The Structure of Fixed–Point Logics

Martin Grohe

Dissertation
zur Erlangung der Doktorwürde
an der Mathematischen Fakultät
der Albert-Ludwigs Universität
Freiburg i. Br.

1994

Dekan der mathematischen Fakultät:
Professor Dr. Victor Bangert

Betreuer der Dissertation und erster Referent:
Professor Dr. Heinz-Dieter Ebbinghaus

Zweiter Referent:
Professor Dr. Jörg Flum

Datum der Promotion:
15. Dezember 1994

Acknowledgements

First of all I would like to thank my supervisor Professor Heinz–Dieter Ebbinghaus. This dissertation would not have been possible without his considerable advice and without all the long discussions we had on almost every detail. Actually I learned a lot beyond this work.

Many others have made valuable contributions to this thesis in discussions on various aspects of it, by detecting mistakes, or by helping me with my English style (it is not their fault that it is still bad enough). In particular I am grateful to Professor Jörg Flum, Anna Fukshansky, Nina Gleichmann, Lauri Hella, Bernhard Herwig, Ivo Herzog, Ian Hodkinson, Henrik Imhof, and Volker Remmert.

Zusammenfassung

Seit Immerman und Vardi [Imm86, Var82] bewiesen haben, daß in der *Logik der kleinsten Fixpunkte* gerade die in polynomialer Zeit berechenbaren Klassen geordneter endlicher Strukturen definierbar sind, spielen Fixpunktlogiken eine zentrale Rolle in der Endlichen Modelltheorie.

Fixpunktlogiken sind Erweiterungen der Logik der ersten Stufe, die die Formalisierung gewisser induktiver Definitionen erlauben. Sie werden gebildet, indem zur Logik der ersten Stufe ein Fixpunktoperator hinzugenommen wird. Verschiedene Varianten eines solchen Operators ergeben verschiedene Logiken, wie zum Beispiel die *Logik der partiellen Fixpunkte*, die bereits erwähnte Logik der kleinsten Fixpunkte oder die *Logik der induktiven Fixpunkte*.

Diese Arbeit beschäftigt sich mit der Frage, inwieweit die syntaktische Form einer Formel der Fixpunktlogiken (einer sogenannten „Fixpunktformel") Aussagen über ihre Ausdrucksstärke zuläßt, das heißt mit der Frage nach Normalformen und Hierarchien innerhalb der Logiken.

Zunächst untersuchen wir die Quantorenstruktur der Fixpunktformeln, insbesondere die Zahl der vorkommenden Fixpunktoperatoren. Wir werden sehen, daß ein einziger Fixpunktoperator genügt, um die volle Ausdrucksstärke zu erhalten, und daß es reicht, diesen auf eine existentielle Formel der ersten Stufe anzuwenden (beziehungsweise auf eine Δ_2–Formel im Fall der Logik der kleinsten Fixpunkte). Diese Ergebnisse sind bekannt, werden aber hier auf neue, sehr einheitliche Art für die verschiedenen Logiken bewiesen.

Ein weiterer syntaktischer Parameter einer Fixpunktformel ist die Anzahl des Auftretens einer einem Fixpunktoperator zugeordneten Relationsvariablen (der sogenannten „Fixpunktvariablen"). Wenn wir an induktive Definitionen denken, spiegelt sich darin die Anzahl der rekursiven Zugriffe wider. Der zentrale Satz des ersten Teils dieser Arbeit zeigt die Existenz einer Normalform für die Logiken der induktiven und partiellen Fixpunkte, in der sowohl die Quantorenstruktur als auch die Anzahl des Auftretens der Fixpunktvariablen optimal reduziert sind. Ähnliche Normalformen erhalten wir für die anderen untersuchten Fixpunktlogiken.

