. Conversely, assume that $W \xrightarrow{\alpha} V$. Then $W \wedge \langle\alpha\rangle V$ is consistent and so is

$$W \wedge \langle\alpha\rangle(V \wedge \langle\alpha^{\smile}\rangle W)$$

by ($\mathbf{df}^{\smile}_{+}$.). Hence, $V \wedge \langle\alpha^{\smile}\rangle W$ is consistent and so $W \xrightarrow{\alpha^{\smile}} V$. The following is due to Dimiter Vakarelov [**217**].

Theorem 10.3.8 (Vakarelov). *$\boldsymbol{PDL}^{\smile}$ has the (global) finite model property.*

Notes on this section. It is known that **PDL** is EXPTIME–complete. That it is EXPTIME–hard has been shown by in M. J. Fischer and R. E. Ladner [**69**]); this follows however also immediately from the fact that the problem '$\varphi \Vdash_{\mathbf{K}} \psi$?' is equivalent to '$[\alpha^*]\varphi \rightarrow \psi \in \mathbf{PDL}$?'. The EXPTIME upper bound has been shown in Vaughan Pratt [**164**]. Giuseppe di Giacomo [**75**] has shown that $\mathbf{PDL}^{\smile}$ can be constructively reduced to **PDL**. This can be used to show that if **PDL** has interpolation, so does $\mathbf{PDL}^{\smile}$. This is given here as an exercise. It was shown by Moshe Vardi and P. Wolper in [**219**] that satisfiability in $\mathbf{PDL}^{\smile}$ can be computed in $O(2^{n^2})$–time. This bound is also easily established using constructive reduction.

Exercise 360. Show Lemma 10.3.2.

Exercise 361. Give a model theoretic proof of the fact that $\mathbf{PDL} = \bigcup_{m,n\in\omega} \mathbf{PDL}(m, n)$.

Exercise 362. Show that $\mathbf{PDL}^{\smile}$ has interpolation if **PDL** has interpolation.

Exercise 363. Let $\alpha^{(n)}$ denote the union of the programs α^i for $i \leq n$. Let $\mathbf{PDL}^n$ be the extension of **PDL** by all axioms of the form $\langle\alpha^*\rangle p \leftrightarrow \langle\alpha^{(n)}\rangle p$. Show that $\mathbf{PDL}^n$

has the finite model property without using the the finite model property of **PDL**. Show furthermore that from the fact that **PDL** has the finite model property one can deduce that

$$\mathbf{PDL} = \bigcap_{n\in\omega} \mathbf{PDL}^n$$

10.4. Regular Languages

The specific power that **PDL** has is that of the star. Without it, no gain in expressive power is reached, and we just have a definitional extension. It is useful to know some basic facts about the star when dealing with **PDL**. There is a well–known theorem by Stephen Kleene saying that a language is regular iff it can be evaluated using a finite state automaton. Since this result is of great significance, we will prove it now, in full generality. For general literature on languages and automata see [**99**], [**105**] or [**176**]. First, let us be given a finite alphabet A, consisting of symbols, denoted here by lower case letters such as a, b etc. A *language* is simply a set of strings over A. An alternative formulation is as follows. Consider the free semigroup generated by A, denoted by $\mathfrak{Fr}_{SG}(A)$. A language is a subset of this semigroup; alternatively, it is a set of terms in the language $\cdot$ (standing for concatenation) and ε (standing for empty word) modulo the following equations.

$$\begin{array}{lcl}
\vec{x}\cdot\varepsilon & \approx & \vec{x} \\
\varepsilon\cdot\vec{x} & \approx & \vec{x} \\
(\vec{x}\cdot\vec{y})\cdot\vec{z} & \approx & \vec{x}\cdot(\vec{y}\cdot\vec{z})
\end{array}$$

We consider the following operations on languages. For two languages L and M, we use the set union $L \cup M$, the composition $L \cdot M = \{\vec{x}\cdot\vec{y} : \vec{x} \in L, \vec{y} \in M\}$, and the Kleene–star $L^* = \bigcup_{n\in\omega} L^n$, where L^n is defined inductively by $L^0 = \{\varepsilon\}$ and $L^{n+1} = L \cdot L^n$. A language is called **regular** if it can be obtained from the languages $\varnothing$, $\{a\}$, where $a \in A$, by use of the operations union, concatenation and star. Another way to express this is as follows. Take the language with the function symbols $\varnothing$, ε, $\cup$, $\cdot$ and *. A **regular** expression is an expression of that language. With each regular expression R over A we associate a language $L(R)$ as follows.

$$\begin{array}{lcl}
L(\varnothing) & := & \varnothing \\
L(\varepsilon) & := & \{\varepsilon\} \\
L(a) & := & \{a\} \\
L(R \cup S) & := & L(R) \cup L(S) \\
L(R \cdot S) & := & L(R) \cdot L(S) \\
L(R^*) & := & L(R)^*
\end{array}$$

We will not always distinguish in the sequel between a regular expression and the language associated with it.

A **finite automaton** is a quadruple $\mathfrak{A} = \langle \mathbb{S}, S_0, F, \tau \rangle$, where $\mathbb{S}$ is a finite set, the set of **states**, $S_0 \in \mathbb{S}$ the **initial state**, $F \subseteq \mathbb{S}$ the set of **accepting states**, and $\tau : \mathbb{S} \times A \to 2^{\mathbb{S}}$ a function, the (nondeterministic) **transition function**. We define the transition function inductively by

$$\begin{array}{lcl} \overline{\tau}(S, \varepsilon) & := & S \\ \overline{\tau}(S, a) & := & \tau(S, a) \\ \overline{\tau}(S, \vec{x} \cdot \vec{y}) & := & \overline{\tau}(\overline{\tau}(S, \vec{x}), \vec{y}) \end{array}$$

$\mathfrak{A}$ **accepts** $\vec{x}$ if $\overline{\tau}(S_0, \vec{x}) \cap F \neq \varnothing$, otherwise $\mathfrak{A}$ **rejects** $\vec{x}$. Obviously, the set of accepted words of $\mathfrak{A}$ is a language, denoted by $\mathsf{L}(\mathfrak{A})$. If $\overline{\tau}(S, \vec{x})$ is a singleton set for any input $\vec{x}$ and state S, then the automaton is called **deterministic**. In that case the transition function can be constructed as a function from pairs of states and letters (or words) to states. A triple $\langle S, a, T \rangle$, also written $S \xrightarrow{a} T$, such that $T \in \tau(S, a)$ is called a **transition** of the automaton. τ is completely determined by its transitions.

Lemma 10.4.1. *For every finite state automaton $\mathfrak{A}$ there exists a deterministic automaton $\mathfrak{B}$ such that $\mathsf{L}(\mathfrak{B}) = \mathsf{L}(\mathfrak{A})$.*

Proof. Let $\mathfrak{A} = \langle \mathbb{S}, S_0, F, \tau \rangle$ be given. Let $\mathbb{T} := 2^{\mathbb{S}}$, $T_0 := \{S_0\}$, $G := \{H \subseteq \mathbb{S} : H \cap F \neq \varnothing\}$ and for $\Sigma \subseteq \mathbb{S}$ and $a \in A$ let $\sigma(\Sigma, a) := \bigcup \langle \tau(S, a) : S \in \Sigma \rangle$. It is verified by induction that for every word $\vec{x} \in L^*$ also

$$\overline{\sigma}(\Sigma, \vec{x}) = \bigcup_{S \in \Sigma} \overline{\tau}(S, \vec{x})$$

Put $\mathfrak{B} := \langle \mathbb{T}, T_0, G, \sigma \rangle$. It is then clear by the definitions that $\mathsf{L}(\mathfrak{B}) = \mathsf{L}(\mathfrak{A})$. □

For any regular language L an automaton $\mathfrak{A}$ can be constructed such that $L = \mathsf{L}(\mathfrak{A})$. Namely, for the one–letter languages $\{a\}$ construct a three–state automaton based on $\mathbb{S} = \{S_0, S_1, S_2\}$, where $F := \{S_1\}$ and such that $\tau(S_0, a) := S_1$, but $\tau(S_i, \ell) := S_2$ if $\ell \neq a$ or $S_i \neq S_0$. For the union assume that $\mathfrak{A}^i = \langle \mathbb{S}^i, S^i, F^i, \tau^i \rangle$ are automata such that $\mathsf{L}(\mathfrak{A}^i) = L^i$, $i = 1, 2$. Then construct the automaton

$$\mathfrak{B} := \langle \mathbb{S}^1 \times \mathbb{S}^2, \langle S^1_0, S^2_0 \rangle, F^1 \times \mathbb{S}^2 \cup \mathbb{S}^1 \times F^2, \tau^1 \times \tau^2 \rangle$$

where $\tau^1 \times \tau^2(\langle S, T \rangle, a) := \langle \tau^1(S, a), \tau^2(T, a) \rangle$. It is easy to check that $\mathsf{L}(\mathfrak{B}) = L^1 \cup L^2$. For the composition $L^1 \cdot L^2$ of the two languages perform the following construction. For each $T \in F^1$ add a copy $\mathfrak{A}^2_T$ of $\mathfrak{A}^2$ to the automaton, identifying the state T with $(S_0)_T$ as follows. For every transition into T add a transition into $(S_0)_T$ and finally remove T. The initial state is S^1_0 and the new accepting states are all states from F^2_T, $T \in F^1$. Finally, for the star just add for each transition $S \xrightarrow{a} T$ where $T \in F$ a transition $S \xrightarrow{a} S_0$, and remove all $T \in F$ different from S_0. The new set of accepting states is $\{S_0\}$. This new automaton accepts $(L^1)^*$. This shows that all regular languages are languages accepted by some finite state automaton.

Theorem 10.4.2 (Kleene). *Let A be a finite alphabet. A language $L \subseteq A^*$ is regular iff there exists a finite state automaton $\mathfrak{A}$ such that $L = \mathsf{L}(\mathfrak{A})$.*

Proof. Let $\mathfrak{A} = \langle \mathbb{S}, S_0, F, \tau\rangle$ be a finite state automaton. We can assume that it is deterministic. Let for simplicity be $\mathbb{S} = \{0, 1, \dots, n-1\}$, and $S_0 = 0$. Take a set X_i of variables ranging over subsets of A^*. The intended interpretation of X_i is that it denotes the set of words $\vec{x}$ such that $\tau(0, \vec{x}) = i$. The automaton is determined by a set of equations over these variables of the following form. For each $i < n$ and $j < n$ let $a_{ij} = \{a : i \xrightarrow{a} j, a \in A\}$. Each a_{ij} is a regular language, being a (possibly empty) union of one–letter languages. We then have

$$\begin{array}{lcllll} X_0 & = & \varepsilon \cup & X_0 \cdot a_{00} & \cup X_1 \cdot a_{10} & \cup \dots \cup X_{n-1} \cdot a_{n-1,0} \\ X_1 & = & & X_0 \cdot a_{01} & \cup X_1 \cdot a_{11} & \cup \dots \cup X_{n-1} \cdot a_{n-1,1} \\ \vdots & & & \vdots & \vdots & \vdots \\ X_{n-1} & = & & X_0 \cdot a_{0,n-1} & \cup X_1 \cdot a_{1,n-1} & \cup \dots \cup X_{n-1} \cdot a_{n-1,n-1} \end{array}$$

It is easy to see that any solution L_0, L_1 etc. for this system of equations has the property that the set of words $\vec{x}$ such that $\tau(0, \vec{x}) = i$ is exactly the solution for X_i. Furthermore, there can be only one such solution. To see this suppose that we have an equation of the form

$$X = X \cdot R \cup Y$$

where R is a regular expression over A containing no variables and Y a regular expression not containing X. This equation has the same solutions as

$$X = Y \cdot R^*$$

Namely, take a word from $Y \cdot R^*$. It is of the form $\vec{y} \cdot \vec{r}_0 \cdot \vec{r}_1 \cdot \vec{r}_2 \dots \vec{r}_{n-1}$, $\vec{r}_i \in R$ and $\vec{y} \in Y$. Then by the first equation $\vec{y} \in X$, since $Y \subseteq X$, and then $\vec{y} \cdot \vec{r}_0 \in X$, since $X \cdot R \subseteq X$. So $\vec{y} \cdot \vec{r}_0 \cdot \vec{r}_1 \in X$, and so on. Hence $\vec{x} \in X$. Conversely, let $\vec{x} \in X$. Then either $\vec{x} \in Y$, or $\vec{x} \in X \cdot R$, that is to say, $\vec{x} = \vec{x}' \cdot \vec{r}_0$ for some $\vec{r}_0 \in R$ and some $\vec{x}' \in X$. Again, either $\vec{x}' \in Y$ or it is of the form $\vec{x}'' \cdot \vec{r}_1$, $\vec{r}_1 \in R$ and $\vec{x}'' \in X$. This process must come to a halt, and this is when we have a decomposition of $\vec{x}$ into $\vec{y} \cdot \vec{r}_{n-1} \cdot \vec{r}_{n-2} \cdot \dots \cdot \vec{r}_0$, $\vec{r}_i \in R$ and $\vec{y} \in Y$. Thus $\vec{x} \in Y \cdot R^*$.

Armed with this reduction we can solve this system of equations in the usual way. We start with X_{n-1} and solve the equation for this variable. We insert this solution for X_{n-1} into the remaining equations, obtaining a new system of equations with less variables. We continue this with X_{n-2} and forth until we have an explicit solution for X_0, containing no variables. This we now insert back again for X_1 obtaining a regular expression for X_1, and so on. Finally, we have $\mathsf{L}(\mathfrak{A}) = \bigcup_{i \in F} X_i$, and inserting the concrete solutions gives us a regular expression for the language accepted by $\mathfrak{A}$. ☐

This theorem has numerous consequences. Let us list a few. Given a word $\vec{x} = a(0) \cdot a(1) \cdot \dots \cdot a(n-1)$, we put $\vec{x}^{\smile} := a(n-1) \cdot \dots \cdot a(1) \cdot a(0)$ and call it the **transpose** of $\vec{x}$. For a language L we write $L^{\smile} := \{\vec{x}^{\smile} : \vec{x} \in L\}$ and call that the **transpose** of L.

Theorem 10.4.3. *The intersection of two regular languages is a regular language. The transpose of a regular language is a regular language.*

The proof is left as an exercise. Given a finite automaton, and two states i and j, let $\mathsf{L}[i,j] := \{\vec{x} : \tau(\vec{x}, i) = j\}$. Then the proof method establishes that all $\mathsf{L}[i,j]$ are regular. The languages $\mathsf{L}[i,j]$ satisfy the equations

$$\mathsf{L}[i,j] = \bigcup_k \mathsf{L}[i,k] \cdot \mathsf{L}[k,j] .$$

Thus $\mathsf{L}[i,k] \cdot \mathsf{L}[k,j] \subseteq \mathsf{L}[i,j]$. For a deterministic automaton $\mathsf{L}[i,j] \cap \mathsf{L}[i,j'] = \varnothing$ whenever $j \neq j'$. Now, define the **prefix closure** L^p of a language L to be the set $\{\vec{x} : (\exists \vec{y})(\vec{x} \cdot \vec{y} \in L)\}$ and the **suffix closure** L^s to be the set $\{\vec{x} : (\exists \vec{y})(\vec{y} \cdot \vec{x} \in L)\}$. If L is regular, so are L^p and L^s. Namely, take an automaton $\mathfrak{A} = \langle \mathbb{S}, S_0, F, \tau \rangle$ recognizing L. L^p is the union of all $\mathsf{L}[S_0, i]$ such that $\mathsf{L}[i,j]$ is not empty for some $j \in F$. This is regular. For L^s, observe that L^s is the union of $\mathsf{L}[i,j]$ such that $j \in F$. All of the latter are regular, and so is L^s. There is an explicit way to compute the suffix closure of a regular expression, which runs as follows.

$$\begin{array}{lcl} a^s & := & \varepsilon \cup a \\ (R \cdot S)^s & := & R^s \cdot S \cup S^s \\ (R \cup S)^s & := & R^s \cup S^s \\ (R^*)^s & := & R^s \cdot R^* \end{array}$$

The correctness of this definition is seen by induction on the definition of the regular expression. Clearly, a suffix of a, $a \in A$, is either a itself or the empty word, so a^s is correctly defined. Now let $\vec{w}$ be suffix of a word $\vec{v}$ of $L \cup M$. Then if $\vec{v} \in L$, $\vec{w} \in L^s$, and if $\vec{v} \in M$ then $\vec{w} \in M^s$. Next let $\vec{w} \in (L \cdot M)^s$, that is, for some $\vec{v}$, $\vec{v} \cdot \vec{w} \in L \cdot M$. There exist $\vec{x} \in L$ and $\vec{y} \in M$ such that $\vec{v} \cdot \vec{w} = \vec{x} \cdot \vec{y}$. Then either $\vec{v}$ is a prefix of $\vec{x}$ or $\vec{x}$ is a prefix of $\vec{v}$. If the first is the case we have a decomposition $\vec{x} = \vec{v} \cdot \vec{x}'$ and so $\vec{v} \cdot \vec{w} = \vec{v} \cdot \vec{x}' \cdot \vec{y}$ from which $\vec{w} = \vec{x}' \cdot \vec{y}$. This means that $\vec{w} \in L^s \cdot M$. Now let $\vec{x}$ be a prefix of $\vec{v}$. Then we have a decomposition $\vec{v} = \vec{x} \cdot \vec{v}'$ and so $\vec{x} \cdot \vec{v}' \cdot \vec{w} = \vec{x} \cdot \vec{y}$ from which $\vec{v}' \cdot \vec{w} = \vec{y}$. Thus $\vec{w} \in M^s$.

