<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.2</td>
<td>Permutability of Congruences</td>
<td>174</td>
</tr>
<tr>
<td>10.3</td>
<td>Congruence Distributive and Arithmetical Varieties</td>
<td>185</td>
</tr>
<tr>
<td>10.4</td>
<td>Congruence Modularity</td>
<td>186</td>
</tr>
<tr>
<td>10.5</td>
<td>Congruence Regularity and Uniformity</td>
<td>197</td>
</tr>
<tr>
<td>10.6</td>
<td>Congruence Identities</td>
<td>211</td>
</tr>
<tr>
<td>10.7</td>
<td>Relationships</td>
<td>226</td>
</tr>
<tr>
<td>10.8</td>
<td>Abelian Varieties</td>
<td>237</td>
</tr>
<tr>
<td>10.9</td>
<td>Notes and Perspectives</td>
<td>248</td>
</tr>
<tr>
<td>11.1</td>
<td>Abstract Clones and their Homomorphisms</td>
<td>251</td>
</tr>
<tr>
<td>11.2</td>
<td>Representations of Clones</td>
<td>275</td>
</tr>
<tr>
<td>11.3</td>
<td>Presentations of Clones</td>
<td>291</td>
</tr>
<tr>
<td>11.4</td>
<td>Interpretations, Clone Homomorphisms and Concrete Functors</td>
<td>302</td>
</tr>
<tr>
<td>11.5</td>
<td>Products and Co-Products of Clones</td>
<td>315</td>
</tr>
<tr>
<td>11.6</td>
<td>$</td>
<td>k</td>
</tr>
<tr>
<td>11.7</td>
<td>Mal’tsev Conditions</td>
<td>334</td>
</tr>
<tr>
<td>11.8</td>
<td>Hypervarieties</td>
<td>335</td>
</tr>
<tr>
<td>11.9</td>
<td>Definable classes of varieties</td>
<td>336</td>
</tr>
<tr>
<td>11.10</td>
<td>History and references for Chapter 11</td>
<td>336</td>
</tr>
</tbody>
</table>
7.1. Elementary Model Theory

\(\langle R, +, \cdot, \leq \rangle\) and \(\langle \omega, +, \cdot, \leq \rangle\) are among the most familiar mathematical structures. Almost at the outset in developing properties of these structures, one encounters sentences like

“For all numbers \(x, y, \text{ and } z\), if \(x + y \leq x + z\), then \(y \leq z\).”

This sentence might be rendered in symbols as follow:

\[\forall x \forall y \forall z [x + y \leq x + z \Rightarrow y \leq z].\]

We see here two departures from the prevailing context of our exposition. First, structures like \(\langle \omega, +, \cdot, \leq \rangle\), while they are very much like algebras, fail to be algebras. Not only does this structure have a universe and some fundamental operations, but it has a fundamental (binary) relation as well. Second, the sentence symbolized above, for all its simplicity, is not an equation. Model theory provides the means to address both of these departures. The central issue in elementary model theory is the connection between such mathematical structures on the one hand and the formulas and sentences of formal languages, like the one displayed above, on the other. While our focus here is openly algebraic, motivations based on model theoretic considerations are unmistakable, and the methods of model theory play an essential role. Indeed, Birkhoff’s HSP Theorem is one of the earliest results in model theory. However, in most contexts, equations offer only meager means of expression, and even the simple cancellation law cited above cannot be formulated by means of equations alone. The real power of model theory emerges once our means of expression are sufficiently rich.

Model theory is an extensively developed part of mathematics closely related to the general theory of algebras. As with lattice theory, many people have made contributions to the development of both fields. This chapter is an introduction to model theory which emphasizes those results most frequently applied in the general theory of algebras. The serious student of our field is well advised to acquire an understanding of the deeper aspects of model theory from one of the excellent books which are available.
DEFINITION 7.1. An elementary language is comprised of a set of finitary operation and relation symbols, a countably infinite sequence \( v_0, v_1, v_2, \ldots \) of distinct variables, and the following symbols which are called logical symbols: \( \approx, \& , \lor, \neg, \Rightarrow, \Leftrightarrow, \exists, \forall, \). These symbols are always taken to be distinct.

Two elementary languages can differ only in their operation and relation symbols. Consequently, we may refer to a particular language by specifying its set of operation and relation symbols. Thus, we take an elementary language \( \mathfrak{L} \) to be the set consisting of all the operation and relation symbols. Each of the operation symbols and each of the relation symbols has a rank which must be a natural number in the case of operation symbols and a positive natural number in the case of relation symbols. We will use operation and relation symbols to name the fundamental operations and relations of given mathematical structures. The variables are intended to range over the elements of a given mathematical structure. This restriction on the variables accounts for our use of the word “elementary”.

The logical symbols have the following meanings:

\[ \approx \quad \text{“equality”} \]

The connectives:

\[ \& \quad \text{“and”} \quad \lor \quad \text{“or”} \]

\[ \Rightarrow \quad \text{“implies”} \quad \Leftrightarrow \quad \text{“if and only if”} \]

\[ \neg \quad \text{“not”} \]

The quantifiers:

\[ \exists \quad \text{“there exists”} \quad \forall \quad \text{“for all”} \]

The parentheses \( ) \) and \( ( \) are used for punctuation. While this set of logical symbols provides a natural way to express many notions, at several places below it proves convenient to rely on a smaller set of quantifiers and connectives:

\[ \approx, \&, \neg, \text{ and, } \exists. \]

The remaining connectives and quantifiers can then be regarded as abbreviations. For example, \( \neg \exists \neg (x \leq x + 1) \) expresses the same notion as \( \forall x (x \leq x + 1) \).

DEFINITION 7.2. Let \( \mathfrak{L} \) be an elementary language. An \( \mathfrak{L} \)-structure is a system \( \mathfrak{A} = \langle A, F \rangle \) where \( A \) is a nonempty set, referred to as the universe of \( \mathfrak{A} \), and \( F \) is a function with domain \( \mathfrak{L} \) such that \( F(Q) \) is an operation on \( A \) with the same rank as \( Q \), for each operation symbol \( Q \) of \( \mathfrak{L} \), and \( F(R) \) is a relation on \( A \) of the same rank as \( R \), for each relation symbol \( R \) of \( \mathfrak{L} \); \( F(Q) \) is also denoted by \( Q^A \) and referred to as a fundamental or basic operation of \( \mathfrak{A} \), while \( F(R) \) is also denoted by \( R^A \) and referred to as a fundamental or basic relation of \( \mathfrak{A} \).
An $\mathcal{L}$-structure is just a nonempty set equipped with a fundamental operation of the right rank for each operation symbol of $\mathcal{L}$ and a fundamental relation of the right rank for each relation symbol of $\mathcal{L}$. If $\mathcal{L}$ has no relation symbols, then $\mathcal{L}$-structures are just algebras. If $\mathcal{L}$ has no operation symbols, then each $\mathcal{L}$-structure is said to be a relational structure—ordered sets and graphs (with vertices and edges) are examples of relational structures. A structure is an $\mathcal{L}$-structure for some $\mathcal{L}$. We frequently display structures as sets equipped with a sequence of fundamental operations and relations. Thus $\langle \mathbb{R}, +, -, \cdot, 0, 1, \leq \rangle$ denotes the ordered field of real numbers. Two structures are said to be similar provided they are both $\mathcal{L}$-structures for the same language.

Let $\mathcal{L}$ and $\mathcal{L}'$ be languages such that $\mathcal{L} \subseteq \mathcal{L}'$. Let $A$ be an $\mathcal{L}$-structure and let $A'$ be an $\mathcal{L}'$-structure. We say that $A'$ is an expansion of $A$ and that $A$ is a reduct of $A'$ provided

\[ A = A', \]
\[ Q^A = Q^{A'} \text{ for every operation symbol } Q \in \mathcal{L} \text{ and} \]
\[ R^A = R^{A'} \text{ for every relation symbol } R^+ \in \mathcal{L}. \]

That is, $A'$ is an expansion of $A$ if and only if $A$ and $A'$ have the same universe and the fundamental operation and relation symbols of $\mathcal{L}$ denote the same fundamental operations and relations in both $A$ and $A'$. So we could obtain $A$ from $A'$ by ignoring some of the fundamental operations and relations.

Let $\mathcal{L}$ be an elementary language and let $A$ and $B$ be $\mathcal{L}$-structures. Several notions of what it might mean for $B$ to be a substructure of $A$ or for a map from $A$ to $B$ to be a homomorphism suggest themselves. We settle on the following alternatives. We say that $B$ is a substructure of $A$, and we write $B \subseteq A$, provided

i. $B \subseteq A$,
ii. $Q^B$ is the restriction to $B$ of $Q^A$, for every fundamental operation symbol, and
iii. $R^B = R^A \cap B^r$, where $r$ is the rank of $R$ and $R$ is any basic relation symbol.

The notion of a subuniverse remains as it was for algebras: a subset of the universe which is closed with respect to all be basic operations—the basic relations play no role.

Now let $h : A \rightarrow B$. We say that $h$ is a homomorphism from $A$ into $B$, and we write $h : A \rightarrow B$, provided

i. $h(Q^A(a_0, a_1, \ldots, a_{r-1})) = Q^B(h(a_0), h(a_1), \ldots, h(a_{r-1}))$ whenever $Q$ is a basic operation symbol, $r$ is its rank, and $a_0, a_1, \ldots, a_{r-1} \in A$, and
ii. if $R^A(a_0, a_1, \ldots, a_{r-1})$, then $R^B(h(a_0), h(a_1), \ldots, h(a_{r-1}))$, whenever $R$ is a basic relation symbol, $r$ is its rank, and $a_0, a_1, \ldots, a_{r-1} \in A$.

Among the notions which might be reasonably called “homomorphism”, the one defined above is weak. This can be an irritation since it produces effects that run counter to those holding for homomorphisms between algebras. We saw some of these effects when we considered isotone maps for lattice-ordered
sets in Chapter 2. In particular, it is possible for \( h \) to be a homomorphism from a structure \( A \) onto several nonisomorphic structures. In the context of arbitrary structures, we cannot expect the close connection between homomorphic images and quotient structures that prevails among algebras according to the Homomorphism Theorem. One must be careful not to leap to the conclusion that \( h(A) \) is a uniquely determined structure. Perhaps more vexing, there can be one-to-one homomorphisms from one structure onto another which should not be called isomorphisms. In fact, we insist that an isomorphism from \( A \) onto \( B \) is a one-to-one homomorphism \( h \) from \( A \) onto \( B \) such that \( h^{-1} \) is a homomorphism from \( B \) onto \( A \).

### Congruence relations

For structures have the same definition as for algebras. In particular, the fundamental relations play no role and, in the absence of any fundamental operations (i.e. for relational structures) every equivalence relation is a congruence relation. Let \( A \) be any structure and let \( \theta \) be any congruence on \( A \). We form the quotient structure \( A/\theta \) by handling the universe and the basic operations exactly as for algebras, and then defining the basic relations of \( A/\theta \) as follows. Let \( R \) be a basic relation symbol and let \( r \) be its rank. Define \( R^{A/\theta} \) on \( A/\theta \) so that for all \( a_0, a_1, \ldots, a_{r-1} \in A \), we have

\[
R^{A/\theta}(a_0/\theta, a_1/\theta, \ldots, a_{r-1}/\theta)
\]

if and only if

\[
R^{A}(b_0, b_1, \ldots, b_{r-1})
\]

for some \( b_0, b_1, \ldots, b_{r-1} \) such that \( b_0 \theta a_0, b_1 \theta a_1, \ldots, b_{r-1} \theta a_{r-1} \). Thus, in the quotient structure only those basic relations are made to hold which are necessary in order that the quotient map \( a \mapsto a/\theta \) be a homomorphism.

Turning to direct products, let \( I \) be any set and let \( A_i \) be an \( L \)-structure for each \( i \in I \). The direct product \( A = \prod_{i \in I} A_i \) is defined as for algebras, with the basic relations handled “coordinatewise”. That is, if \( R \) is a basic relation symbol and \( r \) is its rank, then

\[
R^A(b_0, b_1, \ldots, b_{r-1})
\]

if and only if \( R^{A_i}(b_{0,i}, b_{1,i}, \ldots, b_{r-1,i}) \) for all \( i \in I \).

With this definition of direct product, the projection functions are homomorphisms. To keep the notation simple, we write \( \prod_i A_i \) if it is clear that the index \( i \) ranges over the index set \( I \).

Just as operation symbols name the fundamental operations of a structure, so relation symbols name the fundamental relations. We have seen how the variables and operations symbols can be compounded to form terms and how these terms name the term functions of a given structure. With the essential help of the connectives and quantifiers, we will now see how to make compound expressions called formulas which will serve as names for certain finitary relations on a given structure.

Let \( L \) be an elementary language. An \( L \)-expression is just a finite string of \( L \)-symbols. The simplest \( L \)-formulas are called atomic formulas and they consist of all \( L \)-expressions of the following forms:

i. \( s \approx t \), where \( s \) and \( t \) are any \( L \)-terms, and

ii. \( R t_0 \ldots t_{r-1} \), where \( R \) is any relation symbol of \( L \), \( r \) is the rank of \( R \), and \( t_0, \ldots, t_{r-1} \) are any \( L \)-terms.

The \( L \)-formulas are just those expressions which can be built from the atomic formulas with the help of the quantifiers and connectives.
DEFINITION 7.3. Take $\mathcal{L}$ to be an elementary language. The set of $\mathcal{L}$-formulas is the smallest set $F$ of $\mathcal{L}$-expressions such that

i. Every atomic $\mathcal{L}$-formula belongs to $F$,

ii. If $\varphi, \psi \in F$, then $\neg \varphi$, $(\varphi \land \psi)$, $(\varphi \lor \psi)$, $(\varphi \implies \psi)$, and $(\varphi \iff \psi) \in F$,

iii. If $\varphi \in F$ and $x$ is any variable, then $\exists x \varphi, \forall x \varphi \in F$.

This definition of the notion of formulas, like the definition of the notion of terms given as Definition 4.114 in Volume 1, is recursive: more complicated formulas are built from less complicated formulas. Induction is therefore a powerful gambit for proving assertions about formulas. Ultimately, such inductions rest on the fact that formulas are uniquely readable in the sense that there is only one way to decompose a given formula along the lines of the definition above. We dispense with the precise formulation and proof of this unique readability of formulas, but the interested reader can consult Lemma 4.115 in Volume 1 for the corresponding result for terms.

Our definition of the set of $\mathcal{L}$-formulas leads to a rather awkward notation for formulas. We follow the customary practice and regard formulas like

$$x \leq y \land y \leq z \Rightarrow x \leq z$$

as perfectly acceptable, even though our official definition would yield

$$((x \leq y \land y \leq z) \Rightarrow x \leq z)$$

We omit parentheses when the meaning is clear, we use a mixture of parentheses and brackets to enhance readability, and we insert binary relation symbols between two terms rather than in front of them.

In displaying long formulas, we sometimes write

$$\bigwedge_{i \in I} \varphi_i$$

in place of $\varphi_0 \land \varphi_1 \land \ldots \land \varphi_i \land \ldots$. We use

$$\bigvee_{i \in I} \varphi_i$$

to replace $\varphi_0 \lor \varphi_1 \lor \ldots$.

Let $\mathcal{L}$ be the language appropriate to the ordered field of real numbers. Consider the following formulas:

$$x^2 + y^2 \leq 4$$

$$x^2 + y^2 \leq 4 \land \exists x [x^2 \approx y]$$

$$\forall x \exists y [x^2 + y^2 \leq 4]$$

Here 4 is an abbreviation of $1 + 1 + 1 + 1$ and $x^2$ abbreviates $x \cdot x$. For the ordered field $\langle \mathbb{R}, +, -, \cdot, 0, 1, \leq \rangle$ of real numbers, the first formula defines the closed disc of radius 2, the second formula defines the upper half of the disc of radius 2, and the third formula is a statement which is false in the ordered field of real numbers. The reason for the distinction between the meaning of the
third formula and the other two stems from the fact that all the occurrences of variables in the last formula are bound under the control of the quantifiers, while the first two formulas have occurrences of variables free of such control. The bound variables are very much like the dummy variable \( u \) used in the definite integral \( \int_a^b f(u) \, du \). In the first of the formulas displayed above all occurrences of variables are free. In the second formula all occurrences of \( y \) are free, as is the leftmost occurrence of \( x \), but the remaining occurrences of \( x \) are bound. The last formula has no free occurrences of variables. We formalize this concept.

**Definition 7.4.** Let \( \mathcal{L} \) be an elementary language. Let \( V_a \) denote the function which assigns to each term \( \varphi \) the set \( V_a(\varphi) \) of all variables that occur in \( \varphi \). \( F_v \) is the function with domain the set of all \( \mathcal{L} \)-formulas specified by

\[
F_v(\varphi) = \begin{cases} 
V_a(\varphi) & \text{if } \varphi \text{ is atomic} \\
F_v(\theta) \cup F_v(\psi) & \text{if } \varphi = (\theta \land \psi), (\theta \lor \psi), (\theta \Rightarrow \psi), \text{ or } (\theta \Leftrightarrow \psi) \\
F_v(\psi) & \text{if } \varphi = \neg \psi \\
F_v(\psi) - \{x\} & \text{if } \varphi = \exists x \psi \text{ or } \forall x \psi 
\end{cases}
\]

\( F_v(\varphi) \) is called the **set of free variables** of \( \varphi \). A formula without any free variables is called a **sentence**.

Let \( \varphi \) be a formula. We write \( \varphi(y_0, y_1, \ldots, y_{n-1}) \) to mean that all the free variables of \( \varphi \) occur among \( y_0, y_1, \ldots, y_{n-1} \). More formally, this means \( F_v(\varphi) \subseteq \{y_0, y_1, \ldots, y_{n-1}\} \). Notice that we do not insist of equality, but only on set-inclusion. If \( n \) is unimportant, we write \( \varphi(\bar{y}) \) in place of \( \varphi(y_0, y_1, \ldots, y_{n-1}) \). If it is unimportant to distinguish the variables, we write

\[
\exists \bar{y} \varphi(\bar{x}, \bar{y}) \text{ for } \exists y_0 \exists y_1 \ldots \exists y_{n-1} \varphi(\bar{x}, y_0, y_1, \ldots, y_{n-1})
\]

The universal quantifier \( \forall \) is treated in the same manner.

Now we can specify how formulas define relations on structures and what it means for an elementary sentence to be true in a structure. Because our definition goes by recursion on the complexity of formulas, it is inconvenient to set a bound on the number of variables which occur freely in the formulas. To accommodate this, we employ countably infinite sequences of elements. Also, to simplify this definition and subsequent work, we limit our connectives to \( \neg \) and \( \land \), and we take \( \exists \) as our only quantifier. This represents no loss of generality since the remaining symbols can be regarded as abbreviating various schema built with the help of only these three symbols.

**Definition 7.5.** Let \( A \) be an \( \mathcal{L} \)-structure. For every \( \mathcal{L} \)-formula \( \varphi \), we take \( \varphi^A \) to be a certain \( \omega \)-ary relation on \( A \) by specifying whether \( \varphi^A(\bar{a}) \) [that is, whether \( \bar{a} \in \varphi^A \)] for each \( \bar{a} \in A^\omega \) as follows:

i. \((s \approx t)^A(\bar{a}) \) if and only if \( s^A(\bar{a}) = t^A(\bar{a}) \)

ii. \((R t_0 t_1 \ldots t_{r-1})^A(\bar{a}) \) if and only if \( R^A(t_0^A(\bar{a}), \ldots, t_{r-1}^A(\bar{a})) \)

iii. \((-\psi)^A(\bar{a}) \) if and only if \( \psi^A(\bar{a}) \) fails.
iv. \((\theta \land \psi)^A(\bar{a})\) if and only if \(\theta^A(\bar{a})\) and \(\psi^A(\bar{a})\).

v. \((\exists v. \psi)^A(\bar{a})\) if and only if \(\psi^A(\bar{b})\) for some \(\bar{b} \in A^\omega\) such that \(a_j = b_j\) for all \(j \neq i\).

We use \(A \models \varphi(\bar{a}), \varphi^A(\bar{a}),\) and \(\bar{a} \in \varphi^A\) interchangeably. We read all three of these as "\(\bar{a}\) satisfies \(\varphi\) in \(A\)."

This definition appears rather involved, but it holds no surprises. What we have done is give to the connectives and quantifiers the meanings we intended for them all along. In this definition, even though \(\bar{a}\) is an infinite sequence of elements, it is plain that whether \(\bar{a}\) satisfies \(\varphi\) hinges only on (the indices of) the free variables of \(\varphi\), a finite set. The sets

\[\{\bar{a} : \bar{a} \in A^\omega \text{ and } \varphi^A(\bar{a})\}\]

are infinitary relations which depend essentially on only finitely many coordinates. We regard \(\varphi^A\) as a finitary relation on \(A\), relegating a completely correct formulation of this to the exercises. Those finitary relations on \(A\) which are of the form \(\varphi^A\) for some formula \(\varphi\) are said to be definable in \(A\).