Als Anwendung dieses Satzes konstruieren wir vollständige Probleme für die Fixpunktlogiken unter quantorenfreien Reduktionen. Damit wird eine Verbindung hergestellt zwischen unseren Logiken und den allgemeinen Konzepten von logischen Reduktionen und verallgemeinerten (sogenannten Lindström–) Quantoren, die in den letzten Jahren eine immer größere Rolle in der Endlichen Modelltheorie spielen.

Der zweite Teil dieser Arbeit beschäftigt sich mit der *Stelligkeit* von Fixpunktoperatoren. Denken wir wieder an induktive Definitionen, so entspricht diese der Stellenzahl der induktiv definierten Relation. Im Gegensatz zu den im ersten Teil betrachteten syntaktischen Parametern hat die Stelligkeit großen Einfluß auf die Ausdrucksstärke — die durch Formeln beschränkter Stelligkeit definierten

Fragmente bilden eine echte Hierarchie.

Tatsächlich beweisen wir eine wesentlich stärkere Aussage, die auch das Verhältnis der Ausdrucksstärke von Fragmenten beschränkter Stelligkeit der verschiedenen von uns untersuchten Logiken klärt. Ferner zeigen wir, daß diese Aussage bereits auf der Klasse der Graphen gilt.

Die zum Beweis entwickelte Methode erweist sich als überaus fruchtbar und erlaubt es uns, ähnliche Hierarchieaussagen auch für andere Logiken zu beweisen, so für verschiedene Varianten der Logik der transitiven Hülle und für die Datenbanksprache *Datalog*.

Contents

Introduction

In classical model theory first–order logic has always played an overwhelming role, whereas in finite model theory there has been an emphasis on the study of more expressive logics right from the start. A reason for this is that most classical results and methods fail in the scope of finite models, and first–order logic loses its distinguishing features such as the compactness theorem and the Löwenheim Skolem theorem.

Furthermore, an important direction in finite model theory is the study of the connection between computational complexity and logical expressibility. It is evident that for this purpose the logics need to be able to describe recursion in an appropriate way which first–order logic is not.

So it seems reasonable to strengthen first–order logic by a suitable recursion mechanism. A natural way to do this is to augment it by a fixed–point operator. In fact there are different ways to define such an operator which result in several well–known fixed–point logics such as *partial fixed–point logic* and *least fixed–point logic*.

Logics with a fixed–point operator have already played a role in classical recursion theory as a way to formalize inductive definitions. In particular, Moschovakis devoted a monograph [Mos74] to the study of this subject. Many basic results that have later been applied in finite model theory were given there.

Fixed–point logics became the focus of attention in finite model theory when Immerman and Vardi [Imm86, Var82] proved that a class of ordered finite structures is definable in least fixed–point logic if and only if it is contained in the complexity class **PTIME**, and similarly, a class of ordered finite structures is definable in partial fixed–point logic if and only if it is contained in **PSPACE**.

These results suggested a further investigation of the logics in the scope of finite models. First important results were obtained by Immerman [Imm86] who proved that every least fixed–point formula is equivalent to a formula in which only one fixed–point operator occurs, and by Gurevich and Shelah [GS86] who showed that least fixed–point logic equals *inductive fixed–point logic*.

Among the numerous results that followed maybe the most important one is Abiteboul's and Vianu's theorem [AV91b] that least fixed–point logic equals partial fixed–point logic if and only if **PTIME** equals **PSPACE** (which is *not* an immediate consequence of Immerman's and Vardi's results from descriptive complexity theory mentioned above since they essentially need ordered structures). In a sense this closes a gap between the model theoretic investigation of the logics which naturally encompasses all finite structures and the results in descriptive complexity theory which usually need ordered structures.

A field which is strongly related to finite model theory is relational database theory. The study of fixed–point logics in this context was initiated by Chandra and Harel [CH82]. Later Abiteboul and Vianu [AV91a] showed that the various fixed–point logics coincide with extensions of the database language datalog. This connection to database theory turned out to be useful also under a purely logical point of view; as an immediate consequence of Abiteboul's and Vianu's theorems we obtain normal forms for inductive and partial fixed–point logic.

