Part D

Second-Order Logic

This part of the book is devoted to the study of second-order logic and some of its applications. We discuss the two chapters in the opposite order from that in which they appear.

Chapter XIII is about monadic second-order logic, logic that allows quantification over arbitrary subsets of the domain, but not over arbitrary relations or functions. While this does not make any difference on structures like the natural numbers with plus and times, where sequences can be coded by numbers, it turns out to make an enormous difference in more algebraic settings. In these cases, monadic second-order logic is a good source of theories that are both highly expressive yet manageable. Section 2 illustrates the uses of finite automata and games in the proof of decidability results. It begins with a simple case, the monadic theory of finite chains, which it works out in complete detail, and shows how the method generalizes to a number of results, including one of the most famous, Rabin's theorem on the decidability of the monadic second-order theory of two successor functions. In Section 3 more model-theoretic methods, generalized products, are used to prove some of the same and related results. Some undecidability results are also presented. Proofs of these have to be novel, since we are dealing with theories where one cannot interpret first-order arithmetic.

If we think of monadic second-order logic as the part of second-order logic obtained restricting the quantification in a simple definable manner, we can ask whether there are any other natural sublogics that can be obtained by restricting the second-order quantifiers in some other first-order definable manner. There is one other. Namely, one might quantify not over arbitrary functions, but over permutations of the domain. This is called permutational logic. It arose in Shelah's study of symmetric groups. However, as it turns out, that's all! Up to a strong form of equivalence, the only sublogics of second-order logic given by first-order restricted second-order quantifiers are first-order logic, monadic second-order logic, permutational logic, and full second-order logic. This result, first proved in Shelah [1973c], is established by some new methods in Chapter XII. In addition, a number of newer, related results are presented.

Chapter XII
Definable Second-Order Quantifiers

by J. BALDWIN

In this chapter we investigate the class of second-order quantifiers which are definable in a sense which will shortly be made precise. This subject arose from investigations of the following sort. Let κ be an infinite cardinal and let S_κ denote the symmetric group on κ elements. What can we say about the first-order theories T_κ of the groups S_κ? Isbell showed that there is a sentence in the language of group theory that is true of S_κ just in case $\kappa = \omega$. McKenzie [1971] showed that $T_{\aleph_\alpha} = T_{\aleph_\beta}$ implies α and β are elementarily equivalent as ordered sets. We can describe the Isbell result as asserting that ω is characterized by a sentence of group theory. McKenzie asked whether or not the set of cardinals characterized in this way was the same as the set of second-order definable cardinals. Shelah [1973a] showed that this was not so. McKenzie had also reformulated the notion of characterization so as to make the question more natural. Instead of discussing the first-order theory of the group S_κ, we can discuss the theory of the set κ in a logic allowing quantification over permutations. Shelah [1973a, b] showed that the Hanf number of this logic is $\aleph_{\Omega^\omega}$, where $\Omega = (2^\omega)^+$. This answers McKenzie's questions, since there certainly are larger cardinals that are definable in second-order logic. In his proof, Shelah discussed a similar quantifier: quantification over permutations of order two. The first quantifier is certainly stronger than the second; moreover, it is easy to describe the first quantifier in terms of the second. To see this, we simply replace an arbitrary permutation f by three permutations g, h, j of order two such that on each orbit of f, g fixes "every other" element while, at the same time, h and j are a product of two-cycles. These cycles agree with f on the elements fixed by g and on the elements moved by g, respectively.

Prompted by questions raised by Stavi, Shelah [1973c] addressed the problem of determining which quantifiers the discussion was about and how many of them there were. The main aim of this chapter is to report his answer to this question. That is, that there are four second-order quantifiers (which are definable in the sense of Section 1.2 below): First-order (Q_{I}), monadic second-order (Q_{mon}), permutational (Q_{1-1}), and full second-order (Q_{II}). These quantifiers range over, respectively, elements, subsets, 1–1 functions, and arbitrary relations. In Section 1, we will formulate the entire question more precisely as well as provide some further examples of this class of quantifier. In Section 2 we will prove Shelah's theorem that there are only four second-order quantifiers. The proofs in Sections 1 and 2 focus attention on two ideas. The argument that Q_{II} is not interpretable in Q_{1-1} depends on the computation of the Hanf number of Q_{1-1}. On the other hand, the argument that any quantifier weaker than Q_{II} is interpretable in Q_{1-1} depends on a decomposition theorem. This kind of decomposition or Feferman–Vaught theorem is discussed in Section 3 (see also Chapter XIII) and is applied in Sections 4 and 5. In Section 4, we will explore the requirements on the notion of interpretation that are necessary to give a proof of non-interpretability via the computation of Hanf numbers. Section 5 surveys the classification of first-order theories by the interpretability of second-order quantifiers. This classification naturally falls into the unstable case (discussed in Section 4) and the stable case (discussed in Section 5). Section 6 contains a brief survey of some other generalizations that were found by Shelah.

1. *Definable Second-Order Quantifiers*

1.1. *Logics, Theories, and Quantifiers*

In Chapter II a logic L is defined as a function L which assigns to each vocabulary τ a set of sentences $L(\tau)$ and a semantics $\models_\tau$. In discussing higher-order quantifiers it is natural to examine theories rather than logics. For, the properties of a specific logic—say, monadic logic—vary tremendously depending on the vocabulary involved. In a vocabulary with only unary predicates the Hanf and Löwenheim numbers of monadic logic are $\aleph_0$ and the Feferman–Vaught theorem holds. On the other hand, if the vocabulary contains a binary function symbol f, then, by specifying f to be a pairing function, we extend from monadic logic to full second-order logic and all these pleasant properties are thus destroyed. Notice, however, that we must not only make a binary function symbol available but we must, in addition, specify that it defines a pairing function in order to induce the tragedy. The major results in this chapter concern the relative interpretability of theories in logics with second-order quantifiers.

Following are some notations and conventions which are perhaps peculiar to this chapter. Small Roman letters x, y, z etc. will represent individual variables while small Roman letters r, s, t etc. will represent predicate variables. Similarly, capital Roman letters R, S, T etc. represent relations, and small Roman letters a, b, c etc. individuals. We will use $\bar{a}$ to denote a finite sequence of individuals and $\bar{R}$ for a finite sequence of relations. We will also write $\bar{a} \in A$, and $\bar{R} \in A$, without writing the appropriate exponent on A. If $\phi(\bar{x}, \bar{y}, \bar{s})$ is a formula and A is a structure with $\bar{b} \in A$ and $\bar{S}$ a relation on A, then $\phi(A, \bar{b}, \bar{S}) = \{\bar{a} \in A \colon A \models \phi(\bar{a}, \bar{b}, \bar{S})\}$. We will regard the ordinary equality sign as a logical symbol. For any formula $\phi(x, \bar{y}, \bar{r})$, $(\exists^{<k} x)\phi(x, \bar{y}, \bar{r})$ abbreviates:

$$(x_0), \ldots, (x_k)\left(\bigwedge_{i<k+1} \varphi(x_i, \bar{y}, \bar{r}) \to \bigvee_{i<j<k+1} x_i = x_j \right)$$

1.2. Definable Second-order Quantifiers

For any structure M, let M_n denote the power set of M^n. Now full second-order logic allows quantification over $\bigcup_{n<\omega} M_n$. We could consider restricting our quantification to n-ary relations for a fixed n. More restrictively, we could allow $\exists X^n$ to range only over a specified subset of M_n. If we require that subset to be definable by a formula in pure equality theory, quantifying only over elements of M, we arrive at the class of definable second-order quantifiers. More formally, we have:

1.2.1 Notation. If $\psi(r)$ is a formula whose only non-logical symbol is the n-ary relation r, then for each infinite set A, $\mathscr{R}_\psi(A)$ is the collection of n-ary relations R on A such that $A \models \psi(R)$. We will use the same notation even if ψ contains a finite sequence $\bar{r}$ of relation variables.

1.2.2 Definition. Let $\psi(r)$ be a formula whose only symbols are r, $=$, first-order quantifiers, and propositional connectives. Then $Q_{\psi(r)}$ is the *second-order quantifier* whose semantics are given by:

$$M \models Q_{\psi(r)}\phi(r) \quad \text{iff} \quad (\exists R) \in \mathscr{R}_\psi(M), \quad M \models \phi(R).$$

There is a first-order theory naturally associated with each quantifier Q_ψ, namely the theory, T_ψ, whose only non-logical symbol is R and whose only non-logical axiom is $\psi(R)$. Note, however, that this theory does not contain all the information that the quantifier does. For, expressions in the language with the generalized quantifier can contain more than one instance of R.

Naturally, first-order quantification (Q_{I}) and full second-order quantification (Q_{II}) are definable second-order quantifiers. As we will see in Section 2, the only other examples are:

Monadic Quantification. Let r be unary and let $\phi(r)$ be any valid formula. Then $Q_{\phi(r)}$ is merely another name for the monadic second-order quantifier.

Permutational Quantification. Let r be binary and let $\phi(r)$ assert that r is an equivalence relation such that every class has two elements. We call $Q_{\phi(r)}$ the permutational quantifier. The name "permutational" will be justified shortly.

Note that quantification over L-automorphisms of M is not a definable second-order quantifier, since the assertion that f is an automorphism cannot be given in pure predicate calculus.

1.2.3 Definition. For T a first-order theory and Q_ψ a definable second-order quantifier we write (T, Q_ψ) for *the collection of all Q_ψ sentences in $L(T)$ valid on the models of T.*

Convention. We write Q_ψ for $(\mathrm{Th}(=), Q_\psi)$ where $\mathrm{Th}(=)$ is the theory of equality. We write $(\mathrm{Th}(<), Q_\psi)$ for the Q_ψ theory of order.

1.2.4 Definition. Let r be k-ary. We say $Q_{\psi(r)}$ is *(first-order) interpretable* in $(T, Q_{\psi(s)})$ if the following conditions hold. There exist first-order formulas $\theta_0(x_0, \bar{y}, \bar{s})$, $\theta(x_0, \ldots, x_{k-1}, \bar{y}, \bar{s})$ and $\chi(\bar{x}, \bar{s})$ such that:

(i) If $A \models \chi(\bar{a}_0, \bar{S}_0)$ then $\theta_0(A, \bar{a}_0, \bar{S}_0)$ is infinite, $\bar{S}_0 \in R_\psi(A)$, and

$$(\theta_0(A, \bar{a}_0, \bar{S}_0), \{\theta(A, \bar{a}, \bar{S}) : \bar{a} \in A, \bar{S} \in \mathscr{R}_\psi(A)\})$$
$$= (\theta_0(A, \bar{a}, \bar{S}_0), \mathscr{R}_\phi(\theta_0(A, \bar{a}, \bar{S}_0))).$$

(ii) For every infinite B, there exist A, $\bar{a}_0$, and $\bar{S}_0$ such that $A \models \chi(\bar{a}_0, \bar{S}_0)$ and $\theta_0(A, \bar{a}_0, \bar{S}_0) \approx B$.

Even though ϕ may contain only a single relation symbol r, the interpreting formulas may contain a sequence $\langle s_0, \ldots, s_n \rangle$. Note that by modifying θ_0 we can require, without loss of generality, that each structure $(\theta_0(A), \theta(A))$ satisfies ψ.

In accordance with our convention we will write $Q_{\phi(r)} \leq Q_{\psi(s)}$ whenever T is the theory of equality.

In this definition the theory which is interpreted is in the language with only the equality symbol. No other notion is needed for Section 2. For the discussion in Sections 4 and 5, we will extend the definition to $(T_1, Q_\psi) \leq (T_2, Q_\phi)$ by requiring that, for each relation symbol in the language of T_1, there be an interpreting formula in the language of (T_2, Q_ϕ). We actually employ this more general notion only when T_1 is the theory of order or $T_1 = T_2$.

The major results of this paper deal with the classification of the theories (T, Q_ψ), where T is a first-order theory. Section 2 concerns the case in which T is the theory of equality. It is easy to see that, for any theory T, we have

$$Q_\psi \leq Q_\phi \quad \text{implies} \quad (T, Q_\psi) \leq (T, Q_\phi).$$

Another formulation of this remark is that if $Q_\psi \leq Q_\phi$, then, for every vocabulary L, $L_{\omega,\omega}(Q_\psi) \leq L_{\omega,\omega}(Q_\phi)$, where '$\leq$' is taken in the sense of Chapter II. That is to say, the finitary logic associated with Q_ψ is weaker than that associated with Q_ϕ. Moreover, this result obviously extends to infinitary logics. Thus, the work described in this chapter provides a refinement of the notions in Chapter II.

We will now use this observation to show that the four quantifiers we have discussed are distinct. However, these quantifiers may coalesce on some T. For example, in the presence of a pairing function, Q_{mon} is equivalent to Q_{II}. This phenomena is discussed in detail in Section 5.

One way to show that quantifiers are distinct is to observe that interpretations as defined in Definition 1.2.4 preserve Hanf number. The Hanf number of a theory (T, Q_ψ) is the least cardinal such that any (Q_ψ)-sentence which has a model of at least that cardinality has arbitrarily large models. A number of variants on this notion are discussed in Baldwin–Shelah [1982], and we discuss it in somewhat more detail in Section 4. For the present, however, a quick application of this observation shows the following.

1.2.5 Theorem. *The four quantifiers are distinct*: $Q_{\text{mon}} \not\leq Q_{\text{I}}$, $Q_{1-1} \not\leq Q_{\text{mon}}$, $Q_{\text{II}} \not\leq Q_{1-1}$.

Proof. The class of well-orders is definable in $(\text{Th}(<), Q_{\text{mon}})$ but not in $(\text{Th}(<), Q_{\text{I}})$. Thus, $Q_{\text{mon}} \not\leq Q_{\text{I}}$.

Every sentence in $(\text{Th}(=), Q_{\text{mon}})$, is either true on all infinite sets or is false on infinite sets. Thus, the Hanf number of $(\text{Th}(=), Q_{\text{mon}})$ is $\aleph_0$. As remarked in the introduction, there are Q_{1-1} sentences of equality theory with only uncountable models. Thus, $Q_{1-1} \not\leq Q_{\text{mon}}$.

Shelah [1973a, b] showed the Hanf number of $(\text{Th}(=), Q_{1-1})$ is $\aleph_{\Omega^\omega}$ and thus that $Q_{\text{II}} \not\leq Q_{1-1}$. □

In the introduction we showed that quantification over arbitrary permutations is bi-interpretable in the sense of Definition 1.2.4 with quantification over permutations of order 2. It is clear that quantification over permutations of order 2 is bi-interpretable with the permutational quantification introduced above.