Finally, we say that \(\varphi\) is true in \(A\), and alternatively that \(A\) is a model of \(\varphi\) if and only if \(A \models \varphi(\bar{a})\) for all \(\bar{a} \in A^\omega\). If \(\varphi\) is an equation and \(A\) is an algebra, then this definition agrees with the definition provided in Chapter 4 of Volume 1. We use

\(A \models \varphi\)

to denote that \(\varphi\) is true in \(A\). Observe that for any formula \(\varphi(x_0, \ldots, x_{n-1})\) we have

\(A \models \varphi(x_0, \ldots, x_{n-1})\) if and only if \(A \models \forall x_0 \ldots \forall x_{n-1} \varphi(x_0, \ldots, x_{n-1})\)

Now \(\forall x_0 \ldots \forall x_{n-1} \varphi(x_0, \ldots, x_{n-1})\) is a sentence. Thus, to know all the formulas true in \(A\) it suffices to know all the sentences true in \(A\). So for the most part, we regard \(\models\) as a binary relation between \(L\)-structures and \(L\)-sentences. This relation establishes a Galois connection in a manner similar to the connection between varieties and equational theories described in Chapter 4 of Volume 1. The polarities of the present Galois connection are denoted as follows:

\[\text{Mod } \Sigma = \{A : A \models \varphi \text{ for all } \varphi \in \Sigma\}\]

where \(\Sigma\) is any set of \(L\)-sentences, and

\[\text{Th } \mathcal{K} = \{\varphi : A \models \varphi \text{ for all } A \in \mathcal{K}\}\]

where \(\mathcal{K}\) is any class of \(L\)-structures. Classes of the form \(\text{Mod } \Sigma\) are called elementary classes and sets of the form \(\text{Th } \mathcal{K}\) are called elementary theories. An elementary class \(\mathcal{K}\) is said to be finitely axiomatizable provided \(\mathcal{K} = \text{Mod } \Sigma\) for some finite set \(\Sigma\) of sentences.

Two sets \(\Sigma\) and \(\Gamma\) of \(L\)-sentences are said to be logically equivalent provided \(\text{Mod } \Sigma = \text{Mod } \Gamma\); that is, they have exactly the same models. We say that the sentence \(\theta\) is a logical consequence of the set \(\Sigma\) of \(L\)-sentences, and
we write $\Sigma \models \theta$, if and only if every model of $\Sigma$ is a model of $\theta$. An elementary theory $T$ is said to be finitely axiomatizable provided $T$ is logically equivalent to some finite set $\Sigma$ of elementary sentences.

Two $\mathcal{L}$-structures $A$ and $B$ are elementarily equivalent, rendered in symbols as $A \equiv B$, if and only if $\text{Th} A = \text{Th} B$. Elementary equivalence is an equivalence relation on the class of $\mathcal{L}$-structures which is coarser than the relation of isomorphism. [That is, if $A \equiv B$, then $A \cong B.$] In the next section we will see that every infinite structure in a countable language is elementarily equivalent to structures of arbitrarily large infinite cardinality, so that $\equiv$ is indeed much coarser than $\cong$ among infinite structures. For finite structures these relations are the same.

**THEOREM 7.6.** Let $\mathcal{L}$ be an elementary language and let $A$ and $B$ be $\mathcal{L}$-structures. If $A \equiv B$ and $A$ is finite, then $A \cong B$.

**Proof.** Using just variables and the logical symbols, for each natural number $n$ it is possible to write down a sentence which expresses “there are at least $n$ distinct elements”. For example, here is such a sentence expressing “there are at least three elements”:

$$\exists x_0 \exists x_1 \exists x_2 [\neg(x_0 \approx x_1) \land \neg(x_0 \approx x_2) \land \neg(x_1 \approx x_2)]$$

Call this sentence $\delta_3$. The reader can devise analogous sentences $\delta_n$ for each natural number $n$. The sentence $\neg \delta_{n+1}$ express “there are no more than $n$ elements” and the sentence $\delta_n \land \neg \delta_{n+1}$ expresses “there are exactly $n$ elements.” It now follows that $B$ has the same (finite) cardinality as $A$.

For the sake of contradiction, assume that $A$ and $B$ are not isomorphic. There are only finitely many one-to-one maps from $A$ onto $B$. For each such map pick an operation or relation symbol which the map does not respect. Let $\mathcal{L}'$ be the finite subset of $\mathcal{L}$ comprised of these symbols, and let $A'$ and $B'$ be the reducts of $A$ and $B$ to $\mathcal{L}'$. Then $A' \equiv B'$ and $A'$ is not isomorphic with $B'$. Let $\bar{a} = (a_0, \ldots, a_{n-1})$ be a list of all the distinct elements of $A$. Let

$$\Delta = \{ \varphi(\bar{x}) : \varphi \text{ is an atomic } \mathcal{L}' \text{-formula and } A \models \varphi(\bar{a}) \}$$

$$\cup \{ \neg \varphi(\bar{x}) : \varphi \text{ is an atomic } \mathcal{L}' \text{-formula and } A \models \neg \varphi(\bar{a}) \}$$

Let $\sigma$ be the conjunction of all the formulas in $\Delta$ [that is, the formula resulting from combining all the formulas in $\Delta$ by $\land$]. Let $\psi$ be the formula $\sigma \land \neg \delta_{n+1}$. Then certainly $A' \models \exists x_0 \ldots \exists x_{n-1} \psi$. Hence, $B' \models \exists x_0 \ldots \exists x_{n-1} \psi$. So pick $b_0, \ldots, b_{n-1} \in B$ so that

$$B' \models \psi(b_0, \ldots, b_{n-1}).$$

$\psi$ carries all the information necessary to demonstrate that the map $a_i \mapsto b_i$ for all $i < n$ is an isomorphism from $A'$ onto $B'$. This is a contradiction. $\blacksquare$

It is interesting to note that the formula $\psi$ in the proof just given is, in some sense, a complete description of $\langle a_0, \ldots, a_{n-1} \rangle$ in $A'$. In fact, taking $\pi = \exists x_0 \ldots \exists x_{n-1} \psi$, we have that if $C \models \pi$, then $C \cong A'$, for all $\mathcal{L}'$-structures $C$. Thus, we have the following corollary.
**COROLLARY 7.7.** If $\mathcal{L}$ is a finite elementary language and $A$ is a finite $\mathcal{L}$-structure, then there is a sentence $\pi$ such that $A \equiv B$ if and only if $B \models \pi$.

**COROLLARY 7.8.** If $\mathcal{L}$ is any elementary language and $A$ is a finite $\mathcal{L}$-structure, then $\text{Mod Th} A$, the smallest elementary class containing $A$, consists of precisely the isomorphic images of $A$.

These two corollaries point out the sharp distinction between varieties and equational theories on the one hand, and elementary classes and elementary theories on the other.

The classification of structures into elementary classes is much finer than the classification of algebras into varieties. For example, the class of all fields is an elementary class, as can be seen from the customary axioms of field theory, but it is not a variety. By adding the sentence $1 + 1 \equiv 0$ to the field axioms we obtain the class of all fields of characteristic 2. Similar sentences can be devised for every prime $p$. By adding instead the negations of these sentences, we arrive at the class of algebraically closed fields is an elementary class. However, there are many significant classes which fail to be elementary. The class of Archimedean ordered fields is of this kind.

One way to measure the price that must be paid for this finer classification scheme is to observe that elementary classes in general are not closed under any of the operators $H$, $S$, or $P$. Consequently, the elaboration of elementary model theory relies on the rich means of expression provided by elementary languages more than it relies on the manipulations of homomorphisms, substructures, congruences, or direct products. Nevertheless, these notions still retain a position of importance in model theory generally. Of course, they are never far from the focus of our attention.

To a large extent, our choices for how to extend the algebraic notions of substructure, homomorphism, and direct product to $\mathcal{L}$-structures was motivated by the desire to have atomic formulas in general behave like equations. For example, whenever $B \subseteq A$, an easy induction on the complexity of terms reveals that for any atomic formula $\varphi(\bar{x})$ and any tuple $\bar{b}$ from $B$,

$$B \models \varphi(\bar{b}) \text{ if and only if } A \models \varphi(\bar{b})$$

Likewise, whenever $h$ is a homomorphism from $B$ into $A$, we have

$$B \models \varphi(\bar{b}) \text{ implies } A \models \varphi(h(\bar{b})).$$

By the same sort of reasoning, if $A = \prod(A_i : i \in I)$, then the family $\{p_i : i \in I\}$ of projection functions separates points in the strong sense that if $\varphi(\bar{x})$ is any atomic formula [not just $v_0 \approx v_1$] and $\bar{a}$ is a tuple from $A$ such that $A \not\models \varphi(\bar{a})$, then $A_i \not\models \varphi(p_i(\bar{a}))$, for some $i \in I$. Some of the key notions of model theory arise from the effort to extend these conditions from atomic formulas to all formulas.
DEFINITION 7.9. Let $A$ and $B$ be $\mathcal{L}$-structures and let $h$ be a function from $A$ into $B$. We say that $h$ is an elementary embedding if and only if $A \models \varphi(\bar{a})$ implies $B \models \varphi(h(\bar{a}))$ for all formulas $\varphi$ and all tuples $\bar{a}$ from $A$.

That $h$ is an elementary embedding of $A$ into $B$ means that any property of any finite sequence of elements of $A$ which can be expressed by $\mathcal{L}$-formulas must also be a property of the sequence of elements of $B$ corresponding by the map $h$. As the properties expressed by atomic formulas are respected, we see that elementary embeddings are homomorphisms. To say that $\neg (x \approx y)$ is respected by $h$ is easily the same as saying that $h$ is one-to-one. More generally, since $h$ respects the negations of all atomic formulas, we see that $h$ is an isomorphism from $A$ onto a substructure of $B$. So the use of “embedding” rather than “morphism” is justified. But elementary embeddings are much stronger than ordinary embeddings.

EXAMPLE 7.10. The only elementary embedding of $\langle \omega, < \rangle$ into itself is the identity map.

The essential reason for this is that using $<$ and the logical symbols we can compose a formula $\varphi_n(x)$ which expresses “there are exactly $n$ elements less than $x$”. Since $\langle \omega, < \rangle \models \varphi_n(n)$, we see that $\langle \omega, < \rangle \models \varphi_n(h(n))$ for all $n \in \omega$ and for every elementary embedding $h$. Thus $h(n) = n$ for all $n \in \omega$ and for all elementary embeddings $h$. On the other hand, $\langle \omega, < \rangle$ has plenty of embedding into itself. The successor function is an embedding.

A is said to be an elementary substructure of $B$ (and $B$ is called an elementary extension of $A$) provided $A \subseteq B$ and the inclusion map in an elementary embedding of $A$ into $B$. $A \prec B$ denotes that $A$ is an elementary substructure of $B$. The same concept arises when the notion of substructure is modified by insisting that all formulas be respected rather than just atomic formulas. The structure $\langle \omega, < \rangle$ has no proper elementary substructures. On the other hand, $\langle \omega - \{0\}, < \rangle$ is isomorphic with $\langle \omega, < \rangle$. So we have an example of two structures $A$ and $B$ such that $A \subseteq B$ and $A \equiv B$ (even $A \cong B$), but $A$ is not an elementary substructure of $B$. Here are some easy facts about $\prec$.

THEOREM 7.11. Let $\mathcal{L}$ be an elementary language and let $A$, $B$, and $C$ be $\mathcal{L}$-structures.

i. If $A \prec B$ and $B \prec C$, then $A \prec C$.

ii. If $A \prec B$, then $A \subseteq B$ and $A \equiv B$.

iii. If $A \prec C$ and $B \prec C$ and $A \subseteq B$, then $A \prec B$.

The next theorem supplies a useful test for elementary extensions.

THEOREM 7.12 (The Tarski-Vaught Elementary Extension Test).

Let $\mathcal{L}$ be an elementary language and let $A$ and $B$ be $\mathcal{L}$-structures. $A \prec B$ if and only if both conditions below hold:
7.1 Elementary Model Theory

i. \( A \subseteq B \), and

ii. If \( B \models \exists y \varphi(y, a_0, a_1, \ldots) \), then \( A \models \varphi(a, a_0, a_1, \ldots) \) for some \( a \in A \), for any formula \( \varphi(y, x_0, x_1, \ldots) \) and any \( a_0, a_1, \ldots \in A \).

**Proof.** We deal only with the more difficult direction, so we assume (i) and (ii) and prove \( A \prec B \). We will prove by induction on the complexity of \( \psi \) that for all formulas \( \psi \) and all \( \bar{a} \in A^n \) that \( B \models \psi(\bar{a}) \) if and only if \( A \models \psi(\bar{a}) \). If \( \psi \) is atomic, this is just the content of (i). In the event that \( \psi \) is not atomic there are three cases.

Case \( \neg \psi \) is \( \neg \varphi \) for some formula \( \varphi \).

Just observe

\[
B \models \psi(\bar{a}) \quad \text{if and only if} \quad B \not\models \varphi(\bar{a})
\]

if and only if \( A \not\models \varphi(\bar{a}) \)

if and only if \( A \models \psi(\bar{a}) \).

Case \( \land \psi \) is \( \varphi \land \theta \) for some formulas \( \varphi \) and \( \theta \).

Just observe

\[
B \models \psi(\bar{a}) \quad \text{if and only if} \quad B \models \varphi(\bar{a}) \quad \text{and} \quad B \models \theta(\bar{a})
\]

if and only if \( A \models \varphi(\bar{a}) \) and \( A \models \theta(\bar{a}) \)

if and only if \( A \models \psi(\bar{a}) \).

Case \( \exists \psi \) is \( \exists y \varphi(y, x_0, x_1, \ldots) \).

If \( B \models \psi(\bar{a}) \), then \( A \models \psi(\bar{a}) \) according to condition (ii). For the converse, suppose that \( A \models \psi(\bar{a}) \). Then \( A \models \varphi(c, \bar{a}) \) for some \( c \in A \). The induction hypothesis yields \( B \models \varphi(c, \bar{a}) \). Hence \( B \models \psi(\bar{a}) \). \( \blacksquare \)

**COROLLARY 7.13.** Let \( A \) and \( B \) be \( \Sigma \)-structures such that \( A \subseteq B \). Suppose that for any finite sequence \( a_0, \ldots, a_{n-1} \) of elements of \( A \) and for any \( b \in B \), there is an automorphism of \( B \) leaving each of \( a_0, \ldots, a_{n-1} \) fixed but mapping \( b \) into \( A \). Then \( A \prec B \).

With the help of this corollary it is easy to see that the rationals interior to the unit interval comprise an elementary substructure of \( \langle \mathbb{Q}, < \rangle \), as does the set of all dyadic rationals (those of the form \( m/2^n \)).

Here is our first application to the theory of varieties.

**THEOREM 7.14.** Let \( \mathcal{V} \) be any variety and let \( X \subseteq Y \) with \( X \) infinite. Let \( B \) be the algebra free in \( \mathcal{V} \) over \( Y \) and let \( A \) be the subalgebra of \( B \) generated by \( X \). Then \( A \prec B \).

**Proof.** We invoke Corollary 7.13. Let \( a_0, \ldots, a_{n-1} \in A \) and let \( b \in B \). Let \( X' \) be a large enough finite subset of \( X \) so that \( a_0, \ldots, a_{n-1} \) belong to the subuniverse of \( A \) generated by \( X' \). Let \( Y' \) be a large enough finite subset of \( Y - X' \) so that \( b \) is in the subuniverse generated by \( X' \cup Y' \). Let \( X'' \) be any
subset of \( X - X' \) such that \(|X''| = |Y'|\). Let \( h \) be any permutation of \( Y \) that maps \( Y' \) onto \( X'' \) and leaves every element of \( Y - (Y' \cup X'') \) fixed. Since \( B \) is free in \( V \) over \( Y \), we know that \( h \) extends to an automorphism of \( B \). Under this automorphism \( X' \) is fixed, and so each of \( a_0, \ldots, a_{n-1} \) is also fixed. On the other hand, \( X' \cup Y' \) gets mapped into \( X \). So the image of \( b \) under this automorphism lies in \( A \). Thus, \( A \prec B \).

**COROLLARY 7.15.** Let \( V \) be any nontrivial variety. If \( \kappa \) and \( \lambda \) are any two infinite cardinals, then \( F_V(\kappa) \equiv F_V(\lambda) \).

The next theorem is a descendant of one published by Leopold Löwenheim in 1914—possibly the oldest theorem which belongs openly to model theory. It reveals a key property of elementary model theory.

**THEOREM 7.16 (The Downward Löwenheim-Skolem-Tarski Theorem).**

Let \( L \) be an elementary language and let \( \kappa \) be the cardinality of the set of all \( L \)-formulas. Let \( A \) be an \( L \)-structure and let \( X \subseteq A \) with \(|X| \geq \kappa \). Then \( A \) has an elementary substructure \( B \) such that \( X \subseteq B \) and \(|B| = |X| \).

**Proof.** Let \( B_0 = X \). For each formula \( \varphi(y, \bar{x}) \) and for all appropriate tuples \( \bar{b} \) from \( B_0 \) such that \( A \models \exists y \varphi(y, \bar{b}) \), select \( a \in A \) such that \( A \models \varphi(a, \bar{b}) \). Let \( C_0 \) be the set of all elements \( a \) selected as above as \( \varphi(y, \bar{x}) \) ranges through the formulas and \( \bar{b} \) ranges through all appropriate tuples from \( B_0 \). Let \( B_1 = B_0 \cup C_0 \). Since \( \kappa \leq |B_0| \) we see that \(|B_1| = |B_0| = |X| \). Now we repeat this process on \( B_1 \) to obtain \( C_1 \) and then \( B_2 \). Continuing in this way through countably many steps, we obtain a tower of sets

\[
X = B_0 \subseteq B_1 \subseteq B_2 \subseteq \cdots
\]

Let \( B = \bigcup_{k \in \omega} B_k \). Clearly, \( X \subseteq B \) and \(|B| = |X| \). To see that \( B \) is a subuniverse of \( A \), let \( Q \) be any operation symbol and pick \( b_0, \ldots, b_{r-1} \in B \), where \( r \) is the rank of \( Q \). Pick \( m \) so large that \( b_0, \ldots, b_{r-1} \in B_m \). Now

\[
A \models \exists y[Qb_0 \ldots b_{r-1} \approx y].
\]

By the construction, \( QA(b_0, \ldots, b_{r-1}) \in B_{m+1} \subseteq B \). So \( B \) is a subuniverse of \( A \). That \( B \) satisfies the second condition of the Tarski-Vaught Elementary Extension Test is built into our construction of \( B \). Thus \( B \prec A \).

At this point, the reader without previous exposure to elementary languages should try to express some common mathematical notions in the elementary setting. There are really two sorts of exercises of this kind:

- The discovery of elementary sentences which axiomatize familiar classes of structures.
- Within a given structure, the discovery of elementary formulas which define familiar sets, relations, or operations.

This section concludes with a set of exercises, among which may be found a number of problems of both kinds. After trying a few of these exercises
the reader should have the impression that quite a few common mathematical notions can find elementary expression. On the other hand, the Downward Löwenheim-Skolem-Tarski Theorem informs us that no set of elementary sentences can capture \( \langle \mathbb{R}, +, -, \cdot, 0, 1, \leq \rangle \) up to isomorphism—the familiar axiomatizations of Dedekind to the contrary. The theorems proven in the next section draw much sharper limitations on the classes of structures which can be axiomatized by elementary sentences. Roughly, a class of structures whose ordinary definition in mathematical practice refers only to properties of elements (taken finitely many at a time) of the structures will be an elementary class. Those classes whose ordinary descriptions refer to sets of elements, to sequences of elements, or to external objects like natural numbers are usually not elementary. Of course, just because the customary description of a familiar class of structures appears to be nonelementary, it does not follow that no elementary description exists. The usual way to specify the class of algebraically closed fields involves quantifying over all polynomials. Nevertheless, this class is elementary.

The determination of which sets, relations, and operations are definable by elementary formulas in a given structure is usually very difficult. The following results, listed here without proof, offer some insight into the subtlety of definability questions:

A. The set of natural numbers is definable in the ring of integers.
B. The set of integers is definable in the field of rational numbers.
C. The binary operation \( x \uparrow y = x^y \) of exponentiation is definable in \( \langle \omega, +, \cdot \rangle \).
D. Addition is not definable in \( \langle \omega, \cdot \rangle \) nor is multiplication definable in \( \langle \omega, + \rangle \), but both addition and multiplication are definable in \( \langle \omega, \uparrow \rangle \).
E. The sets definable in the field of real numbers are exactly the unions of finitely many intervals with algebraic endpoints. In particular, the set of rationals is not definable in the field of real numbers.
F. The set of compact elements is definable in the lattice of all equational theories of a given fixed similarity type.

Perhaps (A) is immediate, but only for those who know Lagrange’s theorem to the effect that every natural number is the sum of four squares. (C) and (D) require imagination and persistence. The remaining items are genuinely difficult.

Exercises 7.17

1.
7.2. Reduced Products, Ultraproducts, and the Compactness Theorem

Let $\mathcal{L}$ be any elementary language. General considerations concerning Galois connections lead us to conclude that $\text{Mod Th}$ is a closure operator on the class of all $\mathcal{L}$-structures and that $\text{Mod Th}\mathcal{X}$ is the smallest elementary class including $\mathcal{X}$, for all classes $\mathcal{X}$ of $\mathcal{L}$-structures. In the equational context, Birkhoff’s HSP Theorem (Theorem 4.131 in Volume 1) plays a central role because, in part, it provides an analysis of the corresponding closure operator $\text{Mod Th}$. The attempt to provide a similar analysis of $\text{Mod Th}$ runs into difficulty early. An easy but crucial observation leading to the HSP Theorem is that the truth of equations is preserved under the formation of homomorphic images, subalgebras, and direct products. The corresponding statements are all false when arbitrary elementary sentences are considered in place of equations. Indeed, the sentence asserting “There are exactly two elements.”, which can be formulated in any elementary language, is not preserved under any of the algebraic constructions, apart from isomorphism, which we have so far considered. To make any progress in analyzing $\text{Mod Th}$ we need to consider new algebraic constructions. The ultraproduct construction described in this section will serve our purpose.