Taking up such results, this thesis is mainly concerned with the question of how far the syntactical structure of the fixed–point formulae affects the expressive power of the resulting logics, that is the question for normal forms and hierarchies within the logics. The first part is devoted to normal forms, and in the second part we establish hierarchy results.

To make this more precise let us shortly sketch how a fixed–point formula is built:

Consider a formula $\varphi(\overset{k}{x}, X)$ (where $\overset{k}{x} = x_1 \ldots x_k$ is a tuple of individual variables and X is a k–ary relation variable). For each structure $\mathfrak{A}$ (with universe A) it gives rise to an operator $F^\varphi : \mathrm{Pot}(A^k) \longrightarrow \mathrm{Pot}(A^k)$ defined by

$$F^\varphi(B) := \{\overset{k}{a} \in A^k \mid \mathfrak{A} \models \varphi[\overset{k}{a}, B]\}.$$

Starting with the empty set and iterating this operator we obtain a sequence

$$X_0^{\mathfrak{A}} := \emptyset,\ X_1^{\mathfrak{A}} := F^\varphi(X_0^{\mathfrak{A}}),\ X_2^{\mathfrak{A}} := F^\varphi(X_1^{\mathfrak{A}}), \ldots$$

of subsets of A^k. The *partial fixed–point* of this sequence is defined to be

$$X_\infty^{\mathfrak{A}} := \begin{cases} X_m^{\mathfrak{A}} & \text{where } m = \min\{i \mid X_i^{\mathfrak{A}} = X_{i+1}^{\mathfrak{A}}\} \text{ if it exists} \\ \emptyset & \text{otherwise} \end{cases} .$$

We can now augment first–order logic by a *fixed–point operator* $[FP_{\overset{k}{x},X} \ldots]$ whose syntax and sematics are explained by

$$\mathfrak{A} \models [FP_{\overset{k}{x},X}\varphi(\overset{k}{x}, X)]\ \overset{k}{a} \iff \overset{k}{a} \in X_\infty^{\mathfrak{A}} .$$

The resulting logic is partial fixed–point logic. The other logics mentioned here can be defined as natural fragments of partial fixed–point logic.

Immerman's [Imm86] normal form theorem says that each least fixed–point formula is equivalent to a formula of the form

$$\exists v[FP_{\bar{x},X}\varphi]v \ldots v$$

where φ is first–order; and Abiteboul's and Vianu's [AV91a] results imply the analogous statements for inductive and partial fixed–point logic.

Starting here one may ask for the quantifier–structure of the first–order part. It turns out that an existential formula suffices in the case of partial fixed–point logic and inductive fixed–point logic and that a Δ_2–formula (where the relation variable X only occurs positively) suffices for least fixed–point logic. (These results are not new, cf. [AV91a, Dah87]).

Another question one may ask is how many occurrences of the relation variable X are necessary in φ. An immediate answer can be given by a theorem of Moschovakis [Mos74]: One positive and one negative occurrence suffice for inductive and partial fixed–point logic, one positive occurrence suffices for least fixed–point logic.

Our new contribution here is to give a normal–form where the quantifier structure *and* the number of occurrences of X are restricted. The main result of the first part is that every (inductive, partial) fixed–point formula is equivalent to a formula of the form

$$\exists v[FP_{\bar{x},X}\ \theta_0 \vee \exists \bar{y} \in X \exists \bar{z} \notin X\theta]v \dots v$$

where θ_0 and θ are quantifier–free formulae in which X and v do not occur.

Similar normal forms are given for least fixed–point logic and its existential fragment.

We will see that these normal forms are optimal.

As an application, we give complete problems for the fixed–point logics under quantifier–free reductions. Questions concerned with logical reducibilities and the related concept of Lindström–quantifiers have caught some interest in finite model theory during the last few years, and our result shows that fixed–point logics can also be considered as a part of these more general concepts. A complete problem for least fixed–point logic has already been given by Dahlhaus [Dah87], and recently Hella [Hel94] has found a vectorized Lindström–quantifier for least fixed–point logic augmented by finitely many monotone Lindström–quantifiers.