We will now give a few easy examples to show that a definable second-order quantifier which can define certain kinds of relations must be stronger than our standard examples, monadic and permutational quantification. The key to our argument will be to deal with very simple Q_ψ formulas, namely those of the form $\phi(\bar{x}, \bar{R})$ with $\bar{R} \in \mathscr{R}_\psi(A)$ and ϕ a first-order formula.

1.2.6 Definition. If the relation S on A is defined by $\phi(\bar{x}, \bar{b}, \bar{R})$ with $\bar{R} \in \mathscr{R}_\psi(A)$, where ϕ is of the first-order, then we say S is *simply definable* by Q_ψ.

It is easy to show from the definitions that Q_{II} is maximal among all the definable quantifiers

1.2.7 Proposition. *For any* ϕ, $Q_\phi \leq Q_{\text{II}}$. □

1.2.8 Lemma. *If* Q_ψ *simply defines an infinite, coinfinite set, then* $Q_{\text{mon}} \leq Q_\psi$.

Proof. Consider a definable second-order quantifier Q_ψ, and a structure A. Suppose that for some first-order formula $\phi(x, \bar{a}, \bar{R})$, with $\bar{R} \in \mathscr{R}_\psi(A)$ and $\bar{a} \in A$, both $\phi(A, \bar{a}, \bar{R})$ and $\neg\phi(A, \bar{a}, \bar{R})$ are infinite. We will show that each subset of A is definable by a formula $\theta(x, \bar{a}, \bar{R})$, with $\bar{R} \in \mathscr{R}_\psi(A)$. Call X a regular subset of A if $|X| = |A - X| = |A|$. Since ψ contains no non-logical symbols, the assumption that one regular subset of A is definable by a first-order formula $\phi(\bar{x}, \bar{b}, \bar{R})$ implies that any other regular subset is also. But any subset of A is a boolean combination of regular subsets so that all subsets of A are $Q_{\psi(r)}$ definable. Thus, $Q_{\text{mon}} \leq Q_{\psi(r)}$. □

We can view these remarks from another perspective, one that makes discussion of their consequences more concise. If $\bar{R} \in \mathscr{R}_\psi(A)$, then $(A, \bar{R})$ can be thought of as a model of a first-order theory in a language with non-logical symbols $\bar{R}$ and whose only axiom is $\psi(\bar{R})$. Then our last observation is simply the assertion that every infinite model of this theory is strongly minimal in the sense of Baldwin–Lachlan [1971] that T is strongly minimal if every definable (with parameters) subset is finite or cofinite. Moreover, standard compactness arguments show that this implies that if for each $\bar{b}$, $\phi(A, \bar{b}, \bar{R})$ is finite, then there is a uniform bound on the cardinalities of these sets.

1.2.9 Lemma. *If Q_ψ simply defines an equivalence relation with infinitely many infinite classes and $Q_{\text{mon}} \leq Q_\psi$, then $Q_{1-1} \leq Q_\psi$.*

Proof. Suppose there is a formula $\theta(x, y, \bar{a}, \bar{R})$, an $\bar{R} \in \mathscr{R}_\psi(A)$, and an $\bar{a} \in A$ such that $\theta(x, y, \bar{a}, \bar{R})$ defines on some infinite subset B of A an equivalence relation having infinitely many classes with more than two elements. By shrinking B, we may assume that each class has exactly two elements and that $A - B$ is infinite. By the compactness and Löwenhiem-Skolem theorems, we may assume that every infinite set C contains a regular subset B_C with such a definable equivalence relation. Using again the fact that ψ contains no nonlogical symbols, we see that a similar equivalence relation can be defined on $C - B_C$. But then, since B is simply definable (as every subset is simply definable), we can easily define an equivalence relation on all of C such that each class has exactly two elements. Thus we have defined Q_{1-1}. □

The main result asserts that the four quantifiers we have discussed are (up to bi-interpretability) the only definable second-order quantifiers and that, in fact, they are linearly ordered by interpretation. In fact, the argument shows that we would gain no additional cases by considering definable second-order quantifiers with finite strings of variables (that is, by replacing $Q_{\phi(r)}$ by $Q_{\phi(\bar{r})}$).

Most of the definitions in this section have described definable second-order quantifiers in pure logic. We can, of course, consider the more general situation in which we add definable second-order quantifiers to a non-trivial first-order theory. We will consider this situation in some detail in Section 5.

1.3. Some Conditions for Interpretability

In this section we will describe a few conditions which suffice for interpreting second-order logic into another logic.

We remarked in Section 1.1 that the introduction of a pairing function transforms monadic logic into full second-order logic. We now want to discuss a slightly weaker condition which has the same effect.

1.3.1 Definition. The theory T is *codable* if, for some n and some model M of T, there are infinite sets $\langle B_i : i < n \rangle$ and C contained in M and a first-order formula (possibly with parameters), $\phi(x, y_0, \ldots, y_{n-1})$, which defines a 1–1 map from $B_0 \times \cdots \times B_{n-1}$ onto C.

If T is codable, then, for any cardinal κ, we have a pairing function from two sets of power κ onto a third. We can thus easily code any binary relation on κ in terms of the pairing function and a subset of the third set. This argument is carried out in detail in Section II.2.4 of Baldwin–Shelah [1982]. Formally, we have

1.3.2 Theorem. *If T is codable, then $Q_{\text{II}} \leq (T, Q_{\text{mon}})$.* □

Arguments like those for 1.3.2 show:

1.3.3 Lemma. *If there is a first-order formula $\phi(x, y)$ which defines on some model M of a first-order theory T and on some infinite subset A of M an equivalence relation with infinitely many infinite classes, then $Q_{\mathrm{II}} \leq (T, Q_{1-1})$.* □

This is Chapter II, Section 2.6 of Baldwin–Shelah [1982].

2. *Only Four Second-Order Quantifiers*

In this section we will prove the main result of Shelah [1973c]: that up to interpretation (in the sense defined in Definition 1.2.4) there are only four (definable) second-order quantifiers. In Section 2.1 we will begin by deriving some consequences of Ramsey's theorem and the Δ-system lemma which will be used several times in the proof of the main theorem. That done, we will then show successively in Section 2.2 that if $Q_{\mathrm{mon}} \not\leq Q_\psi$, then $Q_\psi \leq Q_{\mathrm{I}}$; in Section 2.3 that if $Q_{1-1} \leq Q_\psi$, then $Q_\psi \leq Q_{\mathrm{mon}}$; and finally in Section 2.4 that if $Q_{\mathrm{II}} \not\leq Q_\psi$, then $Q_\psi \leq Q_{1-1}$. These three assertions and Proposition 1.2.7 yield the following theorem.

2.0 Theorem. *If Q_ψ is a definable second-order quantifier, then Q_ψ is bi-interpretable with one of Q_{I}, Q_{mon}, Q_{1-1}, or Q_{II}.*

The proof of the first two of the three assertions constituting this theorem is just a reworking of the argument given in Shelah [1973c]. We give the main idea of the proof for the third in Section 2.4. In Sections 2.5 and 2.6 we give alternate arguments for the crucial Theorem 2.4.6. The argument in Section 2.5 is derived from Baldwin–Shelah [1982], while that in Section 2.6 is a modification of the argument given in Shelah [1973c].

The argument for each of the three cases follows the same general line. To show that $Q_\phi \leq Q_\psi$, we first define an appropriate notion of "$\bar{a}$ and $\bar{b}$ are Q_ψ-similar over" respectively a finite set of elements in Section 2.2, a finite set of elements and a finite set of subsets in Section 2.3, and a finite set of elements, a finite set of subsets, and a finite set of 1–1 functions in Section 2.4. We say $\bar{S}$ determines θ if $\bar{a}$ and $\bar{b}$ are Q_ψ similar over some sequence $\bar{S}$ satisfying ψ implies $\bar{a}$ and $\bar{b}$ satisfy the same formulas $\theta(\bar{x}; \bar{R})$, for $\bar{R} \in \mathscr{R}_\phi(A)$. It is easy to see that if $\bar{S}$ determines each θ then $Q_\phi \leq Q_\psi$. The bulk of the argument which differs from case to case consists in showing by induction on $\lg(\bar{x})$ that each $\theta(\bar{x}; \bar{R})$ is so determined.

2.1 Consequences of Some Combinatorial Lemmas. Our first result is an application of Ramsey's theorem to the problem of interpretation.

2.1.1 Lemma. *Let $\theta(z, y, \bar{x}, \bar{R})$ be a first-order formula. Suppose that for every A and every $\bar{R} \in \mathscr{R}_\psi(A)$ and for some $m < \omega$ we have*

$$A \models (\bar{x})(y)(\exists^m z)\theta(y, z, \bar{x}, \bar{R})$$

and

$$A \models (z)(y)\theta(y, z, \bar{x}, \bar{r}) \to z \neq y.$$

Then either

(1) *for some $n < \omega$, we have $A \models (\bar{x})(\exists^n)[(\exists y)\theta(y, z, \bar{x}, \bar{R})]$; or*
2(a) *$Q_{\text{mon}} \leq Q_\psi$ and*
2(b) *$Q_\psi \leq Q_{1-1}$.*

Proof. Assuming that (1) fails, we first show

(*) There are $C = \langle c_i : i < \omega \rangle$ and $B = \langle b_j : j < \omega \rangle$ and $\bar{d}$ such that $B \cap C = \varnothing$; $c_i = c_j$ iff $i = j$; $\models \theta(b_i, c_j, \bar{d}, \bar{R})$ iff $i = j$; and $\models \neg\theta(c_i, c_j, \bar{d}, \bar{R})$ if $i \neq j$.

If, for each $\bar{d}$, there are only finitely many c such that $\models (\exists z)\theta(z, c, \bar{d}, \bar{R})$, then (1) holds by an easy compactness argument and we are finished. If not, then we can certainly find disjoint sets B and C such that $A \models \theta(b_i, c_j, \bar{d}, \bar{R})$ but $A \models \neg\theta(b_i, c_j, \bar{d}, \bar{R})$, for $i < j$. By applying Ramsey's theorem to the partition of pairs $\{i, j\}$ for $i < j < \omega$ induced by whether or not $\theta(b_i, c_j, \bar{d}, \bar{R})$ holds, we can pass to subsets of B and C so that the truth of $\theta(b_i, c_j, \bar{d}, \bar{R})$ depends only on the order of i and j. We know that $\theta(b_i, c_j, \bar{d}, \bar{R})$ fails if $i < j$ and since $(\exists^{<m} z)\theta(z, c_j, \bar{d}, \bar{R})$ and some c_j has more than m predecessors, we must also have $\neg\theta(b_i, c_j, \bar{d}, \bar{R})$ if $i > j$. A similar use of Ramsey's theorem allows us to assume that $\neg\theta(c_i, c_j)$ also if $i \neq j$. This establishes (*).

We will now define a formula $\chi(y, \bar{d}, \bar{R}, \bar{R}')$ such that $A \models \chi(c_i, \bar{d}, \bar{R}, \bar{R}')$ iff $i \equiv 0 \bmod 3$. Since we will have thus defined an infinite and coinfinite set, it will follow by Lemma 1.2.8 that $Q_{\text{mon}} \leq Q_\psi$. We can assume that none of the b_i's or c_j's occur in $\bar{d}$. Let f be the permutation of A which interchanges c_{3i+2} and c_{3i+1} and leaves all other elements of A fixed. Let $\bar{R}'$ be the image of R under f. That is, f is an isomorphism between $(A, \bar{R})$ and $(A, \bar{R}')$. Then

$$A \models (x)[\theta(x, c_j, \bar{d}, \bar{R}) \leftrightarrow \theta(x, c_j, \bar{d}, \bar{R}')]$$

if and only if $j \equiv 0 \bmod 3$. Thus, letting $\chi(y, \bar{d}, \bar{R}, \bar{R}')$ be

$$(x)[\theta(x, y, \bar{d}, \bar{R}) \leftrightarrow \theta(x, y, \bar{d}, \bar{R}')],$$

we have (2a).

To obtain (2b), we note that $\theta(x, y, \bar{d}, \bar{R})$ defines on a subset of $B \cup C$ an equivalence relation having infinitely many classes with two elements. In the light of (2a) and Lemma 1.2.9 we have (2b).

Our next step is an application of a weak version of the Δ-system lemma. The remainder of this section is applied in Lemma 2.3.6 and 2.5.8.

2.1.2 Definition. A *Δ-system with heart H* is a family of sets $\{C_i: i < \kappa\}$ such that if $i \neq j$, then $C_i \cap C_j = H$. We will frequently fix an enumeration $\bar{h}$ of H. Then $\bar{h}$ will be taken to mean either the sequence $\bar{h}$ or the range of that sequence (that is, H), whichever is appropriate.

An easy combinatorial argument establishes

2.1.3 Lemma (The Weak Δ-System Lemma). *If $\langle C_i: i < \omega\rangle$ is a sequence of distinct sets with the same finite cardinality n, then there is a subsequence of the C_i which is a Δ-system with some heart H, and $|H| < n$.*

For our application we want to distinguish the following families of formulas.

2.1.4 Definition. A family of formulas $\{\theta_n(z_0, \ldots, z_{n-1}, \bar{y}, \bar{r}): n < \omega\}$ is *malleable* if

(i) θ_n is predicate of the set $\{z_0, \ldots, z_{n-1}\}$, not the sequence $\bar{z}$.
(ii) If $\{\bar{c}_i: i < \omega\}$ is a Δ-system with heart $H(|\bar{c}_i| = n$ and $|H| = m < n)$ and $A \models \theta_n(\bar{c}_i, \bar{b}, \bar{R})$ for i, then $A \models \theta_m(\bar{h}, \bar{b}, \bar{R})$.

2.1.5 Example. If $\theta_n(z_0, \ldots, z_{n-1}, \bar{y}, \bar{r})$ is

$$(\bar{x})\left[\phi(x_0, \ldots, x_{m-1}, \bar{y}, \bar{r}) \rightarrow \bigwedge_{i<m} \bigvee_{j<n} x_i = z_j\right],$$

then $\{\theta_n: n < \omega\}$ is a malleable family. To see this, we let $\bar{d}$ be a solution of $\phi(\bar{x}, \bar{b}, \bar{R})$ and let $\bar{c}_i$ be a Δ-system of n-tuples such that $\bar{d} \subseteq \bar{c}_i$ for each i. Then $\bar{d}$ is clearly contained in $\bar{h}$.