Ultraproducts arise as special quotients of direct products. Let $A = \prod_I A_i$, where $A_i$ is an $\mathcal{L}$-structure for each $i \in I$. The members of $A$ are $I$-tuples. The intuition in forming the quotient structure is to identify those $I$-tuples which are substantially the same—that is, the indices for which the identified tuples agree form a large subset of $I$. To frame this in a mathematically acceptable way, we have to come to grips with the notion of “largeness”. Consider two examples where $I$ is the unit interval. The subsets of $I$ which have Lebesgue measure 1 certainly seem to be sets which occupy almost all of $I$. An even more restrictive notion is to take as “large” those subsets of $I$ which are cofinite. Both of these notions of “large” possess the following properties:

i. $I$ is large.
ii. If $X$ is large and $X \subseteq Y \subseteq I$, then $Y$ is large.
iii. If $X$ and $Y$ are large, then $X \cap Y$ is large.

Property (iii) is a reasonable attribute of largeness. If $Y$ is large it ought to occupy all but a small portion of $I$, and hence it overlaps all but a small portion of $X$. Since $X$ occupies all but a small portion of $I$, we see that $X \cap Y$ occupies all but two small portions of $I$. For any set $I$, these three properties characterize those families of subsets of $I$ which are filters in the lattice of all subsets of $I$. More specifically, a collection $F$ of subsets of $I$ is said to be a filter on $I$ provided

i. $I \in F$,
ii. if $X \in F$ and $X \subseteq Y \subseteq I$, then $Y \in F$, and
iii. if $X$ and $Y$ belong to $F$, then $X \cap Y \in F$.

Each filter on $I$ poses a concept of largeness. So there is no single notion of largeness, but rather one associated with each of the plethora of filters on $I$. 
There are two ways in which a filter can be trivial. First, it could consist of all subsets of $I$. Such indiscriminate filters are said to be improper. $\emptyset \notin F$ characterizes a filter $F$ as a proper filter. Second, there could be a single subset $X$ of $I$ which traps the whole essence of largeness in the sense that $Y \in F$ if and only if $X \subseteq Y \subseteq I$. Such filters are called principal filters.

Let $I$ be any infinite set. The collection of cofinite subsets of $I$ is a filter on $I$ which is both proper and not principal. It is called the Fréchet filter on $I$.

On the other hand, it is easy to see that any filter which has a finite member must be principal. In particular, every filter on a finite set must be principal.

Before taking up the task of building filters of more diverse kinds, we present the reduced product construction. Let $I$ be any set, let $F$ be a filter on $I$, and let $A_i$ be an $\mathcal{L}$-structure for each $i \in I$. Let $A = \prod I A_i$. Consider $a$ and $a'$ from $A$. So

$$a = \langle a_i : i \in I \rangle \text{ and } a' = \langle a'_i : i \in I \rangle.$$  

Our idea is to regard $a$ and $a'$ as congruent modulo $F$ provided

$$\{ i : a_i = a'_i \text{ and } i \in I \} \in F.$$  

Another way to phrase this condition is

$$\{ i : A_i \models p_i(a) \approx p_i(a') \text{ and } i \in I \} \in F$$

which emphasizes that $a \approx a'$ is true on a “large” set of coordinates. It is fruitful to introduce some notation, since we will meet such phrases frequently.

Let $\varphi(\bar{x})$ be any formula and let $\bar{a}$ be an appropriate tuple from $A$. Thus, if $\bar{x} = \langle x_0, \ldots, x_{n-1} \rangle$, then $\bar{a}$ will be a certain $n$-tuple, say $\bar{a} = \langle a_0, \ldots, a_{n-1} \rangle$.

Recall that each $a_k$ is, in turn, an $I$-tuple. For each $i \in I$ we use $p_i$ to denote the $i$th projection function. Finally, for each $i \in I$ we abuse our notation by setting

$$p_i(\bar{a}) = \langle p_i(a_0), \ldots, p_i(a_{n-1}) \rangle$$

This allows us to define

$$\llbracket \varphi(\bar{a}) \rrbracket = \{ i : A_i \models \varphi(p_i(\bar{a})) \text{ and } i \in I \}$$

Thus, $\llbracket \varphi(\bar{a}) \rrbracket$ is the set of coordinates on which $\varphi(\bar{x})$ is satisfied by the projection of the sequence $\bar{a}$ into the factor structure.

Finally, we define a binary relation $\sim_F$ on $A$, called congruence modulo $F$, so that for all $a, a' \in A$

$$a \sim_F a' \text{ if and only if } \llbracket a \approx a' \rrbracket \in F.$$  

**THEOREM 7.18.** Let $I$ be any set, let $F$ be any filter on $I$, and let $A_i$ be an $\mathcal{L}$-structure for each $i \in I$. Then $\sim_F$ is a congruence relation on $\prod I A_i$.

**Proof.** Let $A = \prod I A_i$. For each $a \in A$, we see that $\llbracket a \approx a \rrbracket = I$. Thus, the reflexivity of $\sim_F$ follows from property (i) of filters. The symmetry of $\sim_F$ follows from the observation that $\llbracket a \approx a' \rrbracket = \llbracket a' \approx a \rrbracket$. To obtain
transitivity, suppose \([ a \approx a' ] \in F\) and \([ a' \approx a'' ] \in F\). By property (iii) of filters

\[
[ a \approx a' ] \cap [ a' \approx a'' ] \in F.
\]

But it is easy to see that \([ a \approx a' ] \cap [ a' \approx a'' ] \subseteq [ a \approx a'' ]\). So, property (ii) of filters yields the transitivity of \(\sim_F\). Therefore, \(\sim_F\) is an equivalence relation on \(A\).

Finally, let \(Q\) be any operation symbol. Let \(r\) denote the rank of \(Q\). Let \(a_0, \ldots, a_{r-1}\) and \(a'_0, \ldots, a'_{r-1}\) be members of \(A\) so that \(a_k \sim_F a'_k\) for all \(k < r\). Using property (iii) of filters repeatedly, we see

\[
[ a_0 \approx a'_0 ] \cap \cdots \cap [ a_{r-1} \approx a'_{r-1} ] \in F.
\]

But this intersection is included in \([ Qa_0 \ldots a_{r-1} \approx Qa'_0 \ldots a'_{r-1} ]\), allowing us to conclude that

\[
[ Qa_0 \ldots a_{r-1} \approx Qa'_0 \ldots a'_{r-1} ] \in F
\]

by property (ii) of filters. Therefore \(\sim_F\) is a congruence relation on \(A\).

To prevent the buildup of notation in dealing with reduced products, we adopt the following conventions:

- For \(a \in A\), we use \(a/F\) in place of \(a/\sim_F\) to denote the congruence class of \(a\) modulo \(\sim_F\).
- For \(\bar{a} = \langle a_0, \ldots, a_{r-1} \rangle\) with \(a_k \in A\) for each \(k < r\), we use \(\bar{a}/F\) in place of \(\langle a_0/F, \ldots, a_{r-1}/F \rangle\).
- We use \(A/F\) to denote the partition of \(A\) into congruence classes modulo \(\sim_F\).

**Definition 7.19.** Let \(\mathcal{L}\) be an elementary language, let \(I\) be any set, let \(F\) be a filter on \(I\), and let \(A_i\) be an \(\mathcal{L}\)-structure for each \(i \in I\). The **reduced product of** \(\langle A_i : i \in I \rangle\) **modulo** \(F\) **is** the \(\mathcal{L}\)-structure \(B\) such that

1. the universe of \(B\) is \(\prod_I A_i/F\);
2. \(Q^B(a_0/F, \ldots, a_{r-1}/F) = Q^A(a_0, \ldots, a_{r-1})/F\), where \(Q\) is any operation symbol of \(\mathcal{L}\), \(r\) is the rank of \(Q\), \(A = \prod_I A_i\), and \(a_0, \ldots, a_{r-1} \in A\), and
3. \(R^B(a_0/F, \ldots, a_{r-1}/F)\) if and only if \([ Ra_0 \ldots a_{r-1} ] \in F\), where \(R\) is any relation symbol of \(\mathcal{L}\), \(r\) is the rank of \(R\), and \(a_0, \ldots, a_{r-1} \in \prod_I A_i\).

The reduced product is denoted by \(\prod_I A_i/F\).

Theorem 7.18 guarantees that condition (ii) is sound: the particular choice of representatives \(a_0, \ldots, a_{r-1}\) from the congruence classes is immaterial. That condition (iii) is also sound is left as an exercise. Observe that the ordinary direct product is a special sort of reduced product which is obtained by setting \(F\) to the principal filter \(\{I\}\). The following theorem is little more than a reformulation of the definition of reduced product.
THEOREM 7.20. Let $\prod_i A_i / F$ be a reduced product and let $\varphi(\bar{x})$ be an atomic formula. For any sequence $\bar{a}$ from $\prod_i A_i$,

$$\prod_i A_i / F \models \varphi(\bar{a}/F) \text{ if and only if } \llbracket \varphi(\bar{a}) \rrbracket \in F.$$ 

Recall that we are trying to find an algebraic construction which will preserve the truth of elementary sentences. The theorem above is a quite provocative step. Reduced products would work if only this theorem were true for all formulas, not just atomic formulas. However, this is too much to hope for. To see why, let $I = \omega$, and for each $i \in I$, let $A_i = \{0, 1\}$. Take $F = \{\omega\}$ and let the language be empty. The resulting reduced product is the same as the direct power of countably many copies of $\{0, 1\}$. It has continuum many elements. On the other hand, the sentence asserting that there are exactly two elements is true in each $A_i$.

Were we to attempt to prove a stronger version of Theorem 7.20, the strategy at hand would be induction on formulas. We would use Theorem 7.20 as the initial step of the induction. The next step would take up negation, and would look something like:

$$\prod_i A_i / F \models \neg \varphi(\bar{a}/F) \text{ if and only if } \prod_i A_i / F \not\models \varphi(\bar{a}/F)$$

if and only if $\llbracket \varphi(\bar{a}) \rrbracket \notin F$

if and only if $\llbracket \neg \varphi(\bar{a}) \rrbracket \in F$

The first equivalence would hold by the definition of satisfaction. The second equivalence would hold by appealing to the induction hypothesis. But the last equivalence might fail. It is easy to see that

$$\llbracket \neg \varphi(\bar{a}) \rrbracket = I - \llbracket \varphi(\bar{a}) \rrbracket,$$

and so if $F$ is proper we can have

$$\llbracket \neg \varphi(\bar{a}) \rrbracket \in F \text{ implies } \llbracket \varphi(\bar{a}) \rrbracket \notin F.$$ 

It is the converse of this implication that need not be true, since it can happen that neither $\llbracket \varphi(\bar{a}) \rrbracket$ nor its complement belongs to $F$. So we are led to the next definition.

DEFINITION 7.21. Let $I$ be any set. $U$ is an ultrafilter on $I$ if and only if $U$ is a proper filter on $I$ such that either $X \in U$ or $(I - X) \in U$, for every $X \subseteq I$. A reduced product modulo an ultrafilter is called an ultraproduct.

It is easy to see that an ultrafilter $U$ is principal if and only if there is $i \in I$ so that $X \in U$ if and only if $i \in X \subseteq I$. Conversely, for each $i \in I$, the collection of all subsets of $I$ which contain $i$ is easily seen to be an ultrafilter. An ultraproduct built using a principal ultrafilter will be isomorphic with the factor structure designated by the element $i$. Leaving aside for the moment the question of whether any nonprincipal ultrafilters exist, we can now present a key result.
THEOREM 7.22 (The Fundamental Theorem for Ultraproducts).
Let \( \prod_i A_i/U \) be an ultraproduct and let \( \varphi(\bar{x}) \) be any formula. For any sequence \( \bar{a} \) from \( \prod_i A_i \) we have
\[
\prod_i A_i/U \models \varphi(\bar{a}/U) \text{ if and only if } \llbracket \varphi(\bar{a}) \rrbracket \in U.
\]

Proof. We proceed by induction on the complexity of \( \varphi(\bar{x}) \). In view of Theorem 7.20 and the remarks following it, only the cases where \( \varphi \) is a conjunction or an existentialization remain.

Case \( \land \land \): \( \varphi(\bar{x}) \) is \( \psi(\bar{x}) \land \land \theta(\bar{x}) \).
This case is easily handled by properties (ii) and (iii) of filters, since
\[
\llbracket \psi(\bar{a}) \land \land \theta(\bar{a}) \rrbracket = \llbracket \psi(\bar{a}) \rrbracket \cap \llbracket \theta(\bar{a}) \rrbracket.
\]

Case \( \exists \): \( \varphi(\bar{x}) \) is \( \exists y \psi(y, \bar{x}) \).
The crucial observation is \( \llbracket \exists y \psi(y, \bar{a}) \rrbracket = \llbracket \psi(b, \bar{a}) \rrbracket \) for some \( b \in \prod_i A_i \).
Indeed, for each \( \bar{a} \in \llbracket \exists y \psi(y, \bar{a}) \rrbracket \), pick \( b_i \in A_i \) so that \( A_i \models \psi(b_i, \bar{a}) \), and for all other \( i \in I \) pick \( b_i \) arbitrarily from \( A_i \). Setting \( b = \langle b_i : i \in I \rangle \), we have
\[
\llbracket \exists y \psi(y, \bar{a}) \rrbracket = \llbracket \psi(b, \bar{a}) \rrbracket.
\]

Now just observe:
\[
\prod_i A_i/U \models \exists y \psi(y, \bar{a}/U) \text{ if and only if } \prod_i A_i/U \models \psi(b/U, \bar{a}/U)
\]
for some \( b \in \prod_i A_i \)
if and only if \( \llbracket \psi(b, \bar{a}) \rrbracket \subseteq U \) for some \( b \in \prod_i A_i \)
if and only if \( \llbracket \exists y \psi(y, \bar{a}) \rrbracket \subseteq U \)
where the last equivalence follows in one direction from property (ii) of filters, and the other direction is the observation above. \( \square \)

COROLLARY 7.23. Let \( \prod_i A_i/U \) be an ultraproduct and let \( \theta \) be a sentence.
\( \prod_i A_i/U \models \theta \) if and only if \( \llbracket \theta \rrbracket \in U \).

So a sentence is true in an ultraproduct if and only if it is true for a large set of the factors (using the ultrafilter to determine largeness). As a consequence, every elementary class is closed with respect to the formation of ultraproducts.

Suppose that \( A_i = B \) for all \( i \in I \) and that \( F \) is a filter on \( I \). We denote \( \prod_i A_i/F \) by \( B^f/F \) and call it a reduced power (or ultrapower if \( F \) is an ultrafilter) of \( B \). In the event that \( C \cong B^f/U \), where \( U \) is an ultrafilter on \( I \), we call \( B \) an ultraroot of \( C \). Clearly, \( B^f/U \models \theta \) if and only if \( \llbracket \theta \rrbracket \in U \) if and only if \( B \models \theta \), for all sentences \( \theta \). This means that every structure is elementarily equivalent to each of its ultrapowers.
COROLLARY 7.24. Let $\mathcal{L}$ be an elementary language and let $\mathcal{K}$ be any class of $\mathcal{L}$-structures. If $\mathcal{K}$ is an elementary class, then $\mathcal{K}$ is closed under isomorphism, the formation of ultraproducts, and the formation of ultraroots.

So we have in hand several algebraic constructions preserving elementary sentences. The corollary above is like the easy direction of Birkhoff’s HSP Theorem. It turns out that the converse of this corollary is also true, although much more difficult to prove, providing us with an analog of the HSP Theorem.

Every structure $A$ can be embedded into any of its nontrivial direct powers $A^I$ via the diagonal map $a \mapsto \{a : i \in I\}$. The composition of the diagonal map with the quotient map results in an elementary embedding of a structure into any of its ultrapowers.

DEFINITION 7.25. Let $A$ be any structure, $I$ be any nonempty set, and $F$ be any ultrafilter on $I$. The canonical embedding of $A$ into $A^I/F$ is the function $\delta : A \rightarrow A^I/F$ such that $\delta(a) = \{a : i \in I\}/F$ for all $a \in A$.

COROLLARY 7.26. If $A^I/U$ is an ultrapower of $A$, then the canonical embedding is an elementary embedding.

Now we have at hand a means to construct elementary extensions of structures. To take advantage of it we have to be able to construct ultrafilters which are not principal.

Let $I$ be a set. A collection $C$ of subsets of $I$ has the finite intersection property provided the intersection of any finite nonempty subcollection of $C$ is nonempty. Every proper filter has the finite intersection property. Moreover, if $C$ has the finite intersection property, then $C$ can be enlarged to a proper filter by first forming $C'$, the collection of all sets arising as intersections of finitely many sets from $C$, and then letting $F$ consist of those sets $X$ such that $Y \subseteq X$ for some $Y \in C'$. $F$ is clearly the smallest filter on $I$ which includes $C$. We say that $C$ generates $F$.

THEOREM 7.27. Let $I$ be any nonempty set, and let $U$ be a proper filter on $I$. The following are equivalent:

i. $U$ is an ultrafilter on $I$.

ii. For all $X, Y \subseteq I$, if $X \cup Y \in U$, then $X \in U$ or $Y \in U$.

iii. For every $X \in U$ and every partition of $X$ into finitely many blocks, exactly one of the blocks belongs to $U$.

iv. For every proper filter $F$ on $I$, if $U \subseteq F$, then $U = F$.

Proof. It is very easy to establish the equivalence of the first three conditions, so we will only argue that conditions (i) and (iv) are equivalent.

(i) $\Rightarrow$ (iv)
Suppose $U$ is an ultrafilter and that $U \subseteq F$ where $F$ is a filter and $F \neq U$. Pick $X \in F - U$. Since $U$ is an ultrafilter we have $(I - X) \in U \subseteq F$. Since $F$ is a filter, we conclude that $\emptyset = X \cap (I - X) \in F$. Hence, $F$ is not proper.

(iv) $\Rightarrow$ (i)
Suppose $U$ is a maximal proper filter. Pick $X \subseteq I$ with $X \notin U$. Since $U$ is maximal, we have that $U \cup \{X\}$ must not generate a proper filter. Hence $U \cup \{X\}$ cannot have the finite intersection property. However, $U$ is closed under finite nonempty intersections, so there must be $Y \in U$ with $Y$ and $X$ disjoint. This entails that $Y \subseteq (I - X) \in U$. Therefore, $U$ is an ultrafilter. ■

Sometimes ultrafilters are called prime filters (and then ultraproducts are called prime products) because condition (ii) in the theorem above asserts that the notion of ultrafilter is dual to the notion of prime ideal in the distributive lattice of all subsets of $I$. This allows us to recite the Prime Ideal Theorem for Distributive Lattices (Theorem 2.60 in Volume 1):

**THEOREM 7.28** (The Ultrafilter Theorem).
Let $I$ be any set.

i. If $N$ is any ideal in the lattice of subsets of $I$ and $F$ is any filter on $I$ such that $N$ and $F$ are disjoint, then there is an ultrafilter $U$ on $I$ disjoint from $N$ such that $F \subseteq U$.

ii. Every proper filter on $I$ is the intersection of all the ultrafilters which include it.

Letting $N$ be the collection whose only element is the empty set, we see that the first part of the Ultrafilter Theorem tells us that every proper filter can be extended to an ultrafilter. Of course, this also means that any collection of subsets of $I$ with the finite intersection property can be extended to an ultrafilter. In particular, if $I$ is infinite the the Fréchet filter on $I$ can be extended to an ultrafilter. Thus there really are nonprincipal ultrafilters.

**THEOREM 7.29.** Let $I$ be any nonempty set, and let $M$ be a collection of nonempty subsets of $I$. If

i. $I \in M$,
ii. For all $X$ and $Y$, if $X \in M$ and $X \subseteq Y \subseteq I$, then $Y \in M$, and
iii. For all $X$ and $Y$, if $X \cup Y \in M$, then $X \in M$ or $Y \in M$,

then there is an ultrafilter $U$ and $I$ such that $U \subseteq M$.

**Proof.** For the purposes of this proof we will call a collection $M$ of nonempty subsets of $I$ prime if $M$ fulfills the three conditions set out in the theorem. We need to show that every prime collection includes an ultrafilter.

First observe that a minimal prime collection is an ultrafilter: Suppose $M$ is minimal, and $X, Y \in M$. We need to prove that $X \cap Y \in M$. Let

$$N = \{Z : Z \subseteq I \text{ and } X \cap Z \in M\}.$$
It is easy to check that $N$ is a prime collection. [For example, let $Z \cup Z' \in N$. Then $(X \cap Z) \cup (X \cap Z') = X \cap (Z \cup Z') \in M$. So $X \cap Z \in M$ or $X \cap Z' \in M$. Thus $Z \in N$ or $Z' \in N$.] It is also clear that $N \subseteq M$. Since $M$ is minimal, we have $N = M$. But $Y \in N$. Hence $X \cap Y \in M$ and $M$ is an ultrafilter.

It remains to argue that every prime collection includes a minimal prime collection. So let $M$ be any prime collection and let

$$\mathcal{K} = \{N : N \subseteq M \text{ and } N \text{ is a prime collection}\}.$$ 

By the Hausdorff Maximality Principle there is a maximal chain $\mathcal{C}$ in $\mathcal{K}$. Let $U = \bigcap \mathcal{C}$. It is a straightforward matter of checking (i)–(iii) to see that $U$ is a prime collection. Since $\mathcal{C}$ is a maximal chain, $U$ must be a minimal prime collection. Thus $U$ is an ultrafilter.

We are now in a position to prove one of the most fundamental and powerful theorems of model theory. One formulation of this theorem read “The lattice of elementary theories is algebraic.”, but such a formulation disguises the real power of the theorem, which resides largely in its usefulness in the construction of models of sets of elementary sentences.