Since least fixed–point logic equals inductive fixed–point logic Dahlhaus' result also provides a complete problem for inductive fixed–point logic. Hence the existence of such a problem for both logics is not new. However, the complete problem we give reflects in a nice way the function of an *inductive* fixed–point operator (while Dahlhaus' result is strongly based on the monotonicity of a *least* fixed–point operator). Seen in connection with the complete problem we give for partial fixed–point logic it also sheds some new light on the relation between inductive and partial fixed–point logic and hence **PTIME** and **PSPACE** (by Abiteboul's and Vianu's result mentioned above).

Another well–known fixed–point logic, *stratified fixed–point logic*, does not admit a normal form with only one fixed–point operator. However, we can still simplify the syntactical structure of stratified fixed–point formulae and show that they are equivalent to *(2,0)–bounded fixed–point formulae.*

This suggests to consider the subclass of $(1, 0)$–bounded formulae which turns out to coincide with a logic that also plays an important role in finite–model theory, *transitive closure logic.*

So it is a natural part of our investigations to study the structure of transitive closure logic. Most questions concerning normal–forms have already been answered by Immerman [Imm87], Grädel, and McColm [Grä91, GM92a, GM92b]. The interesting new results concerning transitive closure logic are the hierarchy results obtained in the second part.

This second part is devoted to another natural syntactical parameter of a fixed–point formula, the *arity*. The arity of a fixed–point operator $[FP_{\bar{x},X}\ldots]$ is defined to be the arity of X, and the arity of a fixed–point formula is the maximal arity of the fixed–point operators that occur in this formula. The k–ary fragment of a fixed–point logic is the fragment whose formulae are those of arity $\leq k$. These fragments (for $k \geq 1$) form a natural hierarchy inside the logic.

It turns out that, in contrast to the parameters we mentioned before, the arity of a formula strongly affects its expressive power.

Our main lemma states that for every $k \geq 1$ there is a graph query that is not definable by a $(k-1)$–ary partial fixed–point formula, but by a k–ary formula. This query is very simple, actually it is the transitive closure of a formula $\text{EDGE}_k(\overset{k}{x}, \overset{k}{y})$ which says that $x_1, \ldots, x_k, y_1, \ldots, y_k$ form a $2k$–clique.

As an immediate consequence of this we see that the arity hierarchy is strict for all fixed–point logics and for transitive closure logic.

The question whether the arity hierarchy is strict for least fixed–point logic was posed by Chandra and Harel [CH82]. There it was announced that Gaifman solved the problem for arities ≤ 3. Another partial result (concerning formulae of the form $[FP_{\bar{x},X}\varphi]\bar{u}$ with a first–order φ) has been given by Dublish and Maheshwari [DM89]. The analogous problem for transitive closure logic was posed by Calo and Makowsky [CM92]; they could prove that the unary and the binary fragment of transitive closure logic are distinct.

Programs in the database language datalog also have a natural notion of arity and thus an arity hierarchy. Our main lemma implies that this hierarchy is strict for pure datalog and various extensions.

Afrati and Cosmadakis [AC89] have already proved that the arity hierarchy is strict for pure datalog, but their proof requires a new signature containing a $2k$–ary relation symbol to prove that the expressive power of k–ary datalog exceeds the expressive power of $(k-1)$–ary datalog, while our result delivers a graph–query expressible in k–ary datalog but not in $(k-1)$–ary datalog (for each $k \geq 1$).

The method we develop to prove the main lemma turns out to be very fruitful and allows several extensions of the results to other logics. We are going to establish a conjecture of Kolaitis [Kol90] concerning implicit definability, and sharpen a result of Hella [Hel92] concerning arities of Lindström–quantifiers.