For $\{\theta_n: n < \omega\}$ a malleable family, we introduce the following notation: θ_n^* denotes $(\exists z_1) \cdots (\exists z_{n-1})\theta_n(z_0, z_1, \ldots, z_{n-1}, \bar{y}, \bar{R})$. $\theta_n'(\bar{z}, \bar{y}, \bar{R})$ denotes the conjunction of $\theta_n(\bar{z}, \bar{y}, \bar{R})$ with the formulas $(z_0), \ldots, (z_{m-1}) \neg\, \theta_m(\bar{z}, \bar{y}, \bar{R})$, for $m < n$.
This definition is designed to yield the following lemma.

2.1.6 Lemma. *Suppose θ_n is a malleable family of formulas such that for every A, every $R \in \mathscr{R}_\psi(A)$, and every $\bar{b}$ in A, there is a finite sequence $\bar{c}$ with $|\bar{c}| < M$ (for some integer M) such that $A \models \theta_{|c|}(\bar{c}, \bar{b}, \bar{R})$. Then*

(i) *There is an integer $n(\bar{b})$ such that: $\theta_{n(\bar{b})}^*(A, \bar{b}, \bar{R})$ is finite and*

$$A \models (\exists \bar{z})\theta_{n(\bar{b})}'(\bar{z}, \bar{b}, \bar{R}) \wedge \bigwedge_{i<n(\bar{b})} \theta_{n(\bar{b})}^*(z_i, \bar{b}, \bar{R}).$$

(ii) *If, in addition,* $Q_{\text{mon}} \leq Q_\psi$ *then there is an integer k and a formula* $\theta^*(z, \bar{y}, \bar{R})$ *such that*:

(a) $A \models (\bar{y})(\exists^{<k} z)\theta^*(z, \bar{y}, \bar{R})$;

(b) $A \models (\bar{y})(\bar{z})[\theta'_m(\bar{z}, \bar{y}, \bar{R}) \to \wedge \{\theta^*(z_i, \bar{y}, \bar{R}): i < m\}]$, *for all* $m < M$.

Proof. (i) Fix $\bar{b}$ and some $\bar{c}$ of smallest cardinality such that $A \models \theta_{|c|}(\bar{c}, \bar{b}, \bar{R})$ and suppose $|\bar{c}| = n$. Suppose $\theta^*_n(A, \bar{b}, \bar{R})$ is infinite. Then there is an infinite family of n-element sets C_i such that if $\bar{c}_i$ is any enumeration of C_i, $A \models \theta(\bar{c}_i, \bar{b}, \bar{R})$. By the Δ-system lemma, we can find a heart $H(\bar{h})$ for the C_i's with $|H| = m < n$. Moreover, by the very definition of malleable family, $A \models \theta_m(\bar{h}, \bar{b}, \bar{R})$. But this contradicts the minimality of n and so yields (i).

(ii) Since $Q_{\text{mon}} \leq Q_\psi$, we know by Lemma 1.2.8 that any $(A, \bar{R})$ with each $R \in \mathscr{R}_\psi(A)$ is strongly minimal. In particular, there is an integer k such that all the sets $\theta^*_{n(\bar{b})}(A, \bar{b}, \bar{R})$ have cardinality $< k$. Recall that by hypothesis all the $n(\bar{b}) < M$; and, furthermore, let θ^* be the formula:

$$\bigvee_{j<M} (\exists u_0) \cdots (\exists u_{j-1})\theta^*_j(\bar{u}, \bar{y}, \bar{R}) \to \theta^*_j(z, \bar{y}, \bar{R}).$$

This formula clearly meets conditions (a) and (b). □

2.2 Lemma. *If* $Q_{\text{mon}} \not\leq Q_\psi$, *then* $Q_\psi \leq Q_I$.

This subsection is devoted to the proof of Lemma 2.2. We will proceed by induction to show that the hypothesis implies that every formula with parameters $\bar{R}$ in $\mathscr{R}_\psi(A)$ and k free variables is expressible in first-order logic. When k reaches the arity of R we must then have the lemma (see Lemma 2.2.2). In addition to the notions from Sections 1 and 2.1, we will require the following concept.

2.2.1 Definition. (i) Let X be a finite set of relation symbols or formulas. By $\text{tp}_X(\bar{a}; B)$ we mean the *collection of formulas* $\psi(\bar{x}; \bar{b})$ such that $\bar{b} \in B$, $\psi(\bar{x}; \bar{y}) \in X$ and $\models \psi(\bar{a}, \bar{b})$. We will simply write, $t_=(\bar{a}; B)$ for $t_{\{=\}}(\bar{a}; B)$.

(ii) Two finite sequences of the same length, $\bar{a}$ and $\bar{b}$, are (first-order) *similar* over B if $\text{tp}_=(\bar{a}; B) = \text{tp}_=(\bar{b}; B)$.

(iii) The set $D \cup \bar{c}$ *determines* $\phi(\bar{x}; \bar{c}, \bar{R})$ if for any sequences $\bar{a}$ and $\bar{b}$ which are similar over $D \cup \bar{c}$: $A \models \phi(\bar{a}, \bar{c}, \bar{R}) \leftrightarrow \phi(\bar{b}, \bar{c}, \bar{R})$

Note. The notion $D \cup \bar{c}$ determines $\phi(\bar{x}; \bar{c}, \bar{R})$ depends not just on the formula $\phi(\bar{x}, \bar{y}, \bar{R})$ but on the partition of the sequence $\bar{x}\bar{y}$.

2.2.2 Lemma. *If, for every formula* $\phi(\bar{x}, \bar{y}, \bar{r})$, *there exists a formula* $\phi^*(z, \bar{y}, \bar{r})$ *and an integer n such that for every A and every* $\bar{b}$ *in A and* $\bar{R}$ *in* $\mathscr{R}_\psi(A)$

(i) $|\phi^*(A, \bar{b}, \bar{R})| \leq n$; *and*

(ii) $\phi^*(A, \bar{b}, \bar{R}) \cup \{\bar{b}\}$ *determines* $\phi(\bar{x}, \bar{b}, \bar{R})$,

then $Q_\psi \leq Q_I$.

Proof. We apply the hypothesis, taking $r(\bar{x})$ as $\phi(\bar{x}; \bar{y}, \bar{r})$. Then $R(\bar{x})$ is determined by the finite set $\phi^*(A, R)$ so that a suitable coding of the equality types over $\phi^*(A, R)$ defines $R(\bar{x})$ as required. □

2.2.3 Definition. The formula $\chi(w)$ is an $=$-diagram (read simply as *equality diagram*) if χ is a maximal consistent conjunction of equalities and inequalities among the w_i.

Shelah [1973] calls χ a complete formula. The following lemma yields Lemma 2.2.

2.2.4 Lemma. *If $Q_{\text{mon}} \not\leq Q_\psi$, then for every formula $\phi(\bar{x}, \bar{y}, \bar{r})$, there is a formula $\phi^*(z, \bar{y}, \bar{r})$ and an integer k such that for every $A, \bar{b}$, and $\bar{R}$ in $\mathscr{R}_\psi(A)$: $|\phi^*(A, \bar{b}, \bar{R})| = k$ and $\phi^*(A, \bar{b}, \bar{R}) \cup \{\bar{b}\}$ determines $\phi(\bar{x}, \bar{b}, \bar{R})$.*

Proof. The proof is by induction on the length of $\bar{x}$ for arbitrary sequences $\bar{y}$ and $\bar{r}$. If $\lg(\bar{x}) = 1$, the result is immediate from the remark following the proof of Lemma 1.2.8.

We now consider a formula $\phi(\bar{x}; \bar{y}, \bar{r})$. Let $\bar{x} = \bar{x}'\ w$ and $\bar{y}' = w\ \bar{y}$. Now, we have $\phi_0 = \phi(\bar{x}; \bar{y}, \bar{r})$ and $\phi_1 = \phi(\bar{x}'; \bar{y}', \bar{r})$ which differ only in the position of the semicolon. Suppose we have constructed by induction a formula $\phi_0^*(z, \bar{y}', \bar{r})$ and an m such that for each $a, \bar{b}$, and $\bar{R} \in \mathscr{R}_\psi(A)$:

(i) $|\phi_0^*(A, a, \bar{b}, \bar{R})| < m$;
(ii) $\phi_0^*(A, a, \bar{b}, \bar{R}) \cup \{a, \bar{b}\}$ determines $\phi(\bar{x}'; a, \bar{b}, \bar{R})$;
(iii) $\phi_0^*(z, w, \bar{b}, \bar{R}) \to z \neq w$.

By explicitly listing $\{a, \bar{b}\}$ in (ii), we are left free to assume that (iii) holds. Now, applying Lemma 2.1.1, we see $A \models (\bar{y})(\exists^{<k} z)(\exists w)(\phi_0^*(z, w, \bar{y}, \bar{R}))$. Let $\phi_1^*(z, \bar{y}, \bar{R})$ be $(\exists w)\phi_0^*(z, w, \bar{y}, \bar{R})$. Then, for each a and $\bar{b}$, it is easy to see that $\phi_1^*(A, \bar{b}, \bar{R}) \cup \{a, \bar{b}\}$ determines $\phi(\bar{x}', a, \bar{b}, \bar{R})$. It remains to remove the dependence on a. To do this, however, we must first look more carefully at how the determination occurs.

Let $\bar{c}$ be an enumeration of $\phi_1^*(A, \bar{b}, \bar{R})$. Fix $\lg(\bar{z}) = \lg(\bar{c})$ and let $\chi_i(\bar{x}'; w, \bar{y}, \bar{z})$ for $i < p$ be a complete list of the equality diagrams in the displayed variables. For each $a \in A$ and each i, we must have either

(i) $A \models (\bar{x}')[\chi_i(\bar{x}', a, \bar{b}, \bar{c}) \to \phi(\bar{x}', a, \bar{b}, \bar{R})]$; or
(ii) $A \models (\bar{x}')[\chi_i(\bar{x}', a, \bar{b}, \bar{c}) \to \neg\phi(\bar{x}', a, \bar{b}, \bar{R})]$.

Now, for each $S \subseteq p$, let $\chi_S(a, \bar{b}, \bar{R})$ hold just if (i) above holds for exactly those $i \in S$. Now, by strong minimality, there is an L (depending on S and $\bar{b}$) such that if $A \models \chi_S(a, \bar{b}, \bar{R})$ for more than L choices of a, then $A \models \chi_S(a, \bar{b}, \bar{R})$ for all but finitely many a. By compactness and the fact that there are only finitely many choices for S, we can choose a single L with this property for all $\bar{b}$ and S. Now, $\phi(\bar{x}; \bar{b}, \bar{R})$ is clearly determined by $\phi_1^*(A, \bar{b}, \bar{R}) \cup C(\bar{b}) \cup \bar{b}$, where $C(\bar{b})$ denotes the set of those a such that $A \models \chi_S(a, \bar{b}, \bar{R}) \to (\exists^{<L} x)\chi_S(x, \bar{b}, \bar{R})$. Moreover, we now see that $\phi_1^*(A, \bar{b}, \bar{R}) \cup C(\bar{b})$ has less than $(k + L)$ elements and is uniformly definable from $\bar{b}$. Thus, we have proven the lemma. □

In Shelah's original proof, the $C(\bar{b})$ are defined by an appeal to Lemma 2.1.6 so that the structure of his argument is actually closer to that which follows in the proof of Lemma 2.3.

2.3 Lemma. *If* $Q_{1-1} \not\leq Q_\psi$ *then* $Q_\psi \leq Q_{\text{mon}}$.

Our proof of this result is parallel to the proof of Lemma 2.2. We will require the following concept—a concept that is analogous to the notion given in Definition 2.2.1.

2.3.1 Definition. (i) Two finite sequences, $\bar{a}$, $\bar{b}$, of the same length, n, are *monadically similar* over $\langle D; C_0, \ldots, C_{m-1}\rangle$ if for any d in D and any $i < n$, $b_i = d$ iff $a_i = d$; and, for $j < m$, $a_i \in C_j$ iff $b_i \in C_j$.

(ii) *A finite equivalence relation* over F is an equivalence relation (on k-tuples, for some k) which is definable with parameters from F and has only finitely many equivalence classes.

(iii) The set D and the finite equivalence relation E *monadically determine* $\phi(\bar{x}, \bar{c}, \bar{R})$ if, for any sequences $\bar{a}$ and $\bar{b}$ of the same length: if $\bar{a}$ and $\bar{b}$ are monadically similar over $\langle D \cup \bar{c}; C_0, \ldots, C_{k-1}\rangle$ where the C_i are the equivalence classes of E, then $\phi(\bar{a}, \bar{c}, \bar{R}) \leftrightarrow \phi(\bar{b}, \bar{c}, \bar{R})$.

2.3.2 Lemma. *If for every formula* $\phi(\bar{x}, \bar{y}, \bar{r})$ *there exist formulas* $\phi^*(x, u, \bar{y}, \bar{r})$ *and* $\theta(z, \bar{y}, \bar{r})$ *such that for every* A, $\bar{c}$, *and* $\bar{R} \in \mathscr{R}_\psi(A)$:

(i) $\theta(A, \bar{c}, \bar{R})$ *is finite*;

(ii) $\phi^*(x, u, \bar{c}, \bar{R})$ *is a finite equivalence relation*;

(iii) $\theta(A, \bar{c}, \bar{R})$ *and* $\phi^*(x, u, \bar{c}, \bar{R})$ *monadically determine* $\phi(\bar{x}, \bar{c}, \bar{R})$, *then* $Q_\psi \leq Q_{\text{mon}}$.

2.3.3 Definition. For any formula $\phi(\bar{x}; \bar{y}, \bar{r})$, any A and $\bar{c}$, any $\bar{R} \in \mathscr{R}_\psi(A)$, and any $C \subseteq A$, define $e(\phi(\bar{x}; \bar{c}, \bar{R}), C, A) = e(\phi, C, A)$ by

$$e(\phi, C, A) = \{\langle a, b\rangle : \text{tp}_{\{=, \phi(\bar{x}, \bar{c}, \bar{r})\}}(a; A - C) = \text{tp}_{\{=, \phi(\bar{x}, \bar{c}, \bar{r})\}}(b; A - C).$$

The formulas in $\text{tp}_{\{=, \phi(\bar{x}, \bar{c}, \bar{r})\}}(a; X)$ are obtained by fixing any entry in $\bar{x}$ for substitution of a and leaving the others for substitutions from X. Note that $e(\phi, C, A)$ is first-order definable (with parameters C, $\bar{c}$ and $\bar{R}$).