Recall from the previous section the definition of the Fischer–Ladner closure. Apart from the usual syntactic closure it involves a certain kind of closure under subprograms. In a sense to be made precise now, the Fischer–Ladner closure of a formula φ involves the closure under suffixes of each program α occurring in φ.

Proposition 10.4.4. *Let $\langle \alpha \rangle p$ be a $\boldsymbol{PDL}$–formula. Then $\langle \alpha^s \rangle p$ is deductively equivalent to the disjunction of all formulae in $FL(\langle \alpha \rangle p)$.*

Proof. We do induction on α. Suppose α is elementary. Then by (fld.) we have $\bigvee FL(\langle \alpha \rangle p) = \langle \alpha \rangle p \vee p = \langle \alpha^s \rangle p$. Next suppose that $\alpha = \beta; \gamma$. Then $\bigvee FL(\langle \alpha \rangle p)$ is equivalent to $\bigvee FL(\langle \beta \rangle \langle \gamma \rangle p)$. Furthermore, $FL(\langle \beta \rangle \langle \gamma \rangle p) = (FL(\langle \beta \rangle q))[\langle \gamma \rangle p/q] \cup FL(\langle \gamma \rangle p)$. By induction hypothesis, the disjunction over the first is nothing but $(\langle \beta^s \rangle q)[\langle \gamma \rangle p/q] = \langle \beta^s \rangle \langle \gamma \rangle p$, which is equivalent in **PDL** to $\langle \beta^s; \gamma \rangle p$. The second is nothing but $\langle\gamma^s\rangle p$. The disjunction is equivalent to $\langle(\beta^s;\gamma)\cup\gamma^s\rangle p$ and that had to be shown. Next consider $\alpha = \beta\cup\gamma$. Then $\bigvee FL(\langle\alpha\rangle p)$ is equivalent to $\bigvee FL(\langle\beta\rangle p)\vee\bigvee FL(\langle\gamma\rangle p)$. By induction hypothesis the latter is equivalent to $\langle\beta^s\rangle p\vee\langle\gamma^s\rangle p$, and so equivalent to $\langle\alpha^s\rangle p$, as promised. Finally, we turn to $\alpha = \beta^*$. We have that $\bigvee FL(\langle\beta^*\rangle p)$ is equivalent to the disjunction of $FL(\langle\beta^*\rangle p)$ and $FL(\langle\beta\rangle\langle\beta^*\rangle p)$. The first does not reduce again except with the rule (fl*.) so that this disjunction is equivalent to $\langle\beta^*\rangle p\vee\bigvee FL(\langle\beta\rangle\langle\beta^*\rangle p)$. Now notice that $FL(\langle\beta\rangle\langle\beta^*\rangle p) = FL(\langle\beta\rangle q)[\langle\beta^*\rangle p/q]\cup\{p\}$. Therefore, the last of the two is equivalent by induction hypothesis to the disjunction of $\langle\beta^s\rangle\langle\beta^*\rangle p$ and $\langle\beta^*\rangle p$. Thus the whole disjunction is equivalent to the formula $\langle(\beta^s;\beta^*)\cup\beta^*\rangle p$, which is nothing but $\langle\alpha^s\rangle p$. □

Exercise 364. Show that the reduction algorithm in the proof of Kleene's Theorem does not depend on the assumption that $\mathfrak{A}$ is deterministic. Thus, another proof is found that languages accepted by nondeterministic finite state automata are the same as those accepted by a finite state deterministic automaton.

Exercise 365. A regular automaton is said to be an automaton with *ε–transitions* if there are transitions of the form $S\xrightarrow{\varepsilon}T$. Show that any language accepted by a finite state automaton with ε–transitions is regular.

Exercise 366. Spell out in detail the constructions to show that $L_1\cdot L_2$ and L_1^* are languages accepted by a finite state automaton whenever L_1 and L_2 are.

Exercise 367. Use the method of solving equations to show that all regular expressions define languages accepted by some finite state automaton. *Hint.* Let R be a regular language and R the term. Then start with the equations $X_1 = X_0\cdot R$, $X_0 = \varepsilon$. Now reduce the regular expression R by introducing new variables. For example, if $R = R'\cdot R''$, then $X_i = R'\cdot R''\cdot X_j\cup Y$. Introduce a variable X'_j and add instead the equations $X_i = R'\cdot X'_j\cup Y$, $X'_j = R''\cdot X_j$. Finally, if the system of equations is such that the expressions $X_i\cdot R$ are of the form where R is simply of the form a, $a\in A$, or a union thereof, see how you can define a finite state automaton accepting R.

Exercise 368. Show Theorem 10.4.3.

Exercise 369. Try to define the intersection and transpose of a language explicitly on the regular expression. *Hint.* This is complicated for the intersection, but should be relatively easy for the transpose.

10.5. An Evaluation Procedure

The problem that we are now going to attack is that of evaluation of **PDL**–formulae in a given model. This will provide the basis for some interesting theorems in dynamic logic. Before we start let us see how we can simplify **PDL**–formulae.

What is particularly interesting is to obtain some canonical form for the programs. First, we define the dynamic complexity of a formula as follows.

Definition 10.5.1. *A formula φ is of* ***dynamic complexity*** *0 if it is a boolean combination of variables. φ is of* ***dynamic complexity*** *$d+1$ if it is not of dynamic complexity $\leq d$ and a boolean combination of formulae of the form $\langle\alpha\rangle\psi$, where ψ is of dynamic complexity $\leq d$, and α is composed from basic programs and tests over formulae of dynamic complexity $\leq d$ using program union, composition and star. We denote the dynamic complexity of a formula φ by $\boldsymbol{dc}(\varphi)$.*

Call a program γ a **chain** if it is composed from basic programs and tests using only composition. A chain is **semiregular** if it is of the form ζ or $\alpha;\gamma'$, where γ' is a chain and where ζ is a basic program, not a test. A chain is called **regular** if it is of the form $\zeta;\varphi?$ for some basic program ζ, or of the form $\zeta;\varphi?;\gamma'$ for some regular chain γ'. (This is obviously well–defined, since γ' is a chain of smaller length than γ.) The first lemma concerns the reduction of star–free programs to regular chains.

Lemma 10.5.2. *Suppose $\chi = \langle\alpha\rangle\varphi$ is a formula where α is free of star. (1.) There exist semiregular chains γ_i, $i < n$, such that χ is equivalent to a disjunction of the formulae $\langle\gamma_i\rangle\varphi$. (2.) There exist regular chains δ_i, $i < m$, and formulae ψ_i of dynamic complexity less than $dc(\chi)$, such that χ is equivalent in $\boldsymbol{PDL}$ to the disjunction of $\psi_i \wedge \langle\delta_i\rangle\varphi$.*

Proof. First, by (df$\cup$.) and (df;.) we are allowed to convert $\langle\alpha;(\beta_1 \cup \beta_2);\gamma\rangle\varphi$ into the disjunction of $\langle\alpha;\beta_1;\gamma\rangle\varphi$ and $\langle\alpha;\beta_2;\gamma\rangle\varphi$. Thus, we can replace a star–free program by the disjunction of chains. Next, we can replace $\langle\alpha;\psi_1?;\psi_2?;\beta\rangle\varphi$ by $\langle\alpha;(\psi_1 \wedge \psi_2)?;\beta\rangle\varphi$. Finally, we can rewrite $\langle\psi?;\alpha\rangle\varphi$ by $\psi \wedge \langle\alpha\rangle\varphi$. Perform these reductions in succession and this gives the desired disjunction into formulae using regular chains. ◻

The star operator introduces some complications. First of all, however, notice the following identities. (Recall that $\mathsf{skip} := \top?$.)

Lemma 10.5.3. *The following holds.*

1. $(\alpha \cup \beta)^* = (\alpha^*;\beta^*)^*$.
2. $(\alpha;\beta)^* = \mathsf{skip} \cup \alpha;(\beta;\alpha)^*;\beta$.
3. $(\varphi?)^* = \mathsf{skip}$.

Using these identities we can do the following. Take the smallest subformula of the form α^*. Then α can be assumed to be a disjunction of semiregular chains. By (1.) of the above proposition, we can remove the disjunction in α^*, so that we are down to the case where α contains a single semiregular chain. If it is not already regular, then it is of the form $\psi?;\gamma$ or it the form $\psi?$. The latter is dealt with using (3.). It can be replaced by skip. Now consider the first case. Use (2.) to rewrite $(\psi?;\gamma)^*$ into $\mathsf{skip} \cup \psi?;(\gamma;\psi?)^*;\gamma$. γ begins with a basic program, so it is a regular chain. After these manipulations we have that the smallest starred subformulae are of the form α^* with α a regular chain. This we rewrite again into $\mathsf{skip} \cup \alpha^+$. Now we continue with the smallest subprogram of the form α^* in the formula thus obtained. It may now contain a program of the form β^+. However, the same manipulations can be performed. Iterating this we get a formula which we call *proper*. A definition recursive in the dynamic complexity runs as follows.

Definition 10.5.4. *A program α is called* **proper** *if it is composed from regular chains using tests on proper formulae, with the help of composition, union, and $^+$. A formula is proper if it is a boolean combination of formulae $\langle\alpha\rangle\psi$, where α and ψ are proper.*

Proposition 10.5.5. *Every formula of $\boldsymbol{PDL}$ is $\boldsymbol{PDL}$–equivalent to a proper formula.*

The use of proper formulae is explained by taking a look at path sets defined by programs. First, notice that programs do not simply define paths in a model, but they define path *sets* since they are regular expressions. We give a syntactic description in terms of chains.

$$\begin{array}{lcll} ch(\gamma) & := & \{\gamma\} & \text{if } \gamma \text{ is a chain} \\ ch(\gamma \cup \delta) & := & ch(\gamma) \cup ch(\delta) & \\ ch(\gamma^+) & := & \bigcup_{0<i<\omega} ch(\gamma^i) & \end{array}$$

To simplify the notation, we can assume that a regular chain is a composition of programs of the form $\gamma_i; \varphi_i?$. If not, then add suitable tests $\top?$ in between basic programs. With such simplification, take a model $\langle f, \sigma, \beta\rangle$ and a point x_0. A **path** of length n starting at x_0 is a sequence $\pi = \langle x_i : i < n+1\rangle$ such that $\pi(0) = x_0$ and for all $i < n$ there exists a basic γ_i such that $x_i \xrightarrow{\gamma_i} x_{i+1}$. Now, a path π is said to **fall under** a regular chain δ, where δ is of the form

$$\delta = \gamma_0; \psi_0?; \gamma_1; \psi_1?; \ldots; \gamma_n; \psi_n?\ ,$$

if $x_i \xrightarrow{\gamma_i} x_{i+1}$ for all $i < n$, and $\langle f, \sigma, \beta, x_{i+1}\rangle \vDash \psi_i$. If π falls under δ we also say that π is a **computation trace** of δ. (This is called a *computation sequence* in [**119**].) This notion is extended to all proper programs as follows. A path **falls under** or is a **computation trace of** a proper program α if it falls under any one of $ch(\alpha)$. Finally, if one wishes, the notion can also be extended to arbitrary α, by saying that π falls under α if it falls under any β which is proper and equivalent to α.

A first application is the proof that **DPDL** has the finite model property. We will show that **DPDL** is constructively reducible to **PDL** using the following global reduction function

$$X_D(\Delta) := \{\neg[\zeta]\chi \to [\zeta]\neg\chi : [\zeta]\chi \in FL[\Delta], \zeta \in \Pi_0\}.$$

These reduction sets split. Now assume that we have a finite model $\langle \mathfrak{f}, w_0\rangle$ such that $X_D(\varphi)$ holds globally while φ holds at w_0. Starting with this configuration, we will produce a finite **DPDL**–model for φ, based on the basic programs occurring in φ.

(So, we assume Π_0 to be finite.) Moreover, we will assume that φ to be proper. To see that this is harmless even for a proof of interpolation observe that by Proposition 10.5.5 we can turn α into a proper program by some equivalence transformations and insertion of some tests $\top$?. The construction is essentially due to Mordechai Ben–Ari, Joseph Halpern and Amir Pnueli [**7**]. First, we will unravel $\mathfrak{f}$ at x. The resulting frame is called $\mathfrak{u}$. The unravelling is strictly speaking unnecessary, we do it only for technical convenience. The construction of the actual model is inductive. The intermediate objects are not models, however, but what we call *pseudomodels*. A **pseudomodel** is a pair $\langle \mathfrak{f}, \mathsf{b}\rangle$ where b assigns a set of formulae to each $x \in f$. $\langle \mathfrak{f}, \mathsf{b}, x\rangle \vDash \varphi$ is defined inductively. However, it is a priori not excluded that the valuation is inconsistent. So, we work with sets of truth values. Given $\mathfrak{M} := \langle \mathfrak{f}, \mathsf{b}\rangle$ we define $v_{\mathfrak{M}}(\chi, x) \subseteq \{0, 1, *\}$ inductively.

$$
\begin{aligned}
1 \in v_{\mathfrak{M}}(p, x) \quad &\Leftrightarrow \quad p \in \mathsf{b}(x), \\
0 \in v_{\mathfrak{M}}(p, x) \quad &\Leftrightarrow \quad \neg p \in \mathsf{b}(x)
\end{aligned}
$$

For $\chi = \neg\chi'$, let $1 \in v_{\mathfrak{M}}(\chi, x)$ iff $0 \in v_{\mathfrak{M}}(\chi', x)$ and $0 \in v_{\mathfrak{M}}(\chi, x)$ iff $1 \in v_{\mathfrak{M}}(\chi', x)$. For $\chi = \chi_1 \wedge \chi_2$ put $v_{\mathfrak{M}}(\chi, x) := \{\flat \cap \sharp : \flat \in v_{\mathfrak{M}}(\chi_1, x), \sharp \in v_{\mathfrak{M}}(\chi_2, x)\}$. Finally, let $\chi = \langle\alpha\rangle\chi'$. We put $1 \in v_{\mathfrak{M}}(\chi, x)$ iff there exists β and γ such that $\langle\beta; \gamma\rangle p \to \langle\alpha\rangle p \in$ **PDL** and a path from x to y falling under β, such that $\langle\gamma\rangle\chi' \in \mathsf{b}(y)$. Otherwise, $0 \in v_{\mathfrak{M}}(\chi, x)$. (Clearly, for $[\alpha]\chi'$ the dual of this condition is chosen.) Finally, we write $\mathfrak{M} \vDash \chi$ if $1 \in v_{\mathfrak{M}}(\chi, x)$. Call a formula a **strict diamond formula** if it is not of the form $\neg\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\langle\varphi?\rangle\psi$ or $[\alpha]\psi$. Say that a pseudomodel is **proper** if $\mathsf{b}(x)$ always is a set of strict diamond formulae.

Lemma 10.5.6. *Let* $\mathfrak{M} = \langle \mathfrak{f}, \mathsf{b}\rangle$ *be proper. Then the sets* $\Delta(x) := \{\chi : v_{\mathfrak{M}}(\chi, x) = 1\}$ *are consistent.*

For a proof of this lemma notice that a strict diamond formula never gets the value $\{0, 1\}$, by definition. So $\Delta(x)$ can be inconsistent iff there exists a variable p such that $p, \neg p \in \mathsf{b}(x)$. If $\mathsf{b}(x)$ only contains variables for each $x \in f$, define $\beta(p) := \{x : p \in \mathsf{b}(x)\}$. Then $\langle \mathfrak{f}, \mathsf{b}, x\rangle \vDash \varphi$ iff $\langle \mathfrak{f}, \beta, x\rangle \vDash \varphi$, as is easily verified.