In the last section we prove, with a bit more effort than before but with the same method, that the arity hierarchy is also strict for the fragment *deterministic* transitive closure logic.

0 Preliminaries

§0.1 Proviso

We are pursuing finite model theory!

- **We only consider finite structures.**
- **All signatures we consider are finite.**

For technical reasons we include the following two conventions. Most of the results also hold without them, or could be slightly modified to hold without them.

- We only consider structures of cardinality ≥ 2.
- The signatures we consider do not contain any function symbols.

§0.2 Notation

Our notation is standard, based on [EFT94].

We only want to mention the following few items here, other notations being explained in the text or in the footnotes.

- We denote tuples in the form $x_1 \dots x_k$ (without brackets or commata) and abbreviate them by $\overset{k}{x}$ or $\bar{x}$. This allows us, for example, to write $\overset{k}{a}\overset{l}{b}$ for the tuple $a_1 \dots a_k b_1 \dots b_l$.
- The notation $\varphi(x, y, X)$ neither means that the (individual) variables x, y and the relation variable X[1] occur freely in the formula φ nor that all free variables of φ are among them, but only gives us a convenient way to denote substitutions.

  For example, $\varphi(y, t, Y)$ is the formula obtained from $\varphi(x, y, X)$ by simultaneously replacing each occurrence of x by y, of y by the term t, and of X by Y.
- If nothing else is mentioned we start counting with 1, for example "there exists an $i \leq n$ ..." means "there exists an i such that $1 \leq i \leq n$ ...".
- $\mathcal{F}$ denotes the class of all (finite) structures.

[1] We allow first–order formulae to have free relation variables.

§0.3 Interpretations

We are often going to work with interpretations (rather than structures) in inductive proofs or definitions.

A σ–interpretation $\mathfrak{I}$ is a pair $(\mathfrak{A}, \alpha)$ where $\mathfrak{A}$ is a σ–structure and α is an assignment that associates an element of A[2] with each individual variable and a k–ary relation on A with each k–ary relation variable.

We extend α to all terms.

An interpretation *for* a formula φ is an interpretation whose signature contains the signature of φ.

§0.4 Queries

A *k–ary σ–query ρ on a class* $\mathcal{C}$ is a mapping on $\mathcal{C}[\sigma]$[3] that associates with each $\mathfrak{A} \in \mathcal{C}[\sigma]$ a k–ary relation $\rho(\mathfrak{A})$ in a way that each isomorphism π between two structures $\mathfrak{A}, \mathfrak{B} \in \mathcal{C}[\sigma]$ is also an isomorphism between $(\mathfrak{A}, \rho(\mathfrak{A}))$ and $(\mathfrak{B}, \rho(\mathfrak{B}))$.

In particular, a 0–ary (or *boolean*) σ–query on $\mathcal{C}$ associates TRUE or FALSE with each structure in $\mathcal{C}[\sigma]$, thus it can be identified with a subclass of $\mathcal{C}[\sigma]$ that is closed under isomorphism (within $\mathcal{C}$).

A *k–ary σ–query* is a k–ary σ–query on $\mathcal{F}$.

If it is clear from the context we usually do not mention σ.

§0.5 Logics

A *logic* $\mathbf{L}$ is a pair $(L, \models_L)$, where L is a class of formulae and $\models_L$ an assignment that associates a σ–query with each formula of signature σ.[4] We call such a query *definable* in $\mathbf{L}$.

For simplification, we often identify a logic $\mathbf{L}$ with the class of queries definable in $\mathbf{L}$. This explains notations like $\mathbf{L} \subseteq \mathbf{L}', \mathbf{L} = \mathbf{L}'$.

We denote first–order logic by $\mathbf{FO} = (FO, \models)$.

QF denotes the class of all quantifier–free formulae, and $\mathbf{QF}$ is the logic $(QF, \models)$.