2.3.4 Lemma. *If* $Q_{1-1} \leq Q_\psi$, *then for every* $A, C, \bar{b}$, *and* $\phi(\bar{x}; \bar{y}, \bar{r})$, $e(\phi, C, A)$ *has only finitely many equivalence classes.*

Proof. By Lemma 2.2, we can assume that $Q_{\text{mon}} \leq Q_\psi$. We first note that by Lemma 1.2.9, since $e(\phi, C, A)$ is definable, it can have only finitely many equivalence classes with two or more elements. Since replacing C by a smaller set refines the equivalence relation, we can, by proper choice of C, assume that each class of $e(\phi, C, A)$ is a singleton. If $e(\phi, C, A)$ has infinitely many classes, we will define in terms of $\bar{R} \in \mathscr{R}_\psi(A)$, an equivalence relation possessing infinitely many classes with two or more elements. We thereby contradict Lemma 1.2.9. For this, fix a permutation f with order 2 of A whose set of fixed points is $(A - C) \cup \{\bar{b}\}$. Let $\bar{R}_1 = f(\bar{R})$. Let S_0 denote the relation defined by $\phi(\bar{x}, \bar{b}, \bar{R})$ and S_1 the relation defined by $\phi(\bar{x}, \bar{b}, \bar{R}_1)$. Let e_1 be the following equivalence relation (this relation is clearly definable from $\bar{R}$, $\bar{R}_1$, C and $\bar{b}$ and therefore by Q_ψ):

$$\{\langle a, c\rangle : \mathrm{tp}_{\{S_0, =\}}(a: (A - C) \cup \{\bar{b}\}) = \mathrm{tp}_{\{S_1, =\}}(c; (A - C) \cup \{\bar{b}\})$$

and

$$\mathrm{tp}_{\{S_0, =\}}(c; (A - C) \cup \{\bar{b}\}) = \mathrm{tp}_{\{S_1, =\}}(a; (A - C) \cup \{\bar{b}\}).$$

Clearly, if $a, c \in C$ and $f(a) = c$, then $\langle a, c\rangle \in e_1$. Now, if $\langle a, c\rangle \in e_1$, then

$$\mathrm{tp}_{\{S_0, =\}}(c; (A - C) \cup \{\bar{b}\}) = \mathrm{tp}_{\{S_1, =\}}(a; (A - C) \cup \{\bar{b}\}) = q.$$

But since $e(\phi(\bar{x}; \bar{b}, \bar{R}), C, A)$ has only singleton equivalence classes, the unique element realizing q in the S_1 interpretation is $f(c)$. So $a = f(c)$. Since e_1 is clearly symmetric, we see that $e_1(a, c)$ if and only if $a = f(c)$. That is, we can define by $e_1(x, y) \vee x = y$ an equivalence relation with infinitely many two element classes.

Note that by invoking the compactness theorem, we can find a uniform n such that, for all C and $\bar{b}$, $e(\phi(\bar{x}; \bar{b}, \bar{R}), C, A)$ has less than n equivalence classes.

The following technical result asserts that if a definable symmetric, reflexive relation has a bounded number of pairwise incomparable elements then its transitive closure also is definable. We need it for the next lemma.

2.3.5 Proposition. *Suppose $\phi(x, y)$ defines a symmetric reflexive relation such that, for some m and for any set of distinct elements $\{a_i: i < m\}$, there are $i \neq j$ such that $\phi(a_i, a_j)$. Then the equivalence relation E which is obtained by forming the transitive closure of the relation defined by $\phi(x, y)$ is itself defined by:*

$$(\exists z_0), \ldots, (\exists z_{2m-3}) \bigwedge_{i<2m-3} \phi(z_i, z_{i+1}) \wedge z_0 = x \wedge z_{2m-3} = y).$$

Proof. Let $\{a_0, \ldots, a_k\}$ be the shortest path connecting a_0 and a_k, and let $k = 2u$ or $k = 2u + 1$, depending on the parity of u. No pair from $\{a_0, \ldots, a_u\}$ satisfies ϕ. Thus, $u \leq m - 1$ which yields the result. ☐

2.3.6 Lemma. *If $Q_{1-1} \leq Q_\psi$, then for each formula $\phi(\bar{x}; \bar{c}, \bar{R})$ there are formulas $\phi^*(x, u, \bar{y}, \bar{R})$ and $\theta^*(z, x, u, \bar{y}, \bar{r})$ such that for every A, $\bar{c}$, a and $\bar{R} \in \mathscr{R}_\psi(A)$:*

(i) *$\phi^*(x, y, \bar{c}, \bar{R})$ defines an equivalence relation with finitely many classes.*
(ii) *If $A \models \phi^*(a, b, \bar{c}, \bar{R})$, then $\langle a, b\rangle \in e(\phi(\bar{x}, \bar{c}, \bar{R}), \theta^*(A, \bar{c}, \bar{R}) \cup \{a, b\}, A)$.*
(iii) *$\theta^*(A, \bar{c}, \bar{R})$ is finite.*

Proof. We first use Proposition 2.3.5 to establish (i) and the weakened version of (ii) which is obtained by replacing $e(\phi(\bar{x}; \bar{c}, \bar{R}), \theta(A, \bar{c}, \bar{R}) \cup \{a, b\}, A)$ by the equivalence relation e^* which holds for two elements if and only if for some finite B, $e(\phi, B, A)$ also holds of those elements. The formula $\phi^*(x, y, \bar{b}, \bar{R})$ defines a finite equivalence relation which refines the finite equivalence relation $e^*(\phi(\bar{x}; \bar{b}, \bar{R}), A, \bar{R})$. Then two applications of Lemma 2.1.1 yield the full result.

For the first step, define for each A, $\bar{b}$ and $\bar{R} \in \mathscr{R}_\psi(A)$ the binary relation $e_n = e_n(\phi(\bar{x}; \bar{b}, \bar{R}), A, \bar{R})$ to hold for $\langle a, b\rangle$ just if for some n-element subset B of A, $\langle a, b\rangle \in e(\phi(\bar{x}, \bar{b}, \bar{R}), B, A)$. Note that e_n is reflexive and symmetric but not transitive. Moreover, there is a formula $\phi_n(x, y, \bar{b}, \bar{R})$ which defines e_n. Finally, e_n refines e_{n+1}. Now, let the equivalence relation $e^* = \bigcup \{e_n: n < \omega\}$. For a fixed m, not depending on B, each $e_n(\phi, B, A)$ has at most m classes so there is no set of $m + 1$ elements, each pair of which does not satisfy ϕ_n. Thus e^* has at most m classes. So for some l, the set of sentences

$$\Gamma_l = \{\psi(R)\} \bigcup_{\substack{n<l \\ 1 \leq i<j \leq m+1}} \{\neg\phi_n(x_i, x_j, \bar{d}, \bar{R})\}$$

is inconsistent. Let p be the least integer such that Γ_p is inconsistent. By Proposition 2.3.5, the transitive closure of $\phi_p(x, y, \bar{d}, \bar{r})$ is definable by a formula ϕ^*, and defines an equivalence relation with at most m classes. ϕ^* clearly satisfies (i) and the weakened form of (ii). Thus, each equivalence class of e^* is a union of ϕ^* equivalence classes.

To establish the full strength of (ii), we define the malleable family of formulas $\theta_n(x, y, \bar{z}, \bar{u}, \bar{r})$ which assert that $\langle x, y\rangle \in e(\phi(\bar{x}; \bar{u}, \bar{r}), \{x, y, z_0, \ldots, z_{n-1}\}, A)$. Taking p for the bound M in the hypothesis of Lemma 2.1.6, we deduce that there is a formula $\theta^*(z, x, y, \bar{u}, \bar{r})$ such that for some k (first a $k(\bar{b})$ but then, by compactness, independent of $\bar{b}$) we have:

(a) $A \models (x)(y)(\exists^{<k}z)\theta^*(z, x, y, \bar{c}, \bar{R})$.
(b) If $A \models \phi^*(a, b, \bar{c}, \bar{R})$ then $\langle a, b\rangle \in e(\phi, \theta^*(A, a, b, \bar{c}, \bar{R}) \cup \{a, b\}, A)$.
(c) If $A \models \neg\phi^*(a, b, \bar{c}, \bar{R})$ then $\theta^*(A, a, b, \bar{c}, \bar{R}) = \emptyset$.

Now, applying Lemma 2.1.1 twice to condition (a) we obtain

$$A \models (\exists^{<k}z)(\exists x)(\exists y)\theta^*(z, x, y, \bar{b}, \bar{R}) \quad \text{so}$$
$$(\exists^{<k}z)(\exists x)(\exists y)\theta^*(z, x, y, \bar{u}, \bar{R}).$$

Now, to complete the proof of Lemma 2.3, we show by induction that every formula is monadically determined.

2.3.7 Lemma. *If $Q_{1-1} \not\leq Q_\psi$, then for any $\phi(\bar{x}; \bar{y}, \bar{r})$, there are formulas $\phi^*(x, u, \bar{y}, \bar{r})$ and $\theta(z, \bar{y}, \bar{r})$ which monadically determine $\phi(\bar{x}; \bar{y}, \bar{r})$.*

Proof. The proof is by induction on $\lg(\bar{x})$. If $\lg(\bar{x}) = 1$, we are merely restating Lemma 2.3.6. Thus, suppose that we have the result if $\lg(\bar{x}) < n$, and consider a formula $\phi(\bar{x}; \bar{y}, \bar{R})$ with $\lg(\bar{x}) = n$. By Lemma 2.3.6 we can find a finite equivalence relation $\phi^*(x, y, \bar{b}, \bar{R})$ and a set $\theta(A, \bar{b}, \bar{R})$ such that if

$$A \models \phi^*(a, c, \bar{b}, \bar{R}), \quad \text{then} \quad \langle a, c \rangle \in e(\phi, \theta^*(A, a, c, \bar{b}, \bar{R}) \cup \{a, c\}, A).$$

This means that the equivalence classes of $e(\phi, \theta^*(A, a, c, \bar{b}, \bar{R}) \cup \{a, c\}, A)$ are finite unions of equivalence classes of ϕ^*. Now, for each element d of $\theta^*(A, \bar{b}, \bar{R}) \cup \{a, c\}$, let $\phi_{d,i}(\bar{x}'; \bar{b}, d, \bar{R})$ be the $(n-1)$-ary relation obtained by substituting d for x_i in ϕ_i. Then, $\phi(\bar{x}; \bar{b}, \bar{R})$ is first-order definable from the equivalence classes of ψ^*, the elements of $\theta(A, \bar{b}, \bar{R}) \cup \{a, c\}$, and the $\phi_{d,i}(\bar{x}', \bar{b}, \bar{R})$. For, if $\bar{a} \cap \theta(A, \bar{b}, \bar{R}) = \emptyset$, then $\phi(\bar{a}; \bar{b}, \bar{R})$ depends only on the ϕ^* equivalence class of the a_i. If $\bar{a} \cap \theta(A, \bar{b}, \bar{R}) \neq \emptyset$, then $\phi(\bar{a}; \bar{b}, \bar{R})$ depends on one of the $\phi_{d,i}$ which are monadically determined by induction. This completes the proof of Lemma 2.3. □

2.4 Lemma. *If $Q_{\mathrm{II}} \not\leq Q_\psi$ then $Q_\psi \leq Q_{1-1}$.*

Once we have established this lemma, we will have completed the proof of the four second-order quantifier theorem. We will first show that a certain decomposition of all structures $(A, \bar{R})$ with $\bar{R} \in \mathscr{R}_\psi(A)$ implies that $Q_\psi \leq Q_{1-1}$. Afterwards, we will show that the hypothesis $Q_{\mathrm{II}} \not\leq Q_\psi$ implies that such a decomposition exists.

An extremely simple example of such a decomposition is the division of models of $\mathrm{Th}(Z, S)$ into connected components. More complicated examples are elaborated in Baldwin–Shelah [1982].

2.4.1 Definition. (i) If E is an equivalence relation then two sequences $\bar{a}$ and $\bar{b}$ are *similar for E* if $\lg(\bar{a}) = \lg(\bar{b}) = k$ and there is a partition of k into, say, n sets $J_0, \ldots, J_{n-1}$ such that for any elements of the sequences a_i, a_j, b_i, b_j we have $a_i E a_j$ if and only if $b_i E b_j$ if and only if i and j are members of the same partition element J_l. We write $\bar{a} = \langle \bar{a}_0, \ldots, \bar{a}_{n-1} \rangle$ where $\bar{a}_j$ is the set of a_i with $i \in J_j$.

(ii) The model M is decomposed over N if there is an equivalence relation E on $M - N$ such that if $\bar{a}$ is similar for E to $\bar{b}$ and, for each $\bar{a}_i, \bar{b}_i$, we have $\mathrm{tp}(\bar{a}_i; N) = \mathrm{tp}(\bar{b}_i; N)$, then for each $\lg(\bar{a})$-ary relation symbol, R, in the vocabulary of the structure $M \models R(\bar{a}) \leftrightarrow R(\bar{b})$. We say E is an *L-congruence.*

(iii) The L-structure M is *strongly decomposed* over N by E if each equivalence class of E has no more than $|L|$ elements.

(iv) The theory T is (*strongly*) *decomposable* if, for each $M \models T$ and each $N \prec M$ with $|N| \leq |T|$, M is (strongly) decomposed over N.

We will show that if $Q_{\mathrm{II}} \not\leq Q_\psi$, then each structure $(A, \bar{R})$ with $\bar{R} \in \mathscr{R}_\psi(A)$ is strongly decomposed by the following natural equivalence relation. (The hypothesis $Q_{\mathrm{II}} \not\leq Q_4$ is required to show the relation is symmetric.)

2.4.2. Definition. (i) For an element a and a set B, we write $a \in \mathrm{cl}(B)$ if, for some formula $\phi(x)$ with parameters from B, $\phi(a)$ holds and ϕ has only finitely many solutions.

(ii) Let $N \prec M$, then for $a, b \in M - N$, $a \sim_N b$ if $a \in \mathrm{cl}(N \cup \{b\})$.

We will show that such a decomposition suffices for the interpretation of Q_ψ in Q_{1-1} and then that the decomposition exists. For the first task we require a few more definitions.

2.4.3 Definition. Let $C_0, \ldots, C_{m-1}$ be a sequence of subsets of A and let $f_0, \ldots, f_{k-1}$ be a sequence of partial 1–1 functions on A. Then

(i) Two finite sequences $\bar{a}$ and $\bar{b}$ of the same length are $1-1$ *similar* over $\langle D; C_0, \ldots, C_{m-1}; f_0, \ldots, f_{k-1}\rangle$ if for any d in D and any $i < n = \lg(\bar{a})$, $b_i = d$ iff $a_i = d$, and for $j < m$, $a_i \in C_j$ iff $b_i \in C_j$ and for $l < k$, $f_l(a_i) = d\ (\in C_j)$ if and only if $f_l(b_i) = d\ (\in C_j)$.