Now we start the construction. Put $P_0 := \{w_0\}$, where w_0 is the root of $\mathfrak{u}$. Furthermore, let $\mathfrak{p}_0$ be the subframe based on P_0 (that is, all relations are empty); and finally,

$$
\begin{aligned}
\mathsf{b}_0(w_0) := \quad &\{\chi \in FL^{\neg}(\varphi) : \chi \text{ a variable or strict diamond formula} \\
&\quad \text{and } \langle \mathfrak{u}, \beta, \langle w_0\rangle\rangle \vDash \chi\}\ .
\end{aligned}
$$

Starting with $\langle \mathfrak{p}_0, \mathsf{b}_0\rangle$ we will construct a sequence of pseudomodels $\langle \mathfrak{p}_n, \mathsf{b}_n\rangle$ and sets L_n such that

1. $\mathfrak{p}_n$ is based on a finite subset P_n of $\mathfrak{u}$.
2. $\mathsf{b}_n(x) \subseteq FL(\varphi)$ for all $x \in P_n$.
3. L_n is the set of all x for which there exists $i \leq n$ such that x is a leaf of $\mathfrak{p}_i$.

4. $\langle \mathfrak{p}_n, \mathsf{b}_n \rangle$ is proper. Moreover, $\mathsf{b}_n(x)$ contains formulae which are not variables only if x has no successors.
5. For $\chi \in FL(\varphi)$ we have $\langle \mathfrak{p}_n, \mathsf{b}_n, x \rangle \vDash \chi$ iff $\langle \mathfrak{n}, \beta, x \rangle \vDash \chi$.
6. $\mathfrak{p}_n$ is a **DPDL**–frame.

These claims are immediate for $\langle \mathfrak{p}_0, \mathsf{b}_0 \rangle$. We will also construct auxiliary sets L_n. $L_0 := \{w_0\}$. Now let $\langle \mathfrak{p}_n, \mathsf{b}_n \rangle$ and L_n already be constructed. Pick a node x in P_n without successors. **Case 1.** Along a path from w_0 to x there exists a $y \in L_n$ and $y \neq x$ such that $\langle \mathfrak{p}_n, \mathsf{b}_n, y \rangle \vDash \chi$ iff $\langle \mathfrak{p}_n, \mathsf{b}_n, x \rangle \vDash \chi$ for all $\chi \in FL(\varphi)$. Then $y \in L_{n-1}$. Put $P_{n+1} := P_n - \{x\}$. For $\zeta \in \Pi_0$ put $z \xrightarrow{\zeta} z'$ in the new frame if either $z \xrightarrow{\zeta} z'$ in the old frame, or $z' = y$ and $z \xrightarrow{\zeta} x$ in the old frame. $\mathsf{b}_{n+1} := \mathsf{b}_n \restriction P_{n+1}$. $L_{n+1} := L_n - \{x\}$. In this case, the properties (1.) – (4.) and (6.) are immediate. We will verify (5.) in Lemma 10.5.8. **Case 2.** Along no path from w_0 to x there is a $y \in L_n$ different from x such that y satisfies the same set of formulae in $FL(\varphi)$ as does x. Pick a formula $\neg[\alpha]\psi$ in $\mathsf{b}_n(x)$. There exist points s_i in $\mathfrak{n}$, $i < k + 1$, and a computation trace

$$\tau_0?; \gamma_0; \tau_1?; \gamma_1; \ldots \tau_{k-1}?; \gamma_{k-1}$$

falling under $\alpha; \neg\psi?$ such that $s_0 = x$, $s_i \xrightarrow{\gamma_i} s_{i+1}$ in $\mathfrak{n}$, $\langle \mathfrak{n}, \beta, s_i \rangle \vDash \tau_i$ for all $i < k$, and $\langle \mathfrak{n}, \beta, s_k \rangle \vDash \psi$. (It is here that we need φ to be proper. This assumption is not essential, but allows for simpler statement of the facts. If φ is not proper there might not be a τ_i? for certain i, but the proof does in no way depend on that.) For each $i < k$ and each basic ζ we let $t(i, \zeta)$ be a point of $\mathfrak{n}$ such that $s_i \xrightarrow{\zeta} t(i, \zeta)$. If $\zeta = \gamma_i$, then $t(i, \zeta) := s_{i+1}$. Put

$$T(x, \neg[\alpha]\psi) := \{s_i : i < k + 1\} \cup \{t(i, \zeta) : i < k, \zeta \in \Pi_0\} .$$

We call $T(x, \neg[\alpha]\psi)$ the **thorn** sprouting at x for $\neg[\alpha]\psi$. We let $\mathfrak{p}_n(x, \neg[\alpha]\psi)$ be the result of adding the thorn sprouting at x for $\neg[\alpha]\psi$ to $\mathfrak{p}_n$. A map c is defined by $\mathsf{c}(x) := \mathsf{b}_n(x) - \{\neg[\alpha]\psi\}$, $\mathsf{c}(z) := \mathsf{b}_n(z)$ if $z \in P_n - \{x\}$, and $\mathsf{c}(z) = \{\neg[\alpha']\psi' : \neg[\alpha']\psi' \in FL(\varphi), \langle \mathfrak{n}, \beta, z \rangle \vDash \varphi\}$ for $z \in T(x, \neg[\alpha]\psi)$ a leaf, $\mathsf{c}(z) := \{p : p \in var(\varphi), z \in \beta(p)\}$, z not a leaf. As one can show, the new pseudomodel satisfies all properties but (4.) in the above list. Therefore, we would like to add all thorns sprouting at x for a (strict diamond) formula of the form $\neg[\alpha]\psi$. However, the resulting frame will not be a **DPDL**–frame. Therefore, some care has to be exercised in defining the thorns.

LEMMA 10.5.7. *Let φ be a formula, and*

$$\langle \mathfrak{f}, \beta \rangle \vDash \{\neg[\zeta]\chi \rightarrow [\zeta]\neg\chi : [\zeta]\chi \in FL(\varphi), \zeta \in \Pi_0\} .$$

Let $\neg[\alpha]\psi$ be in $FL^{\neg}(\varphi)$ and $\langle \mathfrak{f}, \beta, x \rangle \vDash \neg[\alpha]\psi$, $x \xrightarrow{\zeta} y$. If there is a path starting with $x \xrightarrow{\zeta} z$ falling under a computation trace for $\alpha; \neg\psi?$, then there is a formula $\neg[\alpha']\psi$ in $FL^{\neg}(\varphi)$ such that for any path π starting at z falling under a computation trace for $\alpha'; \neg\psi?$, the path $x^\frown\pi$ falls under a computation trace for $\alpha; \neg\psi?$. Furthermore, $\langle \mathfrak{f}, \beta, y \rangle \vDash \neg[\alpha']\psi$ and there is a path falling under a computation trace for $\alpha; \neg\psi?$ starting at x leading through y.

Thus, let $\neg[\alpha']\psi'$ be a formula that holds at x in $\mathfrak{n}$. By repeated use of the lemma one can show that there exists a path falling under a computation trace for $\alpha; \neg\psi?$ that is either fully contained in (i) $T(x, \neg[\alpha]\psi)$ or (ii) leads through a leaf y of $T(x, \neg[\alpha]\psi)$. In case (i), we do nothing. In case (ii), rather than sprouting a thorn for $\neg[\alpha']\psi'$ at x in $\mathfrak{p}_n$, we sprout a thorn at y for some suitable $\neg[\alpha'']\psi''$ in $\mathfrak{p}_n(x, \neg[\alpha]\psi)$. We perform this sprouting for all formulae $\neg[\alpha']\psi'$ in $\mathsf{b}_n(x)$. Let $T(x)$ be the union of the thorns. Let $y \xrightarrow{\zeta} z$ in $T(x)$ iff $y \xrightarrow{\zeta} z$ in $\mathfrak{n}$. This defines the frame $\mathfrak{t}(x)$. It is a **DPDL**–frame; it is a tree. We put $P_{n+1} := P_n \cup T(x)$, and $y \xrightarrow{\zeta} z$ iff $y, z \in P_n$ and $y \xrightarrow{\zeta} z$ in $\mathfrak{p}_n$ or $y, z \in T(x)$ and $y \xrightarrow{\zeta} z$ in $\mathfrak{t}(x)$. Finally, $\mathsf{b}_{n+1}(y) := \mathsf{b}_n(y)$ if $y \in P_n - \{x\}$, $\mathsf{b}_{n+1}(x) := \{p : p \in \mathsf{b}_n(x)\}$, $\mathsf{b}_{n+1}(y) := \{p : \langle \mathfrak{n}, \beta, y\rangle \vDash p\}$ if y is in $T(x)$ but not a leaf of $\mathfrak{t}(x)$, and finally $\mathsf{b}_{n+1}(y) := \{\neg[\alpha]\psi \in FL^{\neg}(\varphi) : \langle \mathfrak{n}, \beta, y\rangle \vDash \neg[\alpha]\psi\}$ if y is a leaf of $\mathfrak{t}(x)$. Finally, $L_{n+1} := L_n \cup \{y : y \text{ a leaf of } \mathfrak{t}(x)\}$.

Now, $\mathfrak{p}_{n+1}$ is finite. Moreover, one can estimate the size of $T(x)$ in the following way. For each $\neg[\alpha]\psi$ we have added a computation trace. Since $\mathfrak{n}$ is the unravelling of $\mathfrak{f}$, a bound on a computation trace for $\alpha; \psi?$ can be given that depends only on the size of f and $\neg[\alpha]\psi$. $\mathsf{b}_{n+1}(y)$ is a subset of $FL^{\neg}(\varphi)$ by construction; moreover, if y is not a leaf, then $\mathsf{b}_{n+1}(y)$ only contains variables. $\mathsf{b}_{n+1}(y)$ contains by construction only formulae of the form $\neg[\alpha]\psi$. Furthermore, $\mathfrak{p}_{n+1}$ is a **DPDL**–frame, as is straightforward to verify. It remains to be seen that $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \chi$ iff $\langle \mathfrak{n}, \beta, y\rangle \vDash \chi$ for all χ in $FL^{\neg}(\varphi)$. This follows from the lemma below.

Lemma 10.5.8. *For all n, all $\chi \in FL^{\neg}(\varphi)$ and all $y \in P_n$:*

$$(\ddagger) \qquad \langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \chi \quad \Leftrightarrow \quad \langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \chi$$

Proof. By induction on the dynamic complexity of χ. For variables this holds by construction, and the only problematic step is for the strict diamond formulae. Let c be the dynamic complexity of $\neg[\alpha]\chi$, $c > 0$, and let the claim have been shown for formulae of complexity $< c$. In the definition of $\mathfrak{p}_{n+1}$ two cases have been distinguished. **Case 1.** $\mathfrak{p}_{n+1}$ is obtained by removing a point (and adding a transition). Let $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \neg[\alpha]\chi$. And assume that $P_{n+1} = P_n - \{x\}$ and that the transition $v \xrightarrow{\zeta} w$ has been added for some w. Then in $\mathfrak{p}_n$, $v \xrightarrow{\zeta} x$, and x and w satisfy the same strict diamond formulae and the same variables. Then $y \neq x$. Consider a path from y falling under a computation trace for $\alpha; \neg\chi?$.

Case 1a. The path does not go through v. Then it is a path in $\mathfrak{p}_n$, by construction. By inductive hypothesis, since the computation trace involves tests for formulae of dynamic complexity $< c$, the path falls under the same computation trace in $\langle \mathfrak{p}_n, \mathsf{b}_n\rangle$. In this case we are done, for then clearly $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$. Now assume $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$. Consider a path π falling under a computation trace in $\langle \mathfrak{p}_n, \mathsf{b}_n\rangle$. Then it is a path in $\mathfrak{p}_{n+1}$, except if the path ends in x. In that case, let π' differ from π only in that the endpoint (which is x) is replaced by w. Otherwise, $\pi' := \pi$. By inductive hypothesis, the path falls under a computation trace for $\alpha; \neg\chi?$ in $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}\rangle$.

Case 1b. Suppose that the path goes through v. Then $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \neg[\alpha]\chi$ by virtue of the fact that either we have (i) $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, v\rangle \vDash \neg\chi$ or (ii) $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, v\rangle \vDash \neg[\alpha']\chi$ for some strict diamond formula $\neg[\alpha']\chi$ in $FL^{\neg}(\varphi)$. Here as in sequel, α' is such that there exists a computation trace σ such that (i) the path from y to v falls under σ and (ii) for every computation trace τ for α' $\sigma; \tau$ is a computation trace for α. In case (i) we are done by inductive hypothesis. In case (ii) observe that then $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, w\rangle \vDash \psi; \neg[\alpha'']\chi$, for some ψ of dynamic complexity $< c$ and some α'' such that for every computation trace τ for α'' the trace $\zeta; \psi?; \tau$ is a computation trace for α'. (α'' may be identical to skip). $w \in L_{n-1}$. Therefore, by construction of the sequence $\langle \mathfrak{p}_i, \mathsf{b}_i\rangle$, there exists a path falling under a computation trace for $\alpha''; \neg\chi?$ inside $\mathfrak{p}_n$. By inductive hypothesis, since it involves tests of degree $< c$ and the fact that the path is a path in $\mathfrak{p}_n$, this is a path falling under a computation trace in $\mathfrak{p}_n$. So, $\langle \mathfrak{p}_n, \mathsf{b}_n, w\rangle \vDash \neg[\alpha'']\chi$. Furthermore, $\langle \mathfrak{p}_n, \mathsf{b}_n, w\rangle \vDash \psi$. It follows that $\langle \mathfrak{p}_n, \mathsf{b}_n, x\rangle \vDash \psi; \neg[\alpha'']\chi$, by construction of $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}\rangle$. Hence $\langle \mathfrak{p}_n, \mathsf{b}_n, v\rangle \vDash \neg[\alpha']\chi$, since $v \xrightarrow{\zeta} x$. From this follows finally that $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$.

Case 2. Let $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \neg[\alpha]\chi$. Consider a path π falling under a computation trace for $\alpha; \chi?$. If that path is inside $\mathfrak{p}_n$ then by inductive hypothesis, $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$. Otherwise, if the path is inside $\mathfrak{t}(x)$, then $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$ be definition, using the previous lemma. Finally, if the path is not properly contained in either, then it leads through x. So, $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \neg[\alpha]\chi$ holds by virtue of the fact that $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, x\rangle \vDash \neg[\alpha']\chi$ for some $\neg[\alpha']\chi$ in $FL^{\neg}(\varphi)$. The latter is equivalent to $\langle \mathfrak{t}(x), \mathsf{b}_{n+1} \restriction T(x), y\rangle \vDash \neg[\alpha']\chi$ which in turn means that $\neg[\alpha']\chi \in \mathsf{b}_n(x)$. By (5.) since $\langle \mathfrak{p}_n, \mathsf{b}_n, x\rangle \vDash \neg[\alpha']\chi$, we have $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$. (Induction on the length of the path leading from y to x.) Now, assume that $\langle \mathfrak{p}_n, \mathsf{b}_n, y\rangle \vDash \neg[\alpha]\chi$. Then if y is not a leaf, there is a path in $\mathfrak{p}_n$ falling under a computation trace for $\alpha; \neg\chi?$. This path falls under the same computation trace for $\alpha; \neg\chi?$ in $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}\rangle$. By inductive hypothesis, therefore, $\langle \mathfrak{p}_{n+1}, \mathsf{b}_{n+1}, y\rangle \vDash \neg[\alpha]\chi$. If y is a leaf different from x, then $\neg[\alpha]\chi \in \mathsf{b}_n(y)$ and so also $\neg[\alpha]\chi \in \mathsf{b}_{n+1}(y)$. Finally, if $y = x$ the claim holds by construction of $\mathfrak{t}(x)$ and $\mathfrak{p}_{n+1}$. ◻

Theorem 10.5.9 (Ben–Ari & Halpern & Pnueli). *$\mathbf{DPDL}$ has the finite model property.*

Now let us return to the question of evaluating a formula in a model. We will propose a special procedure, which we will use in Section 10.7. It is effective, but we make no claims about its efficiency. We may assume that the formula is proper. Thus take a formula φ and a model $\mathfrak{M} = \langle f, \sigma, \beta\rangle$. First, put φ into the form

$$\varphi = \bigvee_{i<k} \bigwedge_{j<k_i} \psi(i, j)$$

where each $\psi(i, j)$ is of the form $\langle\alpha(i, j)\rangle\chi(i, j)$, or $[\alpha(i, j)]\chi(i, j)$. Moreover, we can rewrite these formulae into $\langle\alpha(i, j); \chi(i, j)?\rangle\top$ and $[\alpha(i, j); \neg\chi(i, j)?]\bot$. Without loss of generality, we can assume that $\psi(i, j) = \langle\alpha(i, j)\rangle\top$ or $\psi(i, j) = \neg\langle\alpha(i, j)\rangle\top$. Thus, φ turns into a boolean combination of existence statements for paths. The strategy we are going to use is simply put the following. Enumerate all possible paths in the model, and see whether they fall under one of the descriptions. To make this work, several things have to be assured. First, that we can in principle enumerate all the paths, second that we are able to see whether they fall under one of these descriptions, that is, whether they are of the form $\alpha(i, j)$; and third, since the $\alpha(i, j)$ might use the star, we must find a way to make the procedure finite if the model is finite as well.