**THEOREM 7.30** (The Compactness Theorem).

*Let $\Sigma$ be a set of elementary sentences. If every finite subset of $\Sigma$ has a model, then $\Sigma$ has a model.*

**Proof.** We make the harmless assumption that $\Sigma$ is closed under conjunction—that is, that $\varphi \land \psi \in \Sigma$ whenever both $\varphi$ and $\psi$ belong to $\Sigma$. Thus we can replace the hypothesis that every finite subset of $\Sigma$ has a model by the simpler condition that every sentence in $\Sigma$ has a model. The idea of the proof is to form an ultraproduct of these models to obtain a model of $\Sigma$. So for each $\varphi \in \Sigma$, pick $A_\varphi$ so that $A_\varphi \models \varphi$.

Our desire is to find an ultrafilter $U$ on $\Sigma$ so that

$$\prod_\Sigma A_\varphi/U \models \Sigma.$$ 

We see by the Fundamental Theorem for Ultraproducts that this is equivalent to

$$\llbracket \theta \rrbracket \in U \text{ for all } \theta \in \Sigma.$$ 

So we only need an ultrafilter extending $\{\llbracket \theta \rrbracket : \theta \in \Sigma\}$. Therefore, it suffices to show that this collection has the finite intersection property. To this end let $\theta_0, \ldots, \theta_{n-1} \in \Sigma$. Since

$$\llbracket \theta_i \rrbracket = \{\varphi \in \Sigma \text{ and } A_\varphi \models \theta_i\},$$

we see that

$$\llbracket \theta_0 \rrbracket \cap \cdots \cap \llbracket \theta_{n-1} \rrbracket = \{\varphi \in \Sigma \text{ and } A_\varphi \models \theta_0 \land \cdots \land \theta_{n-1}\}.$$ 

But $\Sigma$ is closed under conjunction, so $(\theta_0 \land \cdots \land \theta_{n-1}) \in \Sigma$ and we have

$$\llbracket \theta_0 \rrbracket \cap \cdots \cap \llbracket \theta_{n-1} \rrbracket = \llbracket \theta_0 \land \cdots \land \theta_{n-1} \rrbracket.$$
By observing that $\theta \in \{\theta\}$ for all $\theta \in \Sigma$, we conclude that the finite intersection property holds.

Some immediate corollaries of the Compactness Theorem are close enough to it that they could justly be called by the same name. We close this section with three such corollaries.

**COROLLARY 7.31.** Let $\Sigma \cup \{\varphi\}$ be any set of elementary sentences. If $\Sigma \models \varphi$, then there is some finite subset $\Sigma' \subseteq \Sigma$ such that $\Sigma' \models \varphi$.

**COROLLARY 7.32.** Let $\Sigma$ be an elementary language. The lattice of all elementary theories of $\Sigma$ is algebraic.

**COROLLARY 7.33.** Let $\Sigma$ be any set of elementary sentences. If $\Sigma$ is logically equivalent to some finite set of sentences, then $\Sigma$ is logically equivalent to some finite subset of itself.

It follows from this last corollary that if a variety can be axiomatized by a finite set of elementary sentences, then it can also be axiomatized by a finite set of equations.

**Exercises 7.34**

1. 
7.3. Applications of Ultraproducts and Compactness

Like the ultraproduct construction, the Compactness Theorem can be used to build elementary extensions of structures. Suppose we want to construct $A$ such that $B \subseteq A$ for some given $B$. If $B$ is finite and the language is finite, then we saw how to formulate a single elementary sentence that describes $B$ to within isomorphism. This sentence asserts the existence of $|B|$ elements which satisfy various atomic and negated atomic formulas, and then concludes by asserting that no further elements exist. If the last phrase in omitted, what remains would be a sentence such that any model of it would have a substructure isomorphic to $B$. If $B$ is infinite there is more difficulty, since we cannot assert the existence of infinitely many elements in a single sentence, except in very restricted cases. So we cannot talk about how the individual elements of $B$ are interrelated. Fortunately, a simple device circumvents this difficulty: we add a name for each element of $B$ to the language.

**DEFINITION 7.35.** Let $\mathcal{L}$ be an elementary language, and let $A$ be an $\mathcal{L}$-structure. The **diagram language** of $A$ is $\mathcal{L} \cup \{c_a : a \in A\}$ where the $c_a$'s are new distinct constant symbols. The structure $A_A$ is the expansion of $A$ to the diagram language such that $c_a$ is interpreted in $A_A$ as $a$, for each $a \in A$. The **diagram** of $A$ is the set of all atomic and negated atomic sentences of the diagram language true in $A_A$. The **elementary diagram** of $A$ is $\text{Th}_{A_A}$ and it is denoted by $D_A$.

The next corollary can be proven by checking the definitions involved.

**COROLLARY 7.36.** Let $A$ and $B$ be $\mathcal{L}$-structures.

i. $B$ is isomorphic to a substructure of $A$ if and only if $A$ is the reduct to $\mathcal{L}$ of a model of $\Delta_B$.

ii. $B$ can be elementarily embedded into $A$ if and only if $A$ is the reduct to $\mathcal{L}$ of a model of $D_B$.

Using the Compactness Theorem and elementary diagrams, it is easy to produce elementary extensions of infinite structures.

**THEOREM 7.37** (The Upward Löwenheim-Skolem-Tarski Theorem).

*Let $A$ be an $\mathcal{L}$-structure which is infinite. Let $\kappa$ be any cardinal at least as large as $|A|$ and at least as large as the number of $\mathcal{L}$-formulas. $A$ has an elementary extension of cardinality $\kappa$.***

**Proof.** Let $\mathcal{L}'$ be the language obtained by adjoining $\kappa$ new distinct constant symbols, $d_\alpha$ for each $\alpha \in \kappa$, to the diagram language of $A$. Let

$$\Sigma = D_A \cup \{\neg(d_\alpha \approx d_\beta) : \alpha < \beta < \kappa\}.$$ 

Any model of $\Sigma$ will be a model of $D_A$, and so its reduct to $\mathcal{L}$ will be (isomorphic to) an elementary extension of $A$. On the other hand such a model must give
each $d_\alpha$ a distinct interpretation, and so the model must have cardinality at least $\kappa$. Finally, an application of the Downward Löwenheim-Skolem-Tarski Theorem produces an elementary extension of $A$ of the desired cardinality, in view of Theorem 7.11. So it is enough to prove that $\Sigma$ has a model. But any given finite subset $\Sigma'$ of $\Sigma$ involves only finitely many of the $d_\alpha$'s. Hence $A$ can be easily expanded to a model of $\Sigma'$. So by the Compactness Theorem, $\Sigma$ has a model. ■

Some of the limitations of elementary languages should now be apparent. Such familiar structures as $\langle \omega, +, \cdot, \leq \rangle$ and $\langle \mathbb{R}, +, \cdot, \leq \rangle$ cannot be described to within isomorphism by any sets of elementary sentences. The well-known axiom systems of Peano and Dedekind for these structures make references to arbitrary sets of numbers (in the induction and completeness axioms). We see now, from simple cardinality considerations alone, that these references cannot be eliminated in favor of elementary sentences, however cleverly we may attempt to choose them.

Another limitation on the expressive power of elementary sentences is provided by the next theorem. We give two proofs in order to demonstrate how proofs using the Compactness Theorem and proofs using ultraproducts are related.

**THEOREM 7.38.** Let $\Sigma$ be a set of elementary sentences. If $\Sigma$ has arbitrarily large finite models, the $\Sigma$ has an infinite model.

**Proof.**

Using Ultraproducts:

Let $I$ be an infinite set of natural numbers such that for every $n \in I$ there is a model $A_n$ of $\Sigma$ of cardinality $n$. Let $U$ be any ultrafilter on $I$ which extends the Fréchet filter on $I$. Let $B = \prod_I A_n/U$. By the Fundamental Theorem of Ultraproducts, $B \models \Sigma$. To see that $B$ is infinite, let $\delta_k$ be a sentence expressing “there are at least $k$ elements”. Plainly, $\{n : A_n \models \delta_k\}$ is a cofinite subset of $I$ for each natural number $k$. So $\{n : A_n \models \delta_k\} \in U$. But this means $B \models \delta_k$ for all natural numbers $k$. Hence $B$ must be infinite.

Using the Compactness Theorem:

Since $\Sigma$ has arbitrarily large finite models, we see that every finite subset of $\Sigma \cup \{\delta_k : k \in \omega\}$ has a model. So by the Compactness Theorem $\Sigma \cup \{\delta_k : k \in \omega\}$ has a model. Such a model must be infinite. ■

Actually, the ultraproduct proof produces a slightly sharper result: If $\mathcal{K}$ is a class of similar structures such that $\mathcal{K}$ contains arbitrarily large finite structures and $\mathcal{K}$ is closed under ultraproducts, then $\mathcal{K}$ must contain an infinite structure.

**THEOREM 7.39.** Every structure $A$ is isomorphic to a substructure of an ultraproduct of finitely generated substructures of $A$. 
Proof. Let $\Sigma$ be the set of all sentences which are conjunctions of finitely many sentences each belonging to $\Delta_A$. Thus $\Sigma$ and $\Delta_A$ have the same models. Our idea is to follow the proof of the Compactness Theorem to produce a model of $\Sigma$ which is an ultraproduct. Note that each $\varphi \in \Sigma$ mentions only finitely many of the new constants used to name elements of $A$. Let $B_\varphi$ be the substructure of $A$ generated by these elements and let $B'_\varphi$ be some expansion of $B_\varphi$ to the diagram language of $A$ that interprets the new constants correctly in so far as that is possible. As in the proof of the Compactness Theorem, $\{ \llbracket \theta \rrbracket : \theta \in \Sigma \}$ has the finite intersection property. Using the Ultrafilter Theorem, obtain an ultrafilter $U$ on $\Sigma$ extending this set. Then $\prod_{\Sigma} B'_\varphi/U$ is a model of $\Sigma$, and hence of $\Delta_A$. But then $A$ is isomorphic to a substructure of $\prod_{\Sigma} B/U$, since this last ultraproduct is just the reduct to $L$ of $\prod_{\Sigma} B'/U$. ■

The attempt to replace the diagram of $A$ by the elementary diagram of $A$ in the proof just given, leads to the following result.

THEOREM 7.40 (Frayne’s Theorem). $A \equiv B$ if and only if $A$ is isomorphic to an elementary substructure of some ultrapower of $B$.

Proof. If $A \preceq B^I/U$, where $U$ is an ultrafilter on $I$, then $A \equiv B$, since the canonical embedding of $B$ into $B^I/U$ is elementary.

For the converse, let $\Sigma = D_A$. Note that $\Sigma$ is closed under finite conjunctions. Each $\varphi \in \Sigma$ is of the form $\psi(\bar{c}_a)$, where $\psi(\bar{x})$ is a formula in the language of $A$ and $A \models \psi(\bar{a})$. Thus, for $\psi(\bar{c}_a) \in \Sigma$, we have $A \models \exists \bar{x}\psi(\bar{x})$. Since $A \equiv B$, we get $B \models \exists \bar{x}\psi(\bar{x})$. So there is a tuple $\bar{b}$ from $B$ such that $B \models \psi(\bar{b})$. But this is just another way of saying that for each $\varphi \in \Sigma$, we can expand $B$ to $B_\varphi$ so that $B_\varphi \models \varphi$. Now let

$$\llbracket \theta \rrbracket = \{ \varphi : \varphi \in \Sigma \text{ and } B_\varphi \models \varphi \}.$$

A routine proof shows that $\{ \llbracket \theta \rrbracket : \theta \in \Sigma \}$ has the finite intersection property. Let $U$ be an ultrafilter extending this set. By the Fundamental Theorem of Ultraproducts $\prod_{\Sigma} B_\varphi/U \models D_A$. Since $B^\Sigma/U$ is the reduct of this ultraproduct to the language of $A$, we conclude that $A$ is isomorphic to an elementary substructure of $B^\Sigma/U$. ■

While Frayne’s Theorem is a characterization of elementary equivalence, it is one that still refers to the sentences and formulas of elementary languages. The next theorem provides a characterization which relies only on the notions of isomorphism and ultraproduct.

THEOREM 7.41 (The Keisler-Shelah Isomorphism Theorem). $A \equiv B$ if and only if $A^I/U \cong B^I/U$ for some nonempty set $I$ and some ultrafilter $U$ on $I$. 
The proof of this theorem is rather intricate, and it makes use of a number of features of the arithmetic of infinite cardinals. We refer the reader to [?] for a proof.

Let $\mathcal{K}$ be any class of $\mathcal{L}$-structures. We introduce the following notation to supplement $H$, $S$, and $P$:

$A \in P_u\mathcal{K}$ provided $A$ is isomorphic to an ultraproduct of members of $\mathcal{K}$.

$A \in P_f\mathcal{K}$ provided $A$ is isomorphic to a reduced product of members of $\mathcal{K}$.

$A \in \sqrt{\mathcal{K}}$ provided $A$ is isomorphic to an ultraroot of a member of $\mathcal{K}$.

$A \in S^{\sim}\mathcal{K}$ provided $A$ is elementarily embeddable into some member of $\mathcal{K}$.

The next theorem concerns how the operators $P_u$ and $P_f$ are related to the operators $H$, $S$, and $P$. The theorem is stated in a compact form which hides some information. Roughly speaking, all the parts of this theorem are proven by using the natural constructions, and the fact that it is the natural constructions which work is, itself, occasionally useful.

**THEOREM 7.42.** Let $\mathcal{K}$ be a class of $\mathcal{L}$-structures.

1. $\mathcal{K} \subseteq P_u\mathcal{K} \subseteq P_f\mathcal{K} \subseteq HP\mathcal{K}$ and $P\mathcal{K} \subseteq P_f\mathcal{K}$.
2. $P_uS\mathcal{K} \subseteq SP_u\mathcal{K}$ and $P_fS\mathcal{K} \subseteq SP_f\mathcal{K}$.
3. $P_uHP\mathcal{K} \subseteq HP_u\mathcal{K}$ and $P_fHP\mathcal{K} \subseteq HP_f\mathcal{K}$.
4. $P_uP_u\mathcal{K} = P_u\mathcal{K}$ and $P_fP_f\mathcal{K} = P_f\mathcal{K}$.
5. $P_f\mathcal{K} \subseteq SPP_u\mathcal{K}$ and $P_uP\mathcal{K} \subseteq SPP_u\mathcal{K}$.

**Proof.**

(i) All of these inclusions are easy. For instance, $A \cong A^I/U$ if $U$ is a principal ultrafilter on $I$. Hence $\mathcal{K} \subseteq P_u\mathcal{K}$. $
\prod_I A_i \cong \prod_I A_i/F$, if $F = \{I\}$. Hence, $P\mathcal{K} \subseteq P_f\mathcal{K}$.

(ii) Suppose $F$ is any filter on $I$ and $B_i \subseteq A_i$ for all $i \in I$. It is straightforward to check that $\prod_I B_i/F \subseteq \prod_I A_i/F$.

(iii) Suppose $F$ is any filter on $I$ and $h_i : A_i \rightarrow B_i$ for all $i \in I$. Define $h : \prod_I A_i/F \rightarrow \prod B_i/F$ so that

$$h((a_i : i \in I)/F) = (h_i(a_i) : i \in I)/F.$$ 

It is routine to verify that this definition is sound and that $h$ is a homomorphism from $\prod_I A_i/F$ onto $\prod_I B_i/F$.

(iv) This one takes more work. Suppose that $F$ is a filter on $I$ and, for each $i \in I$, that $G_i$ is a filter on $J_i$. For each $i \in I$ and each $j \in J_i$, suppose $A_{ij} \in \mathcal{K}$. Let $B = \prod_J (\prod_{i \in J} A_{ij}/G_i)/F$. So $B$ is a typical member of $P_fP_f\mathcal{K}$. It does not harm the generality of our conclusion to assume that the $J_i$'s are pairwise disjoint. This allows us to simplify the notation, replacing $A_{ij}$ by $A_j$, since $j$ now determines $i$. Let $J = \bigcup_J J_i$, define $D$ as follows: for any $X \subseteq J$

$$X \in D$$

if and only if \{i : X \cap J_i \in G_i\} \in F.

Thus, $D$ consists of those sets $X$ which “meet a large set of coordinate-sets in a large set.” We will prove that $D$ is a filter on $J$ (even an ultrafilter if $F$ and each $G_i$ are ultrafilters) and then that

$$\prod_{j} A_j/D \cong \prod_{i} \left( \prod_{j \in I_i} A_{ij}/G_i \right) / F.$$ 

CLAIM 1. $D$ is a filter on $J$ and, moreover, if $F$ is an ultrafilter and each $G_i$ is an ultrafilter, then $D$ is an ultrafilter.

Proof of Claim 1. If $J \in D$ since $J \cap J_i = J_i \in G_i$ for all $i \in I$. If $X \in D$ and $X \subseteq Y \subseteq J$, then $X \cap J_i \in G_i$ implies $Y \cap J_i \in G_i$, for all $i \in I$. So we obtain $Y \in D$. Now suppose $X, Y \in D$. Then $\{i : X \cap J_i \in G_i\} \subseteq F$ and $\{i : Y \cap J_i \in G_i\} \subseteq F$. So $\{i : X \cap J_i \in G_i \text{ and } Y \cap J_i \in G_i\} \subseteq F$. Consequently, $\{i : (X \cap Y) \cap J_i \in G_i\} \subseteq F$. Therefore, $X \cap Y \in D$. We conclude that $D$ is a filter. To finish the Claim, suppose $F$ and all the $G_i$’s are ultrafilters. Then

$$(X \cup Y) \in D \implies \{i : (X \cup Y) \cap J_i \in G_i\} \subseteq \{i : X \cap J_i \in G_i\} \cup \{i : Y \cap J_i \in G_i\} \subseteq \{i : X \cap J_i \in G_i\} \cup \{i : Y \cap J_i \in G_i\}.$$ 

But each $G_i$ is an ultrafilter, so either $X \cap J_i \in G_i$ or $Y \cap J_i \in G_i$, for each $i$ in the set above. This means that

$$\{i : (X \cap J_i) \cup (Y \cap J_i) \in G_i\} \subseteq \{i : X \cap J_i \in G_i\} \cup \{i : Y \cap J_i \in G_i\}$$

Consequently,

$$(X \cup Y) \in D \implies \{i : X \cap J_i \in G_i\} \cup \{i : Y \cap J_i \in G_i\} \subseteq \{i : X \cap J_i \in G_i\} \cup \{i : Y \cap J_i \in G_i\}.$$ 

Thus, $D$ is an ultrafilter. $\blacksquare$

CLAIM 2. $\prod_{j} A_j/D \cong \prod_{i} \left( \prod_{j \in I_i} A_{ij}/G_i \right) / F$.

Proof of Claim 2. The isomorphism is the natural one: for $a \in \prod_{j} A_j$ define

$$h(a/D) = \langle a^i/G_i : i \in I \rangle / F$$

where $a^i = \langle a_j : j \in J_i \rangle$ for all $i \in I$. Checking the details involves no new ideas. $\blacksquare$

(v) The second inclusion follows from the first with the help of (i) and (iv). So we prove only $P_r \mathcal{K} \subseteq \text{SPP}_a \mathcal{K}$.

Let $\prod_{j} A_j/F$ be a reduced product of structures in $\mathcal{K}$. Let $J = \{U : U$ is an ultrafilter of $I$ extending $F\}$. Let $B_U = \prod_{j} A_j/U$ for each $U \in J$. Both $\sim_F$ and $\sim_U$ are congruences on $\prod_{j} A_j$, and $\sim_F$ is included in $\sim_U$ for all $U \in J$. It follows that the map $h_U$ which sends $a/F \mapsto a/U$ for all $a \in \prod_{j} A_j$,
is a homomorphism from $\prod I A_i / F$ onto $B_U$. Now let $B = \prod J B_U$. We can embed $\prod I A_i / F$ into $B$ using the map $h$ which is defined so that

$$h(a/F) = \langle a/U : U \in J \rangle$$

for all $a \in \prod I A_i$. That $h$ is a homomorphism follows directly from the fact that $h_U$ is a homomorphism for each $U \in J$. To see that $h$ is an embedding we must verify the following claim. (Recall that one-to-one homomorphisms need not be embeddings.)

**CLAIM 3.** For each atomic formula $\varphi(x)$ and each sequence $\bar{a}$ from $\prod I A_i$, if $\prod I A_i / F \models \neg \varphi(\bar{a}/F)$, then $\prod I A_i / U \models \neg \varphi(\bar{a}/U)$, for some $U \in J$.

**Proof of Claim 3.** Since $\prod I A_i / F \models \neg \varphi(\bar{a}/F)$, we know that $\llbracket \varphi(\bar{a}) \rrbracket \notin F$. According to the Ultrafilter Theorem, there is an ultrafilter $U$ extending $F$ such that $\llbracket \varphi(\bar{a}) \rrbracket \notin U$. But for this $U$ we have $\llbracket \neg \varphi(\bar{a}) \rrbracket \in U$. Thus, $\prod I A_i / U \models \neg \varphi(\bar{a}/U)$, by the Fundamental Theorem of Ultraproducts. ■

This completes the proof of the theorem. ■

Further relationships similar to those put forward in the theorem above, but focussing on the ultraroot operator and other constructions can be found among the exercises. An intriguing relationship which is, in essence, the Keisler-Shelah Isomorphism Theorem, is $S_K \subseteq \sqrt{P_u K}$, for any class $K$ of $\Sigma$-structures.

We are now in a position to write down a theorem that does for elementary class what Birkhoff’s HSP Theorem does for varieties.

**THEOREM 7.43** (The Elementary Closure Theorem).

Let $K$ be any class of $\Sigma$-structures, where $\Sigma$ is an elementary language.

i. $\text{Mod Th} K = \{ A : A \equiv B \text{ for some } B \in P_u K \}$.

ii. $\text{Mod Th} K = S^* P_u K$.

iii. $\text{Mod Th} K = \sqrt{P_u K}$.