(ii) The sequence $\langle D; C_0, \ldots, C_{m-1}; f_0, \ldots, f_{k-1}\rangle$ 1–1 *determines* $\phi(\bar{x}, \bar{c}, \bar{R})$ if for any sequences $\bar{a}$ and $\bar{b}$ of the same length we have that if $\bar{a}$ and $\bar{b}$ are 1–1 similar over $\langle D; C_0, \ldots, C_{m-1}; f_0, \ldots, f_{k-1}\rangle$, then $\phi(\bar{a}, \bar{c}, \bar{R}) \leftrightarrow \phi(\bar{b}, \bar{c}, \bar{R})$.

2.4.4 Definition. A formula $\phi(x, y, \bar{n})$ is called a *binding-formula* if, for some integer k, $\models (x)(\exists^{<k}y)\phi(x, y, \bar{n}) \wedge (y)(\exists^{<k}x)\phi(x, y, \bar{n})$.

Note that if M is strongly decomposed via $\sim_N$, then for any pair of elements $a, b \in M - N$, if a and b are equivalent, then for some binding formula M $\phi(x, y, \bar{n})$ with the $\bar{n}$ from N: $\models \phi(a, b, \bar{n})$. Moreover, if $\bar{a}$ is a sequence of equivalent elements from $M - N$, $t(\bar{a}; N)$ is implied by the union of the types $t(a_i; N)$, for $i < n$ with the binding formulas which relate the a_i. Finally, if $\bar{a}$ is a sequence from $M - N$ involving elements from different equivalence classes, then $t(\bar{a}; N)$ is implied by the types of the singleton a_i, the binding formulas which tie together the elements from the same classes and the negations of all binding formulas which might relate pairs that are not in the same class. With this in mind, we will establish a final lemma and complete the proof of the theorem.

2.4.5 Theorem. *If for every infinite A and every $\bar{R} \in \mathscr{R}_\psi(A)(A, \bar{R})$ is strongly decomposable by $\sim_N$, for some proper elementary submodel N of $(A, \bar{R})$, then $Q_\psi \leq Q_{1-1}$.*

Proof. Let $M = (A, \bar{R})$ be strongly decomposed over N. Note that for any $M^* \succ N$, M^* is also strongly decomposed over N. Thus, for any model M of $\mathrm{Th}(N)$ and any $\bar{a} \in M - N$, there is a type $q(\bar{a})$ such that each formula in q contains only one a_i, or is a binding formula, or the negation of a binding formula and is such that $q \vdash t(\bar{a}; N)$ and if $M \models R(\bar{a})$, then $t(\bar{a}; N) \vdash R(\bar{x})$. (The existence of this type is guaranteed by the discussion preceding this lemma.) Now, a standard "double compactness" argument shows that $R(\bar{x})$ is equivalent to a disjunction of formulas over a finite set N_0 and that each of these formulas either contains at most one x_i, or is a binding formula, or is the negation of a binding formula. Now, if D is N_0, C_i picks out the solution set of the ith disjunct with only one x_i; and, for each binding formula $\phi_i(x, y, \bar{d})$, the functions f_i^j for $j < k$ (the number of solutions of $\phi_i(a, x, \bar{n})$) are defined so that $\{f_i^j(a): j < k\} = \{b: \phi_i(a, b, \bar{d})\}$). Then R is 1-1 determined by $D, C_0, \ldots, C_p$ and f_i^j for $i < m$ and $j < k$ (for appropriate p, k, m). ∎

We will complete the proof of Lemma 2.4 by establishing in the rest of Section 2:

2.4.6 Theorem. *If $Q_{\mathrm{II}} \not\leq Q_\psi$, then for every $(A, \bar{R})$ with $\bar{R} \in \mathscr{R}_\psi(A)$ and for some elementary submodel N of $(A, \bar{R})$, A is strongly decomposed by $\sim_N$.* ∎

We will explain two proofs of the above result. The first is both the most natural and the most useful. We will continue to use its methods later in the paper. However, it requires a minimal knowledge of stability theory (for instance, the first half of Lascar–Poizat [1979]) so for those who might be unfamiliar with those basic facts, we have included in Section 2.6 an *ad hoc* but self-contained proof of Theorem 2.4.6.

2.5 Theorem. *If $Q_{\mathrm{II}} \not\leq Q_\psi$, then for every A and every $\bar{R} \in \mathscr{R}_\psi(A)$, $(A, \bar{R})$ is strongly decomposable.* (1st Proof).

We first observe

2.5.1 Lemma. *$Q_{\mathrm{II}} \not\leq Q_\psi$ implies T is stable.*

We give two arguments for this. Note that T being unstable implies there is a definable linear ordering of n-tuples. In Chapter VIII of Baldwin–Shelah [1982] it is shown that in any theory with a definable linear order on n-tuples one can monadically define a linear order on singletons. From this one constructs an equivalence relation with infinitely many infinite classes and finishes by Lemma 1.3.3. Alternatively, we use more of the machinery set up in Section 2.6 and deduce directly from the definable linear order on n-tuples the existence of a definable equivalence relation on n-tuples with infinitely many non-pseudofinite (see Definition 2.6.3) classes which contradicts Lemma 2.6.4. ∎

2.5.2 Definition (The Fundamental Equivalence Relation). Let $N \prec M$ and M a model of a stable theory. We define a relation E_N on $M - N$ by aE_Nb just if $t(a; N \cup b)$ forks over N.

Now the standard properties of forking in a stable theory assure us that E is reflexive and symmetric. In general, E is not transitive. However, in our situation we obtain this and more.

2.5.3 Lemma. *If T is stable, $N \prec M$ and E is the fundamental equivalence relation then M is decomposed over N by E.*

Proof. We must show that E is an equivalence relation and that condition (ii) of Definition 2.4.1 is satisfied. We will give a brief outline of the argument.

2.5.4 Lemma. *Suppose that in a model of T, there exists an element a and* $B = \langle b_i : i < \omega \rangle$ *and* $C = \langle c_j : j < \omega \rangle$ *such that*

(i) *B is a set of indiscernibles;*
(ii) *C is a set of indiscernibles over B and there is a formula* $\phi(x, y, z)$ *such that* $\models \phi(a, b_i, c_j)$ *if and only if* $i = j$.

Then T is codable.

This lemma is an easy reworking of the definition of codable given in Definition 1.3.1. Its proof as well as that of the following lemma are detailed as Sections IV.2.4 and IV.2.6 of Baldwin–Shelah [1982]. The following lemma is a fairly routine calculation using the properties of the forking relation and Lemma 2.5.4.

2.5.5 Lemma. *If T is stable and either:*

(i) *There exists a subset A of a model of T and elements a, b, c such* $t(a; A \cup b)$ *forks over A and* $t(b; A \cup c)$ *forks over A, but* $t(a; A \cup c)$ *does not fork over A, or*
(ii) *There exists a subset A of a model of T and elements* $a, b_1, \ldots, b_n$ *such that for each i* $t(a; A \cup b_i)$ *does not fork over A but* $t(a; A \cup \{b_1, \ldots, b_n\})$ *forks over A.*

Then T is codable.

This result shows that if $Q_{\text{II}} \not\leq Q_{\text{mon}}$ and $\bar{R} \in \mathscr{R}_\psi(A)$, then $\text{Th}(A, \bar{R})$ is decomposable (see Baldwin–Shelah [1982]). In order to show that it is actually strongly decomposable, we will need one further fact from stability theory.

2.5.6 Lemma. *If T is stable and there exist* $a, b \in A \models T$ *and* $B \subseteq A$ *such that* $t(a; B \cup b)$ *forks over B but* $t(b; B \cup a)$ *is not algebraic, then on some subset of a model of T there is a definable equivalence relation which has infinitely many infinite classes.*

(This result is Lemma VI.1.1 of Baldwin–Shelah [1982].) Now by Definition 1.2.4 we see the conclusion of Lemma 2.5.6 cannot hold unless $Q_{\text{II}} \leq Q_\psi$ (as we have $Q_{1-1} \leq A_\psi$). Thus, we have established Theorem 2.5.

We turn now to the other proof of Theorem 2.5.

2.6 Theorem. *If* $Q_{\text{II}} \not\leq Q_\psi$, *then for every A and every* $\bar{R} \in \mathscr{R}_\psi(A)$, $(A, \bar{R})$ *is strongly decomposable.* (2nd Proof)

We first use an argument similar to the ones given in Baldwin–Shelah [1982] to show that for any model N, $\sim_N$ is symmetric and thus is an equivalence relation. For this, we will require a few other concepts. The first is given in

2.6.1 Definition. Let $N \subseteq A$ and $p \in S(A)$, then p is *finitely satisfied in N* if every formula in p has a solution in N.

Using compactness it is easy to see that if $A \subseteq B \subseteq C$ and $p \in S(B)$ is finitely satisfied in A, then p extends to a $p' \in S(C)$ which is also finitely satisfied in A. Next, we consider

2.6.2 Lemma. *If $Q_{\mathrm{II}} \not\leq Q_\psi$ and $\bar{R} \in \mathscr{R}_\psi(A)$, then for any $N \prec (A, \bar{R})$ if neither a nor b is algebraic in N and $t(a, N \cup b)$ is algebraic, then $t(b, N \cup a)$ is also algebraic.*

Proof. Suppose not and choose b_i for $i < \omega$, which are distinct, with $t(b_i; N \cup a) = t(b; N \cup a)$. Let $\bar{c}_0 = \bar{b}a$ and choose $\bar{c}_i$ for $i < \omega$ such that $t(\bar{c}_i; C_i) = t(\bar{c}_{i+1}; C_i)$ and $t(\bar{c}_{i+1}; N \cup C_{i+1})$ is finitely satisfied in N. Here $C_i = N \cup \{\bar{c}_j : j < i\}$. Thus $\bar{c}_i$ has the form $\langle b_{i,j} : \text{for } j < \omega\rangle a_i$. Clearly a_i is algebraic in each $b_{i,j}$ by the same formula ψ. But no a_j is algebraic in $N \cup b_{i,k}$ with $i > j$. For, if it were, we could, by finite satisfiability, find a $b' \in N$ with a_i algebraic in $N \cup b'$ and hence, in N, also, which is impossible. But no a_i can be algebraic in $b_{j,k}$, with $j < i$ since all a_l with $l > i$ realize the same type as a_i over $N \cup b_{j,k}$. Thus, by adding predicates A and B to pick out the a's and b's, we can define an equivalence relation on B with infinitely many infinite classes by $E(x, y) \leftrightarrow (\exists z)\phi(x, z) \wedge \phi(y, z) \wedge A(z)$. This contradicts Lemma 1.3.3 and establishes the lemma. □

We now want to show that if N is chosen appropriately, then $\sim_N$ actually determines a strong decomposition of $(A, \bar{R})$. To accomplish this, we return to the original Shelah argument. We will proceed by extending the properties of strongly minimal sets to finite sequences. We will accordingly arrive at a notion reminiscent of the weakly minimal formulas that are examined in Shelah [1974a].

2.6.3 Definition. The family $F = \{\bar{f}_i : i < \alpha\}$ is *pseudo-finite*, if there is a finite set C such that for every i, $C \cap \bar{f}_i \neq \varnothing$.

The formula $\phi(\bar{x}, \bar{a}, \bar{R})$ is *pseudo-algebraic* in $(A, \bar{R})$ if its solution set is pseudo-finite. The sequence $\bar{a}$ is *pseudo-algebraic over B*, if for some formula $\phi(\bar{x})$ with parameters from B, $\models \phi(\bar{a})$ and ϕ is pseudoalgebraic.

Note that $\bar{a}$ is *not-pseudo-finite* over B means that we can find arbitrarily many disjoint finite sequences which realize $t(\bar{a}; B)$.

2.6.4 Lemma. *If $Q_{\mathrm{II}} \not\leq Q_\psi$, then for any A and any $\bar{R} \in \mathscr{R}_\psi(A)$, there is no formula $\phi(\bar{x}, \bar{y}, \bar{c}, \bar{R})$ which defines an equivalence relation with infinitely many non-pseudo-finite equivalence classes.*

Proof. If $\lg(\bar{x}) = \lg(\bar{y}) = 1$, then this assertion is only Lemma 1.3.3. Using $Q_{1\text{-}1} \leq Q_\psi$, we will reduce the case $n > 1$ to the case $n = 1$ and thus finish the argument. By induction choose sequences $\bar{a}_{i,j}$ such that $\bar{a}_{i,j}$ is equivalent to $\bar{a}_{k,l}$ just if $i = k$ and such that the $\bar{a}_{i,j}$ having distinct indices are pairwise disjoint and all are disjoint from c. Now, define for each $m < n$ a permutation f_m of A which exchanges the first and mth members of each sequence $\bar{a}_{i,j}$ and which fixes all other elements of A. Let B^* consist of the first coordinates of the $\bar{a}_{i,j}$. Now the formula

$$\phi^*(x, y, \bar{c}, \bar{R}, f_1, \ldots, f_n) = \phi(f_1(x), \ldots, f_n(x), f_1(y), \ldots, f_n(y), \bar{c}, \bar{R})$$

defines on B^* an equivalence relation with infinitely many infinite equivalence classes. This is, of course, contrary to Lemma 1.3.3 and we are done.

2.6.5 Definition. The formula $\psi(\bar{x}, \bar{c}, \bar{R})$ is $\phi(\bar{x}, \bar{y}, \bar{r})$-minimal, if ψ is not pseudo-finite but for every $\bar{d}$ either $\psi(\bar{x}) \wedge \phi(\bar{x}, \bar{d}, \bar{R})$ or $\psi(\bar{x}) \wedge \neg\phi(\bar{x}, \bar{d}, \bar{R})$ is pseudo-finite.

The search for a ϕ-minimal formula is similar to the search for a strongly minimal formula in an ω-stable theory. We will show that we cannot build a complete binary tree of instances of ϕ and negations of ϕ such that each path is not pseudo-finite. The main step for this is

2.6.6 Lemma. *If $Q_{\mathrm{II}} \not\leq Q_\psi$ then there are no A and $\bar{R} \in \mathscr{R}_\psi(A)$ such that there exist a $\phi(\bar{x}, \bar{y}, \bar{R})$ and $\bar{a}_n$ for $n < \omega$ so that for each $n < \omega$, the formula*

$$\theta_n = \bigwedge_{m<n} \phi(\bar{x}, \bar{a}_m, \bar{R}) \wedge \neg\phi(\bar{x}, \bar{a}_n, \bar{R})$$

is not pseudo-algebraic.