We will deal with these problems in the following way. Suppose that we have a path π of length n. Then there are basic programs $\zeta(i)$ such that $x_i \xrightarrow{\zeta(i)} x_{i+1}$. Given $\alpha(i, j)$, under what condition does π fall under $\alpha(i, j)$? If $\alpha(i, j)$ is free of stars, this is easy. For then $\alpha(i, j)$ is a disjunction of chains, each chain being composed from basic programs and tests. Moroever, we can assume that it is a composition of programs of the form $\zeta_i; \psi_i?$, where ζ_i is basic. Then check whether $x_i \xrightarrow{\zeta_i} x_{i+1}$ for all $i < n$, and whether or not $\langle f, \sigma, \beta, x_{i+1}\rangle \vDash \psi_i$. The latter is a task similar to the evaluation of the main formula φ; however, ψ_i has dynamic complexity less than the complexity of φ. Assuming that the latter task can be achieved by the same method as we describe now at least for formulae of lesser dynamic complexity that φ, we have succeeded. Of course, the case that the ψ_i have dynamic complexity 0 is granted to us. We just need to see whether a boolean combination of variables holds at a node. We use β to tell us so. Now, if the program contains stars, we can do the following. Let us assume that the frame has at most p points. Then any pair of points related via α^* can be related via $\alpha^{\leq p}$, which is the union of all α^n such that $n \leq p$. For any sequence of length $> p$ must contain a repetition, which we could have avoided. Thus, we can simply replace the star by a suitable disjunction. This solves the problem of deciding whether a path falls under a path description.

The next problem is that of enumerating the paths. Recall that we were able to replace the star, so we are down to a finite disjunction of chains. This problem is therefore solved if for given maximum length c of a chain we can successfully enumerate all possible paths of length $\leq c$. The choice of c makes sure we really enumerate all paths that possibly stand a chance of falling under a description $\alpha(i, j)$. Now to enumerate these paths starting at a given point x_0, let us first of all assume that given a node w and a basic program ζ, the ζ–successors of w are ordered, or ranked. Moreover, we consider the basic programs ordered, say by their enumeration as ζ_0, ζ_1 etc. Then, the first degree successors of w are ordered as follows. y precedes z if either $w \xrightarrow{\zeta_i} y$ and there is no $j \leq i$ such that $w \xrightarrow{\zeta_j} z$, or else $w \xrightarrow{\zeta_i} y, z$ and y is ranked higher than z as a ζ_i–successor. Finally, we rank the paths starting at x_0 as follows. $\pi \gg \rho$ if either (i) ρ is a prefix of π or (ii) there exists a largest i such that $\pi(i) = \rho(i)$ and $\pi(i+1)$ precedes $\rho(i+1)$ as a first degree successor of $\pi(i)$ $(= \rho(i))$. This ranking corresponds to a depth–first search. Starting at x_0 we always pick the successors of highest priority, going as deep as we can, but at most c steps deep. After that we do what is known as backtracking. To get the next path we go back to the last point $\pi(i)$

(in the numeration of the path, i. e. we choose the highest i with this property) where we could have chosen a successor of lower priority rather than $\pi(i+1)$ and there we go instead to the point immediately lower in priority than $\pi(i+1)$. After this choice we start picking successors of highest rank again, up to depth c, or as deep as we can otherwise. If there exists no point lower in priority that $\pi(i+1)$, the new path ends at $\pi(i)$, that is, has length i. In this way we really enumerate all paths starting at x_0 of length $\leq c$.

The problem of deciding whether φ is accepted at x_0 in a model of size p is now solved as follows. Replace all programs α^* by $\alpha^{\leq p}$. Compute the constant c, the maximum nesting of basic programs in φ. Now open a *table* for φ. We understand a table for φ to be a bit–vector consisting of one bit per subformula $\psi(i, j) = (\neg)\langle\alpha(i, j)\rangle\top$. The table is initialized by putting $t(i, j) = 0$, if $\psi(i, j)$ is of the form $\langle\alpha(i, j)\rangle\top$, and by putting $t(i, j) = 1$ if $\psi(i, j)$ is of the form $\neg\langle\alpha(i, j)\rangle\top$. Now start generating all paths of length $\leq c$ and check whether they fall under one of the $\alpha(i, j)$. This may require a recursive step, that is, opening up tables for formulae of lower rank, but this is immaterial. Now suppose we have generated a new path π. Then for each index (i, j) we check whether or not the path falls under $\alpha(i, j)$. If so, we overwrite the entry $t(i, j)$ to 1; if not, $t(i, j)$ remains untouched. At the end of this procedure, $\langle\alpha(i, j)\rangle\top$ comes out true if $t(i, j) = 1$, because the latter means $t(i, j)$ has been overwritten at least once, so must have found a path falling under $\alpha(i, j)$. It comes out false if $t(i, j) = 0$, because that means we have never touched $t(i, j)$, which in turn means that no path has been found. Then for $\neg\langle\alpha(i, j)\rangle\top$ we also know whether or not it is accepted. The number of tables that we need to keep at a given moment of time is bounded by the dynamic complexity of φ. This means that apart from the procedure that generates paths we only need a finite memory to do the bookkeeping.

If φ is proper we can deduce a number of useful properties of the computation procedure. Suppose that we want to evaluate φ at x_0. The path generator starts with generating paths going through the successor of φ with highest priority. Call it x_1. Then it chooses a successor of highest priority, x_2. And so on. This defines first of all a linear order of the paths generated, which we can also represent as an enumeration function $\ell : j \to f$, j some natural number. Say that the computation **exits** u at i, $i < j$, if the end point of $\ell(i)$ is u, and for no larger i' u is the endpoint of $\ell(i')$. Now, define a priority $\sqsubset$ over the transit of x_0 by saying that $v \sqsubset w$ if the computation exits v at i and the computation exits w at i' and $i < i'$. This is defined only for the main computation, generating paths from the point x_0, at which we want to evaluate the main formula φ. However, in the course of evaluating φ the procedure generates a path with intermediate points and calls on subroutines to check certain subformulae of lower dynamic complexity at certain nodes u. This procedure may by itself also start generating paths. The set of **active nodes** of a computation at time t is the set of nodes at the main path active at t, together with the set of nodes active in the subroutine just at work at time t, plus whatever subroutines are called at t. The set of active nodes is a union of at most d paths, d the dynamic complexity of φ. Now say

— finally — that the overall computation **totally exits** x at t if x is active in it at t, but for no point of time $t' > t$, x is active at t'.

Notes on this section. Already in the abovementioned paper, Ben–Ari, Halpern and Pnueli have shown that **DPDL** is EXPTIME–complete. This can be established using constructive reduction. The procedure described at the end of this section is called a *model checking procedure*. Model checking, especially its complexity, is an important topic in temporal logic (see E. A. Emerson [**54**]). In general, the model checking problem is as follows: given a finite model $\mathfrak{M}$ and a formula φ, is $\mathfrak{M}$ a model for φ? Of course, this is in general a decidable problem, so one is interested in the number of steps needed to compute the answer, or alternatively, in the amount of storage space. This number may depend both on the size of $\mathfrak{M}$ as well as the length of φ. The procedure described above is not very efficient. If we want to save space, we should precompute the relations corresponding to the programs occurring in φ. This can be done iteratively, by induction on $FL(\varphi)$. In tandem, we shall evaluate the assignments of the subformulae. All this needs space quadratic in the size of the model and linear in the size of $FL(\varphi)$. It is easy to see that $card(FL(\varphi))$ is linear in the length of φ. The time is a polynomial in the sum of the size of the model and the length of the formula (this was already observed in the paper by M. J. Fischer and R. E. Ladner [**69**]). Notice that despite these low bounds, satisfiability of formulae in **PDL** takes exponentially many steps to compute (see Section 10.3). For the models to be built are in the worst case exponential in the length of φ.

Exercise 370. Show Lemma 10.5.3.

Exercise 371. Show Lemma 10.5.7.

Exercise 372. Show that **DPDL** with converse does not possess the finite model property. (This is due to Joseph Halpern.)

10.6. The Unanswered Question

The remaining two sections will deal mainly with the problem of interpolation for **PDL**. This is one of the major open problems in this area. Twice a solution has been announced, in [**138**] and in [**33**], but in neither case was it possible to verify the argument. The argument of Leivant makes use of the fact that if $\varphi \vdash_{\mathbf{PDL}} \psi$ then we can bound the size of a possible countermodel so that the star α^* only needs to search up to a depth d which depends on φ and ψ. Once that is done, we have reduced **PDL** to **EPDL**, which definitely has interpolation because it is a notational variant of polymodal **K**. However, this is tantamount to the following. Abbreviate by $\mathbf{PDL}^n$ the strengthening of **PDL** by axioms of the form $[\alpha^*]p \leftrightarrow [\alpha^{\leq n}]p$ for all α. Then, by the finite model property of **PDL**, **PDL** is the intersection of the logics $\mathbf{PDL}^n$. Unfortunately, it is not so that interpolation is preserved under intersection. A counterexample is the logic **G.3**, which fails to have interpolation while all proper extensions have interpolation, since they have all constants, by Theorem 1.6.4. We have not been able to decide the question of interpolation for **PDL**. But some answers can be given that point to the fact that **PDL** does indeed have interpolation. Also, we wish to show that no significant fragment of **PDL** has interpolation. The picture that emerges is this. If we start with a polymodal language $\mathbf{K}_n$, then interpolation obtains, because the language is not so strong. As soon as we add just one more operator, the star closure of the basic programs, we can regain interpolation only if we add at least all test–free programs of **PDL**. We believe that this latter fragment of **PDL** does in fact have interpolation, and show moreover that if it does, **PDL** must as well have interpolation.

Let us have the basic programs $\zeta_0, \zeta_1, \ldots, \zeta_{n-1}$. Put $\gamma := \zeta_0 \cup \zeta_1 \cup \ldots \cup \zeta_{n-1}$. We will show first that if a fragment of **PDL** contains at least the program γ^*, then it has interpolation only if it is closed under union, composition and star. This generalizes an observation of Maksimova in [**152**]. To understand this result, let us call a **fragment** of **PDL** a modal logic which contains some subset of the programs definable from Π_0 plus the relevant axioms. There are various interesting fragments of **PDL**. One is the fragment consisting of the basic programs and the star closure of the basic programs, another is test–free **PDL**, where we close Π_0 only under union, composition and star.

Theorem 10.6.1. *Let $\boldsymbol{PDL}^-$ be a fragment of $\boldsymbol{PDL}$ containing at least the star closure of the basic programs. Then it has interpolation only if it is at least the fragment of test–free $\boldsymbol{PDL}$.*

Proof. We show that it is possible to give an implicit definition of any given regular language, so that **PDL**$^-$ can have the global Beth–property only if it contains the closure of the basic programs under union, composition and star. For the proof, let L be a regular language, definable by a regular expression β over the basic modalities, ζ_i, $i < m$. There is a finite automaton with m states recognizing $L^{\smile}$. As in the proof of Kleene's Theorem we can describe the automaton recognizing L with a system of equations.

$$\begin{array}{lclllll}
X_0 & = & \epsilon \cup & X_0 \cdot a_{00} & \cup X_1 \cdot a_{10} & \cup \ldots \cup X_{n-1} \cdot a_{n-1,0} \\
X_1 & = & & X_0 \cdot a_{01} & \cup X_1 \cdot a_{11} & \cup \ldots \cup X_{n-1} \cdot a_{n-1,1} \\
\vdots & & & \vdots & \vdots & \vdots \\
X_{n-1} & = & & X_0 \cdot a_{0,n-1} & \cup X_1 \cdot a_{1,n-1} & \cup \ldots \cup X_{n-1} \cdot a_{n-1,n-1}
\end{array}$$

If F is the set of accepting states, $L = \bigcup_{i \in F} X_i$. Now, take a propositional letter q_i for each state i, and one more letter, q^*. Let $A(\vec{q}, q^*)$ be the conjunction of the following set of formulae.

$$\begin{array}{llllll}
q_0 & \leftrightarrow & b_0 \vee & \langle a_{00}\rangle q_0 & \vee \langle a_{10}\rangle q_1 & \vee \ldots \vee \langle a_{n-1,0}\rangle q_{n-1} \\
q_1 & \leftrightarrow & b_1 \vee & \langle a_{01}\rangle q_0 & \vee \langle a_{11}\rangle q_1 & \vee \ldots \vee \langle a_{n-1,1}\rangle q_{n-1} \\
\vdots & & & \vdots & \vdots & \vdots \\
q_{n-1} & \leftrightarrow & b_{n-1} \vee & \langle a_{0,n-1}\rangle q_0 & \vee \langle a_{1,n-1}\rangle q_1 & \vee \ldots \vee \langle a_{n-1,n-1}\rangle q_{n-1}
\end{array}$$

Here, $b_i := q^*$ if $i \in F$ and $b_i := \bot$ else. In addition, let $B(\vec{q}, q^*)$ be the conjunction of

$$q_i \to \bigwedge_{i \neq j} \neg q_j, \quad \bigvee_{i<n} q_i, \quad q^* \to \bigwedge_{j<m} [\zeta_j] q^*, \quad [\gamma^*]\langle\gamma^*\rangle q^*, \quad \bigwedge_{i<n} (\langle\gamma\rangle q_i \to [\gamma] q_i) \,.$$

Put $C(\vec{q}, q^*) := A(\vec{q}, q^*) \wedge B(\vec{q}, q^*)$. The proof is complete if we prove the following three things. (i) $C(\vec{q}, q^*)$ is a global implicit definition of q_0, (ii) An explicit definition is $q_0 \leftrightarrow \langle\beta\rangle q^*$, (iii) No explicit definition can be found if β is not definable by a formula in $\mathbf{PDL}^-$. For (i), notice that since **PDL** has the finite model property, so does $\mathbf{PDL}^-$. Now take a finite model for $C(\vec{q}, q^*)$. We show that the values of the q_i are completely determined by the values of q^*. First of all, $B(\vec{q}, q^*)$ is chosen so that if q^* is true at a point, it remains true throughout the transit of that point. Second, if there is a point x and a one–step ζ_i–successor y satisfying q_i, then all one–step successors satisfy q_i. Now we show the following. $x \vDash q_i$ iff there exists a path w from x to y where $y \vDash q^*$, and $\vec{w} \in \mathsf{L}[j, i]^{\smile}$ for some accepting state j. First of all, observe that for each x there exists a i such that $x \vDash q_i$, and this q_i is unique. Now we do induction of the smallest path $\vec{w}$ from x to a point $y \vDash q^*$. Suppose, $\vec{w}$ is of length 0. Then $x = y$ and so $x \vDash q^*$ iff $y \vDash q^*$, and $\epsilon \in \mathsf{L}[j, j]^{\smile}$ for an accepting state j. Now let $\vec{w} = \zeta_k \cdot \vec{w}'$ and assume $x \xrightarrow{\zeta_k} x' \xrightarrow{\vec{w}'} y$ for some $y \vDash q^*$. Then $x' \vDash q_s$ for some s. Then any ζ_k–successor of x satisfies q_s, and among them we choose the one through which the minimal path $\vec{w}'$ goes. By induction hypothesis there exists a path $\vec{v}$ to a point y' such that $y' \vDash q^*$ and $\vec{v} \in \mathsf{L}[j, s]^{\smile}$ for some accepting state j. Then $\zeta_k \cdot \vec{v} \in \mathsf{L}[i, s]^{\smile}$, by the fact that the automaton has a transition $s \xrightarrow{k} i$. Hence, the q_i are implicitly defined, and equivalent to $\langle\beta_i\rangle q^*$, where β is the regular expression belonging to $\bigcup_{j \in F} \mathsf{L}[i, j]^{\smile}$. The claim follows in the particular case of $i = 0$. Finally, for (iii) notice that there exist the following models. Take any word w in the alphabet $A = \{\zeta_i : i < n\}$. Let the frame consist of the prefixes of $\vec{w}$ and put $\vec{v} \xrightarrow{\zeta_i} \vec{v}'$ iff $\vec{v} = \vec{v}' \cdot \zeta_i$. The frame codes nothing but $\vec{w}$ in reverse order. On this frame, put $\beta(q^*) := \{\epsilon\}$, ϵ the empty word. Then there is a unique valuation β^+ making $C(\vec{q}, q^*)$ globally true. At any point $\vec{v}$ in this frame, $\vec{v} \vDash q_i$ exactly if $\vec{v}^{\smile} \in \mathsf{L}[j, i]$ for some accepting state j. Therefore, no simpler definition for q_0 can be given. ◻

On other hand, if test–free **PDL** has interpolation, then full **PDL** also has interpolation. This is the content of the next theorem.