**Proof.**

(i) The inclusion $\text{Mod Th} K \supseteq \{ A : A \equiv B \text{ for some } B \in P_u K \}$ follows from the Fundamental Theorem for Ultraproducts. For the reverse inclusion suppose $A \in \text{Mod Th} K$. Let $\theta$ be a sentence true in $A$. Hence $\neg \theta \notin \text{Th} K$, since $A \models \text{Th} K$. So pick $B_\theta \in K$ such that $B_\theta \models \theta$. Now let $I = \text{Th} A$. For each $\theta \in I$ we have $\llbracket \theta \rrbracket = \{ \varphi : \varphi \in I \text{ and } B_\varphi \models \theta \}$. As in several earlier proofs the collection

$$\{ \llbracket \theta \rrbracket : \theta \in I \}$$

has the finite intersection property. Let $U$ be an ultrafilter on $I$ extending this collection. That $A \equiv \prod I B_\theta / U$ follows by that Fundamental Theorem for Ultraproducts, since $\llbracket \theta \rrbracket \in U$ for all $\theta \in \text{Th} A = I$.

(ii) This follows from (i) by Frayne’s Theorem and $P_u P_u K = P_u K$. 

(iii) this follows from (i) by the Keisler-Shelah Isomorphism Theorem and $\mathcal{P}_u \mathcal{P}_u \mathcal{K} = \mathcal{P}_u \mathcal{K}$. ■

COROLLARY 7.44. Let $\mathcal{K}$ be any class of $\mathcal{L}$-structures, where $\mathcal{L}$ is any elementary language. The following are equivalent.

i. $\mathcal{K}$ is an elementary class.

ii. $\mathcal{K} = \mathcal{S}^e \mathcal{P}_u \mathcal{K}$.

iii. $\mathcal{K} = \mathcal{S}^e \mathcal{P}_u \mathcal{K}'$, for some class $\mathcal{K}'$ of $\mathcal{L}$-structures.

iv. $\mathcal{K} = \mathcal{P}_u \mathcal{K} = \mathcal{S}^e \mathcal{K}$.

v. $\mathcal{K} = \mathcal{P}_u \mathcal{K}$ and $\mathcal{K}$ is closed with respect to elementary equivalence.

vi. $\mathcal{K} = \mathcal{P}_u \mathcal{K}$.

vii. $\mathcal{K} = \mathcal{P}_u \mathcal{K}'$ for some class $\mathcal{K}'$ of $\mathcal{L}$-structures.

viii. $\mathcal{K} = \mathcal{P}_u \mathcal{K} = \sqrt{\mathcal{K}}$.

COROLLARY 7.45. Let $\mathcal{K}$ be a finite set of finite $\mathcal{L}$-structures, and let $\mathcal{I}\mathcal{K}$ be the class of isomorphic images of members of $\mathcal{K}$. Then $\mathcal{I}\mathcal{K} = \mathcal{P}_u \mathcal{K}$ and $\mathcal{I}\mathcal{K}$ is an elementary class.

Proof. According the Theorem 7.6, $\mathcal{I}\mathcal{K}$ is closed under elementary equivalence. To see that $\mathcal{I}\mathcal{K}$ is also closed with respect to the formation of ultraproducts, let $I$ be any nonempty set, let $U$ be any ultrafilter on $I$, and let $A_i \in \mathcal{K}$ for each $i \in I$. Now partition $I$ by placing $i$ and $j$ in the same block provided $A_i \equiv A_j$. Since $\mathcal{K}$ is finite, we have partitioned $I$ into finitely many blocks. Since $U$ is an ultrafilter on $I$, exactly one of these blocks belongs to $U$. Let us say it is the block to which $j$ belongs. By the Fundamental Theorem of Ultraproducts, $\prod_{A_i \in \mathcal{K}} A_i / U \equiv A_j$, and so the ultraproduct belongs to $\mathcal{I}\mathcal{K}$, since we already know that $\mathcal{I}\mathcal{K}$ is closed with respect to elementary equivalence. ■

Recall that an elementary class $\mathcal{K}$ is said to be finitely axiomatizable provided there is a finite set $\Sigma$ of sentences such that $\mathcal{K} = \text{Mod} \Sigma$. Likewise, an elementary theory is said to be finitely axiomatizable provided it is logically equivalent to some finite set of sentences. For any class $\mathcal{K}$ of $\mathcal{L}$-structures, we use $\mathcal{K}^c$ to denote the complement of $\mathcal{K}$. That is

$$\mathcal{K}^c = \{ A : A \text{ is an } \mathcal{L}\text{-structure and } A \notin \mathcal{K} \}.$$ 

Ultraproducts offer a technique for dealing with finite axiomatizability.

THEOREM 7.46 (The Finite Axiomatizability Theorem).

Let $\mathcal{L}$ be an elementary language, and let $\mathcal{K}$ be any class of $\mathcal{L}$-structures. The following are equivalent.

i. $\mathcal{K}$ is a finitely axiomatizable elementary class.

ii. $\mathcal{K}^c$ is a finitely axiomatizable elementary class.

iii. Both $\mathcal{K}$ and $\mathcal{K}^c$ are elementary classes.

iv. Both $\mathcal{K}$ and $\mathcal{K}^c$ are closed under isomorphism and the formation of ultraproducts.
Proof. (i)$\iff$(ii)$\Rightarrow$(iii)$\Rightarrow$(iv) are all easy, since any finite set of sentences is equivalent to a single sentence, namely the conjunction of the set. The negation of this single sentence axiomatizes the complement of the original class of models. (iv)$\Rightarrow$(iii) follows immediately from the Keisler-Shelah Isomorphism Theorem. To prove (iii)$\Rightarrow$(i) we invoke the Compactness Theorem. Let $\Sigma$ and $\Gamma$ be sets of sentences so that $\mathcal{K} = \text{Mod}\Sigma$ and $\mathcal{K}^c = \text{Mod}\Gamma$. Then $\Sigma \cup \Gamma$ has no model since $\mathcal{K} \cap \mathcal{K}^c$ is empty. By the Compactness Theorem there is a finite set $\Sigma' \subseteq \Sigma$ and a finite set $\Gamma' \subseteq \Gamma$ so that $\Sigma' \cup \Gamma'$ has no model. We claim that $\mathcal{K} = \text{Mod}\Sigma'$. Clearly $\mathcal{K} \subseteq \text{Mod}\Sigma'$. For the reverse inclusion, let $A \in \text{Mod}\Sigma'$. Then $A \notin \text{Mod}\Gamma'$ since $\Sigma' \cup \Gamma'$ has no models. But then $A \notin \text{Mod}\Gamma$, and so $A \notin \mathcal{K}^c$. Therefore $A \in \mathcal{K}$ as desired. Hence $\mathcal{K}$ is finitely axiomatizable.

The most common use of the Finite Axiomatizability Theorem is to prove that certain classes are not finitely axiomatizable. In the typical case, a class $\mathcal{K}$ of structures is given, then certain structures in $\mathcal{K}^c$ are selected and an ultraproduct of these is shown to be in $\mathcal{K}$. As a consequence, $\mathcal{K}$ is not finitely axiomatizable.

EXAMPLE 7.47. The class of fields of characteristic zero is not finitely axiomatizable.

Proof. Let $I$ be the set of prime numbers and let $\mathcal{K}$ be the class of fields of characteristic zero. For each $p \in I$, let $F_p$ be the field with $p$ elements. Let $U$ be any nonprincipal ultrafilter on $I$. Using the Fundamental Theorem for Ultraproducts, it is easy to see that $\prod_I F_p/U$ is a field of characteristic zero. Hence $\mathcal{K}^c$ is not closed under ultraproducts. Consequently, $\mathcal{K}$ is not finitely axiomatizable. As an aside, observe that $\mathcal{K}$ is an elementary class, since the customary field axioms are openly elementary and supplementing this (finite) set of axioms with infinitely many sentences of the form $\neg(1 + \cdots + 1 \approx 0)$ produces a set of elementary sentences axiomatizing the class of fields of characteristic zero.

For varieties of algebras, it is possible to take advantage of a knowledge of subdirectly irreducible algebras in the variety to frame conditions sufficient for finite axiomatizability. Recall that a variety $\mathcal{V}$ is said to be finitely based provided $\mathcal{V}$ is finitely axiomatizable.

THEOREM 7.48. Let $\mathcal{V}$ be a variety. Let $\mathcal{U}$ be a finitely axiomatizable elementary class with $\mathcal{V} \subseteq \mathcal{U}$. Let $\mathcal{W}$ be a class closed with respect to the formation of ultraproducts. If

i. every subdirectly irreducible algebra in $\mathcal{U}$ belongs to $\mathcal{V} \cup \mathcal{W}$, and
ii. $\mathcal{V} \cap \mathcal{W}$ is finitely axiomatizable,

then $\mathcal{V}$ is finitely based.
7.3 Applications of Ultraproducts and Compactness 115

Proof. We will assume that $V$ is not finitely based and that (i) holds, and prove that $V \cap W$ is not finitely axiomatizable.

Since $V$ is not finitely axiomatizable, there must be an algebra $B \in V$ such that $B = \prod_{i} A_i/U$, where $U$ is an ultrafilter on $I$ and $A_i \notin V$ for all $i \in I$. Let $\delta$ be a sentence which axiomatizes $U$. Then $B \models \delta$ and so $\{ \delta \} \in U$ by the Fundamental Theorem for Ultraproducts. Now for each $i \in I$ select an algebra $A'_i$ according to the following stipulations: if $i \in \{ \delta \}$, then take $A'_i = A_i$, otherwise choose $A'_i$ arbitrarily from $U - V$. Let $B' = \prod_{i} A'_i/U$. The Fundamental Theorem for Ultraproducts easily yields $B' \equiv B$. So $B' \in V$, and $A'_i \in U - V$ for all $i \in I$. Since $V$ is a variety, by Birkhoff’s Subdirect Representation Theorem (Theorem 4.44 in Volume 1) select, for each $i \in I$, a subdirectly irreducible factor $C_i$ of $A'_i$ so that $C_i \notin V$. Let $B'' = \prod_{i} C_i/U$. Since $C_i$ is a homomorphic image of $A'_i$ for each $i \in I$, we see that $B''$ is a homomorphic image of $B'$. Consequently, $B'' \in V$. Hence, $B'' \models \delta$, and the Fundamental Theorem for Ultraproducts tells us that $\{ i : C_i \models \delta \} \in U$. As above, pick $D_j$ so that $D_j = C_j$ if $j \in \{ i : C_i \models \delta \}$, and otherwise choose $D_j$ to be any subdirectly irreducible algebra not in $V$. Finally, let $B''' = \prod_{i} D_i/U$. As before, $B''' \equiv B''$ by the Fundamental Theorem for Ultraproducts, and so $B''' \in V$. Now $\{ i : i \in I$ and $D_i \in U \} \in U$, so, in view of condition (i) $\{ i : i \in I$ and $D_i \in V \cup W \} \in U$. However, the $D_i$’s were all chosen to lie outside of $V$. Hence, $\{ i : i \in I$ and $D_i \in W \} \in U$. Since $W$ is closed under ultraproducts, we infer $B''' \in W$. So $B''' \in V \cap W$. Since $B'''$ is an ultraproduct of algebras not in $V \cap W$ (actually, not even in $V$), we conclude that $V \cap W$ is not finitely axiomatizable, by the Finite Axiomatizability Theorem.

The proof above is nonconstructive in the sense that it provides no method to find a finite base for $V$. With some additional restrictions, however, such a constructive method is available. Suppose, in addition to the stipulations of the theorem, all the following hold:

a. The sentence $\theta$ axiomatizes $U$.
b. The sentence $\varphi$ axiomatizes $W$.
c. The sentence $\psi$ axiomatizes $V \cap W$.
d. There is an algorithm for enumerating the equations true in $V$.

Then notice that the sentence $\theta \land (\varphi \Rightarrow \psi)$ is true in $V$. By Birkhoff’s HSP Theorem, $V$ is axiomatized by some set of equations. Therefore, in view of the Compactness Theorem, there is a finite set $\Sigma$ of equations true in $V$ such that $\Sigma \models \theta \land (\varphi \Rightarrow \psi)$. According to the well-known Completeness Theorem of elementary logic, given any finite set of sentences, like $\Sigma$, the elementary sentences which are its logical consequences can be enumerated algorithmically. So if we first begin enumerating the equations true in $V$ to obtain finite sets of equations, and then for each finite set spawn a subprocess to enumerate its elementary logical consequences, we must eventually discover a finite set $\Sigma$ such that $\Sigma \models \theta \land (\varphi \Rightarrow \psi)$. This $\Sigma$ must be a base for $V$, since any subdirectly irreducible model $B$ of $\Sigma$ belongs to $U$ (since $\Sigma \models \theta$). But then $B \in V$ or $B \in W$. In the latter case, we have $B \models \varphi$. But since $B \models \varphi \Rightarrow \psi$, we also conclude that $B \models \psi$ and hence that $B \in V \cap W$. So in any event, $B \in V$. 
Thus every subdirectly irreducible model of $\Sigma$ is in $\mathcal{V}$. It follows from Birkhoff’s Subdirect Representation Theorem, that $\Sigma$ is a base for $\mathcal{V}$.

The restrictions (a) and (c) enumerated above do not substantially alter the sense of Theorem 7.48. Restriction (b) does openly limit the range of applicability, but at least it is a reasonable restriction to impose in order to make the theorem constructive. Restriction (d), at first glance, seems to be rather severe. However, in practice it is frequently met. For example, $\mathcal{V}$ may be described by an infinite but computationally recognizable list of axioms, or $\mathcal{V}$ may be the variety generated by a given finite set of finite algebras. In all cases of this kind, the restriction (d) will be satisfied.

The method described for obtaining a finite basis for $\mathcal{V}$, under the additional restrictions, is effective. However, it depends on algorithms for listing all the logical consequences on a given finite set of sentences. Such algorithms are notoriously slow and expensive in computational resources. It would be useful to have more efficient algorithms for constructing finite basis of varieties, at least in some special cases.

**Exercises 7.49**

1.
7.4 Jónsson’s Lemma for Congruence Distributive Varieties

A variety \( \mathcal{V} \) is **congruence distributive** provided \( \text{Con} \ A \) is a distributive lattice for each \( A \in \mathcal{V} \). Our focus in this section is a very powerful theorem that goes by the name “Jónsson’s Lemma” which concerns subdirectly irreducible algebras in congruence distributive varieties.

By a **subdirect representation** of \( A \) with factors \( A_i \) for \( i \in I \), we mean an embedding \( f : A \to \prod_i A_i \) such that \( p_i \circ f \) maps \( A \) onto \( A_i \) for every \( i \in I \), where \( p_i \) denotes the projection function from \( \prod_i A_i \) onto \( A_i \) for each \( i \in I \). \( A \) is said to be a **subdirect product** of \( \langle A_i : i \in I \rangle \) when there is such a subdirect representation \( f \). The subdirect representation \( f \) is **trivial** provided \( p_i \circ f \) is an isomorphism (\( \Leftrightarrow \) is one-to-one) for some \( i \in I \). \( A \) is **subdirectly irreducible** if and only if every subdirect representation of \( A \) is trivial. Subdirectly irreducible algebras and subdirect products were discussed in some detail in Chapter 4 of Volume 1. For any class \( \mathcal{K} \) of algebras we use \( \mathcal{K}_{si} \) to denote the class comprised of all subdirectly irreducible algebras in \( \mathcal{K} \).

An algebra \( A \) is said to be **finitely subdirectly irreducible** if and only if every subdirect representation of \( A \) over a finite system of (possibly infinite) algebras is trivial. Evidently, every subdirectly irreducible algebra is finitely subdirectly irreducible. It is an easy exercise to prove the \( A \) is finitely subdirectly irreducible if and only if \( A \) has more than one element and the least element \( 0_A \) of \( \text{Con} \ A \) is meet irreducible (i.e. the meet of any two nontrivial congruences of \( A \) must be nontrivial). For any class \( \mathcal{K} \) of algebras we use \( \mathcal{K}_{fsi} \) to denote the class comprised of all finitely subdirectly irreducible algebras in \( \mathcal{K} \).

**Theorem 7.50** (Jónsson’s Lemma).  

i. Suppose \( B \) has a distributive congruence lattice. If \( B \subseteq \prod_i A_i \), and \( \theta \in \text{Con} \ B \) such that \( B/\theta \) is finitely subdirectly irreducible, then there is an ultrafilter \( U \) on \( I \) such that the restriction of \( \sim_U \) to \( B \) is included in \( \theta \).

ii. Let \( \mathcal{V} \) be a congruence distributive variety and suppose \( \mathcal{V} = \mathcal{V}_{K} \). Then \( \mathcal{V}_{fsi} \subseteq \text{HSP}_{u} \mathcal{K} \), and therefore \( \mathcal{V} = \text{P}_{s}\text{HSP}_{u} \mathcal{K} \).

**Proof.** Statement (i) requires us to find a certain ultrafilter \( U \). Every subset \( X \) of \( I \) which belongs to this ultrafilter is required to have the following property:

\[ \star \quad \text{For any } b, b' \in B, \text{ if } X \subseteq \{ b \approx b' \}, \text{ then } b \theta b'. \]

It is just property (\( \star \)) which will ensure that \( \sim_U \) restricted to \( B \) is included in \( \theta \), as desired.

So for every \( X \subseteq I \) define \( \varphi_X \), a binary relation on \( B \), by

\[ b \varphi_X b' \text{ if and only if } X \subseteq \{ b \approx b' \}. \]

Then \( \varphi_X \) is easily seen to be a congruence relation on \( B \). Notice \( X \) satisfies (\( \star \)) if and only if \( \varphi_X \subseteq \theta \).
Now let $M$ be the collection of all subsets $X$ of $I$ which satisfy $(\ast)$. What we need is an ultrafilter included in $M$. Theorem 7.29 supplies the following sufficient conditions for $M$ to include an ultrafilter:

a. $I \in M$.

b. If $X \subseteq Y \subseteq I$ and $X \in M$, then $Y \in M$.

c. If $X \cup Y \in M$, then either $X \in M$ or $Y \in M$.

d. $\emptyset \not\in M$.

Now (a) holds since $\phi_I$ is the identity relation, and (b) holds since $\phi_Y \subseteq \phi_X$, when $X \subseteq Y$. To deduce (c) suppose $X \cup Y \in M$. Observe first that $\phi_{X \cup Y} = \phi_X \cap \phi_Y$. Since $X \cup Y \in M$, we have

$$\theta = \phi_{X \cup Y} \lor \theta = (\phi_X \land \phi_Y) \lor \theta = (\phi_X \lor \theta) \land (\phi_Y \lor \theta)$$

where the last equality holds because $\textbf{Con} B$ is distributive. Since $B/\theta$ is finitely subdirectly irreducible, we know that $\theta$ is meet irreducible. Therefore, either $\phi_X \lor \theta = \theta$ or $\phi_Y \lor \theta = \theta$. But this entails that either $X \in M$ or $Y \in M$. Therefore (c) holds. For (d), just note that $\phi_\emptyset = B \times B$. Since $B/\theta$ is finitely subdirectly irreducible it must have more than one element. Therefore $\theta \neq B \times B$, and so $\emptyset \not\in M$.

Thus, $M$ includes an ultrafilter $U$ by Theorem 7.29. Evidently, $\sim_U$ is included in $\theta$.

To obtain (ii), suppose that $C$ is a finitely subdirectly irreducible member of $V$. By Birkhoff’s HSP Theorem $C \in \text{HSP} K$. So we select a set $I$, and for each $i \in I$ an algebra $A_i \in K$, so that $\prod_i A_i$ has a subalgebra $B$, and $B$ has a congruence $\theta$ such that $C \cong B/\theta$. By statement (i) let $U$ be an ultrafilter on $I$ so that $\sim_U$ restricted to $B$ is included in $\theta$. Let $h$ be the quotient map from $\prod_i A_i$ onto $\prod_i A_i/U$ and let $h'$ be the restriction of $h$ to $B$. Then the kernel of $h'$ is the restriction of $\sim_U$ to $B$. Hence, the kernel of $h'$ is included in $\theta$. So by the Second Isomorphism Theorem (Theorem 4.10 in Volume 1), we have

$$C \cong B/\theta \cong (B/\ker h')/(\theta/\ker h')$$

and

$$B/\ker h' \cong h(B) \subseteq \prod_i A_i/U.$$ 

So $C \in \text{HSP}_u K$, as desired. That $V = \text{P}_u \text{HSP}_u K$ follows from Birkhoff’s Subdirect Representation Theorem.

At first glance Jónsson’s Lemma—which gives $V = \text{P}_u \text{HSP}_u K$ for congruence distributive varieties $V$—seems less strong than Birkhoff’s HSP Theorem—which gives $V = \text{HSP} K$ for all varieties $V$. A closer look reveals something different. Because ultraproducts preserve elementary properties, we typically have much better information about the algebras in $\text{P}_u K$ than we do about the algebras in $\text{P} K$. In many cases, this applies to algebras in $\text{HSP}_u K$ as well, and this last class includes of the subdirectly irreducible algebras in $V$. Another way to compare Jónsson’s Lemma with Birkhoff’s HSP Theorem is to notice
that the HSP Theorem focuses on building algebras free in \( V \) early (in \( SP^K \)), whereas Jónsson’s Lemma is based on getting the subdirectly irreducible algebras early (in \( HSP_u^K \)). Although as a general matter, the description of the free algebras and the description of the subdirectly irreducible algebras in a variety both pose overwhelming difficulties, there are quite a few varieties where the subdirectly irreducible algebras are more accessible than the free algebras. The varieties of lattices are congruence distributive. Jónsson’s Lemma has proven to be a key tool for the investigation of these varieties.