Proof. Assume that the lemma fails. By the compactness theorem, we can assume that each θ_n is satisfied by more than $2^{\aleph_0}$ disjoint sequences. Let B be the collection of elements which appear in any of the parameter sequences $\bar{a}_n$. Define two sequences $\bar{b}, \bar{c}$ from A to be e equivalent just if for every $\bar{a}$ from B $\phi(\bar{b}, \bar{a}, \bar{R}) \leftrightarrow \phi(\bar{c}, \bar{a}, \bar{R})$. Now, for each n and m, if $n \neq m$ a sequence satisfying θ_m and a sequence satisfying θ_n are not equivalent so that e has infinitely many classes. But each of these classes is not pseudo-finite. For, there are more than $2^{\aleph_0}$ disjoint sequences satisfying θ_n and at most (since B is countable) $2^{\aleph_0}$ classes of e so that some e-class intersects θ_n in uncountably many disjoint sequences and thus that class is not pseudo-finite. Thus, for each n, we find a distinct class of the definable equivalence relation e which is not pseudo-finite. By Lemma 2.6.4, $Q_{\mathrm{II}} \leq Q\psi$. □

2.6.7 Lemma. *If $Q_{\mathrm{II}} \not\leq Q_\psi$, then for any $\phi(\bar{x}, \bar{y}, \bar{r})$ there is an integer $m(\phi)$ and there are formulas $\chi_i(\bar{x}, \bar{z}, \bar{r})$ (depending on ϕ) for $i < m(\phi)$, such that for any A and any $\bar{R} \in \mathscr{R}_\psi(A)$, there is a $\bar{c} \in A$ such that the formulas $\chi_i(\bar{x}, \bar{c}, \bar{R})$ partition A and each $\chi(\bar{x}, \bar{c}, \bar{R})$ is ϕ-minimal.*

Proof. Build a binary tree of instances of $\phi(\bar{x}, \bar{y}, \bar{R})$ and its negation. Either, for some n, each path of length n defines a ϕ-minimal set; or, for arbitrary k, we can find $\bar{a}_i$ for $i < k$ such that taking $\lambda_i(\bar{x}, \bar{y}, \bar{R})$ as $\phi(\bar{x}, \bar{y}, \bar{R})$ or $\neg\phi(\bar{x}, \bar{y}, \bar{R})$ (depending on i) $\wedge \{\lambda_i(\bar{x}, \bar{a}_i, \bar{R}) : i < k\}$ is not pseudo-algebraic. If $k = 2m + 2$, the formula θ_m from Lemma 2.6.6 is not pseudo-algebraic and we violate Lemma 2.6.6.

We will need one more nice property of pseudo-algebraic formulas to complete the proof.

2.6.8 Lemma. If $\bar{a} = \langle a_0, \ldots, a_n \rangle$ *is pseudo-algebraic over B, then some a_i is algebraic over B.*

Proof. Let $\phi(\bar{x}, \bar{b}, \bar{R})$ be a pseudo-algebraic formula satisfied by $\bar{a}$. Let C be a set with minimal cardinality n such that if $A \models \phi(\bar{a}', \bar{b}, \bar{R})$, then $\bar{a}^* \cap C \neq \emptyset$. Recall from Example 2.1.5 that if $\theta_n(z_0, \ldots, z_{n-1}, \bar{y}, \bar{r})$ is

$$(x)\left[\phi(x_0, \ldots, x_{m-1}, \bar{y}, \bar{r}) \rightarrow \bigvee_{\substack{m<n \\ j<n}} x_i = y_j\right],$$

then $\{\theta_n: n < \omega\}$ is a malleable family. Now, by applying Lemma 2.1.6, we see that some component of $\bar{a}$ satisfies the algebraic formula $\theta^*(x, \bar{b}, \bar{R})$ and we are done. ∎

2.6.9 Theorem. *If $Q_{\mathrm{II}} \not\leq Q_\psi$, then for any A and any $\bar{R} \in \mathscr{R}_\psi(A)$, there is an elementary submodel N of $(A, \bar{R})$ such that $\sim_N$ strongly decomposes $(A, \bar{R})$ over N.*

Proof. For each $\phi(\bar{x}, \bar{y}, \bar{r})$, choose a sequence $\bar{c}$ and formulas χ_i as in Lemma 2.6.7 and let N contain all the $\bar{c}$. By induction on n we will prove that if $\bar{a}$ and $\bar{b}$ with length n are similar for $\sim_N$ and for each $\bar{a}_i, \bar{b}_i$ (see notation in Definition 2.4.1) $t(\bar{a}_i; N) = t(\bar{b}_i; N)$, then $t(\bar{a}; N) = t(\bar{b}; N)$. If $n = 1$, this assertion is tautogical. Suppose that we have proved the claim for n. To prove it for $n + 1$, we consider a formula $\phi(x, \bar{y}, \bar{z}, \bar{r})$ with $\lg(\bar{y}) = n$, and let $\bar{n}$ be in N. If all elements of $\bar{a}$ are in the same $\sim_N$ equivalence class, then there is nothing to prove. Let $\bar{a}_1$ be a maximal pairwise equivalent subsequence of $\bar{a}$—as is indeed implied by our notation. Then, if we let $\bar{a}'$ (respectively $\bar{b}'$) denote $\bar{a}$ without $\bar{a}_1$ (respectively $\bar{b}$ without $\bar{b}_1$), no component of $\bar{a}'$ (respectively $\bar{b}'$) is algebraic in $N \cup \bar{a}_1$ (respectively in $N \cup \bar{b}_1$) and thus $\bar{a}'$ ($\bar{b}'$) is not pseudo-algebraic in $N \cup \bar{a}_1$ (in $N \cup \bar{b}_1$), (by Lemma 2.6.8).

We must prove that for any $\bar{n} \in N$, $A \models \phi(\bar{a}_1, \bar{a}', \bar{n}, \bar{R}) \leftrightarrow \phi(\bar{b}_1, \bar{b}', \bar{n}, \bar{R})$. By Lemma 2.6.7 and the choice of N, we can find a $\bar{d} \in N$ and a ϕ-minimal $\chi(\bar{x}, \bar{d}, \bar{R})$ such that $A \models \chi(\bar{a}', \bar{d}, \bar{R})$. By the definition of ϕ-minimality, one of

$$\chi(\bar{x}, \bar{d}, \bar{R}) \wedge \phi(\bar{a}_1, \bar{x}, \bar{c}, \bar{R}) \quad \text{and} \quad \chi(\bar{x}, \bar{d}, \bar{R}) \wedge \neg\phi(\bar{a}_1, \bar{x}, \bar{c}, \bar{R})$$

is pseudo-algebraic. Without loss of generality, we can take it to be the second one. By a simple application of compactness, this means that for some $m_1(\phi)$, the formula is satisfied by no more than $m_1(\phi)$ pairwise disjoint sequences. As $\bar{a}'$ is not pseudo-algebraic over $\bar{a}_1$, we have $A \models \phi(\bar{a}_1, \bar{a}', \bar{n}, \bar{R}')$. By induction, $\bar{a}'$ and $\bar{b}'$ have the same type over N so $A \vdash \chi(\bar{b}', \bar{d}, \bar{R})$. Since $\bar{a}_1$ and $\bar{b}_1$ have the same type over N, $\chi(\bar{x}, \bar{d}, \bar{R}) \wedge \neg\phi(\bar{b}_1, \bar{x}, \bar{n}, \bar{R})$ is not satisfied by more than $m_1(\phi)$ pairwise disjoint sequences. Since $\bar{b}'$ is not pseudo-algebraic over $N \cup \bar{b}_1$, we thus have $A \models \phi(\bar{b}_1, \bar{b}', \bar{n}, \bar{R})$ as was required. ∎

3. *Infinitary Monadic Logic and Generalized Products*

Our primary focus so far has been on the classification of theories of equality, Q_ψ. Now we will consider the following question: What are the possibilities for theories of the form (T, Q_ψ), where T is a complete first-order theory and Q_ψ is one of the four second-order quantifiers? The notion of a decomposable model is a key tool in the proof of Lemma 2.4. We will develop a generalization of this idea and use it, for example, to compute the Hanf numbers of some logics (see Sections 4.5 and 5.2). The major device for these computations is a Feferman–Vaught type theorem for monadic logic. As Gurevich pointed out to me, this is a natural development of the original Feferman–Vaught theorem which described the first-order properties of a generalized product of a family $\{M_i: i \in I\}$ in terms of the first-order theory of the factors and the monadic theory of the index set (enriched by unary predicates which pick out the indices whose models have the same theory). The material in this section is largely taken from Shelah [1975e] and Gurevich [1979a].

In many cases, it is artificial to consider the first-order monadic theory of a class of structures, because this theory already encodes a certain amount of information that we would normally think of as "$L_{\omega_1,\omega}$" information. For example, we can monadically define the closure of a subset of a group. Or, consider the class of all structures containing two infinite classes, P_0, P_1, and a binary extensional relation, E, between them. (That is to say, one is the set of subsets of the other). Now, if T is the monadic theory of this class, any model of the monadic sentence

$$(X) \subseteq P_0(\exists y) \in P_1(z) \in P_0(z \in X \leftrightarrow z \in y)$$

has models only of power $\geq \beth_1$ This kind of argument shows that the Hanf number of $L_{\omega,\omega}(Q_{\mathrm{mon}}) \geq$ the Hanf number of $L_{\omega_1,\omega}$; furthermore, it leads us to consider infinitary monadic logic. We are going to prove a Feferman–Vaught type theorem by way of a back-and-forth argument. This requires some means of handling variables. Rather than deal with variables explicitly we will expand the language by adding additional constant symbols. Since this is monadic logic, we must add not only names for individuals but for subsets as well. We want to describe a specific sentence in $L_{\infty,\lambda}(Q_{\mathrm{mon}})$ which contains the information we need in order to make our induction. Individuals are considered to be subsets with only one element. Note that if (A, R) and (B, R) are equivalent for existential first-order sentences, then R is a singleton in A iff it is a singleton in B.

This section repeats the discussion in Section 3 of Chapter XIII in a superficially more general situation. The chief differences here are that Chapter XIII restricts itself to finitary logic and, for expository purposes, merely works out the preservation theorem for ordered sums. Here, however, we will give an abstract notion of product in Section 3.4, a notion which focuses attention on exactly those properties (for example, of the ordered sum construction) which allow the argument for the preservation theorem to go through. In Chapter XIII monadic logic is interpreted into a first-order logic; here, on the other hand, the monadic logic is taken as basic. The following glossary connects the two chapters.

Chapter XIII	Chapter XII
a sequence ξ	an ordinal $\alpha = \lg(\xi)$
an l-tuple of elements	a λ-tuple of elements
ξ – $\mathrm{Th}(M, a_1, \ldots, a_l)$	$t_{\alpha,\lambda}(M, \bar{Q})$
ξ – l-Box	$t_{\alpha,\lambda}(L)$

Observe that the correspondence suggested by the tabular arrangement is not exact since a $\xi - l$-Box depends on a (suppressed) theory T.

$X_1, \ldots, X_l$	$\langle Q_t : t \in \lambda \rangle$,
$P(\xi, X, t)$	$Q_t(I)$,
$P(\xi, X)$	$\langle Q_t(I): t \in t_{\alpha, \lambda}(L) \rangle$.

Another difference in the presentation of results arises from the fact that one chapter emphasizes decidability results, while the other stresses preservation results. In Chapter XIII, the bounded theories are viewed as objects in their own right and the ξ-theory of the product is computed from the $H(\xi, l)$ theory of the index set. In this chapter, however, the bounded theories are viewed as properties of structures and the theorem has the following form: If *the bounded theories of two index structures are the same, then so are the theories of the product structures.*

3.1 Definition. We define by induction the set of formulas $t_{\alpha, \lambda}(M)$ as follows:

(i) For any L-structure M, let $t_{0, \lambda}(M) = \{\theta : M \models \theta\}$.

Here θ ranges over existential first-order formulas with at most λ variables. Note that $t_{0, \lambda}(M)$ is the same for all infinite λ. We would just say the existential theory of M, but the decidability results require that if λ is finite, then so is $t_{0, \lambda}$. We require existential rather than quantifier-free formulas in $t_{0, \lambda}(M)$ in order that we may know the cardinality (mod $\aleph_0$) of every subset of M defined by a boolean combination of unary predicates.

Now, for any α and λ, we define $t_{\alpha, \lambda}(M)$ as follows: $t_{\alpha+1, \lambda}(M) = \{t_{\alpha, \lambda}(M, \bar{Q}): \lg(Q) = \lambda\}$

$$t_{\delta, \lambda}(M) = \bigcup \{t_{\alpha, \lambda}(M): \alpha < \delta\}, \text{ if } \delta \text{ is a limit ordinal.}$$

(ii) For any α and λ, let $t_{\alpha, \lambda}(L)$ denote $\{t_{\alpha, \lambda}(M): M \text{ is an } L\text{-structure}\}$.

Thus, $t_{\alpha+1, \lambda}(M)$ describes the $L^{\alpha}_{\infty, \lambda}(Q_{\mathrm{mon}})$-theory of the expansion of M by λ unary predicates. Similarly, $t_{\alpha, \lambda}(L)$ denotes the set of all possible $L^{\alpha}_{\infty, \lambda}(Q_{\mathrm{mon}})$-theories.

Observe here that if α, λ, and L are finite, then so is $t_{\alpha, \lambda}(L)$. Moreover, for each L-structure M, $t_{\alpha, \lambda}(M)$ is equivalent (that is, it holds of the same structures) to a sentence in $L_{\beth(\alpha, \lambda + |L|), \lambda}$. The following lemma illustrates the expressive power of the $t_{\alpha, \lambda}(M)$. And, interestingly enough, it also provides the key technical step for our Feferman–Vaught like theorem.