Theorem 10.6.2. *$\mathbf{PDL}$ has interpolation if test–free $\mathbf{PDL}$ has interpolation.*

Proof. We use constructive reduction. The argument is therefore of a more general character. First of all, we can restrict ourselves to finite Π_0, so that we have weak transitivity. Then global reduction sets can be reduced to local reduction sets. Take a set Δ and let

$$X_?(\Delta) := \{\langle\psi?\rangle\chi. \leftrightarrow .\psi \wedge \chi : \langle\psi?\rangle\chi \in FL(\Delta)\} .$$

It is enough to show that these sets are global reduction sets. For they split, and therefore interpolation can be deduced for **PDL** as follows. Full **PDL** is the logic which is obtained from test–free **PDL** with infinitely many basic programs by adding the test axioms for all programs $[\varphi?]$. (This is to say, from the basic modalities we select some to play the role of tests, and call them $[\varphi?]$. For exactly those modalities, the test axioms are added.) Now take a model $\langle\mathfrak{f},\beta\rangle$ for φ in which the reduction formulae hold globally. We can assume $\mathfrak{f}$ to be a dynamic frame. The only problem with that frame is that the test–programs are interpreted freely. Let $\mathfrak{g}$ differ from $\mathfrak{f}$ in that $x \stackrel{\psi?}{\to} y$ iff $x = y$ and $\langle\mathfrak{f},\beta\rangle \vDash \psi$, for all $\psi \in FL(\varphi)$. We show that for all formulae $\chi \in FL(\varphi)$, and all x,

$$(\ddagger) \qquad \langle\mathfrak{f},\beta, x\rangle \vDash \chi \qquad \text{iff} \qquad \langle\mathfrak{g},\beta, x\rangle \vDash \chi$$

After that we can actually drop the assignment of the test programs, and obtain a full dynamic model for φ. But now for the proof of (‡). Clearly, for variables and boolean junctors there is nothing to show. So, let us take the case of a formula $\langle\alpha\rangle\omega$. If $\alpha = \beta \cup \gamma$, or $\alpha = \beta;\gamma$ or $\alpha = \beta^*$, then we can also use the induction hypothesis in a straightforward way. There remain the cases $\alpha = \zeta_i$ and $\alpha = \psi?$. The first is also straightforward since the interpretation of the ζ_i has not changed. So let us proceed to the really critical case, $\chi = \langle\psi?\rangle\omega$. Here, if $\langle\mathfrak{f},\beta, x\rangle \vDash \langle\psi?\rangle\omega$ then also $\langle\mathfrak{f},\beta, x\rangle \vDash \psi;\omega$. By induction hypothesis, $\langle\mathfrak{g},\beta, x\rangle \vDash \psi;\omega$ and so $\langle\mathfrak{g},\beta, x\rangle \vDash \langle\psi?\rangle\omega$. And conversely. ∎

We will now present some particular cases of formulas where interpolants can be found. In view of the preceding result it is enough to consider test–free formulae, and this is what we will do now. The method is a rather explicit construction of the interpolant, using a method analogous to the one described at the end of Chapter 3.8. We will illustrate it with the case where $\varphi \vdash \psi$ and φ is of dynamic complexity 1, that is, of the form $\bigvee \varphi_i$, each φ_i a conjunction of fomulas of the form $\langle\beta\rangle\chi$, $[\beta]\chi$, where χ is nonmodal. In this case we must have $\varphi_i \vdash \psi$ for all i, so it is enough if interpolants can be produced for each φ_i individually. Now recall the strategy of the Chapter 3.8. Let us define $\varphi^\top$ to be the result of replacing p by $\top$ where it occurs positively, and by $\bot$ where it occurs negatively. Then we have $\varphi \vdash \varphi^\top$ by construction. It remains to be verified that $\varphi^\top \vdash \psi$. To have that it is enough to show that if we have a model for $\varphi^\top;\neg\psi$ then we can also produce a model for $\varphi;\neg\psi$. Since the latter does not obtain, $\varphi^\top \vdash \psi$ is proved. To make that work, φ has to be carefully prepared before the elimination of the variable p takes place. The reason is that forgetting p can lead to a failure in the strategy, because $\varphi^\top$ now allows for essentially more models than

φ itself. A case in point is when we have subformulae of the form $\Diamond p \wedge \Box\neg p$, or $\Box p \wedge \Box\neg p$. In the first case we have $(\Diamond p \wedge \Box\neg p)^{\top} = \Diamond\top \wedge \Box\top$, which is equivalent to $\Diamond\top$. On the other hand, the original formula is simply false, that is, deductively equivalent to $\bot$. It is the latter formula that must be chosen as an interpolant, and not $\Diamond\top$. This problem does not arise with the formula $\Diamond p \wedge \Diamond\neg p$. Why is this so? The reason is that in the first two cases we have a formula that speaks over *all* successors of a point. Hence, if another formula also speaks about successors, then the valuation on the successors must be matched with the requirements of the first formula. If a point accepts $\Box p$, then all successors must satisfy p, and so $\Diamond\neg p$ cannot hold at that point. However, if we have formulas $\Diamond p \wedge \Diamond\neg p$ then no conflict arises, because we can always arrange it that a point has two different successors, one satisfying p, the other satisfying $\neg p$.

Now return to the case where φ is a conjunction of formulae of the form $\langle\beta\rangle\chi$ or $[\beta]\chi$, χ nonmodal, β test–free. We want to rewrite β in a similar way as we have done with polymodal formulae. However, this time matters are even more complex. For example, the programs $(\alpha^2)^+$ and $(\alpha^3)^+$, although different, give rise to subtle interactions. From $[(\alpha^2)^+]p$ and $\langle(\alpha^3)^+\rangle\neg p$ we can deduce that $\neg\langle\alpha^6\rangle\top$. Hence we must reckon beforehand with a new operator, α^6. To care for this, we analyse the possible intersections of programs. Recall that the regular languages over a finite alphabet are closed under intersection. Therefore, let us take a second look at φ. Suppose that the regular expressions occurring in φ are β_i, $i < n$. Then for each subset $S \subseteq n$ we let γ_S be the regular expression corresponding to

$$\bigcap_{i\in S}\beta_i \cap \bigcap_{i\notin S} -\beta_i$$

γ_S exists by the results of Section 9.4. The following is then clear. The languages corresponding to the γ_S are mutually disjoint, and

$$\beta_i = \bigcup_{i\in S}\gamma_S$$

We now change the 'program basis' in φ by replacing talk of β_i by talk of γ_S. Hence, we can assume that φ is a conjunction of formulae of the form $\langle\gamma_i\rangle\chi$, $[\gamma_i]\chi$, χ nonmodal, γ_i test–free and mutually disjoint. Moreover, we assume that for no i, ϵ falls under γ_i, that is, we assume that the programs are proper. With this given, we can compute the interpolant in the same way as for polymodal $\mathbf{K}$. In fact, let us make the reduction as follows. Each $[\gamma_i]$ is regarded now as a primitive modality, with dual operator $\langle\gamma_i\rangle$. Then compute the interpolant $\varphi^{\top}$ as if working in polymodal $\mathbf{K}$, letting $\Diamond_i$ replace $\langle\gamma_i\rangle$. This is possible for the following reason.

LEMMA 10.6.3. *Let Π_0 be finite, and γ_i, $i < n$, be regular test–free programs over Π_0 such that (i.) no path falls under both γ_i and γ_j, $i \neq j$, (ii.) ϵ does not fall under any γ_j, $j < n$. Let $\varphi \in \boldsymbol{K}_n$ be a formula of degree 1, and $p(\varphi)$ be the result of replacing each occurrence of $\Box_i$ by $[\gamma_i]$, for all $i < n$. Then $\varphi \in \boldsymbol{K}_n$ iff $p(\varphi) \in \boldsymbol{PDL}$.*

Proof. Assume that $\varphi \notin \mathbf{K}_n$; then there is a finite Kripke–frame $\mathfrak{f}$ and β and w_0 such that $\langle \mathfrak{f}, \beta, w_0\rangle \models \neg\varphi$. Since φ is of degree 1, it is enough to assume that f consists of the 1–transit of w_0. Now fix for every $i < n$ a finite sequence $\sigma(i) \subseteq \Pi_0^*$ falling under γ_i. Let $\sigma(i)$ have length $\ell(i)$. Now define

$$f^p := \{w_0\} \cup \{\langle \tau, i, x\rangle : \tau \text{ a prefix of } \sigma(i), w_0 \lhd_i x\} \ .$$

Further, let $\langle \epsilon, i, x\rangle := w_0$. Then

$$\langle \tau, i, x\rangle \xrightarrow{\zeta} \langle \tau', i', x'\rangle \qquad \text{iff} \qquad \begin{cases} & x = x', i = i', \tau' = \tau^\frown\alpha \\ \text{or} & \tau = \epsilon, \tau' = \zeta \end{cases}$$

(Here, ζ denotes as usual an elementary program.) Fix for $i < n$ a world $y(i)$ such that $w_0 \lhd_i y(i)$. This defines the dynamic Kripke–frame $\mathfrak{f}^p$. Now

$$\begin{aligned} \gamma(p) := \quad & \{\langle \sigma(i), i, x\rangle : x \in \beta(p)\} \\ \cup \quad & \{\langle \epsilon, i, x\rangle : w_0 \in \beta(p)\} \\ \cup \quad & \{\langle \tau, i, x\rangle : \tau \text{ falls under } \gamma_i, \tau \neq \sigma(i), y(i) \in \beta(p)\} \end{aligned}$$

This is well–defined since the i such that τ falls under γ_i is unique if it exists. Moreover, ϵ does not fall under any γ_i. We claim that $\langle \mathfrak{f}^p, \gamma, w_0\rangle \models p(\varphi)$. To that end, we prove that (1.) for a formula $\Box_i\mu$, μ nonmodal, $\langle \mathfrak{f}^p, \gamma, w_0\rangle \models \Box_i\mu$ iff $\langle \mathfrak{f}, \beta, w_0\rangle \models [\gamma_i]\mu$; that (2.) for a nonmodal formula μ, $\langle \mathfrak{f}^p, \gamma, w_0\rangle \models \mu$ iff $\langle \mathfrak{f}, \beta, w_0\rangle \models \mu$. (2.) is immediate from the definition of γ. For (1.) let $\langle \mathfrak{f}^p, \gamma, w_0\rangle \models \Box_i\mu$. Take a x such that $w_0 \lhd_i x$. Then $\langle \sigma(i), i, x\rangle \in \overline{\gamma}(\mu)$ and so, by definition of γ, $x \in \overline{\beta}(\mu)$. Hence $\langle \mathfrak{f}, \beta, w_0\rangle \models \Box_i\mu$. Assume $\langle \mathfrak{f}, \beta, w_0\rangle \models \Box_i\mu$. Take a point $\langle \tau, i, x\rangle$ such that $w_0 \xrightarrow{\gamma_i} \langle \tau, i, x\rangle$. Then $\tau \neq \epsilon$. **Case 1.** $\tau = \sigma(i)$. Then $w_0 \lhd_i x$ and by definition of γ, $\langle \sigma, i, x\rangle \in \overline{\gamma}(\mu)$. **Case 2.** $\tau \neq \sigma(i)$. Let τ fall under γ_j. Then $\langle \tau, i, x\rangle \in \overline{\gamma}(\mu)$ iff $\langle \sigma(j), j, x\rangle \in \overline{\gamma}(\mu)$ iff $y(j) \in \overline{\beta}(\mu)$. By choice of $y(j)$, $w_0 \lhd_j y(j)$. Hence $y(j) \in \overline{\beta}(\mu)$ and therefore $\langle \sigma(j), j, x\rangle \in \overline{\gamma}(\mu)$ and this shows that $\langle \tau, i, x\rangle \in \overline{\gamma}(\mu)$. Hence, $\langle \mathfrak{f}^p, \gamma, w_0\rangle \models \neg p(\varphi)$ and so $p(\varphi) \notin$ **PDL**. Conversely, assume that $p(\varphi) \notin$ **PDL**. Then there exists a finite dynamic Kripke–frame $\mathfrak{g}$ and a model $\langle \mathfrak{g}, \gamma, w_0\rangle \models \neg p(\varphi)$. Put $y \lhd_i z$ iff $y = w_0$ and $w_0 \xrightarrow{\gamma_i} z$. This defines $\mathfrak{f}$. Let $\beta(p) := \gamma(p)$. We claim that $\langle \mathfrak{f}, \beta, w_0\rangle \models \neg\varphi$. To that end, observe that for a nonmodal formula μ, $\langle \mathfrak{f}, \beta, w_0\rangle \models \mu$ iff $\langle \mathfrak{g}, \gamma, w_0\rangle \models \mu$. Moreover, for a formula $\Box_i\mu$, μ nonmodal, $\langle \mathfrak{f}, \beta, w_0\rangle \models \Box_i\mu$ iff $\langle \mathfrak{g}, \gamma, w_0\rangle \models [\gamma_i]\mu$. For $w_0 \lhd_i x$ iff $w_0 \xrightarrow{\gamma_i} x$. ∎

Proposition 10.6.4. *$\boldsymbol{PDL}$ has interpolants for $\varphi \vdash \psi$ if φ and ψ are dynamic complexity* ≤ 1.

Proof. It is enough to show this for formulae without test, by Theorem 10.6.2. Let $\varphi \vdash \psi$ and let φ and ψ be of dynamic complexity ≤ 1. Then φ and ψ are locally equivalent to formulae $\widehat{\varphi}$ and $\widehat{\psi}$ which are of dynamic complexity ≤ 1 and a boolean combination of nonmodal formulae and formulae $[\gamma_i]\mu$, $i < n$, such that μ is nonmodal; moreover, the γ_i do not contain the empty path, and the path sets of γ_i and γ_j are disjoint for $i \neq j$. Then $\widehat{\varphi} = p(\tau)$ and $\widehat{\psi} = p(\upsilon)$ for some $\tau, \upsilon \in \mathbf{K}_n$. By the previous lemma $\tau \vdash \upsilon$. Both τ and υ are of degree ≤ 1. There exists an interpolant

ρ in $\mathbf{K}_n$ of modal degree ≤ 1. Then, again by the previous lemma, $p(\tau) \vdash p(\rho)$ and $p(\rho) \vdash p(\upsilon)$. So, $p(\tau)$ is an interpolant of φ and ψ, since $var(p(\rho)) = var(\rho)$; for the γ_i are free of tests. □

Notice that the same argument does not work if we iterate the programs. There are interactions between the γ_i, but they are not noticeable in the first iteration. For example, if we have a single program, α^*, then taking this as a primitive program is fine unless we study iterations, such as $\alpha^*; \alpha^*$, which is — namely — the same as α^*.

Exercise 373. Show that $\mathbf{PDL}(\omega, \omega)$ has interpolation if it holds for every n that $\mathbf{PDL}(\omega, n)$ has interpolation.

10.7. The Logic of Finite Computations

Finally, we want to study the logic of finite computations, which we call **PDL.f**. It is the logic of all those structures in which no computation can run forever. We have encountered in the monomodal case an axiom that ensures such a property, the axiom **G**. Suppose that we have finitely many basic programs, $\Pi_0 := \{\zeta_i : i < n\}$, then putting $\gamma := \bigcup_{i<n} \zeta_i$ we obtain $\mathbf{PDL.f}(\omega, n)$ by adding to $\mathbf{PDL}(\omega, n)$ the axiom

$$[\gamma^+]([\gamma^+]p \to p) \to [\gamma^+]p \ .$$

If we have infinitely many basic programs, **PDL.f** is obtained by adding all postulates of its weak fragments. Thus, $\mathbf{PDL.f} := \bigcup_{n\in\omega} \mathbf{PDL.f}(\omega, n)$. We will prove two facts in this section. First, that **PDL.f** has the finite model property, from which it follows by constructive reduction that **DPDL.f** has the finite model property as well. Second, that **DPDL.f** does not have the global Beth–property, from which the same follows for **PDL.f**. The argument will be a constructive reduction to plain **PDL**. It does not use splitting reduction sets, and so the negative example on the Beth–property does not transfer to **PDL**. It follows from the fact that **PDL.f** has the finite model property that it is the logic of finite computations, whence the title of this section. It also follows that **DPDL.f** is the logic of finite computations of deterministic basic programs.