**COROLLARY 7.51.** If \( K \) is a finite set of finite algebras, and \( V_K \) is congruence distributive, then all the subdirectly irreducible algebras in \( V_K \) belong to \( HS^K \). Consequently, there is a finite upper bound on the cardinality of the subdirectly irreducible members of \( V_K \) and \( V_K = P^HS^K \).

**COROLLARY 7.52.** If \( K \) is a finite set of finite algebras and \( V_K \) is congruence distributive, then \( V_K \) has only finitely many subvarieties.

**COROLLARY 7.53.** Let \( V \) be a congruence distributive variety, and suppose that \( A \) and \( B \) are finite subdirectly irreducible members of \( V \). \( A \cong B \) if and only if \( \Theta A = \Theta B \).

This last corollary can be viewed as a variation on the theme set out in Theorem 7.6: Two finite algebras \( A \) and \( B \) are isomorphic if and only if \( Th A = Th B \). In the case of finite subdirectly irreducible algebras belonging to a congruence distributive variety, it is enough that their equational theories are the same.

**COROLLARY 7.54.** If \( V \) is a congruence distributive variety for a finite language, and \( V \) is generated by a finite set of finite algebras, then the class of subdirectly irreducible members of \( V \) is a finitely axiomatizable elementary class, as is the class of finitely subdirectly irreducible members of \( V \).

**COROLLARY 7.55.** Let \( V \) be a congruence distributive variety. The map defined on the set of subvarieties of \( V \) which sends \( U \mapsto U_{si} \) for all subvarieties \( U \) of \( V \) is an embedding of the lattice of subvarieties of \( V \) into a lattice of subclasses of \( V \), where the lattice operations are intersection and union.

The essence of this last corollary is \( (U \lor W)_{si} = U_{si} \cup W_{si} \).

Jónsson’s Lemma, the Fundamental Theorem of Ultraproducts, and the Finite Axiomatizability Theorem can be used jointly to show that certain varieties are not finitely based. The general strategy is

1. use Jónsson’s Lemma to obtain a complete description of the subdirectly irreducible algebras in the variety \( V \), and then
2. find infinitely many subdirectly irreducible algebras in \( V^c \) whose ultraalgebra belongs to \( V \).
EXAMPLE 7.56. Let $L_f$ denote the lattice displayed in Figure 1. Let $V$ be the variety generated by $L_f$. Then $V$ is not finitely based.

Proof. Now $L_f$ is a lattice of length 3 consisting of a largest element, a smallest element, a set of atoms, and a set of coatoms. In $L_f$ every atom is covered by exactly two coatoms and every coatom covers exactly two atoms. These features of $L_f$ can all be expressed with elementary sentences. Certain finite lattices have these properties as well. The lattice $L_{c_3}$ is displayed in Figure 2. This lattice has three atoms and three coatoms. For $n \geq 3$, the lattice $L_{c_n}$ resembles $L_{c_3}$, except it has $n$ atoms and $n$ coatoms. When $n = 2$ the corresponding ordered set fails to be a lattice. For each $n \geq 3$, it is easy to invent an elementary sentence that is true in $L_f$ but fails in $L_{c_n}$. The ordered set that remains after the largest and smallest elements of $L_f$ are removed is sometimes called a (two-away) infinite fence. When the largest and smallest elements of $L_{c_n}$ are removed, the ordered set remaining is called an $n$-crown.

Jónsson’s Lemma tells us that every subdirectly irreducible lattice in $V$ belongs to $HSP_u L_f$. Our ambition is to establish

a. $L_{c_n} \notin HSP_u L_f$,

b. $L_{c_n}$ is subdirectly irreducible, so $L_{c_n} \notin V$, and

c. a nonprincipal ultraproduct of the $L_{c_n}$’s belongs to $SP_u L_f$. 

Figure 1. The lattice $L_f$

Figure 2. The lattice $L_{c_3}$
With these three properties in hand, the Finite Axiomatizability Theorem tells us that \( \mathcal{V} \) is not finitely based.

To establish (a), suppose, to the contrary, that \( \mathcal{L}_{c_n} \in HSP_u \mathcal{L}_f \). Because \( \mathcal{L}_{c_n} \) is finite, there must be a finitely generated sublattice \( F \) of an ultrapower of \( \mathcal{L}_f \) and a congruence \( \theta \) on \( F \) so that \( \mathcal{L}_{c_n} \cong F/\theta \). Now by the Fundamental Theorem for Ultraproducts, every elementary sentence true in \( \mathcal{L}_f \) must be true in every ultrapower of \( \mathcal{L}_f \). Thus an ultrapower of \( \mathcal{L}_f \) resembles \( \mathcal{L}_f \) itself, except instead of one infinite fence across the middle the ultrapower is provided with a (usually large infinite) pairwise disjoint collection of infinite fences. Since \( F \) is a finitely generated sublattice of such an ultrapower, it is not hard to see that it must be finite, and, moreover, it is isomorphic to a sublattice of \( \mathcal{L}_f \) itself. Roughly speaking, since \( F \) must have at least \( 2n + 2 \) elements (with \( n \geq 3 \)), \( F \) looks like what would result by erasing connected stretches of the fence in \( \mathcal{L}_f \) leaving only finitely many connected pieces. Apart from lattices of cardinality 4 or less, the homomorphic images of such lattices are again lattices of the same kind. Hence \( \mathcal{L}_{c_n} \not\in HSP_u \mathcal{L}_f \).

A condition stronger than (b) holds:

\[ \mathcal{L}_{c_n} \text{ is simple, for each } n > 2. \]

To see this, let \( a_0, a_1, \ldots, a_{n-1} \) be a list of the atoms, let \( c_0, c_1, \ldots, c_{n-1} \), be a listing of the coatoms, and suppose matters are arranged so that \( a_i \lor a_{i+1} = c_i \) holds for all \( i \) modulo \( n \). Then \( c_i \land c_{i+1} = a_{i+1} \) will also hold for all \( i \) modulo \( n \). To demonstrate that \( \mathcal{L}_{c_n} \) is simple, it is only necessary to show that \( Cg^{\mathcal{L}_{c_n}}(a_1, c_1) \) and \( Cg^{\mathcal{L}_{c_n}}(c_0, 1) \) both collapse 0 to 1. Indeed, since \( c_0 = a_0 \lor a_1 \) and \( 1 \lor c_1 = 1 \), it follows that \( Cg^{\mathcal{L}_{c_n}}(c_0, 1) \subseteq Cg^{\mathcal{L}_{c_n}}(a_1, c_1) \). So we only need to prove that \((0, 1) \in Cg^{\mathcal{L}_{c_n}}(c_0, 1)\). The following reasoning establishes this. Let \( \theta = Cg^{\mathcal{L}_{c_n}}(c_0, 1) \).

\[
\begin{align*}
0 &= c_0 \land a_2 \theta \land a_2 = a_2 \\
a_0 &= a_0 \lor 0 \theta a_0 \lor a_2 = 1 \\
a_0 &= a_0 \land c_0 \theta 1 \land c_0 = c_0 \\
0 &= a_0 \lor a_1 \theta c_0 \land a_1 = a_1 \\
c_2 &= c_2 \lor 0 \theta c_2 \lor a_1 = 1 \\
0 &= c_0 \land c_2 \theta c_0 \land 1 = c_0 \\
0 \theta c_0 \theta 1
\end{align*}
\]

Finally, to establish (c), let \( I = \{3, 4, 5, \ldots\} \) and let \( U \) be any nonprincipal ultrafilter on \( I \). Let \( \mathcal{L} = \prod_I \mathcal{L}_{c_n}/U \). For each \( n \in I \), let \( \sigma_n \) be a sentence which assert that every atom is in the middle of a fence of length \( 2n + 1 \). For instance, \( \sigma_3 \) would assert, "For every atom \( a_3 \), there are six additional atoms \( a_0, a_1, a_2, a_4, a_5, \) and \( a_6 \) and seven distinct coatoms \( c_0, c_1, c_2, c_3, c_4, c_5, \) and \( c_6 \) so that \( a_i \lor a_{i+1} = c_i \) for all \( i < 7 \) and \( c_i \land c_{i+1} = a_{i+1} \) for all \( i < 7 \)." Since each of these sentences is true in all but finitely many of the lattices \( \mathcal{L}_{c_n} \), the Fundamental Theorem of Ultraproducts tells us that each of these sentences in true in \( \mathcal{L} \). Consequently, \( \mathcal{L} \) might be described as a certain (infinite) antichain
of disjoint two-way infinite fences with a top and bottom element adjoined. In
other words, it looks like an ultrapower of $L_f$. We leave it as an exercise to
show, in fact, that $L = \prod L_{c_n}/U$ is embeddable into $L_f/U$.

In consequence of (a), (b), and (c) we conclude that $\mathcal{V}$ is not finitely based.

\section*{EXAMPLE 7.57.} Let $\mathcal{K}$ be the class of all lattices of length at most 4, and
let $\mathcal{V}$ be the variety generated by $\mathcal{K}$. Then $\mathcal{V}$ is not finitely based.

\textbf{Proof.} It is clear that $\mathcal{K}$ is closed with respect to $H$, $S$, and $P_u$. So by
Jónsson’s Lemma, every subdirectly irreducible algebra in $\mathcal{V}$ belongs to $\mathcal{K}$.

For each natural number $n$, let $L_n$ be the lattice illustrated in Figure 3. $L_n$
has $4n + 16$ elements. Arranged circularly, it resembles the lattice $L_{c_n}$. Once
again, we will have established that $\mathcal{V}$ is not finitely based if we prove
\begin{enumerate}[a.]
\item $L_n \notin \mathcal{K}$,
\item $L_n$ is subdirectly irreducible, so $L_n \notin \mathcal{V}$, and
\item a nonprincipal ultraproduct of the $L_n$’s belongs to $\mathcal{V}$.
\end{enumerate}

\begin{figure}[h]
\centering
\caption{The lattice $L_n$}
\end{figure}

Now (a) is obvious, since the length of $L_n$ is 5 for every natural number $n$.

To prove (b), arguments like the one in the preceding example show that $Con L_n$
is a three-element chain. Hence, $L_n$ is subdirectly irreducible.

To prove (c), let $U$ be any nonprincipal ultrafilter on the set $\omega$ of natural
numbers. Add three constants $a, b,$ and $c$ to the language of lattices. Let $L'_n$
be the expansion of $L_n$ which assigns these new constant symbols as names
to the three elements labeled in Figure 3. Let $L' = \prod L'_n/U$ and let $L$
be the reduct of $L'$ to the language of lattices. Now in the middle of each $L_n$
there is a unique configuration comprised of 10 elements which, with the top
and bottom elements of the lattice adjoined, includes all the chains of length
5 which occur in $L_n$. The fact that there is exactly one configuration like this
can be expressed by an elementary sentence. By the Fundamental Theorem
for Ultraproducts, this sentence is also true for $L$. It follows that $L$ can be
roughly described as being comprised of a top element, a bottom element, and
an (infinite) antichain of components each of which is a "doubled two-way
infinite fence”, with the exception of exactly one component which has our 10-configuration inserted into the middle of a “two-way infinite doubled fence”.

Let $\varphi$ be the principal congruence relation on $L$ which collapses the elements denoted in $L'$ by $a$ and $b$. Let $\psi$ be the principal congruence relation on $L$ which collapses the elements denoted in $L'$ by $b$ and $c$. It is easy to show that $\varphi \cap \psi = 0_L$. It follows that $L$ is embeddable into $L/\varphi \times L/\psi$. As it is also clear that both $L/\varphi$ and $L/\psi$ are lattices of length 4, we conclude that $L \in V$. ■
7.5. Baker’s Finite Basis Theorem

In the last section we saw that certain varieties of lattices, varieties in fact generated by lattices with easily discerned structure, turn out not to be finitely based. This section is devoted to proving Baker’s Finite Basis Theorem, which asserts, in part, that any congruence distributive variety generated by a finite algebra with only finitely many fundamental operations must be finitely based.

In order to take real advantage of Jónsson’s Lemma it is crucial to gain a sharper understanding of the elementary properties of congruence relations. The Congruence Generation Theorem (Theorem 4.19 in Volume 1) provides a characterization of \( \langle a, b \rangle \in \text{Cg}^A(X) \) that seems to just miss being elementary. Here we formulate a slightly sharper version of this theorem.

**DEFINITION 7.58.** Let \( A \) be an algebra. By a **basic translation** on \( A \) we mean a function of the form \( Q^A(a_0, a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_{r-1}) \) where \( Q \) is a basic operation symbol of \( A \), \( r \) is the rank of \( Q \), \( 0 \leq i < r \), and \( a_0, \ldots, a_{r-1} \in A \). For each natural number \( k \), by a \( k \)-translation on \( A \) we mean a composition of \( k \) or fewer basic translations. A translation on \( A \) is simply a \( k \)-translation on \( A \) for some natural number \( k \). The identity map on \( A \) is the only 0-translation. Every basic translation is a 1-translation.

The notion of translation can be relativized to a set \( D \) of terms. A **basic translation relative to** \( D \) is a function of the form \( t^A(a_0, a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_{r-1}) \) where \( a_0, \ldots, a_{r-1} \in A \), and \( t \) is either a basic operation symbol of \( A \), \( r \) is the rank of \( t \), \( 0 \leq i < r \), or \( t(x_0, \ldots, x_{r-1}) \in D \). A \( k \)-translation relative to \( D \) is a composition of \( k \) or fewer basic translations relative to \( D \). Notions defined on the basis of translations can be modified to obtain relativized versions.

Every translation on \( A \) is a unary polynomial on \( A \), but there may easily be unary polynomials which are not translations. For instance, take \( A = \langle \omega, + \rangle \).

For any set \( X \) we use \( \binom{X}{2} \) to denote the collection of all two-element subsets of \( X \). Let \( A \) be an algebra and let \( \{a, b\}, \{c, d\} \in \binom{A}{2} \). We write \( \{a, b\} \rightarrow_k \{c, d\} \) to denote that there is a \( k \)-translation \( p(x) \) on \( A \) such that \( \{c, d\} = \{p(a), p(b)\} \). \( \{a, b\} \rightarrow \{c, d\} \) means that \( \{a, b\} \rightarrow_k \{c, d\} \) for some natural number \( k \).

**THEOREM 7.59 (The Congruence Generation Theorem).** Let \( A \) be an algebra and let \( X \subseteq A^2 \). For all \( a, b \in A \), \( \langle a, b \rangle \in \text{Cg}^A(X) \) if and only if there is some finite sequence \( c_0, \ldots, c_n \in A \) such that

i. \( a = c_0 \) and \( c_n = b \), and

ii. for all \( i < n \) there is \( \{x, y\} \in X \) with \( x \neq y \) so that \( \{x, y\} \rightarrow \{c_i, c_{i+1}\} \).
7.5 Baker’s Finite Basis Theorem

The proof of this more specific statement of the Congruence Generation Theorem is not substantially different from the proofs of Theorem 4.18 and Theorem 4.19 in Volume 1.

Now consider the 4-ary relation \( \{ \langle x, y, z, w \rangle : \langle x, y \rangle \in Cg^A(\langle z, w \rangle) \} \). The Congruence Generation Theorem gives us a definition of this relation which fails to be elementary in two ways. First it quantifies over finite sequences of elements of \( A \). Another way to say this is that the condition is really a countably infinite disjunction with each alternative disjunct asserting the existence of some fixed number of elements satisfying certain further conditions. The other way this definition fails to be elementary is that it quantifies over translations, since \( \{ x, y \} \maps \{ c_i, c_i+1 \} \) really reads “there is a translation \( p(x) \) such that . . . .” In cases where one or both of these difficulties can be overcome, a very useful understanding of congruence relations is usually at hand.

Let \( A \) be an algebra. \( \text{Eqv}^A \) is the lattice of all equivalence relations on \( A \). We know that \( \text{Con}^A \) is a sublattice of \( \text{Eqv}^A \). Define \( \mathcal{O} : \text{Eqv}^A \to \text{Eqv}^A \) by setting \( \mathcal{O} \alpha \) to be the equivalence relation generated by

\[ \{ (p(a), p(b)) : (a, b) \in \alpha \text{ and } p \text{ is a 1-translation on } A \} \]

for every \( \alpha \in \text{Eqv}^A \). \( \mathcal{O} \) depends on the algebra \( A \). When this dependence is important we use \( \mathcal{O}_A \). The theorem below assembles some of the easy consequences of this definition. Its proof is left as an exercise.

**Theorem 7.60.** Let \( A \) be an algebra.

i. \( \alpha \subseteq \mathcal{O} \alpha \) for all \( \alpha \in \text{Eqv}^A \).

ii. \( \mathcal{O} \) is isotone.

iii. \( \mathcal{O} \) is a complete join-endomorphism of \( \text{Eqv}^A \).

iv. \( \mathcal{O}(\alpha \cap \beta) \subseteq \mathcal{O} \alpha \cap \mathcal{O} \beta \), for all \( \alpha, \beta \in \text{Eqv}^A \).

v. The fixed points of \( \mathcal{O} \) are exactly the congruence relations of \( A \).

vi. \( Cg^A \alpha = \bigcup_{k<\omega} \mathcal{O}^k \alpha \) for all \( \alpha \in \text{Eqv}^A \).

vii. For any unary polynomial \( p \) of \( A \), there is a natural number \( d \) such that for every \( \alpha \in \text{Eqv}^A \), if \( \langle a, b \rangle \in \alpha \), then \( \langle p(a), p(b) \rangle \in \mathcal{O}^d \alpha \).

The number least number \( d \) fulfilling (vii) is called the nesting depth of the polynomial \( p \). More generally, we define nesting depth of terms as follows:

i. Every variable and every constant symbol has nesting depth 0.

ii. Every term of the form \( Qt_0 t_1 \ldots t_{r-1} \), where \( Q \) is an operation symbol of rank \( r > 0 \), has nesting depth \( d+1 \) where \( d \) is the maximum of the nesting depths of the terms \( t_0, t_1, \ldots, t_{r-1} \).

Assertion (vi) of the theorem above is just a version of the Congruence Generation Theorem. This leads us to the following definition, which honors A. I. Mal’tsev’s recognition of the importance of the Congruence Generation Theorem.
DEFINITION 7.61. An algebra \( A \) has **Maltsev depth at most** \( m \) provided \( \forall^{m+1} \alpha = \forall^m \alpha \) for all \( \alpha \in \text{Eqv} \ A \). A class \( \mathcal{K} \) of algebras has **Maltsev depth at most** \( m \) provided every algebra in \( \mathcal{K} \) does. An algebra or class of algebras has **bounded Maltsev depth** if and only if it has Maltsev depth at most \( m \), for some natural number \( m \).

In varieties with bounded Maltsev depth one of the causes for the nonelementary character of congruences vanishes.

THEOREM 7.62. Let \( V \) be a nontrivial variety with a finite similarity type, and let \( m \) be any natural number. If \( V \) has Maltsev depth at most \( m \), then there is a finite set \( M \) of equations such that

i. \( V \models M \), and

ii. if \( A \models M \), then \( A \) has Maltsev depth at most \( m \).

Proof. Call a term \( t \) **singular** provided no variable occurs in \( t \) more than once. Every translation on \( A \) has the form \( t^A(x, \bar{a}) \), where \( t(x, \bar{w}) \) is a singular term and \( \bar{a} \) is a tuple from \( A \) of the appropriate length. \( t^A(x, \bar{a}) \) is a \( k \)-translation if \( t \) has nesting depth no more than \( k \); every \( k \)-translation arises in this way from some singular term of nesting depth at most \( k \).

Since there are only finitely many operation symbols, there is a finite set \( S \) of singular terms of nesting depth \( m + 1 \) such that every \( m + 1 \)-translation on any algebra arises, as described in the last paragraph from a term in \( S \). Let \( x, w_0, w_1, \ldots, w_{q-1} \) be a listing of the variables that actually occur in terms belonging to \( S \). We make the harmless assumption that \( S \) is rich enough so that for every algebra \( A \), each \( m + 1 \)-translation of \( A \) has the form \( t^A(x, \bar{a}) \) for some \( t(x, \bar{w}) \in S \).

Now let \( F \) be the algebra in \( V \) free over the set \( \{ h_0, h_1, g_0, \ldots, g_{q-1} \} \), which has \( q + 2 \) elements. Let \( \delta \) be the equivalence relation on \( F \) whose only nontrivial block is \( \{ h_0, h_1 \} \). Let \( s(x, w_0, \ldots, w_{q-1}) \in S \). So

\[
\{ s^F(h_0, \bar{g}), s^F(h_1, \bar{g}) \} \in C^F_{g} \delta.
\]

Now suppose \( V \) has Maltsev depth at most \( m \). Then \( C^F_{g} \delta = \forall^{m} \delta \). This means that there is a finite sequence \( t_0(x, y, \bar{w}), t_1(x, y, \bar{w}), \ldots, t_p(x, y, \bar{w}) \) of terms on \( q + 2 \) variables such that

\[
\begin{align*}
& s^F(h_0, \bar{g}) = t_0^F(h_0, h_1, \bar{g}) \\
& s^F(h_1, \bar{g}) = t_p^F(h_0, h_1, \bar{g}) \\
& \{ h_0, h_1 \} \equiv_{m} \{ t_{i}^F(h_0, h_1, \bar{g}), t_{i+1}^F(h_0, h_1, \bar{g}) \} \text{ for all } i < p
\end{align*}
\]

For each \( i < p \), pick a singular term \( q_i(z, \bar{w}) \) of nesting depth at most \( m \) and, for each \( k < q \) a term \( u_k(x, y, \bar{w}) \) so that

\[
\{ q_{i}^F(h_0, r_0, \ldots, r_{q-1}), q_{i}^F(h_1, r_0, \ldots, r_{q-1}) \} = \{ t_{i}^F(h_0, h_1, \bar{g}), t_{i+1}^F(h_0, h_1, \bar{g}) \}.
\]

where \( r_k = u_k(h_0, h_1, \bar{g}) \) for all \( k < q \).
So the following equations are true in $V$:

$$s(x, \bar{w}) \approx t_0(x, y, \bar{w})$$
$$s(y, \bar{w}) \approx t_p(x, y, \bar{w})$$

and for each $i < p$, either

$$q_i(x, u_0(x, y, \bar{w}), \ldots, u_{q-1}(x, y, \bar{w})) \approx t_i(x, y, \bar{w})$$

or

$$q_i(y, u_0(x, y, \bar{w}), \ldots, u_{q-1}(x, y, \bar{w})) \approx t_{i+1}(x, y, \bar{w})$$

Let $\Sigma$ be the set comprised of all such equations true in $V$ obtained in this way as $s$ runs through $S$. Evidently, $V \models \Sigma$.