3.2 Lemma. *Let λ, λ' and κ be cardinals with $\lambda + \lambda' \leq \kappa$. Let $\mathscr{I}$ and $\mathscr{J}$ be structures (having, for the sake of simplicity, a finite language) and universes I and J respectively. Suppose the sets I and J are partitioned by the sequences $Q_t(I)$, $Q_t(J)$, respectively, for $t \in \lambda$, and suppose further that*

$$t_{\alpha+1, \kappa}(\langle \mathscr{I}, Q_t(I): t \in \lambda \rangle) = t_{\alpha+1, \kappa}(\langle \mathscr{J}, Q_t(J): t \in \lambda \rangle).$$

If $\langle X_i: i \in \lambda' \rangle$ *is a partition of* I *refining the partition* $\langle Q_t(I): t \in \lambda \rangle$, *then there exists* $\langle Y_i: i \in \lambda' \rangle$, *a partition of* J, *such that*:

$$t_{\alpha,\kappa}(\langle \mathscr{I}, Q_t(I), X_i: t \in \lambda, i \in \lambda' \rangle) = t_{\alpha,\kappa}(\langle \mathscr{J}, Q_t(J), Y_i: t \in \lambda, i \in \lambda' \rangle). \quad \square$$

3.3 Generalized Products. We begin our treatment of the Feferman–Vaught theorem by giving a rather "soft" definition of a generalized product. This notion differs from that in Feferman–Vaught in several respects. Perhaps the most basic is that it is designed to describe only operations taking a set of L-structures to an L-structure. Thus, the definition focuses on the relation between the truth of basic relations in the language L (as opposed to arbitrary definable relations) in the factor structure and the product structure. The intent of this definition is to emphasize those properties of the definition of the basic relations in the product structure which allow the assertion, "truth of basic relations depends on truth in the factors" to propagate to, "truth of all sentences in first-order logic (in infinitary monadic logic) depends on their truth in the factors". This definition is abstracted from the accounts of the monadic preservation theorem in Shelah [1975e] and Gurevich [1979b]. The emphasis here differs from that in Feferman [1972] where the role of functors from one similarity type to another is of central importance.

Examples of the notion of generalized product defined here—not of minor modifications of it—include direct product, disjoint union, ordinal sum of linear orderings, ultraproduct, and reduced product. Observe that in the last two, the language for the index set contains symbols binding subsets. Note also that the notion we are here examining does not include the concept of a sheaf over a boolean space.

Following is the key idea of the definition. Since we are going to give a proof by induction on quantifiers, we must describe how the product operation behaves with respect to structures obtained by naming elements and—since we will work in monadic logic—subsets. In fact, the notion of projection which we formulate below would be harder to explicate if we were to deal with elements rather than with sets since (for example, in disjoint unions) we frequently want to project to the empty set.

3.4 Definition. A *generalized product* is a function (or a family of functions) which, to each language L and each sequence $\langle A_i: i \in I \rangle$ of L-structures, assigns an L-structure $F(\langle A_i: i \in I \rangle) = A^*$ satisfying the following conditions:

(i) For each i there exists a function $\rho_i: \mathscr{P}(A^*) \to \mathscr{P}(A_i)$ such that if $\bar{P}$ is a sequence of subsets of A^*, then

$$F(\langle (A_i, \rho_i(\bar{P})) : i \in I \rangle) = \langle F(A_i: i \in I), \bar{P} \rangle.$$

(ii) For any sequence $\bar{a}$ (of arbitrary length $< |A^*|$) and for each L-symbol R, letting $K_R(\bar{a}) = \{i: A_i \models R(\rho_i(\bar{a}))\}$ and analogously in B^*, if $t_{0,\lambda}(\langle \mathscr{I}, K_R(\bar{a}) \rangle) = t_{0,\lambda}(\langle \mathscr{J}, K_R(\bar{b}) \rangle)$, then $A^* \models R(\bar{a})$ if and only if $B^* \models R(\bar{b})$.

Here and below, when $\bar{a}$ is a sequence of individuals, we will simply write $\rho_i(\bar{a})$ for $\langle \rho_i(\{a_0\}), \ldots, \rho_i(\{a_{k-1}\})\rangle$. Each $\rho(\{a_i\})$ has cardinality at most 1. Now we can state our version of the Feferman–Vaught theorem. The proof is similar to that of Theorem 2 of Chapter XIII.

3.5 Theorem (Preservation Theorem). *Suppose F is a generalized product operation and suppose also that $\langle A_i: i \in I\rangle$ and $\langle B_j: j \in J\rangle$ are families of L structures. For $t \in t_{\alpha,\lambda}(L)$, let $Q_t(I) = \{i: t_{\alpha,\lambda}(A_i) = t\}$ and let $Q_t(J) = \{j: t_{\alpha,\lambda}(B_j) = t\}$. Moreover, let $W = \{t_{\alpha,\lambda}(A_i): i \in I\} \cup \{t_{\alpha,\lambda}(B_j): j \in J\}$. There exists a $\kappa = \kappa(\alpha, |W|)$ such that if*

$$t_{\alpha,\kappa}(\langle \mathscr{I}, Q_t(I): t \in W\rangle) = t_{\alpha,\kappa}(\langle \mathscr{J}, Q_t(J): t \in W\rangle)$$

then

$$t_{\alpha,\lambda}(A^*) = t_{\alpha,\lambda}(B^*). \quad \square$$

As a corollary, we get a result mentioned in Chapter IX.

3.6 Corollary. *If κ is strongly inaccessible, then $L_{\kappa,\lambda}$-equivalence is preserved by generalized product.*

Proof. If $\phi \in L_{\kappa,\lambda}(Q_{\text{mon}})$ then for some $\mu < \kappa$, $\phi \in L_\mu$. But then $\phi \in L^{\alpha}_{\mu,\lambda}(Q_{\text{mon}})$ where $\alpha < \mu^+$ (this is a straightforward computation). Thus, the truth of ϕ in M is determined by $t_{\alpha,\lambda}(M)$ which is equivalent to a formula in $L_{\kappa,\lambda}$ since κ is strongly inaccessible. $\square$

This argument also yields the results in 3.3.4, Corollary 2.3.5, and 3.3.6 of Chapter IX.

We will now describe a generalization of disjoint union which is the example of generalized product that is of most use in the study of second-order quantifiers. This is a generalization of the notion of decomposition that was employed in Lemma 2.4. If we form a disjoint union, no relation holds between sequences $\bar{a}$, $\bar{b}$ from different constituents of the union. We want to allow such relations to hold but we also want to require that whether $R(\bar{a}, \bar{b})$ holds shall depend only the separate properties of $\bar{a}$ and $\bar{b}$. To make this notion precise, we require several preliminary definitions.

3.7 Definition. (1) If $\langle M_i: i \in I\rangle$ is a sequence of L-structures with $M_i \cap M_j - N$, we call the M_i a *sequence with heart N*.

(2) Let $\langle M_i: i \in I\rangle$ be a sequence with heart N. To define *the free union (with respect to σ) over N of the M_i*, we first need the following auxilliary notions:

(i) An n-condition τ is a pair $\langle P, \langle \phi_0, \ldots, \phi_{k-1}\rangle\rangle$ consisting of a partition, P, of n into sets $P_0, \ldots, P_{k-1}$ and a k-tuple of first order formulas such that ϕ_i has $|P_i|$ free variables.

(ii) σ is a map which assigns to each m-ary relation symbol R of L a finite set of m-conditions.

(iii) Let M be $\bigcup\{M_i: i \in I\}$. If $\bar{a} \in M$, then $\bar{a}$ satisfies the n-condition $\langle P, \langle \phi_0, \ldots, \phi_{k-1}\rangle\rangle$ if for some $M_{i_0}, \ldots, M_{i_{k-1}}$, we have $P_j = \{m: a_m \in M_{i_j}\}$; and, letting $\bar{a}_j = \{a_m: m \in P_j\}$, taken in increasing order of subscript, $M_{i_j} \models \phi_j(\bar{a}_j)$.

Now the free union of the M_i over N is the structure whose universe is $\cup \{M_i: i \in I\}$, where $R^M = \{\bar{a}: \bar{a}$ satisfies an m-condition in $\sigma(R)\}$.

It is easy to see that such a free union satisfies the definition of generalized product. Technically, we note that one must make allowance for the amalgamation, but this is straightforward. The details of Theorem 3.3.5 are, in this special case, carried out in III.1.13 in Baldwin–Shelah [1982]. In that paper, the free union is defined in terms of $t(\bar{a}; N)$. An easy application of compactness shows that when every model of T containing N can be decomposed in the sense of Definition 2.4.1, then each such model can, in fact, be written as free union over N in the sense of 3.7.3.

4. *The Comparison of Theories*

This section discusses a nuance in Shelah's argument, reported in Theorem 1.2.5, that $Q_{\mathrm{II}} \not\leq Q_{1-1}$. Namely, we consider the exact rôle of the assertion that interpretations preserve Hanf number. We show that a similar in form but technically easier argument shows $Q_{\mathrm{II}} \not\leq (\mathrm{Th}(<, Q_{\mathrm{mon}})$, the monadic theory of order. This last remark is apparently paradoxical in the light of the proof (Gurevich–Shelah [1982]) that it is consistent to interpret Q_{II} into $(\mathrm{Th}(<), Q_{\mathrm{mon}})$. To resolve this paradox we must distinguish the usual notion of interpretation from the stronger notions used in this paper.

4.1 Definition. The theory T_1 is *syntactically-interpretable* in the logic T_2 if there is a map* assigning to each T_1-sentence ϕ a T_2-sentence ϕ^* such that $T_1 \vdash \phi$ iff $T_2 \vdash \phi^*$.

Clearly, if* is recursive the Turing degree of T_1 is less than or equal to the Turing degree of T_2. However, this map need not preserve model-theoretic properties. Thus, using the Feferman–Vaught theorem for monadic logic, we will show that the Hanf number for monadic sentences on linearly ordered models (the Hanf number of the monadic theory of order) is strictly less than the Hanf number of second-order logic. It is easily seen that this implies that there can be no strong interpretation (in the sense of Definition 3.2) of Q_{II} into $\mathrm{Th}(<)$, Q_{mon}, (see Baldwin–Shelah [1982, VIII.2.12]). Nevertheless, Gurevich–Shelah [1982] have shown that it is consistent—indeed, it follows from the GCH—that there be a syntactic interpretation of Q_{II} into the monadic theory of order. The reader should consult Chapter XIII for more details on the monadic theory of order.

Several variants on the notion of interpretation and their roles are discussed in Baldwin–Shelah [1982]. We will use here only interpretations which satisfy the following condition.

4.2 Definition. The logic T_1 is *semantically interpretable* in the logic T_2 if there exist a pair of maps (both denoted by *) taking T_1-sentences to T_2-sentences and the models of T_2 onto the models of T_2 such that:

(i) $M \subseteq M^*$;
(ii) $M \models \phi$ iff $M^* \models \phi^*$.

If, in addition, we have

(iii) $|M^*|$ can be computed from $|M|$,

then we say T_1 is *strongly semantically interpretable* into T_2.

We will now show how bounds on the Hanf number of a theory can be used to show that there is no strong semantic interpretation of one theory into another. This, however, requires the technical notion given in

4.3 Definition. We say that the Hanf number of T_1 is bounded in terms of the Hanf number of T_2 and write $B(T_1, T_2)$ if there is a second-order definable function $f(x)$ such that $H(T_1) \leq f(H(T_2))$.

Observe that this relation is obviously transitive. Now, if $B(Q_{\mathrm{II}}, T)$, it is fairly easy to see that there can be no strong semantic interpretation of Q_{II} into T. Since our notion of $\leq$ is a strong semantic interpretation, this gives a more general explanation for Theorem 1.2.5. We will now show that that theorem can be extended to the monadic theory of order.

In some respects, Silver [1971] begins this program with his explicit computation of an upper bound for the Hanf number for logic with the well-ordering quantifier (Chapter XVII). This shows that fewer classes of cardinals are characterized as cardinals (that is, as, well-ordered sets) in the monadic theory of order than in second-order logic. This leaves open the possibility that we might be able to characterize the missing classes as sets of cardinals in which a sentence in the monadic theory of order has a model (although not necessarily a well-ordered one).

We use the following notation.

4.4 Notation. We denote the Hanf number of (Th(<), Q_{mon}), the monadic theory of well-orderings, and Q_{II} respectively by HL, HW, and H_{II}.

We write $H(T)$ for the Hanf number of theory T. If $H(T)$ can be bounded by a cardinal definable in second-order logic (for example, HW), then H_{II} cannot be bounded in terms of $H(T)$. As, we would then have a second-order definable bound on H_{II}, which is clearly impossible. Thus, the assertion $HL < H_{\mathrm{II}}$ follows immediately from the next lemma.

4.5 Lemma. *HL is bounded in terms of HW.*

Proof. Specifically, we will show that $HL \leq \Sigma\,\{2^\lambda : \lambda < HW\}$. Let $(M, <)$ be a linear order and suppose that λ can be embedded in M as $\langle a_i : i < \lambda \rangle$. Now, M is a free union (in the sense of Theorem 2.6) of the intervals determined by the a_i. For any fixed monadic sentence ϕ, say with quantifier depth m, we can find a k such that $t_{2,k}(\lambda, Q_t(\lambda))$, where t ranges over the finitely many monadic theories of quantifier depth n, determines whether M satisfies ϕ. Since $\lambda > HW$, we can replace λ with an arbitrarily large λ' with $t_{2,k}(\lambda, Q_t(\lambda)) = t_{2,k}(\lambda, A_t)$ for appropriate subsets A_t of λ'. But it is an easy matter to find an M' such that M' is a free union of intervals indexed by λ' and so that $Q_t(\lambda') = A_t$. But then $M' \models \phi$.

Since, for any linear order M, if $|M| > 2^\lambda$, there is an order embedding of either λ or λ^* into M—and since the preceding argument works equally well for λ^*—we see $HL \leq \Sigma\{2^\lambda: \lambda < HW\}$.

Clearly, if Q_{II} could be strongly interpreted in $(\mathrm{Th}(<), Q_{mon})$, then H_{II} would be bounded in terms of HL. Thus, we have

4.6 Theorem. *There is no strong semantic interpretation of* Q_{II} *into* $(\mathrm{Th}(<), Q_{mon})$. □

5. *The Classification of Theories by Interpretation of Second-Order Quantifiers*

We will not investigate the partial order of interpretability among theories (T, Q_ψ). That order refines the interpretability order of among first-order theories and so defies model-theoretic analysis. Rather, we will discuss the following question for a given first-order theory T: Do the four second-order quantifiers coalesce when restricted to models of T? The answer to this question can be viewed either as a comment on the quantifiers or as a comment on the theory T. We will adopt the latter viewpoint here. The non-interpretability of second-order logic imposes an extremely strong structure theory on the models of T. This structure theory and some of its consequences are outlined below. In particular, we measure the complexity of (T, Q_ψ) by computing Hanf and Löwenheim numbers.

5.1. *Outline of the Classification*

In making such a classification, we consider those theories for which (T, Q_{mon}) interprets Q_{II} as being beyond analysis. The remainder can then be divided into four classes as follows. Assume $Q_{II} \not\leq (T, Q_{mon})$.