Before we enter the first proof, let us get some intuition about **PDL.f**. Clearly, if $[\gamma^+]$ satisfies the **G**–postulate, then the γ^+–transitions in a Kripke–frame must be without circles. Since γ is the union of all ζ_i, it follows that the ζ_i must be irreflexive. The converse does not hold, however. Now, take a formula φ. Let $\mathfrak{Fl}(\varphi)$ be the collection of all atoms in the boolean algebra generated by $FL(\varphi)$. We propose as reduction sets the set of all formulae

$$[\gamma^*](\langle\gamma^+\rangle A \to \langle\gamma^+\rangle(A \wedge \neg\langle\gamma^+\rangle A))$$

where $A \in \mathfrak{Fl}(\varphi)$. Now let $\langle\mathfrak{f}, \beta, x\rangle$ be a finite **PDL**–model for φ and the reduction formulas above. Define a new model in as follows. Call a point x **maximal** if there exists an atom A such that $x \vDash A \wedge \neg\langle\gamma^+\rangle A$. Then g is defined to be the set of all maximal points. Moreover, we put $x \xrightarrow{\zeta_i} y$ for two maximal points if there exists a y^- in the original model such that $x \xrightarrow{\zeta_i} y^- \xrightarrow{\gamma^*} y$ and y^- and y satisfy the same atom in $\mathfrak{Fl}(\varphi)$. This defines the frame $\mathfrak{g}$. This new frame contains no cycles. For each point has been chosen to satisfy a formula $A \wedge \neg\langle\gamma^+\rangle A$ in $\mathfrak{f}$, so each chain $y_0 \xrightarrow{\gamma^+} y_1 \xrightarrow{\gamma^+} y_2 \ldots y_{n-1} \xrightarrow{\gamma^+} y_n$ is without cycles.

Now $\mathfrak{g}$ is based on a subset of f, so we take the valuation β restricted to g, denoted also by β. We show now

(†) $\quad \langle \mathfrak{g}, \beta, y\rangle \vDash \chi$ iff $\langle \mathfrak{f}, \beta, y\rangle \vDash \chi$

for all $\chi \in FL(\varphi)$ and $y \in g$. The proof is by induction on the number of γ^+–successors of x in $\mathfrak{g}$. Assume x has no γ^+–successors in $\mathfrak{g}$. Then it actually has no γ^+–successors in $\mathfrak{f}$ either. Namely, if x has a successor in $\mathfrak{f}$, y, then y satisfies an atom A that x does not satisfy (by choice of x) and has a maximal successor y^+ satisfying A, and so $x \xrightarrow{\gamma^+} y^+$ in $\mathfrak{f}$ and so $x \xrightarrow{\gamma^+} y^+$ in $\mathfrak{g}$. Thus, the generated subframe of x in $\mathfrak{g}$ is isomorphic to the generated subframe generated in $\mathfrak{f}$; the models defined on them are as well. Hence, the two models satisfy the same formulae. Now let us assume that all successors of x satisfy (†). Then (†) holds for x and $\chi = p$, p a variable. The induction steps for $\wedge$ and $\neg$ are straightforward. The only critical step is $\chi = \langle\alpha\rangle\psi$. Again, many cases are easy. If $\alpha = \beta \cup \gamma$, $\alpha = \psi?$, $\alpha = \beta;\gamma$ or $\alpha = \beta^*$ then we can reduce the problem once more. The case $\alpha = \zeta_i \in \Pi_0$ remains. From left to right let $\langle \mathfrak{g}, \beta, x\rangle \vDash \langle\zeta_i\rangle\chi$. Then there is a ζ_i–successor y of x in $\mathfrak{g}$ such that $\langle \mathfrak{g}, \beta, y\rangle \vDash \chi$. By induction hypothesis, since y has less γ^+–successors, $\langle \mathfrak{f}, \beta, y\rangle \vDash \chi$. By construction, there exists a y^- such that in $\mathfrak{f}$ $x \xrightarrow{\zeta_i} y^- \xrightarrow{\gamma^*} y$, and y and y^- satisfy the same atom. Hence $\langle \mathfrak{f}, \beta, y^-\rangle \vDash \chi$ and so $\langle \mathfrak{f}, \beta, x\rangle \vDash \langle\zeta_i\rangle\chi$. This shows one direction. For the other, assume that $\langle \mathfrak{f}, \beta, x\rangle \vDash \langle\zeta_i\rangle\chi$. Thus there exists a y such that $x \xrightarrow{\zeta_i} y$ and $\langle \mathfrak{f}, \beta, y\rangle \vDash \chi$. Either y is already maximal for its atom A or else — by choice of the reduction sets — $\langle \mathfrak{f}, \beta, y\rangle \vDash \langle\gamma^+\rangle(A \wedge \neg\langle\gamma^+\rangle A)$. Hence we can find a y^+ such that $y \xrightarrow{\gamma^*} y^+$ satisfying A and being maximal for A. By construction of $\mathfrak{g}$ we have $x \xrightarrow{\zeta_i} y^+$ in $\mathfrak{g}$. By induction hypothesis, since y^+ has less sucessors than x, $\langle \mathfrak{g}, \beta, y^+\rangle \vDash A$, and so $\langle \mathfrak{g}, \beta, y^+\rangle \vDash \chi$. From this we deduce $\langle \mathfrak{g}, \beta, x\rangle \vDash \langle\zeta_i\rangle\chi$, as desired.

Theorem 10.7.1. *The logic* $\boldsymbol{PDL.f}$ *has the finite model property.*

This proof method is basically the same as in the monomodal case for **G**, and so the expectation is that it can be generalized. In the exercises, some of these generalization are considered. One consequence is

Corollary 10.7.2. *The logic* $\boldsymbol{DPDL.f}$ *has the finite model property.*

The proof is by constructive reduction. Use the formulae as described in the last section. The proof procedure works exactly in the same way. This last result will be used in the remaining part, where we show that **PDL.f** fails to have the Beth property. For notice that the reduction sets for **DPDL** split, so that the failure of **DPDL.f** is passed down to **PDL.f**.

Now we proceed to the demonstration that the Beth–property fails for the logics of finite computation. The formula we consider is the following. It uses three elementary programs, ζ_0, ζ_1 and ζ_2, and $\gamma := \zeta_0 \cup \zeta_1 \cup \zeta_2$.

$$A(p,q) := p. \leftrightarrow .(q \wedge [\gamma]\bot) \vee \left\{ \begin{array}{cc} & \langle\zeta_0\rangle p \wedge \langle\zeta_1\rangle p \\ \vee & \langle\zeta_1\rangle p \wedge \langle\zeta_2\rangle p \\ \vee & \langle\zeta_2\rangle p \wedge \langle\zeta_0\rangle p \end{array} \right\}$$

Lemma 10.7.3. *$A(p,q)$ is an implicit global definition of p in* $\boldsymbol{DPDL.f}$.

Proof. We have to show that

$$A(p,q); A(r,q) \Vdash p \leftrightarrow r$$

To see that, we have to check only the finite models of **DPDL.f**. Now, suppose that we have a global model for $A(p,q)$ rooted at x_0. We show the uniqueness of the valuation for p, given the valuation on q. Namely, the valuation of p can be reconstructed by induction on the number of γ^+–successors as follows. At points without γ^+–successors, p is true if and only if q is true. Now at a point x with γ^+–successors, p is true iff it is true at exactly two immediate successors. Since the immediate successors of x have less γ^+–successors, this is actually well–defined and unique. Therefore we have an implicit definition. ◻

Theorem 10.7.4. *The logic* $\boldsymbol{DPDL.f}$ *fails to have the global Beth–property.*

Proof. We aim to show that there is no explicit definition for p in $A(p,q)$. In view of the results of Section 10.5 it is enough if we show that there exists no evaluation procedure using finite memory which can tell us whether or not p holds at a given point. For any formula $B(q)$ can be evaluated with such a procedure. We work on special frames, which are ternary branching trees of depth δ, coded as sequences of numbers $0, 1, 2$ of length $\leq \delta$. The root is the empty sequence, ϵ. Moreover, for sequences s_1 and s_2 we put $s_1 \overset{\zeta_i}{\to} s_2$ iff $s_2 = s_1 \cdot i$. An evaluation procedure for a formula $B(q)$ will start enumerating the paths giving priority to 0 over 1 and to 1 over 2. It generates the path $\langle \epsilon, 0, 00, \ldots\rangle$ up to length δ. Let $\delta = 3$. Then after $\langle \epsilon, 0, 00, 000\rangle$ the next path is $\langle \epsilon, 0, 00, 001\rangle$ and then $\langle \epsilon, 0, 00, 002\rangle$ after which comes $\langle \epsilon, 0, 01, 010\rangle$. And so on. Recall that in addition $B(q)$ calls subroutines which themselves generate paths. The set of active nodes, however, is connected. Consider now what happens if the main subroutine exits the point 0. Then it next goes to $\langle \epsilon, 1, 10, 100\rangle$ and starts working there. No subroutines will ever look into the structure generated by 0. So the computation totally exits that structure. Therefore, since the valuation of p at the root (= the truth value of $B(q)$ at ϵ) depends on the value of p at 0, we need to remember this value. Thus we need a memory of bit–size at least 1. Next consider the point 10. When the computation exits 10 it needs to store the result that it computed for p at 10 to correctly compute the value of p at 1. Hence we need an additional bit of memory, raising the size to 2. Likewise, for the points 110, 1110 and so on. Since the size of the tree is not bounded a priori, we cannot say how large the memory is that we need. To put it differently, if we have a memory of μ states, then the computation will give wrong results for trees whose depth exceeds $\log_2 \mu$. □

The argument actually has an information theoretic flavour. It does not really matter whether the memory encodes exactly the fact that the computation yielded the value 'yes' or 'no' at the node 0. What is important is how many computation histories the memory can discriminate. If we have a memory of size μ then at most μ histories can be discriminated. We have shown, however, that in the case at hand no upper bound can be given on the number of computation histories that must be distinguished. This proves the theorem. We can note a number of consequences. First of all, $A(p, q)$ is of the form $p \leftrightarrow D(p, q)$ where p occurs inside a modal operator. Thus, we have no analogue of the fixed–point theorems of Chapter 3.7 for **PDL.f**. The strategy of this example is interesting in many respects. It has been shown by Harvey Friedman in [**71**] that an unbounded memory really increases the power of algorithmic procedures. One example is the task of generating uniform two–branching trees, which cannot be achieved with bounded memory. This has been used by Jerzy Tiuryn [**213**] to show that the expressive power of programming languages is increased if an unbounded program memory is allowed. The example that we have produced here uses only propositional dynamic logic, not first–order dynamic logic, but the same result is obtained, using a variable q. The property *is true at a binary branching subtree whose leaves are q* cannot be checked using finite memory. However, as the proof indicated, there is way to check this property using a *memory stack* of numbers from 0 to 2.

Exercise 374. Let $\mathbf{PDL.f}^{\smile}$ be the extension of **PDL.f** by the axiom of finite computations in one direction. Show that $\mathbf{PDL.f}^{\smile}$ fails to have the finite model property.

Exercise 375. Let $\mathbf{PDL}^{\smile}.\mathbf{f}^{\pm}$ be the extension of **PDL.f** with converse operator, where both the programs γ^{+} and $\gamma^{\smile +}$ satisfy the **G**–axiom. Show that this logic fails to have the finite model property.

Exercise 376. (Continuing the previous exercise.) Add to $\mathbf{PDL}^{\smile}.\mathbf{f}^{\pm}$ the axiom $\langle\gamma^{\smile}\rangle p \rightarrow [\gamma^{\smile}]p$. Show that this logic has the finite model property. What finite frames does this logic describe?

Exercise 377. The example formula $A(p, q)$ given above is not the most economical one. Name an implicit definition of p that uses only two basic modalities. This reduces the failure to the case of binary branching trees. This seems to be the best possible result. Unary branching trees are just strings, so here the argument must break down.

Exercise 378. Let $\mathbf{PDL}^{\smile}.\mathbf{f}$ be the logic of finite computations, where programs are allowed to be executed forwards and backwards. Show that this logic is decidable. *Hint.* Do not take this too seriously.

Index

List of Symbols

List of Logics

List of Names

List of Terms

Bibliography

[1] Martin Amerbauer. Cut–free tableau calculi for some propositional modal logics. *Studia Logica*, 57:359 – 371, 1996.

[2] K. A. Baker. Finite equational bases for finite algebras in a congruence–distributive equational class. *Advances in Mathematics*, 24:207 – 243, 1977.

[3] Philippe Balbiani and Andreas Herzig. A translation from the modal logic of provability into **K4**. *Journal of Applied Non–Classical Logics*, 4(1):73 – 78, 1994.

[4] Fabio Bellissima. An effective representation for finitely generated free interior algebras. *Algebra Universalis*, 20:302 – 317, 1985.

[5] Fabio Bellissima. On the lattice of extensions of the modal logics $K.Alt_n$. *Archive for Mathematical Logic*, 27:107 – 114, 1988.

[6] Fabio Bellissima. Post complete and 0–axiomatizable modal logics. *Annals of Pure and Applied Logic*, 47, 1990.

[7] M. Ben-Ari, J. I. Halpern, and A. Pnueli. Deterministic propositional dynamic logic; finite models, complexity and completeness. *Journal of Computer and System Sciences*, 25:402 – 417, 1982.

[8] Johan van Benthem. Modal reduction principles. *Journal of Symbolic Logic*, 41:301 – 312, 1976.

[9] Johan van Benthem. Syntactic aspects of modal incompleteness theorems. *Theoria*, 45:63 – 77, 1979.

[10] Johan van Benthem. *Modal and Classical Logic*. Bibliopolis, 1983.

[11] Johan van Benthem. Correspondence theory. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 167 – 247. Reidel, 1984.

[12] Johan van Benthem. Notes on Modal Definability. *Notre Dame Journal of Formal Logic*, 39:20 – 39, 1989.

[13] Johan van Benthem. Beyond Accessibility: Functional Models for Modal Logics. In Maarten de Rijke, editor, *Diamonds and Defaults*, pages 1 – 18. Kluwer, Dordrecht, 1993.

[14] Johan van Benthem and Wim Blok. Transitivity follows from Dummett's axiom. *Theoria*, 44:117 – 118, 1978.

[15] Johan van Benthem and I. L. Humberstone. Halldén–completeness by gluing Kripke–frames. *Notre Dame Journal of Formal Logic*, 24:426 – 430, 1983.

[16] E. W. Beth. On padoa's method in the theory of definition. *Indagationes Mathematicae*, 15:330 – 339, 1953.

[17] Garreth Birkhoff. Subdirect products in universal algebra. *Bulletin of the American Mathematical Society*, 50:767 – 768, 1944.

[18] Garreth Birkhoff. *Lattice Theory*. American Mathematical Society Colloquium Publications, Rhode Island, 1973.

[19] Patrick Blackburn, Wilfried Meyer-Viol, and Maarten de Rijke. A Proof System for Finite Trees. In H. Kleine Büning, editor, *Computer Science Logic '95*, number 1092 in Lecture Notes in Computer Science, pages 86 – 105. Springer, 1996.

[20] Wim Blok. Pretabular varieties of modal algebras. *Studia Logica*, 39:101 – 124, 1880.

[21] Wim Blok. *Varieties of interior algebras*. PhD thesis, University of Amsterdam, 1976.

[22] Wim Blok. The free closure algebra on finitely many generators. *Indagationes Mathematicae*, 39:362 – 379, 1977.

[23] Wim Blok. The lattice of modal algebras is not strongly atomic. *Algebra Universalis*, 11:285 – 294, 1980.

[24] Wim Blok. The lattice of modal logics: An algebraic investigation. *Journal of Symbolic Logic*, 45:221 – 236, 1980.

[25] Wim Blok. On the degree of incompleteness in modal logics and the covering relation in the lattice of modal logics. Technical Report 78 – 07, Department of Mathematics, University of Amsterdam, 1980.

[26] Wim Blok, Peter Köhler, and Don Pigozzi. On the structure of varieties with equationally definable principal congruences, II. *Algebra Universalis*, 18:334 – 379, 1984.

[27] Wim Blok, Peter Köhler, and Don Pigozzi. On the structure of varieties with equationally definable principal congruences, IV. *Algebra Universalis*, 31:1 – 35, 1994.

[28] Wim Blok and Don Pigozzi. On the structure of varieties with equationally definable principal congruences, I. *Algebra Universalis*, 15:195 – 227, 1982.