§0.6 Binary Relations

We always let E be a binary relation symbol. From time to time we refer to the following kinds of $\{E\}$–structures:

---

[2] We always denote the universe of a structure $\mathfrak{A}$ by A.

[3] For a class $\mathcal{C}$ of structures and a signature σ we denote the class of all σ–structures in $\mathcal{C}$ by $\mathcal{C}[\sigma]$.

[4] We do not explain what a formula is, it suffices here to know that each formula is associated with a signature. $L[\sigma]$ denotes the class of all L–formulae whose signature is σ.

- A *di(rected)graph* $\mathfrak{G}$ is an $\{E\}$–structure where E is antireflexive.

  A vertex b is a *successor* of a vertex a if $E^{\mathfrak{G}}ab$, and b is a *predecessor* of a if $E^{\mathfrak{G}}ba$. The *in–degree* of a node is the number of predecessors, the *out–degree* the number of successors.

  A *leaf* of $\mathfrak{G}$ is a vertex with out–degree 0, a *root* of $\mathfrak{G}$ is a vertex with in–degree 0.

- A *forest* is an acyclic digraph where each node has in–degree at most 1.

- A *tree* is a connected forest.

- A *network* is an acyclic connected digraph with exactly one leaf and one root.

- A *graph* is a digraph where E is symmetric.

Moreover, a *partial order* is a binary relation that is transitive and anti–reflexive; and a *linear order* is a total partial order.

Part I
Normal Forms

1 Fixed–Point Logics

§ 1.1 Partial and Inductive Fixed–Point Logic

1.1 Definition : The class *FP* of *fixed–point* formulae is given by the following calculus:

$$(A)\quad \frac{}{\varphi} \quad \text{where } \varphi \text{ is an atomic formula,}$$

$$(\vee)\quad \frac{\varphi,\psi}{\varphi\vee\psi}\quad,\quad (\wedge)\quad \frac{\varphi,\psi}{\varphi\wedge\psi}\quad,\quad (\neg)\quad \frac{\varphi}{\neg\varphi}\quad,$$

$$(\exists)\quad \frac{\varphi}{\exists x\varphi}\quad,\quad (\forall)\quad \frac{\varphi}{\forall x\varphi},\qquad \text{(the \textit{first–order rules})}$$

and

$$(FP)\quad \frac{\varphi}{[FP_{\overset{k}{x},X}\ \varphi]\ \overset{k}{u}}\quad \text{where } k\geq 1,\ \overset{k}{u} \text{ is a tuple of terms, } X \text{ is } k\text{-ary.}$$

$[FP_{\overset{k}{x},X}\ \ldots]$ is called a *fixed–point operator.*

The free (first and second order) variables of a fixed–point formula are defined in the standard way adding the clause

$$\mathrm{free}(\ [FP_{\overset{k}{x},X}\ \varphi]\ \overset{k}{u}) = (\mathrm{free}(\varphi)\setminus\{x_1,\ldots,x_k,X\})\cup\mathrm{free}(\overset{k}{u}).$$

□

There are two ways to define the semantics. We start with the more general way, the *partial fixed–point semantics*:
For each fixed–point formula $\chi =\ [FP_{\overset{k}{x},X}\ \varphi]\ \overset{k}{u}$ and interpretation $\mathfrak{I} = (\mathfrak{A},\alpha)$ we define a sequence $(X_i^{\mathfrak{I}})_{i\geq 0}$ of subsets of A^k by

$$\begin{aligned}
X_0^{\mathfrak{I}} &:= \emptyset\\
X_{i+1}^{\mathfrak{I}} &:= \{\overset{k}{a}\in A^k \mid \mathfrak{I}\models\varphi[\overset{k}{a},X_i^{\mathfrak{I}}]\}\\
X_\infty^{\mathfrak{I}} &:= \begin{cases} X_m^{\mathfrak{I}} & \text{where } m=\min\{i \mid X_i^{\mathfrak{I}} = X_{i+1}^{\mathfrak{I}}\} \text{ if it exists}\\ \emptyset & \text{otherwise}\end{cases}\quad .
\end{aligned}$$