	$Q_{II} \leq (T, Q_{1-1})$	$Q_{II} \not\leq (T, Q_{1-1})$
$\mathrm{Th}(<, Q_{mon}) \leq (T, Q_{mon})$ (unstable)	prototype $(\mathrm{Th}(<, Q_{mon})$	impossible
$\mathrm{Th}(<, Q_{mon}) \not\leq (T, Q_{mon})$ (stable)	tree decomposable prototypes $\lambda^{\leq\omega}, \lambda^{<\omega}$	strongly decomposable

We could discuss the desirable properties of a particular entry of this table in two ways. We could prove a specific theorem (for example, that: the Löwenheim number of a countable theory such that $Q_{\mathrm{II}} \not\leq (T, 1\text{-}1)$ is $\aleph_0$). Even when such precise information cannot be obtained, we may be able to reduce such questions to the computation of, for instance, Löwenheim numbers for a specific theory T_0 by showing, for example, that if $Q_{\mathrm{II}} \not\leq (T, Q_{\mathrm{mon}})$, then (T, Q_{mon}) is bi-interpretable with the models of T_0. In some cases, we will prove a slightly weaker reduction than the second alternative: We will replace the theory T_0 by a class of structures which is not first-order definable. In some respects, of course, such a reduction is actually stronger than proving a particular theorem since it provides a "normal form" for models of T; the strength of the reduction depends on how well we are able to analyze the class to which we reduce.

The first line of the table distills an argument for the importance of studying the monadic theory of order. First, interpretability of the monadic theory of order is related to the important distinction between stable and unstable first-order theories.

5.1.1 Lemma. *If the complete first-order theory T is unstable, then* $(\mathrm{Th}(<), Q_{\mathrm{mon}}) \leq (T, Q_{\mathrm{mon}})$.

This result is proven in detail in Baldwin–Shelah [1982]. In outline, the proof proceeds by noticing (see Shelah [1978a]) that T is unstable iff T admits a definable linear ordering of an infinite set of n-tuples. A fairly complicated analysis of order indiscernibles (see Baldwin–Shelah [1982, VIII.1.3] shows that with additional unary predicates a linear ordering of a definable subset can be specified.

A second reason for the intensive study of the monadic theory of order as opposed to $(\mathrm{Th}(<), Q_\psi)$, for some other ψ, is that no other ψ is really possible. We have already shown in Section 2 that the only possibility for Q_ψ is Q_{1-1}. The next theorem rules out even that. It is fairly easy to deduce from Lemma 1.2.9 that $Q_{\mathrm{II}} \leq (\mathrm{Th}(<), 1\text{-}1)$. Combining this result with Section 5.1, we obtain

5.1.2 Theorem. *If T is unstable, then $Q_{\mathrm{II}} \leq (T, Q_{1-1})$.*

Further expansion of the argument that $\mathrm{Th}(<, Q_{\mathrm{mon}})$ is the prototype for those monadic unstable theories which can be analyzed occurs in Shelah [198?b, 198?d].

We will now discuss the situation characterized by bottom line of this table: The situation in which T is stable. In Section 2.5, we outlined the argument that if $Q_{\mathrm{II}} \not\leq (T, Q_{1-1})$ and T is stable, then T is strongly decomposable. If $Q_{\mathrm{II}} \leq (T, Q_{1-1})$, the argument that the fundamental equivalence relation is the same as algebraic closure and thus that each class is small (see Lemma 2.5.6) does not apply so that the classes may indeed be large. In this case, we iterate the procedure by choosing submodels inside each equivalence class and decomposing the class over this model. Since T is stable, this process cannot be iterated more than $|T|$ times (see Baldwin–Shelah [1982, IV.2.1]). This decomposes each model by a tree of height $\leq |T|$ in the sense of the following definition.

Before we examine the definition in detail, observe that the notation τ^- denotes the result of deleting the last symbol from a sequence.

5.1.3 Definition. The model M is *tree-decomposed* by the tree I of sequences of length at most κ if there exist models $\{\langle M_\eta, N_\eta \rangle : \eta \in I\}$ such that:

(i) $|N_\eta| = |T|$ for every η.
(ii) If $\eta \subseteq \rho$ then $N_\eta \subseteq N_\rho \subseteq M_\rho \subseteq M_\eta$.
(iii) For each $\tau \in I$ there are index sets J and functions σ such that:
 (a) M_τ is the free union of the $\{M_{\tau^\frown j} : j \in J\}$ amalgamated over N_τ and taken with respect to σ.
 (b) M is the free union over N_τ (with respect to σ) of $\{M_{\tau^\frown j} : j \in J\} \cup \{M_\rho \cup N_\tau : \rho \neq \tau \text{ but } \rho^{--} = \tau^-\}$;
(iv) $M_{<>} = M$; if $\lg(\eta)$ is a limit ordinal then $N_\eta = \bigcup \{N_\tau : \tau \subseteq \eta\}$, $M_\eta = \bigcap \{M_\tau : \tau \subseteq \eta\}$;
(v) $M = \bigcup \{N_\tau : \tau \in I\}$.

If a theory is κ tree-decomposable (that is, every model of T is tree-decomposed by a tree of height κ), then the models of T are short in the sense that no matter how large a model is, complete information about a finite sequence of elements from the model depends only on the less than κ elements which precede it in the tree.

5.1.4 Definition. The theory T is *shallow* if every model of T can be tree-decomposed by a well-founded tree. Otherwise T is *deep.*

If T is shallow, then we assign a rank to models of T, namely, the ordinal rank of the tree.

Now we can describe our prototypes.

5.1.5 Notation. Let K_0 be the class of all trees $\{\lambda^{<\omega} : \lambda \in \text{Ord}\}$ and K_1 the class of all trees $\{\lambda^{\leq\omega} : \lambda \in \text{Ord}\}$. If $Q_{\text{II}} \leq (T, Q_{\text{mon}})$, then the models of T are very closely tied to the trees which arise as skeletons when the models are tree-decomposed. Specifically, we have

5.1.6 Theorem. (i) *If T is a countable superstable deep theory and $Q_{\text{II}} \leq (T, Q_{\text{mon}})$, then $(T, L_{\omega_1,\omega}(Q_{\text{mon}})$ and $(K_0, L_{\omega_1,\omega})$ are bi-interpretable.*
(ii) *If T is a countable stable but not superstable theory and $Q_{\text{II}} \not\leq (T, Q_{\text{mon}})$, then $(T, L_{\omega,\omega}(Q_{\text{mon}})$ and $(K_1, L_{\omega,\omega})$ are bi-interpretable.*

This is Theorem VII.2.1 of Baldwin–Shelah [1982].

5.2. Computations of Hanf and Löwenheim Numbers

In this section we will briefly discuss the results on Hanf and Löwenheim numbers which can be derived from the preceding classification. We will then indicate how such computations are made. For the sake of simplicity, we will discuss only the case of countable languages here. The results extend to uncountable languages and such extensions are considered in Baldwin–Shelah [1982].

5.2.1. Finitary Monadic Logic

	Löwenheim Number	Hanf Number
$\lambda^{\nleq\omega}$	?	$(\beth_\omega)^+$
$\lambda^{<\omega}$ deep	*	$(\beth_\omega)^+$
$\lambda^{<\omega}$ shallow depth $= \beta$	$\min(\beth_\beta, \beth_\omega)$	$\min(\beth_\beta, \beth_\omega)$
strongly decomposable	$\beth_1$	$\beth_1$

* Shelah [1983b] has shown that there are superstable deep theories such that the Löwenheim number of (T, Q_{mon}) is (assuming $V = L$) the same as that of second-order logic.

This table and the one in 5.2.3 reports the Hanf number for sets of sentences. For a single sentence the '$^+$' can be dropped in some cases. See III.2 of Baldwin–Shelah [1982].

In order to completely determine the Löwenheim number, we must consider one further property. This we do in

5.2.2 Definition. The free union of $\langle M_i : i \in I \rangle$ over N is *nice* if for each i there exist finite subsets H_i of N and U_i of M_i such that for any $m \in M$, $t(m; H_i \cup U_i) \vdash t(m; N)$.

If the decomposition is nice, then the Löwenheim number of a shallow theory is $\aleph_0$; otherwise, it is $2^{\aleph_0}$. Details on this nicety are given in VI.2 of Baldwin–Shelah [1982].

5.2.3 Infinitary Monadic Logic ($L^\alpha_{\infty,\omega}$). For the sake of simplicity, assume that $\alpha \geq \omega_1$, then the following arrangement is possible.

	Löwenheim Number	Hanf Number
$\lambda^{\leq\omega}$	?	$\beth^+_{2+1}$
$\lambda^{<\omega}$ deep	*	$\beth^+_{\alpha+1}$
$\lambda^{<\omega}$ shallow shallow: depth $= \beta$	$(\beth_\beta)^+$	$(\beth_\beta)^+$
strongly decomposable	$(\beth_1)^+$	$(\beth_1)^+$

* Shelah [198?b] has shown that for every superstable deep theory such that the Löwenheim number in infinitary logic of (T, Q_{mon}) is, assuming that $V = L$, the same as that of second-order logic.

5.2.4 Outline of the Argument. These computations depend on (i) the decomposition of the models; (ii) the generalized Feferman–Vaught theorems; and (iii) the computation of the cardinality of $t_{\alpha,\lambda}(L)$. The general program is simply this: to decompose a model as free union of structures N_i for $i \in I$. Suppose we are trying to extend (Hanf number) or restrict (Löwenheim number) M for a sentence with λ

quantifiers (either individual or monadic) and α alternations. Let $W = \bigcup_{i \in I} t_{\alpha,\lambda}(N_i)$. Then, by Theorem 2.5, we can find a κ such that $t_{\alpha,\lambda}(M)$ is determined by $t_{\alpha,\kappa}(\langle \mathscr{I}, Q_t(I)\colon t \in W\rangle)$. Thus, if we can guarantee the cardinality of I to be sufficiently greater than $|W|$, there will be a large number of indices with the "same theory". We can then expand or contract this set at will. The full details are given in Chapters III, VI, and VII of Baldwin–Shelah [1982]. One sample is perhaps instructive. If T is strongly decomposable, then each model is a free union of countable structures. Since there are only $\beth_1$ possible $L_{\omega,\omega}(Q_{\text{mon}})$ theories of a countable structure, this reduces both the Hanf and Löwenheim numbers of (T, Q_{mon}) to $\beth_1$ precisely. In fact, for theories with a nice decomposition these numbers can be reduced to $\aleph_0$.

The situation when T is only tree-decomposable is somewhat more subtle. We can compute the Hanf number for $L_{\infty,\lambda}$ by noting that if $|M| > \beth_{\alpha+1}$ somewhere in the tree, we have a free union with more that $|t_{\alpha,\lambda}(L)|$ factors and then extend M. But this argument yields no information on the Löwenheim number. If T is shallow and β is the sup of the ranks of models of T, then we obtain the bound $\min(\beth_\beta, \beth_\omega)$ for both the Hanf and Löwenheim numbers by induction on this rank.

6. Generalizations

This work can be extended in several directions. In particular, the results in Section 5 can be sharpened, and the notion of quantifier can be extended. With respect to the first direction, Shelah [198?d] confirms the close connection between Hanf number and interpretability by showing

6.1 Theorem. *For any first-order theory T the Hanf number of (T, Q_{mon}) is at most H_{II} iff $Q_{\text{II}} \not\leq (T, Q_{\text{mon}})$.* □

In the other direction, we again return to the definition of a second-order quantifier.

1.2.2 Definition. If $\psi(\bar{r})$ is a formula of pure identity theory, then $Q_\psi(\bar{r})$ is the second-order quantifier whose semantics are given by:

$$M \models Q_\psi(\bar{r})\phi(\bar{r}) \quad \text{iff for some sequence} \quad \bar{R} \in \mathscr{R}_\psi(M),\ M \models \phi(\bar{R}).$$

There are several ways to extend this definition. Perhaps the most obvious one is to replace the requirement that ϕ be a first-order formula by introducing a parameter for the language. Thus, we have been discussing first-order definable second-order quantifiers. One could discuss infinitarily definable second-order quantifiers, or second-order quantifiers defined in stationary logic, or second-order definable second-order quantifiers etc. *ad nauseum.* A second possibility is to partition the variables $\bar{r}$ into a sequence $\bar{s}\bar{t}$. Then, by freezing the $\bar{s}$, we move out of pure logic and are thus able to discuss automorphisms, congruences and other algebraic concepts. Finally, we could remove the restriction that the relations $\bar{r}$ be subsets of A_n, for some n, and allow them, for example, to be families of subsets. Thus, we would obtain definable third-order quantifiers. At this level, we spread our net to include $L(\mathrm{aa})$. Another approach is to relax the definability requirement and allow the class of subsets defining a quantifier to be any class that is closed under isomorphism. This is the line adopted by Shelah [1983a]. Thus, we identify a quantifier with a class K of subsets of $\bigcup A^n$. Naturally, we may also deal with a finite sequence of quantifiers (classes) $\bar{K} = \langle K_0, \ldots, K_n \rangle$.

In discussing this widened class of quantifiers, Shelah weakens the notion of interpretability somewhat.

6.2 Definition. We say $\bar{K}$ is *expressible* in K if for each $R \in K$ there is a formula $\phi(\bar{x}, \bar{r})$ (with quantifiers over the K_i) such that for some $R_0, \ldots, R_n$ each in one of the K_i, $R(\bar{x}) \leftrightarrow \phi(\bar{x}, \bar{R})$. The problem—already hinted at in Shelah [1973c]—was finally addressed in Shelah [1983a], and it asks the following: Is every quantifier (that is, class K) bi-interpretable with a finite sequence $\bar{K}$, where each K_i is an equivalence relation? The main result on this is given in

6.3 Theorem (Expressibility with Equivalence Relations). (i) If $V = L$, then every K is bi-expressible with an equivalence relation (see Shelah [1983c]; p. 53).

(ii) It is consistent that there is a K which is not biexpressible with an equivalence relation. (Shelah [1983c]; pp. 48–57).

There is still another way these methods might be used. In many of the technical successes of stability theory over the last few years—for example, Vaught's conjecture for ω-stable T (Harrington–Makkai–Shelah [1983]) and the solution by Shelah [1982f, 198?c] of Morley's conjecture that (with the obvious exception) the spectrum function is increasing—the part of the proof showing there are many models can be viewed as an interpretation of Q_{II} into the $L_{\omega_1,\omega}(Q_{\mathrm{mon}})$ theory of T.