[29] Wim Blok and Don Pigozzi. *Algebraizable Logics*. Number 396 in Memoirs of the American Mathematical Society. 1989.

[30] Wim Blok and Don Pigozzi. On the structure of varieties with equationally definable principal congruences III. *Algebra Universalis*, 32:545 –608, 1994.

[31] George Boolos. *The Logic of Provability*. Cambridge University Press, Cambridge, 1994.

[32] George Boolos and Giovanni Sambin. Provability: The Emergence of a Mathematical Modality. *Studia Logica*, 50:1 – 23, 1991.

[33] Manfred Borzechowski. Tableau–Kalkül für PDL und Interpolation. Master's thesis, Department of Mathematics, FU Berlin, 1988.

[34] Robert A. Bull. That all normal extensions of S4.3 have the finite model property. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 12:341 – 344, 1966.

[35] Robert A. Bull and Krister Segerberg. Basic modal logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 1 – 88. Reidel, 1984.

[36] John P. Burgess. Basic tense logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 2, pages 89 – 133. Reidel, 1984.

[37] Stanley Burris and Sankappanavar H. P. *A Course in Universal Algebra*. Number 78 in Graduate Texts in Mathematics. Springer, 1981.

[38] Rudolf Carnap. *Meaning and Necessity*. The University of Chicago Press, Chicago, 1947.

[39] Alexander Chagrov. The undecidability of the tabularity problem for modal logic. In *Logic Colloquium '94*, page 34, 1994.

[40] Alexander Chagrov and Michael Zakharyaschev. The Disjunction Property of Intermediate Propositional Logics. *Studia Logica*, 50:189 – 216, 1991.

[41] Alexander Chagrov and Michael Zakharyaschev. The undecidability of the disjunction property and other related problems. *Journal of Symbolic Logic*, 58:967 – 1002, 1993.

[42] Alexander Chagrov and Michael Zakharyaschev. On the Independent Axiomatizability of Modal and Intermediate Logics. *Journal of Logic and Computation*, 5:287 – 302, 1995.

[43] Alexander Chagrov and Michael Zakharyaschev. *Modal Logic*. Oxford University Press, Oxford, 1997.

[44] Lilia A. Chagrova. Undecidable problems in correspondency theory. *Journal of Symbolic Logic*, 56:1261 – 1272, 1991.

[45] C. C. Chang and H. Jerome Keisler. *Model Theory*. North–Holland, Amsterdam, 3 edition, 1990.

[46] Brian F. Chellas and Krister Segerberg. Modal Logics with the MacIntosh Rule. *Journal of Philosophical Logic*, 23:67 – 86, 1994.

[47] S. A. Cook. The complexity of theorem proving procedures. In *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing*, pages 151 – 158, New York, 1971. Association of the Computing Machinery.

[48] Max Creswell. An incomplete decidable logic. *Journal of Symbolic Logic*, 49:520 – 527, 1984.

[49] Janusz Czelakowski. Equivalential logics (I). *Studia Logica*, 40:227 – 236, 1981.

[50] Janusz Czelakowski. Equivalential logics (II). *Studia Logica*, 40:335 – 372, 1981.

[51] Marcello d'Agostino. Are Tableaux an Improvement on Truth–Tables? Cut–free Proofs and Bivalence. *Journal of Logic, Language and Information*, 1:235 – 252, 1993.

[52] B. A. Davey and H. A. Priestley. *Lattices and Order*. Cambridge University Press, Cambridge, 1990.

[53] Dick de Jongh and Albert Visser. Explicit Fixed Points in Interpretability Logic. *Studia Logica*, 50:39 – 49, 1991.

[54] E. A. Emerson. Temporal and Modal Logic. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science, Vol. B, Formal Models and Semantics*, pages 996 – 1072. Elsevier, Amsterdam, 1990.

[55] N. M. Ermolaeva and Albert A. Muchnik. Modal logics defined by endomorphisms of distributive lattices (Russian). In A. D. Bochvar and Grishin, editors, *Issledovaniya po Teorii Mnozhestv i Neklassicheskim Logikam. Sbornik Trudov. (Investigations on Set–Theory and Non–Classical Logics)*, pages 229 – 246. 1976.

[56] N. M. Ermolaeva and Albert A. Muchnik. Pretabular temporal logic (Russian). In A. I. Mikhailov, editor, *Issledovaniya po Neklassicheskim Logikam i Teorii Mnozhestv (Investigations on Non–Classical Logics and Set–Theory)*, pages 288 – 297. 1979.

[57] L. Esakia and V. Meskhi. Five critical systems. *Theoria*, 40:52 – 60, 1977.

[58] Leo Esakia. Topological Kripke Models (Russian). *Soviet Mathematical Doklady*, 15:147 – 151, 1974.

[59] Solomon Feferman. Persistent and invariant formulas for outer extensions. *Compositio Mathematica*, 20:29 – 52, 1969.

[60] Kit Fine. The logics containing S4.3. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 17:371 – 376, 1971.

[61] Kit Fine. An ascending chain of S4 logics. *Theoria*, 40:110 – 116, 1974.

[62] Kit Fine. An incomplete logic containing S4. *Theoria*, 40:23 – 29, 1974.

[63] Kit Fine. Logics containing K4, Part I. *Journal of Symbolic Logic*, 39:229 – 237, 1974.

[64] Kit Fine. Normal forms in modal logic. *Notre Dame Journal of Formal Logic*, 16:31 – 42, 1975.

[65] Kit Fine. Some connections between elementary and modal logic. In Stig Kanger, editor, *Proceedings of the Third Scandinavian Logic Symposium*, pages 15 – 31. North–Holland, 1975.

[66] Kit Fine. Logics containing K4, Part II. *Journal of Symbolic Logic*, 50:619 – 651, 1985.

[67] Kit Fine and Gerhard Schurz. Transfer theorems for stratified modal logics. In *Proceedings of the Arthur Prior Memorial Conference*, Christchurch, New Zealand, 1991.

[68] Marcello Finger and Dov M. Gabbay. Adding a Temporal Dimension to a Logic System. *Journal of Logic, Language and Information*, 1:203 – 233, 1993.

[69] M. J. Fischer and R. E. Ladner. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18:194 – 211, 1979.

[70] Melvin Fitting. *Proof Methods for Modal and Intuitionistic Logic*. Number 169 in Synthese Library. Reidel, Dordrecht, 1983.

[71] Harvey Friedman. Algorithmic procedures, generalized Turing algorithms, and elementary recursion theory. In *Logic Colloquium 1969*, pages 361 – 390, Amsterdam, 1971. North–Holland.

[72] Dov M. Gabbay. Expressive Functional Completeness in Tense Logics. In Uwe Mönnich, editor, *Aspects of Philosophical Logic*, pages 91 – 117. Reidel, Dordrecht, 1981.

[73] Michael R. Garey and David S. Johnson. *Computers and Intractability. A Guide to the Theory of NP–Completeness*. W. H. Freeman and Company, New York, 1979.

[74] Silvio Ghilardi and M. Zawadowski. Undefinability of propositional quantifiers in modal system s4. *Studia Logica*, 55:259 – 271, 1995.

[75] Giuseppe di Giacomo. Eliminating "Converse" from Converse PDL. *Journal of Logic, Language and Information*, 5:193 – 208, 1996.

[76] Zachary Gleit and Warren Goldfarb. Characters and fixed points in provability logic. *Notre Dame Journal of Formal Logic*, 31:26 – 26, 1990.

[77] Robert I. Goldblatt. First–order definability in modal logic. *Journal of Symbolic Logic*, 40:35 – 40, 1975.

[78] Robert I. Goldblatt. Metamathematics of modal logics, Part I. *Reports on Mathematical Logic*, 6:41 – 78, 1976.

[79] Robert I. Goldblatt. Metamathematics of modal logics, Part II. *Reports on Mathematical Logic*, 7:21 – 52, 1976.

[80] Robert I. Goldblatt. *Topoi — The Categorial Analysis of Logic*. Number 98 in Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1984.

[81] Robert I. Goldblatt. *Logics of Time and Computation*. Number 7 in CSLI Lecture Notes. CSLI, 1987.

[82] Robert I. Goldblatt. Varieties of complex algebras. *Annals of Pure and Applied Logic*, 38:173 – 241, 1989.

[83] Robert I. Goldblatt and S. K. Thomason. Axiomatic classes in propositional modal logic. In J. Crossley, editor, *Alebra and Logic*, number 450 in Springer Lecture Notes in Mathematics. Springer, 1974.

[84] Martin Goldstern and Haim Judah. *The Incompleteness Phenomenon*. AK Peters, Wellesley, Massachussetts, 1995.

[85] Valentin Goranko. Refutation Systems in Modal Logic. *Studia Logica*, 53:299 – 324, 1994.

[86] Valentin Goranko and George Gargov. Modal logic with names. *Journal of Philosophical Logic*, 22:607 – 636, 1993.

[87] Valentin Goranko and Solomon Passy. Using the universal modality: Gains and Questions. *Journal of Logic and Computation*, 2:5 – 30, 1992.

[88] Rajeev Goré. Cut–free sequent and tableau systems for propositional Diodorean modal logics. *Studia Logica*, 53:433 – 457, 1994.

[89] George Grätzer. *Universal Algebra*. van Nostrand, New York, 1968.

[90] George Grätzer. *Lattice Theory: First Concepts and Distributive Lattices*. Freeman, 1971.

[91] Carsten Grefe. Modale Logiken funktionaler Frames. Master's thesis, Department of Mathematics, FU Berlin, 1994.

[92] Jeroen Groenendijk and Martin Stokhof. Dynamic Predicate Logic. *Linguistics and Philosophy*, 14:39 – 100, 1991.

[93] Sören Halldén. On the semantic non–completeness of certain Lewis calculi. *Journal of Symbolic Logic*, 16:127 – 129, 1951.

[94] Paul Halmos. Algebraic logic I. monadic boolean algebras. *Compositio Mathematica*, 12:217 – 249, 1955.

[95] Joseph Halpern and Y. Moses. A guide to the modal logics of knowledge and belief. *Artificial Intelligence*, 54:319 – 379, 1992.

[96] Bengt Hansson and Peter Gärdenfors. Filtrations and the finite frame property in boolean semantics. In Stig Kanger, editor, *Proceedings of the Third Scandinavian Logic Symposium*, pages 32 – 39. North–Holland, 1975.

[97] D. Harel. Dynamic logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, volume 2. Reidel, 1984.

[98] David Harel, Amir Pnueli, and Jonathan Stavi. Propositional dynamic logic of nonregular programs. *Journal of Computer and Systems Sciences*, 26:222 – 243, 1983.

[99] Michael A. Harrison. *Introduction to Formal Language Theory*. Addison Wesley, Reading (Mass.), 1978.

[100] R. Harrop. On the existence of finite models and decision procedures for propositional calculi. *Proceedings of the Cambridge Philosophical Society*, 54:1 – 13, 1958.

[101] Horst Herrlich and George E. Strecker. *Category Theory*. Allyn and Bacon, 1973.

[102] Burghard Herrmann and Wolfgang Rautenberg. Finite replacement and finite Hilbert–style axiomatizability. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 38:327 – 344, 1992.

[103] Burghard Herrmann and Frank Wolter. Representations of algebraic lattices. *Algebra Universalis*, 31:612 – 613, 1994.

[104] Jaakko Hintikka. *Knowledge and Belief. An Introduction into the logic of the two notions*. Cornell University Press, Ithaca, 1962.

[105] John E. Hopcroft and Jeffrey D. Ullman. *Formal Languages and their Relation to Automata*. Addison Wesley, Reading (Mass.), 1969.

[106] George E. Hughes and Max J. Cresswell. *A Companion to Modal Logic*. Methuen, London, 1984.

[107] Stephen Isard. A finitely axiomatizable undecidable extension of K. *Theoria*, 43:195 – 202, 1977.

[108] V. A. Jankov. The construction of a sequence of strongly independent superintuitionistic propositional calculi. *Soviet Mathematics*, 9:806 – 807, 1968.

[109] V. A. Jankov. Conjunctively indecomposable formulas in propositional calculi. *Izvestija Akad. Nauk. SSSR*, 33:18 – 33, 1969.

[110] Peter T. Johnstone. *Stone Spaces*. Number 3 in Cambridge studies in advanced mathematics. Cambridge University Press, 1982.

[111] Bjarni Jónsson. Algebras whose lattice of congruences is distributive. *Mathematica Scandinavica*, 21:110 – 121, 1967.

[112] Bjarni Jónsson. On the canonicity of Sahlqvist identities. *Studia Logica*, 53:473 – 491, 1994.

[113] Bjarni Jónsson and Alfred Tarski. Boolean algebras with operators. *American Journal of Mathematics*, 73:891 – 939, 1951.

[114] Winfried Just and Martin Weese. *Discovering Modern Set Theory. I. The Basics*, volume 8 of *Graduate Studies in Mathematics*. AMS, 1996.

[115] Winfried Just and Martin Weese. *Discovering Modern Set Theory. II. Set–Theoretic Tools for Every Mathematician*, volume 18 of *Graduate Studies in Mathematics*. AMS, 1997.

[116] Steven C. Kleene. *Introduction to Metamathematics*. North–Holland, Amsterdam, 1964.

[117] Sabine Koppelberg. General theory of boolean algebras. In J. Monk and S. Koppelberg, editors, *Handbook of Boolean Algebras. I*. Elsevier, 1988.

[118] Dexter Kozen and Rohit Parikh. An elementary proof of the completeness of PDL. *Theoretical Computer Science*, 14:113 – 118, 1981.

[119] Dexter Kozen and Jerzy Tiuryn. Logics of Programs. In Jan van Leeuwen, editor, *Handbook of Theoretical Computer Science, Vol. B, Formal Models and Semantics*, pages 789 – 840. Elsevier, Amsterdam, 1990.

[120] Marcus Kracht. An almost general splitting theorem for modal logic. *Studia Logica*, 49:455 – 470, 1990.

[121] Marcus Kracht. *Internal Definability and Completeness in Modal Logic*. PhD thesis, FU Berlin, 1991.

[122] Marcus Kracht. A solution to a problem by Urquhart. *Journal of Philosophical Logic*, 20:285 – 286, 1991.

[123] Marcus Kracht. Even more about the lattice of tense logics. *Archive of Mathematical Logic*, 31:243 – 357, 1992.

[124] Marcus Kracht. How completeness and correspondence theory got married. In Maarten de Rijke, editor, *Diamonds and Defaults*, Synthese, pages 175 – 214. Kluwer, 1993.

[125] Marcus Kracht. Prefinitely Axiomatizable Modal and Intermediate Logics. *Mathematical Logic Quarterly*, 39:301 – 322, 1993.

[126] Marcus Kracht. Splittings and the finite model property. *Journal of Symbolic Logic*, 58:139 – 157, 1993.

[127] Marcus Kracht. Highway to the danger zone. *Journal of Logic and Computation*, 5:93 – 109, 1995.

[128] Marcus Kracht. Power and Weakness of the Modal Display Calculus. In Heinrich Wansing, editor, *Proof Theory of Modal Logic*, pages 95 – 120. Oxford University Press, 1996.