Now suppose $A \models \Sigma$. Let $a, b, c, d \in A$ and suppose that $\{a, b\} \equiv_{m+1} \{c, d\}$. So there is a term $s(x, \bar{w}) \in S$ so that $\{c, d\} = \{s^A(a, \bar{e}), s^A(b, \bar{e})\}$. The equations in $\Sigma$ allow us to conclude that $\{c, d\} \in \theta^m$, where $\theta$ is the equivalence relation on $A$ whose only nontrivial block is $\{a, b\}$. Since every element of $\text{Eqv}A$ is a join of atoms, and since $\theta$ is a complete join-endomorphism, we conclude that $A$ has Maltsev depth at most $m$.

It should be clear at this point that if the similarity type is finite, $D$ is some finite set of terms, and $r$ is a natural number, then there is an elementary formula $\gamma(x, y, z, w)$ such that for any algebra $A$ and any $a, b, c, d \in A$,

$$A \models \gamma(a, b, c, d)$$

if and only if $\{a, b\} \equiv_r \{c, d\}$ in $A$ relative to $D$.

We define a second operator $\theta^{-1}$ on $\text{Eqv}A$ by setting $\theta^{-1}\alpha$ to be the largest equivalence relation $\beta \in \text{Eqv}A$ such that $\theta\beta \subseteq \alpha$. Hence,

$$\theta^{-1}\alpha = \bigvee\{\theta : \theta \in \text{Eqv}A \text{ and } \theta\beta \subseteq \alpha\}.$$

Observe, despite the notation, that $\theta^{-1}$ need not be the inverse of $\theta$. The theorem below assembles some the easy consequences of this definition.

**THEOREM 7.63.** Let $A$ be an algebra.

1. $\theta^{-1}\alpha \subseteq \alpha$ for all $\alpha \in \text{Eqv}A$.
2. $\theta^{-1}$ is isotone.
3. $\theta^{-1}$ is a complete meet-endomorphism of $\text{Eqv}A$.
4. $\theta^{-1}\alpha \vee \theta^{-1}\beta \subseteq \theta^{-1}(\alpha \vee \beta)$.
5. The fixed points of $\theta^{-1}$ are exactly the congruence relations of $A$.
6. $\theta\theta \subseteq \alpha$ if and only if $\theta \subseteq \theta^{-1}\alpha$, for all $\theta, \alpha \in \text{Eqv}A$.
7. $\theta^{-1}\alpha = \bigvee\{\theta : \theta \in \text{Eqv}A \text{ and } \theta\theta \subseteq \alpha\}$ for all $\alpha \in \text{Eqv}A$.

There is one more ingredient that we need. For easy reference, we restate here Theorem 4.144 from Volume 1.
THEOREM 7.64. A variety \( \mathcal{V} \) is congruence distributive if and only if there exists a natural number \( n \) and 3-ary terms \( d_0, \ldots, d_n \) such that the following equations hold in \( \mathcal{V} \):

\[
\begin{align*}
i. \quad & d_0(x, y, z) \approx x. \\
ii. \quad & d_i(x, y, x) \approx x \text{ for all } i \leq n. \\
iii. \quad & d_i(x, y, y) \approx d_{i+1}(x, y, y) \text{ for all even } i < n. \\
iv. \quad & d_i(x, x, y) \approx d_{i+1}(x, x, y) \text{ for all odd } i < n. \\
v. \quad & d_n(x, y, z) \approx z
\end{align*}
\]

The system \( d_0, \ldots, d_n \) of terms mentioned in this theorem is called a system of Jónsson terms for \( \mathcal{V} \), and the equations enumerated in the theorem are called Jónsson equations for \( \mathcal{V} \). The Jónsson depth of a congruence distributive variety \( \mathcal{V} \) is the least natural number \( j \) so that \( j \) is the maximum of the nesting depths of some system of Jónsson terms for \( \mathcal{V} \).

The Congruence Generation Theorem involves finite sequences of elements. The Jónsson terms and the Jónsson equations offer a means to manipulate such sequences. Let \( \mathcal{V} \) be a congruence distributive variety with Jónsson terms \( d_0, \ldots, d_n \). Let \( A \in \mathcal{V} \) and let \( a = c_0, c_1, \ldots, c_q = b \) be a sequence of elements of \( A \) (we will say that this is a sequence from \( a \) to \( b \)). The derived sequence has length \((n+1)(q+1)\). It is displayed in the following zigzag array.

\[
\begin{array}{ccccc}
d_0(a, c_0, b) & d_1(a, c_0, b) & d_2(a, c_0, b) & \ldots & d_n(a, c_0, b) \\
d_0(a, c_1, b) & d_1(a, c_1, b) & d_2(a, c_1, b) & \ldots & d_n(a, c_1, b) \\
d_0(a, c_2, b) & d_1(a, c_2, b) & d_2(a, c_2, b) & \ldots & d_n(a, c_2, b) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
d_0(a, c_q, b) & d_1(a, c_q, b) & d_2(a, c_q, b) & \ldots & d_n(a, c_q, b)
\end{array}
\]

The arrows in this array convey the arrangement of the derived sequence. Notice that the arrows point down for columns with even indices and up for columns with odd indices. The heads of the arrows were left off the last column because \( n \) can be either even or odd. In the presence of the Jónsson equations, all the entries in column 0 are \( a \) and all the entries in column \( n \) are \( b \). Also the entries linked by the horizontal arrow \( \rightarrow \) are equal. In the event that the sequence \( a = c_0, c_1, \ldots, c_q = b \) actually arises as in the Congruence Generation Theorem, further conclusions can be drawn concerning entries in the derived sequence which are linked by vertical arrows.

THEOREM 7.65. Let \( \mathcal{V} \) be a congruence distributive variety with Jónsson depth \( j \). For any \( A \in \mathcal{V} \) and any \( \alpha, \beta_0, \ldots, \beta_{p-1} \in \text{Eqv} \ A \) where \( p \) is any positive integer, the following inclusions hold in the lattice \( \text{Eqv} \ A \):

\[
\begin{align*}
i. \quad & \alpha \land (\beta_0 \lor \cdots \lor \beta_{p-1}) \subseteq (\bigvee \alpha \land \bigvee \beta_0) \lor \cdots \lor (\bigvee \alpha \land \bigvee \beta_{p-1}). \\
ii. \quad & \bigvee^{p+1}(\alpha \lor \beta_0 \land \cdots \land (\alpha \lor \beta_{p-1}]) \subseteq \bigvee \alpha \lor (\bigvee \beta_0 \land \cdots \land \bigvee \beta_{p-1}).
\end{align*}
\]
**Proof.** Let \( d_0, \ldots, d_n \) be a system of Jónsson terms for \( V \) that witnesses the Jónsson depth \( j \).

For (i) let \( \langle a, b \rangle \in \alpha \cap (\beta_0 \vee \cdots \vee \beta_{p-1}) \). So \( \langle a, b \rangle \in \alpha \) and we can pick \( c_0, \ldots, c_q \) so that \( a = c_0, c_q = b \), and for all \( i < q \) there is \( k < p \) so that \( \langle c_i, c_{i+1} \rangle \in \beta_k \). Hence for all \( i < q \) and all \( \ell < n \), there is \( k < p \) so that

\[
\delta_\ell(a, c_i, b) \subseteq \delta_\ell(a, c_i, a) \subseteq \delta_\ell(a, c_{i+1}, a) \subseteq \delta_\ell(a, c_{i+1}, b).
\]

Therefore for all \( i < q \) and all \( \ell < n \), there is \( k < p \) so that

\[
\delta_\ell(a, c_i, b) \subseteq \delta_\ell(a, c_i, a) \subseteq \delta_\ell(a, c_{i+1}, a) \subseteq \delta_\ell(a, c_{i+1}, b).
\]

Consequently, in the derived sequence from \( a \) to \( b \) each of the vertical arrows witnesses a relation of the form \( \mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_k \). Hence, \( \langle a, b \rangle \in (\mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_k) \vee \cdots \vee (\mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_{p-1}) \), as desired.

For (ii), let \( \rho_i = \mathcal{O}^i \beta_0 \cap \cdots \cap \mathcal{O}^i \beta_{i-1} \) for \( i < p \). We take \( \rho_0 = 1_A \), the largest equivalence relation on \( A \). To verify the inclusion in (ii), it is enough to show that if \( \mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_k \) for all \( k < p \), then \( \theta \subseteq \mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_k \). To this end, assume \( \mathcal{O}^i \alpha \cap \mathcal{O}^j \beta_k \) for all \( k < p \).

**CLAIM 1.** \( \mathcal{O}^i \alpha \cap (\mathcal{O}^j \beta \cap \mathcal{O}^{k+1} \beta) \subseteq \mathcal{O}^i \alpha \cap (\mathcal{O}^{k+1} \beta \cap \mathcal{O}^{k+1} \beta) \) for all \( k < p - 1 \).

**Proof of Claim 1.** Observe

\[
\mathcal{O}^j \beta \cap \mathcal{O}^{k+1} \beta = (\mathcal{O}^j \beta \cap \mathcal{O}^{k+1} \beta) \cap (\alpha \vee \beta_{k+1}) \text{ by our assumption}
\]

\[
\subseteq [\mathcal{O}^j (\mathcal{O}^j \beta \cap \mathcal{O}^{k+1} \beta) \cap (\alpha \vee \beta_{k+1})] \text{ by (i)}
\]

\[
\subseteq \mathcal{O}^j \alpha \cap [\mathcal{O}^{j+1} \beta \cap \mathcal{O}^j \beta_{k+1}]
\]

\[
\subseteq \mathcal{O}^j \alpha \cap [\mathcal{O}^{j+1} \beta \cap \mathcal{O}^{k+1} \beta_{k+1}]
\]

The claim now follows by joining \( \mathcal{O}^j \alpha \) to both sides. ■

It follows from Claim 1 that

\[
\mathcal{O}^j \alpha \cap (\mathcal{O}^j \beta \cap \mathcal{O}^{k+1} \beta) \subseteq \mathcal{O}^j \alpha \cap (\mathcal{O}^{k+1} \beta \cap \mathcal{O}^{k+1} \beta)
\]

The left side of this inclusion is \( \mathcal{O}^j \alpha \cap \theta \) and the right side is included in \( \mathcal{O}^j \alpha \cap \mathcal{O}^{k+1} \beta \). Since \( \theta \subseteq \mathcal{O}^j \alpha \cap \theta \), we can draw the desired conclusion \( \theta \subseteq \mathcal{O}^j \alpha \cap \mathcal{O}^{k+1} \beta \). ■

**THEOREM 7.66.** If \( V \) is a locally finite congruence distributive variety of Jónsson depth \( j \), and the class of finite subdirectly irreducible members of \( V \) has Maltsev depth at most \( m \), then \( V \) has Maltsev depth at most \( m + j \).

**Proof.** Let \( A \in V \) and let \( a, b \in A \) with \( a \neq b \). Let \( \delta \) be the equivalence relation on \( A \) whose only nontrivial block is \( \{a, b\} \). Our first goal is to prove that \( Cg^A(a, b) = \mathcal{O}^{m+j} \delta \). The nonobvious inclusion is \( Cg^A(a, b) \subseteq \mathcal{O}^{m+j} \delta \). So suppose \( \langle c, d \rangle \in Cg^A(a, b) \). In view of the Congruence Generation Theorem, there
is a finite sequence of translations which bears witness to \( \langle c, d \rangle \in Cg^A(a, b) \). Let \( B \) be the subalgebra of \( A \) generated by \( a, b, c, d \), and all the elements of \( A \) appearing in the translations. So \( B \) is finitely generated. Because \( V \) is locally finite, we have that \( B \) is finite. Abusing notation, we also let \( \delta \) denote the equivalence relation on \( B \) whose only nontrivial block is \( \{ a, b \} \).

Pick finitely many finite subdirectly irreducible algebras \( C_0, \ldots, C_{p-1} \) so that \( B \) is a subdirect product of these algebras. Let \( f_k \) denote the projection map from \( B \) onto \( C_k \) and let \( \beta_k = \ker f_k \) for each \( k < p \). Consider any \( k < p \). Then

\[
f_k(Cg^B(a, b)) = f_k(\bigcup_{\ell < \omega} \mathcal{O}_\ell^\delta) = \bigcup_{\ell < \omega} f_k(\mathcal{O}_\ell^\delta)
= \bigcup_{\ell < \omega} \mathcal{O}_\ell f_k(\delta)
= \mathcal{O}_m f_k(\delta)
= f_k(\mathcal{O}_m^\delta)
\]

where \( f_k(\delta) \) is the smallest equivalence relation on \( C_k \) containing \( \langle f_k(a), f_k(b) \rangle \). Consequently, \( Cg^B(a, b) \subseteq \mathcal{O}_m^\delta \lor \ker f_k = \mathcal{O}_m^\delta \lor \beta_k \). Since this is true for each \( k < p \), we have

\[
Cg^B(a, b) \subseteq (\mathcal{O}_m^\delta \lor \beta_0) \cap \cdots \cap (\mathcal{O}_m^\delta \lor \beta_{p-1}).
\]

Hence, by (ii) of Theorem 7.65, we obtain

\[
Cg^B(a, b) \subseteq \mathcal{O}_{m+j}^\delta \lor (\beta_0 \cap \cdots \cap \beta_{p-1}),
\]

noting that \( Cg^B(a, b) \) and each \( \beta_k \) are congruences and hence fixed by \( \mathcal{O} \) and \( \mathcal{O}^{-1} \). But now notice that \( \beta_0 \cap \cdots \cap \beta_{p-1} = 0_B \) the least equivalence relation on \( B \). It follows that \( Cg^B(a, b) \subseteq \mathcal{O}_{m+j}^\delta \). Therefore, we can find a sequence of \( (m + j) \)-translations of \( B \) which witness that \( \langle c, d \rangle \in Cg^B(a, b) \). These same translation witness that \( \langle c, d \rangle \in Cg^A(a, b) \) equally well.

At this point we know that \( \mathcal{O}_{m+j}^\delta = \mathcal{O}_{m+j+1}^\delta \) for every atom \( \delta \) of \( \text{Eqv} A \). Since every element of \( \text{Eqv} A \) is a join of atoms, and \( \mathcal{O} \) is a complete join-endomorphism of \( \text{Eqv} A \), it follows that \( \mathcal{O}_{m+j+1}^\alpha = \mathcal{O}_{m+j+1}^\alpha \) for every \( \alpha \in \text{Eqv} A \).

**DEFINITION 7.67.** An algebra \( A \) has weak projective radius at most \( r \) provided for any \( \{a, b\}, \{a', b'\} \in (A)^2 \), if there is \( \{c, d\} \in (A)^2 \) so that \( \{a, b\} \triangleright \{c, d\} \) and \( \{a', b'\} \triangleright \{c, d\} \), then there is \( \{c', d'\} \in (A)^2 \) so that \( \{a, b\} \triangleright_r \{c', d'\} \) and \( \{a', b'\} \triangleright_r \{c', d'\} \). A class \( K \) of algebras has weak projective radius at most \( r \) provided every algebra in \( K \) does. An algebra or a class of algebras has a bounded weak projective radius if and only if it has weak projective radii at most \( r \), for some natural number \( r \).
The notion of weak projective radius can be easily relativized to any set $D$ of terms. For congruence distributive varieties, it is useful to relativize it to a set $D$ of Jónsson terms.

The notion of weak projective radius is related to the notion of Maltsev depth. Roughly speaking, while the notion of Maltsev depth grows out of the 4-ary relation:

$$\{⟨x, y, z, w⟩ : ⟨x, y⟩ ∈ Cg^A(z, w)\},$$

the notion of weak projective radius grows out of the 4-ary relation:

$$\{⟨x, y, z, w⟩ : Cg^A(x, y) ∩ Cg^A(y, z) > 0_A\}.$$

This makes the weak projective radius especially germaine to the investigation of $V_{bi}$.

Let $A$ be an algebra and let $\mathcal{C}$ be a collection of two-element subsets of $A$. We say $\{c, d\} ∈ (\frac{1}{2})$ is a **common $k$-bound** of $\mathcal{C}$ provided $\{a, b\} \models_k \{c, d\}$ for all $\{a, b\} ∈ \mathcal{C}$. This notion can be relativized to any set $D$ of terms. Let $⟨c_0, \ldots, c_p⟩$ be a finite sequence of elements of $A$. By a **link** of this sequence we mean a set $\{c_i, c_{i+1}\}$ for some $i < p$.

**THEOREM 7.68 (The Multisequence Lemma).**

Let $\mathcal{V}$ be a congruence distributive variety with a set $D$ of Jónsson terms. Let $A ∈ \mathcal{V}$, let $\{a, b\} ∈ (\frac{1}{2})$, and let $n$ be a natural number. Given $n$ finite sequences from $a$ to $b$, there is a set $\mathcal{C}$ consisting of one link from each sequence so that $\mathcal{C} ∪ \{\{a, b\}\}$ has a common $n$-bound relative to $D$.

**Proof.** The proof is by induction on $n$. The initial step, when $n = 0$, is vacuously true. However, the case $n = 1$ will be used to establish the inductive step, so we prove it next.

**Case** $n = 1$: Let $⟨c_0, \ldots, c_p⟩$ be a sequence of distinct elements from $a$ to $b$. (We leave aside the reduction to sequences of distinct elements.) The derived sequence also runs from $a$ to $b$. Since $a \neq b$, at least one link of the derived sequence must be a two-element set. Let $\{d_ℓ(a, c_i, b), d_ℓ(a, c_{i+1}, b)\}$ be the first such link. This means that $a ∈ \{de(a, c_i, b), de(a, c_{i+1}, b)\}$. The other case being handled in a similar manner, we will suppose $a = t_ℓ(a, c_{i+1}, b)$. Now we have

$$\{c_i, c_{i+1}\} \models_n \{d_ℓ(a, c_i, b), d_ℓ(a, c_{i+1}, b)\}$$

via the relative $1$-translation $d_ℓ(a, x, b)$. But also

$$\{a, b\} \models_n \{de(a, c_i, b), de(a, c_{i+1}, b)\}$$

via the relative $1$-translation $d_ℓ(a, c_i, x)$.

**The Inductive Step:** Here we suppose that we have $n + 1$ sequences connecting $a$ to $b$. Let $⟨c_0, \ldots, c_p⟩$ be one of the sequences. It is harmless to suppose its elements are distinct. By the inductive hypothesis assemble a collection $\mathcal{D}$ consisting of a link from each of the remaining $n$ sequences and pick $\{c, d\}$ as a common relative $n$-bound for $\mathcal{D} ∪ \{\{a, b\}\}$. Let $p(x)$ be a relative $n$-translation taking $\{a, b\}$ to $\{c, d\}$. Setting $p(c_i) = c'_i$ for each $i ≤ p$ we obtain a sequence connecting $c$ to $d$. Using the $n = 1$ case already established there is a link
\{c_i', c_{i+1}'\} and a set \{c', d'\} so that the ⇔ relations in the diagram below all hold:

\[
\begin{align*}
&\{a, b\} \quad ⇔_n \quad \{c, d\} \quad ⇔_{r+1} \quad \{c', d'\} \\
&\{c_i, c_{i+1}\} \quad ⇔_n \quad \{c_i', c_{i+1}'\} \quad ⇔_{r+1} \quad \{c', d'\}
\end{align*}
\]

Thus \(\mathcal{D} \cup \{\{c_i, c_{i+1}\}\}\) is the desired collection of links and \{\{c', d'\}\} is the common \(n + 1\)-bound relative to \(D\).

**THEOREM 7.69.** Let \(\mathcal{V}\) be a congruence distributive variety with a set \(D\) of Jónsson terms, let \(A \in \mathcal{V}\), and let \(a_0, b_0, a_1, b_1 \in A\).

\[
Cg^A(a_0, b_0) \cap Cg^A(a_1, b_1) > 0_A
\]

if and only if

\{\{a_0, b_0\}\} and \{\{a_1, b_1\}\} have a common bound relative to \(D\).

**Proof.** First suppose \((c, d) \in Cg^A(a_0, b_0) \cap Cg^A(a_1, b_1)\) with \(c \neq d\). This means that there are two sequences connecting \(c\) to \(d\) so that the links along the first sequence are each translates of \{\{a_0, b_0\}\}, while the links along the second are each translates of \{\{a_1, b_1\}\}. By Theorem 7.68 we can find a link from each, say \{\{e_0, g_0\}\} and \{\{e_1, g_1\}\} and a relative bound \{\{c', d'\}\} so that the following relations all hold:

\[
\begin{align*}
\{a_0, b_0\} &\quad ⇔ \quad \{e_0, g_0\} \quad ⇔_2 \quad \{c', d'\} \\
\{a_1, b_1\} &\quad ⇔ \quad \{e_1, g_1\} \quad ⇔_2 \quad \{c', d'\}
\end{align*}
\]

Thus \{\{c', d'\}\} can serve as the common relative bound we seek.

For the converse, it is clear that a common relative bound must belong to the intersection of the two principal congruences.