[129] Marcus Kracht. Modal Logics that Need Very Large Frames. manuscript, 1998.
[130] Marcus Kracht. Reducing Modal Consequence Relations. manuscript, 1998.
[131] Marcus Kracht. Lattices of Modal Logics and Their Groups of Automorphisms. *Journal of Pure and Applied Logic*, 1999. to appear.
[132] Marcus Kracht and Frank Wolter. Properties of independently axiomatizable bimodal logics. *Journal of Symbolic Logic*, 56:1469 – 1485, 1991.
[133] Marcus Kracht and Frank Wolter. Normal monomodal logics can simulate all others. *Journal of Symbolic Logic*, 64, 1999.
[134] Saul Kripke. Semantic analysis of modal logic, Part I. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9:67 – 96, 1963.
[135] J. B. Kruskal. Well–quasi–orderings, the Tree Theorem and Vazsonyi's Conjecture. *Transactions of the American Mathematical Society*, 95:210 – 225, 1960.
[136] A. H. Lachlan. A note on Thomason's refined structures for tense logics. *Theoria*, 40:117 – 120, 1974.
[137] R. E. Ladner. The computational complexity of provability in systems of modal logic. *SIAM Journal of Computing*, 6:467 – 480, 1977.
[138] Daniel Leivant. Proof theoretic methodology for propositional dynamic logic. In *Formalization of Programming Concepts, Proceedings of the International Colloquium, Peniscola*, number 107 in Lecture Notes in Computer Science, pages 356 – 373. Springer, 1981.
[139] E. J. Lemon. A note on Halldén–incompleteness. *Notre Dame Journal of Formal Logic*, 7:296 – 300, 1966.
[140] J. Łos and Roman Suszko. Remarks on sentential logics. *Indagationes Mathematicae*, 20:177 – 183, 1958.
[141] J. Łukasiewicz and Alfred Tarski. Untersuchungen über den aussagenkalkül. *Comptes Rendues des Séances de la Societé des Sciences et des Lettres de Varsovie*, 23:30 – 50, 1930.
[142] Robert C. Lyndon. Identities in 2–valued calculi. *Transactions of the Americal Mathematical Society*, 71:457 – 465, 1951.
[143] Saunders MacLane. *Categories for the Working Mathematician*. Number 5 in Graduate Texts in Mathematics. Springer, Heidelberg, 1971.
[144] S. Maehara. On the interpolation theorem of Craig (japanese). *Sugaku*, 12:235 – 237, 1960.
[145] David Makinson. A normal modal calculus between T and S4 without the finite model property. *Journal of Symbolic Logic*, 34:35 – 38, 1969.
[146] David C. Makinson. Some embedding theorems for modal logic. *Notre Dame Journal of Formal Logic*, 12:252 – 254, 1971.
[147] Larisa L. Maksimova. Pretabular extensions of Lewis S4. *Algebra i Logica*, 14:28 – 55, 1975.
[148] Larisa L. Maksimova. Interpolation theorems in modal logic and and amalgamated varieties of topoboolean algebras. *Algebra and Logic*, 1979.
[149] Larisa L. Maksimova. Definability theorems in normal extensions of the provability logic. *Studia Logica*, 48:495 – 507, 1989.
[150] Larisa L. Maksimova. The Beth Properties, interpolation, and amalgamability in varieties of modal algebras. *Soviet Mathematical Doklady*, 44:327 – 331, 1992.
[151] Larisa L. Maksimova. Definability and interpolation in classical modal logics. *Contemporary Mathematics*, 131:583 – 599, 1992.
[152] Larisa L. Maksimova. Temporal logics with "the next" operator do not have Interpolation or the Beth Property. *Siberian Mathematical Journal*, 32:989 – 993, 1992.
[153] Larisa L. Maksimova. On variable separation in modal and superintuitionistic logics. *Studia Logica*, 55:99 – 112, 1995.
[154] Larisa L. Maksimova and Vladimir V. Rybakov. Lattice of normal modal logics. *Algebra i Logica*, 13:188 – 216, 1974.
[155] A. I. Malcev. On the general theory of algebraic systems. *Math. Sbornik*, 35:3 – 20, 1954.

[156] A. A. Markov. Impossibility of Certain Algorithms in the Theory of Associative Systems (Russian). *Doklady Akademia Nauk SSSR*, 77:953 – 956, 1951.

[157] Maarten Marx, László Pólos, and Michael Masuch. *Arrow Logic and Multi–Modal Logic*. Studies in Logic, Language and Information. CSLI, Stanford, 1996.

[158] M. Mortimer. On languages with two variables. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 21:135 – 140, 1975.

[159] Solomon Passy and Tinko Tinchev. An essay in combinatory dynamic logic. *Information and Computation*, 93:263 – 332, 1991.

[160] Brigitte Penther. A Dynamic Logic of Action. *Journal of Logic, Language and Information*, 3:169 – 210, 1994.

[161] A. Pitts. On an interpretation of second order quantification in first order intuitionistic logic. *Journal of Symbolic Logic*, 57:33 – 52, 1992.

[162] W. A. Pogorzelski. Structural completeness of the propositional calculus. *Bulletin de l'Académie Polonaise des Sciences, Séries des sciences mathématiques, astronomiques et physiques*, 19:439 – 451, 1971.

[163] Emil L. Post. Recursive Undecidability of a Problem of Thue. *Journal of Symbolic Logic*, 12:1 – 11, 1947.

[164] Vaughan Pratt. Models of program logics. In *Proceedings of the 20th IEEE Symposium on Foundations of Computer Science*, pages 115 – 122, 1979.

[165] Vaughan Pratt. Dynamic Algebras: Examples, Constructions, Applications. *Studia Logica*, 50:571 – 605, 1991.

[166] T. Prucnal and A. Wroński. An algebraic characterization of the notion of structural completeness. *Bulletin of Section Logic of the Polish Academy of Sciences*, 3:20 – 33, 1974.

[167] Michael O. Rabin. Recursive Unsolvability of Group Theoretic Problems. *Annals of Mathematics*, 67:172 – 194, 1958.

[168] Wolfgang Rautenberg. Der Verband der normalen verzweigten Modallogiken. *Mathematische Zeitschrift*, 156:123 – 140, 1977.

[169] Wolfgang Rautenberg. *Klassische und nichtklassische Aussagenlogik*. Vieweg, Braunschweig/Wiesbaden, 1979.

[170] Wolfgang Rautenberg. Splitting lattices of logics. *Archiv für Mathematische Logik*, 20:155 – 159, 1980.

[171] Wolfgang Rautenberg. Modal tableau calculi and interpolation. *Journal of Philosophical Logic*, 12:403 – 423, 1983.

[172] Wolfgang Rautenberg. Applications of Weak Kripke Semantics to Intermediate Consequence Relations. *Studia Logica*, 45:119 – 134, 1986.

[173] Wolfgang Rautenberg. Axiomatizing Logics Closely Related to Varieties. *Studia Logica*, 50:607 – 622, 1991.

[174] Wolfgang Rautenberg. Strongly finitely based equational theories. *Algebra Universalis*, 28:549 – 558, 1991.

[175] Lisa Reidhaar-Olson. A new proof of the fixed–point theorem of provability logic. *Notre Dame Journal of Formal Logic*, 31:37 –43, 1990.

[176] György E. Révész. *Introduction to Formal Languages*. McGraw–Hill, New York, 1985.

[177] Maarten de Rijke. *Extending Modal Logic*. PhD thesis, Universiteit van Amsterdam, 1993.

[178] Vladimir V. Rybakov. Admissible rules of pretabular modal logics. *Algebra and Logic*, 20:291 – 307, 1981.

[179] Vladimir V. Rybakov. Problems of substitution and admissibility in the modal systems grz and intuitionistic calculus. *Annals of Pure and Applied Logic*, 50:71 – 106, 1990.

[180] Vladimir V. Rybakov. Rules of inference with parameters for intuitionistic logic. *Journal of Symbolic Logic*, 57:912 – 923, 1992.

[181] Vladimir V. Rybakov. Criteria for Admissibility of Inference Rules. Modal and Intermediate Logics with the Branching Property. *Studia Logica*, 53:203 – 226, 1994.

[182] Vladimir V. Rybakov. *Admissibility of Inference Rules*. Number 136 in Studies in Logic. Elsevier, Amsterdam, 1997.

[183] Hendrik Sahlqvist. First and second order semantics for modal logic. In Stig Kanger, editor, *Proceedings of the Third Scandinavian Logic Symposium*, pages 15 – 31. North-Holland, 1975.

[184] Giovanni Sambin. An effective fixed–point theorem in intuitionistic diagonalizable algebras. *Studia Logica*, 35:345 – 361, 1976.

[185] Giovanni Sambin. Subdirectly irreducible modal algebras and initial frames. *Studia Logica*, 62, 1999.

[186] Giovanni Sambin and Virginia Vaccaro. Topology and duality in modal logic. *Annals of Pure and Applied Logic*, 37:249 – 296, 1988.

[187] Giovanni Sambin and Virginia Vaccaro. A new proof of Sahlqvist's theorem on modal definability and completeness. *Journal of Symbolic Logic*, 54:992 – 999, 1989.

[188] Giovanni Sambin and S. Valentini. The modal logic of provability. the sequential approach. *Journal of Philosophical Logic*, 11:311 – 342, 1982.

[189] Walter J. Savitch. Relationship between nondeterministic and deterministic tape complexities. *Journal of Computer and Systems Sciences*, 4:177 – 192, 1970.

[190] George Schumm. Bounded properties in modal logic. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 27:197 – 200, 1981.

[191] George F. Schumm. Why does Halldén–completeness matter? *Theoria*, 59:192 – 206, 1993.

[192] Dana Scott. Advice on Modal Logic. In Lambert, editor, *Philosophical Problems in Logic. Some Recent Developments*, pages 143 – 174. Reidel, Dordrecht, 1970.

[193] Krister Segerberg. An essay in classical modal logic. Mimeograph, Uppsala, 1971.

[194] Krister Segerberg. That all extensions of S4.3 are normal. In Stig Kanger, editor, *Proceedings of the Third Scandinavian Logic Symposium*, pages 194 – 196. North-Holland, 1975.

[195] Krister Segerberg. A completeness theorem in the modal logic of programs. *Notices of the AMS*, 24(6):A – 522, 1977.

[196] Krister Segerberg. *Classical Propositional Operators*. Number 5 in Oxford Logic Guides. Oxford University Press, 1982.

[197] Krister Segerberg. Modal logics with functional alternative relations. *Notre Dame Journal of Formal Logic*, 27:504 – 522, 1986.

[198] Valentin Shehtman. Undecidable propositional calculi (russian). In *Problemy kibernetiki. Neklassicheskije logiki i ikh primenenije*. Akademia Nauk SSSR, 1982.

[199] Joseph A. Shoenfiled. *Mathematical Logic*. Addison Wesley, Reading, Massachussetts, 1967.

[200] Craig Smoryński. Beth's Theorem and Self–referential sentences. In *Logic Colloquium 77*, pages 253 – 261, Amsterdam, 1978. North–Holland.

[201] Robert Solovay. Provability interpretations of modal logic. *Israel Journal of Mathematics*, 25:287 – 304, 1976.

[202] Edith Spaan. *Complexity of Modal Logics*. PhD thesis, Department of Mathematics and Computer Science, University of Amsterdam, 1993.

[203] L. J. Stockmeyer and A. R. Meyer. Word problems requiring exponential time. In *Proceedings of the 5th Annual ACM Symposium on Theory of Computing*, pages 1 – 9, New York, 1973. Association of Computing Machinery.

[204] Timothy Surendonk. *Canonicity for Intensional Logics*. PhD thesis, Automated Reasoning Project, Australian National University, 1998.

[205] Oswald Teichmüller. Braucht der algebraiker das auswahlaxiom? *Deutsche Mathematik*, 4:567 – 577, 1939.

[206] S. K. Thomason. Semantic analysis of tense logic. *Journal of Symbolic Logic*, 37:150 – 158, 1972.

[207] S. K. Thomason. An incompleteness theorem in modal logic. *Theoria*, 40:30 – 34, 1974.

[208] S. K. Thomason. Reduction of tense logic to modal logic I. *Journal of Symbolic Logic*, 39:549 – 551, 1974.

[209] S. K. Thomason. Reduction of second-order logic to modal logic. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 21:107 – 114, 1975.
[210] S. K. Thomason. Reduction of tense logic to modal logic II. *Theoria*, 41:154 – 169, 1975.
[211] S. K. Thomason. Independent propositional modal logics. *Studia Logica*, 39:143 – 144, 1980.
[212] S. K. Thomason. Undecidability of the completeness problem of modal logic. In *Universal Algebra and Applications, Banach Center Publications Vol. 9*, Warsaw, 1982. PNW–Polish Scientific Publishers.
[213] Jerzy Tiuryn. Unbounded Program Memory Adds to the Expressive Power of First–Order Programming Logic. *Information and Control*, 60:12 –35, 1984.
[214] M. Tokarz. Connections between some notions of completeness of structural propositional calculi. *Studia Logica*, 32:77 – 89, 1973.
[215] J. W. Tukey. *Convergence and Uniformity in Topology*, volume 2 of *Annals of Mathematical Studies*. Princeton University Press, Princeton NJ, 1940.
[216] Alasdair Urquhart. Decidability and the finite model property. *Journal of Philosophical Logic*, 10:367 – 370, 1981.
[217] Dimiter Vakarelov. Filtration theorem for dynamic algebras with tests and inverse operator. In A. Salwicki, editor, *Logics of programs and their applications (Poznan 1980)*, number 148 in Springer Lecture Notes in Computer Science, pages 314 – 324, 1983.
[218] Jan van Eijck and Fer-Jan de Vries. Reasoning about update logic. *Journal of Philosophical Logic*, 24:19 – 45, 1995.
[219] Moshe Y. Vardi and P. Wolper. Automata Theoretic Techniques for Modal Logics of Programs. Technical report, IBM Research Report, 1984.
[220] Robert L. Vaught. *Set Theory. An Introduction*. Birkhäuser, Basel, 2 edition, 1995.
[221] Yde Venema. *Many-Dimensional Modal Logics*. PhD thesis, Universiteit van Amsterdam, 1991.
[222] Yde Venema. Canonical Pseudo–Correspondence. In Maarten de Rijke, Heinrich Wansing, and Michael Zakharyaschev, editors, *Advances in Modal Logic 98*. CSLI, 1998. to appear.
[223] Albert Visser. Uniform interpolation and layered bisimulation. In Peter Hajek, editor, *Gödel 96*, pages 139 – 164. Springer, 1996.
[224] Heinrich Wansing, editor. *Proof Theory of Modal Logic*, number 2 in Applied Logic Series. Kluwer, 1996.
[225] Titus Weider. *Simulation kristalliner Strukturen unter dem Einfluß einer Scherströmung*. Wissenschaft und Technik Verlag, Berlin, 1996.
[226] Timothy Williamson. An alternate rule of disjunction in modal logic. *Notre Dame Journal of Formal Logic*, 33, 1992.
[227] Timothy Williamson. Some Admissible Rules in Nonnormal Modal Logics. *Notre Dame Journal of Formal Logic*, 34:378 – 400, 1993.
[228] Timothy Williamson. Non–genuine Macintosh Logics. *Journal of Philosophical Logic*, 23:87 – 101, 1994.
[229] Ryszard Wójcicki. Locical matrices stronlgy adequate for structural sentential calculi. *Bulletin de l'Académie Polonaise des Sciences, Séries des sciences mathématiques, astronomiques et physiques*, 6:333 – 335, 1969.
[230] Ryszard Wójcicki. Some remarks on the consequence operation in sentential logics. *Funadmenta Mathematicae*, 68:269 – 279, 1970.
[231] Ryszard Wójcicki. Matric approach in methodology of sentential calculi. *Studia Logica*, 32:7 – 37, 1973.
[232] Ryszard Wójcicki. *Theory of Logical Calculi*. Kluwer, Dordrecht, 1988.
[233] Frank Wolter. Eigenschaften unabhängig axiomatisierbarer polymodaler Logiken. Master's thesis, II. Mathematisches Institut, FU Berlin, 1991.
[234] Frank Wolter. *Lattices of Modal Logics*. PhD thesis, FU Berlin, 1993.
[235] Frank Wolter. Solution to a problem of Goranko and Passy. *Journal of Logic and Computation*, 4:21 – 22, 1994.

[236] Frank Wolter. Tense Logic, Part II: Decidability. unpulished, 1994.

[237] Frank Wolter. What is the Upper Part of the Lattice of Bimodal Logics? *Studia Logica*, 53:235 – 242, 1994.

[238] Frank Wolter. The finite model property in tense logic. *Journal of Symbolic Logic*, 60:757 – 774, 1995.

[239] Frank Wolter. A Counterexample in Tense Logic. *Notre Dame Journal of Formal Logic*, 37:167 – 173, 1996.

[240] Frank Wolter. Properties of tense logics. *Mathematical Logic Quarterly*, 42:481 – 500, 1996.

[241] Frank Wolter. Tense logic without tense operators. *Mathematical Logic Quarterly*, 42:145 – 171, 1996.

[242] Frank Wolter. Completeness and Decidability of Tense Logics closely related to Logics above K4. *Journal of Symbolic Logic*, 62:131 – 159, 1997.

[243] Frank Wolter. Independent fusions of modal logics revisited. In Marcus Kracht, Maarten de Rijke, Heinrich Wansing, and Michael Zakharyaschev, editors, *Advances in Modal Logic*, pages 361 – 379, Stanford, 1997. CSLI.

[244] Frank Wolter. The structure of lattices of subframe logics. *Annals of Pure and Applied Logic*, 86:47 – 100, 1997.

[245] Michael Zakharyaschev. Canonical formulas for K4, Part I: Basic results. *Journal of Symbolic Logic*, 57:1377 – 1402, 1992.

[246] Michael Zakharyaschev. Canonical formulas for K4, Part II: Cofinal Subframe Logics. *Journal of Symbolic Logic*, 61:421 – 449, 1996.

[247] Michael Zakharyaschev and Alexander Alekseev. That All Finitely Axiomatizable Normal Extensions of **k4.3** are Decidable. *Mathematical Logic Quarterly*, 41:15 – 23, 1995.