**THEOREM 7.70.** If \(\mathcal{V}\) is a congruence distributive variety with a set \(D\) of Jónsson terms, and \(\mathcal{V}_{a_1}\) has weak projective radius at most \(r\) relative to \(D\), then \(\mathcal{V}\) has weak projective radius at most \(r + 2\) relative to \(D\).

**Proof.** Let \(D = \{d_0, \ldots, d_n\}\) be the set of Jónsson terms. All the translations mentioned in this proof are understood to be translations relative to \(D\). Let \(A \in \mathcal{V}\), and suppose that \{\{a, b\}, \{a', b'\}, \{c, d\}\} is in \(\binom{\mathcal{A}}{2}\) with \{\{a, b\} \quad ⇔ \quad \{c, d\}\} and \{\{a', b'\} \quad ⇔ \quad \{c, d\}\}. We must find \{\{c', d'\}\} in \(\binom{\mathcal{A}}{2}\) so that \{\{a, b\} \quad ⇔_{r+2} \quad \{c', d'\}\} and \{\{a', b'\} \quad ⇔_{r+2} \quad \{c', d'\}\}.

Pick a subdirectly irreducible algebra \(B\) and a homomorphism \(f\) from \(A\) onto \(B\) so that \(f(c) \neq f(d)\). It follows that \{\{f(a), f(b)\}, \{f(a'), f(b')\}\} is in \(\binom{\mathcal{B}}{2}\) and that

\[
\begin{align*}
\{f(a), f(b)\} &\quad ⇔ \quad \{f(c), f(d)\} \quad \text{and} \quad \{f(a'), f(b')\} \quad ⇔ \quad \{f(c), f(d)\}.
\end{align*}
\]

Now \(B\) has relative weak projective radius no more than \(r\), so pick \{\{u, v\}\} in \(\binom{\mathcal{B}}{2}\) so that

\[
\begin{align*}
\{f(a), f(b)\} &\quad ⇔_r \quad \{u, v\} \quad \text{and} \quad \{f(a'), f(b')\} \quad ⇔_r \quad \{u, v\}.
\end{align*}
\]
This allows us to find $c_0, d_0, c_1, d_1 \in A$ such that $f(c_0) = u = f(c_1), f(d_0) = v = f(d_1)$, and
\[
\{a, b\} \models_r \{c_0, d_0\} \text{ and } \{a', b'\} \models_r \{c_1, d_1\}.
\]

Now we contend that there is $\ell < n$ so that $d\ell(c_0, c_1, d_0)$ and $d\ell(c_0, d_1, d_0)$ are distinct elements of $A$. Were this not so, then in $B$ we would have $d\ell(u, u, v) = d\ell(u, v, v)$ for all $\ell < n$. These equations, together with the Jónsson equations, easily entail that $u = v$, a contradiction. So let $c^* = d\ell(c_0, c_1, d_0)$ and $d^* = d\ell(c_0, d_1, d_0)$ with $\ell$ chosen so that $c^* \neq d^*$. From the terms used to define $c^*$ and $d^*$ and the Jónsson equations, we see that
\[
\{c_1, d_1\} \models_{c^*} \{c^*, d^*\}, \quad \{c_0, d_0\} \models_{c^*} \{c^*, c_0\}, \text{ and } \{c_0, d_0\} \models_{c^*} \{c^*, d^*\}.
\]

By Theorem 7.68 applied to the single sequence $(c^*, c_0, d^*)$, we must have $(c^*, d^*) \in \binom{A}{2}$ so that either
\[
(c^*, d^*) \models_{1} \{c', d'\} \text{ and } (c^*, c_0) \models_{1} \{c', d'\}
\]
or
\[
(c^*, d^*) \models_{1} \{c', d'\} \text{ and } \{c_0, d^*\} \models_{1} \{c', d'\}.
\]
In either case, it follows that $(c', d') \in \binom{A}{2}$ so that $(a, b) \models_{r+2} \{c', d'\}$ and $(a', b') \models_{r+2} \{c', d'\}$, as desired.

**THEOREM 7.71.** Let $\mathcal{W}$ be an elementary subclass of an congruence distributive variety $\mathcal{V}$ of finite type and let $D$ be a set of Jónsson terms for $\mathcal{V}$. $\mathcal{W}_{\text{fin}}$ has a bounded weak projective radius relative to $D$ if and only if $\mathcal{W}_{\text{fin}}$ is an elementary class.

**Proof.** Suppose first that $r$ is the weak projective radius of $\mathcal{W}_{\text{fin}}$ relative to $D$. Let $\gamma(x, y, z, w)$ be an elementary formula defining \"$\{x, y\} \models_r \{z, w\}$ relative to $D\"$ in every algebra. Then $\mathcal{W}_{\text{fin}}$ is an elementary class since $\mathcal{W}$ is elementary and the sentence
\[
\forall x, y, x', y' \exists z, w[\neg(\approx x \approx y) \land \neg(\approx x' \approx y') \Rightarrow (\gamma(x, y, z, w) \land \gamma(x', y', z, w))]
\]
characterizes $\mathcal{W}_{\text{fin}}$ has a subclass of $\mathcal{W}$, by Theorem 7.69.

For the converse, suppose that $\mathcal{W}_{\text{fin}}$ does not have a bounded weak projective radius relative to $D$. Then for every $r > 0$, there is $B_r \in \mathcal{W}_{\text{fin}}$ with $a_r, b_r, a'_r, b'_r \in B_r$ so that $\{a_r, b_r\}$ and $\{a'_r, b'_r\}$ are not relatively $r$-bounded. Now expand the language by adding four new constant symbols $a, b, a'$, and $b'$. Then \"$\{a, b\}$ and $\{a', b'\}$ are not relatively $r$-bounded\" can be expressed as an elementary sentence $\sigma_r$. Each of the algebras $B_r$ can be expanded to $B'_r$, an algebra for the expanded language, so that $B'_r \models \sigma_r$ for each $r$. Let $I$ be the set of positive integers and let $U$ be any nonprincipal ultrafilter on $I$. Let $B' = \prod_f B'_f/U$. Since $\sigma_r \Rightarrow \sigma_s$ is logically valid whenever $r \geq s$, it follows from the Fundamental Theorem of Ultraproducts, that $B' \models \sigma_r$ for all $r > 0$. Thus, in $B'$ the elements denoted by $a, b, a'$, and $b'$ establish two two-element sets without any common relative bound. By Theorem 7.69 $B'$ is not finitely subdirectly irreducible. Let $B$ be the reduct of $B'$ to the original language.
Thus, \( B = \prod_{I} B_{r}/U \) is not finitely subdirectly irreducible. Consequently, \( W_{\text{fsi}} \) is not closed under ultraproducts, and so it is not an elementary class.  

**Corollary 7.72.** If \( V \) is a congruence distributive variety of finite type, and \( V_{\text{fsi}} \) is an elementary class, then there is a formula \( \varphi(x, y, z, w) \) such that for all \( A \in V \) and all \( a, b, c, d \in A \)

\[
A \models \varphi(a, b, c, s) \text{ if and only if } Cg_{A}(a, b) \cap Cg_{A}(c, d) > 0_{A}.
\]

**Theorem 7.73 (Baker’s Finite Basis Theorem).**

Let \( V \) be a congruence distributive variety in a finite similarity type.  

i. If \( V_{\text{fsi}} \) is a finitely axiomatizable elementary class, then \( V \) is finitely based.  

ii. If \( V \) is generated by finitely many finite algebras, then \( V \) is finitely based.

**Proof.** Let \( D = \{d_{0}, d_{1}, \ldots, d_{n}\} \) be a set of Jónsson terms for \( V \) and take \( j \) to bethe maximum nesting depth of the terms in \( D \). Let \( \Delta \) be the corresponding set of Jónsson equations.

Our strategy for proving (i) is to invoke Theorem 7.48. To this end, we need a finitely axiomatizable class \( U \) and a class \( W \) closed under ultraproducts so that

a. \( V \subseteq U \) and every subdirectly irreducible algebra in \( U \) belongs to \( V \cup W \), and

b. \( V \cap W \) is finitely axiomatizable.

Now since \( V_{\text{fsi}} \) is elementary, we know that there is a natural number \( r \) so that \( V_{\text{fsi}} \) has weak projective radius no more than \( r \) relative to \( D \). So \( V \) has weak projective radius no more than \( r + 2 \) relative to \( D \). Let \( \rho_{r+2} \) be the sentence which asserts that the weak projective radius is no more than \( r + 2 \) relative to \( D \).

Therefore \( V \models \rho_{r+2} \). We let \( U \) be the class of all models of \( \Delta \cup \{\rho_{r+2}\} \). Plainly, \( V \subseteq U \) and \( U \) is finitely axiomatizable. Let \( W = U_{\text{fsi}} \). Since \( U \) is elementary and \( U_{\text{fsi}} \) has weak projective radius no more than \( r + 2 \) relative to \( D \), we know that \( W(= U_{\text{fsi}}) \) is elementary, and therefore closed under ultraproducts. (We even know a finite set of axioms for \( W \).) Evidently, (a) holds for this choice of \( U \) and \( W \). Since \( V \cap W = V_{\text{fsi}} \), and \( V_{\text{fsi}} \) is finitely axiomatizable by hypothesis, (b) is also fulfilled. Consequently, \( V \) is finitely based.

The assertion (ii) is an immediate consequence of (i) by way of Jónsson’s Lemma and Corollary 7.7. We provide here a second proof of (ii) which has the virtue of giving a more explicit construction of a finite equational base for \( V \). Roughly speaking, the role played by the weak projective radius in the proof completed above, will now be played by Maltsev depth. This entails a number of other changes in the proof. The proof could still be modelled after Theorem 7.48, but in the end there will be no need to invoke that theorem.

So let \( \mathcal{K} \) be a finite set of finite algebras which generates \( V \). let \( K \) be the maximum of the cardinalities of the algebras in \( \mathcal{K} \). According to Jónsson’s Lemma, all the subdirectly irreducible algebras in \( V \) actually belong to \( \text{HSX} \).
We can take $\mathcal{W}$ to be the class of all algebras of cardinality no more than $K$. Then $\mathcal{V} \cap \mathcal{W}$ is finitely axiomatizable since it is, up to isomorphism, just a finite set of finite algebras. The construction of the class $\mathcal{U}$ is more difficult.

Now $\mathcal{HSK}$ has Maltsev depth at most $m$, for some natural number $m$ which can be calculated by examining the finite set $\mathcal{X}$ of finite algebras. Thus $\mathcal{V}$ has Maltsev depth at most $m + j$. By Theorem 7.62 we can find a finite set $M_{m+j}$ of equations true in $\mathcal{V}$ so that, if $A \models M_{m+j}$, then $A$ has Maltsev depth at most $m + j$. The last set we need is a finite set $\Gamma_K$ of equations true in $\mathcal{V}$ such that any subdirectly irreducible model of $\Delta \cup M_{m+j} \cup \Gamma_K$ must have cardinality no more than $K$. Had we such a set, then we could let $\mathcal{U}$ to be the variety based on $\Delta \cup M_{m+j} \cup \Gamma_K$.

To build the set $\Gamma_K$ we introduce, for every $k \leq n$, a new binary operation $*_{k}$ on $A^2$ for every algebra $A$. For $(a, b), (c, d) \in A^2$ we set

$$\langle a, b \rangle *_{k} \langle c, d \rangle = \langle d^A_k(a, c, b), d^A_k(a, d, b) \rangle.$$

**Lemma 7.74.** Let $A \models \Delta \cup M_{m+j}$ and let $\langle a_0, b_0 \rangle, \ldots, \langle a_p, b_p \rangle$ be any finite sequence of elements of $A^2$. Define $J_p$ so that for all $\langle r, s \rangle \in A^2$

$$\langle r, s \rangle \in J_p$$

if and only if

$$\langle r, s \rangle = (\ldots ((\langle t_0(a_0), t_0(b_0) \rangle *_{k_1} \langle t_1(a_1), t_1(b_1) \rangle ) *_{k_2} \ldots ) *_{k_p} \langle t_p(a_p), t_p(b_p) \rangle)$$

for some $k_1, \ldots, k_p \leq n$ and some $(m+2j)$-translations $t_0(x), \ldots, t_p(x)$

Then

a. if $a_i = b_i$ some some $i \leq p$, then $r = s$ for all $\langle r, s \rangle \in J_p$, and

b. $C^A_{g}(a_0, b_0) \cap \cdots \cap C^A_{g}(a_p, b_p) = \bigvee_{\langle r, s \rangle \in J_p} C^A_{g}(r, s)$.

**Proof of Lemma 7.74.** For (a) observe that

$$\langle a, a \rangle *_{k} \langle e, f \rangle = \langle d^A_k(a, e, a), d^A_k(a, f, a) \rangle = \langle a, a \rangle$$

by the Jónsson equations. Observe also

$$\langle a, b \rangle *_{k} \langle c, e \rangle = \langle d^A_k(a, e, b), d^A_k(a, e, d) \rangle$$

is a diagonal pair. It follows that all pairs in $J_p$ must be diagonal.

To establish (b), we define a schema of sets which resemble $J_p$. For each $\langle c, d \rangle \in C^A_{g}(a_0, b_0) \cap \cdots \cap C^A_{g}(a_p, b_p)$, let $H_p(c, d)$ be defined so that for all $\langle r, s \rangle \in A^2$

$$\langle r, s \rangle \in H_p(c, d)$$

if and only if

$$\langle r, s \rangle = (\ldots ((\langle c, d \rangle *_{k_0} \langle t_0(a_0), t_0(b_0) \rangle ) *_{k_2} \ldots ) *_{k_p} \langle t_p(a_p), t_p(b_p) \rangle)$$

for some $k_0, \ldots, k_p \leq n$ and some $(m+j)$-translations $t_0(x), \ldots, t_p(x)$

**Claim 1.** $C^A_{g}(c, d) = \bigvee_{\langle r, s \rangle \in H_p(c, d)} C^A_{g}(r, s)$. 

Proof of Claim 1. Our proof is by induction on \( p \).

For the initial step, we have \( \langle c, d \rangle \in Cg^A(a_0, b_0) \). Since \( A \) has Maltsev depth at most \( m + j \) there is some sequence \( \langle e_0, \ldots, e_q \rangle \) from \( c \) to \( d \) so that

\[
\{ a_0, b_0 \} \cap_{m+j} \{ e_i, e_{i+1} \} \text{ for all } i \leq q.
\]

Now the links of the derived sequence of \( \langle e_0, \ldots, e_q \rangle \) are just those pairs \( \langle c, d \rangle \ast_k \langle e_i, e_{i+1} \rangle \), as \( k \) runs from 0 to \( n \). Since every element of the derived sequence has the form \( d_k^A(c, e_i, d) \), we know that \( Cg^A(c, d) \) collapses the whole sequence, in view of Jónsson's equations. On the other hand, collapsing each link of the derived sequence individually certainly collapses \( c \) to \( d \). Thus, \( Cg^A(c, d) = \bigvee_{\langle r, s \rangle \in H_{r,s}(c, d)} Cg^A(r, s) \), since \( e_i = t_i(a_0) \) or \( e_i = t_i(b_0) \) for some suitably chosen \((m+j)\)-translation \( t_i \).

For the inductive step, we assume that

\[
Cg^A(c, d) = \bigvee_{\langle r', s' \rangle \in H_{r,s}(c, d)} Cg^A(r', s')
\]

and that \( \langle c, d \rangle \in Cg^A(a_{\ell+1}, b_{\ell+1}) \), and argue that

\[
Cg^A(c, d) = \bigvee_{\langle r, s \rangle \in H_{r,s}(c, d)} Cg^A(r, s).
\]

Consider a pair \( \langle r', s' \rangle \in H_{\ell}(c, d) \). Then \( \langle r', s' \rangle \in Cg^A(c, d) \subseteq Cg^A(a_{\ell+1}, b_{\ell+1}) \).

Now applying the initial step with \( \langle r', s' \rangle \) in place of \( \langle c, d \rangle \) we obtain

\[
Cg^A(r', s') = \bigvee_{\langle r'', s'' \rangle \in H''} Cg^A(r'', s''),
\]

where the members of \( H' \) have the form \( \langle r', s' \rangle \ast_k (t(a_{\ell+1}), t(b_{\ell+1})) \) for \( k \leq n \) and \( t \) some \( m+j \)-translation. In this way the induction is complete. \( \blacksquare \)

Now observe

\[
\langle c, f \rangle \ast_i \langle g, h \rangle = (d_i^A(c, e, f), d_i^A(e, h, f)) \in Cg^A(c, f) \cap Cg^A(g, h)
\]

by the Jónsson equations. This observation supports an easy inductive proof that \( J_p \subseteq Cg^A(a_0, b_0) \cap \cdots \cap Cg^A(a_p, b_p) \).

Let \( \theta \) denote \( Cg^A(a_0, b_0) \cap \cdots \cap Cg^A(a_p, b_p) \). We can now establish (b) as follows:

\[
Cg^A(a_0, b_0) \cap \cdots \cap Cg^A(a_p, b_p) = \bigvee_{\langle c, d \rangle \in \theta} Cg^A(c, d) = \bigvee_{\langle c, d \rangle \in \theta} \bigvee_{\langle r, s \rangle \in H_{r,s}(c, d)} Cg^A(r, s)
\]

\[
\subseteq \bigvee_{\langle r', s' \rangle \in J_p} Cg^A(r', s') \subseteq Cg^A(a_0, b_0) \cap \cdots \cap Cg^A(a_p, b_p).
\]

The second equality is just Claim 1. The first inclusion is just a consequence of

\[
\langle c, d \rangle \ast_k \langle t_0(a_0), t_0(b_0) \rangle = \langle d_k^A(c, t_0(a_0), d), d_k^A(c, t_0(b_0), d) \rangle.
\]
For each \(i < \ell\) occurring in \(x, w\) constructed in the proof of Theorem 7.62. Say CLAIM 2.

\(\text{r equation uses each of the equations } e^K\)

Now form any left associated \(\text{K equation in } \Gamma\). Since \(\text{A}\) has at most \(K\) elements, we know that \(a_i = a_{\ell}\) for some \(i < \ell \leq K\). Thus the equation \(e_{i,\ell}\) holds under this assignment. By part (a) of Lemma 7.74, \(r \approx s\) must hold under this assignment as well. Since the assignment was arbitrary, \(r \approx s\) must be true in \(A\).

Let \(\Gamma_K\) be the set of all such equations \(r \approx s\).

Now suppose that \(A\) is subdirectly irreducible, that \(A \models \Delta \cup M_{m+j}\), and that \(A\) has \(K+1\) distinct elements \(a_0, a_1, \ldots, a_K\). We argue that \(A\) is not a model of \(\Gamma_K\). By part (b) of Lemma 7.74, setting \(p = K(K+1)/2\), we have

\[
\bigcap_{i < \ell \leq p} Cg^A(a_i, a_{\ell}) = \bigvee_{\langle r, s \rangle \in J_p} Cg^A(r, s).
\]

Since \(A\) is subdirectly irreducible, the left side of the equation above is not \(0_A\). So there must be \(\langle r, s \rangle \in J_p\) with \(r \neq s\). But this entails that there is an equation in \(\Gamma_K\) which fails to hold in \(A\).

Let \(\mathcal{U}\) be the variety of all models of the set \(\Delta \cup M_{m+j} \cup \Gamma_K\) of equations. Evidently, \(V \subseteq \mathcal{U}\) and every subdirectly irreducible algebra in \(\mathcal{U}\) has cardinality no greater than \(K\). Thus, while we could invoke Theorem 7.48 at this stage, it is more direct to observe that, up to isomorphism, there are only finitely many subdirectly irreducible algebras \(B\) that belong to \(\mathcal{U}\) but not to \(V\). Let \(\Psi\) be a set consisting of one equation for each such \(B\) which is true in \(V\) but fails in \(B\). Then \(\Delta \cup M_{m+j} \cup \Gamma_K \cup \Psi\) is a finite equational base for \(V\).
Consider the task of actually finding a finite equational base for the congruence distributive variety $\mathcal{V}$ in Baker’s Finite Basis Theorem, given an elementary axiom for $\mathcal{V}_{\text{fsi}}$. The remarks following Theorem 7.48 indicate how that theorem may be cast in a constructive manner, provided that finite axiom sets for $\mathcal{U}$, $\mathcal{W}$, and $\mathcal{V} \cap \mathcal{W}$ are available. The ingredient that is missing is the value $r$ so that $\mathcal{V}_{\text{fsi}}$ has weak projective radius no more than $r$ relative to $D$. We know that one of the sentences $\rho_r$ is true in $\mathcal{V}_{\text{fsi}}$. By enumerating the sentences true in $\mathcal{V}_{\text{fsi}}$ (using the given finite set of axioms for $\mathcal{V}_{\text{fsi}}$), we will eventually find such a value for $r$. While this way to produce a finite base for $\mathcal{V}$ is effective, it is costly since algorithms for enumerating all the logical consequences of certain finite sets of sentences must be invoked at several points. Such algorithms are known to be necessarily costly in computational resources.

The second proof offered for part (ii) of Baker’s Finite Basis Theorem avoids those costly enumerations of logical consequences. However, that construction depended on producing the set $M_{m+j}$. That construction depends on the manipulation of the free algebra in $\mathcal{V}$ on a number of generators depending linearly on $m + j$. Computing in this free algebra could also be costly. Alternatively, one may search among the candidates for $M_{m+j}$ until a usable one is found.

While the second proof for part (ii) of Baker’s Finite Basis Theorem has advantages, there seems to be no obvious way to adapt it to prove part (i).
Sections Yet to Come

7.6. Universal Classes and Quasivarieties

7.7. The Fefermann-Vaught Theorem

7.8. Boolean Powers

7.9. Boolean Products