The sets $X_i^{\mathfrak{I}}$ ($i\geq 0$) are called the *stages* of the fixed–point process. We say that the process (or the sequence $(X_i^{\mathfrak{I}})_{i\geq 1}$) *reaches a fixed–point* if there exists an $m\geq 1$ such that $X_m^{\mathfrak{I}} = X_{m+1}^{\mathfrak{I}}$. A bit sloppily, we call $X_\infty^{\mathfrak{I}}$ the *fixed–point* of χ (or of the sequence $(X_i^{\mathfrak{I}})_{i\geq 1}$) in $\mathfrak{I}$, even if the sequence does not reach a fixed–point (i.e. $X_\infty^{\mathfrak{I}} = \emptyset$).

Now we can define the satisfaction–relation $\models_P$ for the class *FP* inductively with respect to the given calculus, letting

$$\mathfrak{I} \models_P \ [FP_{\overset{k}{x},X}\ \varphi]\overset{k}{u} \quad :\Longleftrightarrow \quad \alpha(\overset{k}{u}) \in X_\infty^{\mathfrak{I}}$$

in the fixed–point case.

The other way is the *inductive fixed–point semantics*. Again we define a sequence of subsets of A^k:

$$\begin{aligned} X_0^{I,\mathfrak{I}} &:= \emptyset \\ X_{i+1}^{I,\mathfrak{I}} &:= X_i^{I,\mathfrak{I}} \cup \{\overset{k}{a} \in A^k \mid \mathfrak{I} \models \varphi[\overset{k}{a}, X_i^{I,\mathfrak{I}}]\}. \end{aligned}$$

Since here we have $X_i^{I,\mathfrak{I}} \subseteq X_{i+1}^{I,\mathfrak{I}}$ for all $i \geq 0$ the process always reaches a fixed–point and we simply let

$$X_\infty^{I,\mathfrak{I}} := \bigcup_{i \geq 0} X_i^{I,\mathfrak{I}} .$$

The satisfaction–relation $\models_I$ for *FP* is also defined inductively, letting

$$\mathfrak{I} \models_I \ [FP_{\overset{k}{x},X}\ \varphi]\overset{k}{u} \quad :\Longleftrightarrow \quad \alpha(\overset{k}{u}) \in X_\infty^{I,\mathfrak{I}} .$$

1.2 Definition : *Partial fixed–point logic* is the logic **PFP** $:= (FP, \models_P)$; *inductive fixed–point logic* is defined to be **IFP** $:= (FP, \models_I)$. □

To see that **IFP** $\subseteq$ **PFP** we define:

1.3 Definition : A fixed–point formula χ is called *inductive* if for every subformula of the form $[FP_{\overset{k}{x},X}\ \varphi]\overset{k}{u}$ we have $X_i^{\mathfrak{I}} \subseteq X_{i+1}^{\mathfrak{I}}$ for all interpretations $\mathfrak{I}$ and $i \geq 0$.

IFP is the class of inductive fixed–point formulae. □

Using Trakhtenbrot's Theorem it can be easily seen that the class *IFP* is not recursive. (The sentence $\forall v[FP_{x,X}\ \varphi \lor \exists y \neg Xy]v$ is inductive if and only if the first–order sentence φ is valid.)

Let $\mathfrak{I}$ be an interpretation. Clearly for each $\chi \in IFP$ we have

$$\mathfrak{I} \models_I \chi \quad \Longleftrightarrow \quad \mathfrak{I} \models_P \chi .$$

Moreover, for an arbitrary φ the formula $[FP_{\bar{x},X}\ X\bar{x} \lor \varphi]\bar{u}$ is inductive and

$$\mathfrak{I} \models_I [FP_{\bar{x},X}\ X\bar{x} \lor \varphi]\bar{u} \quad \Longleftrightarrow \quad \mathfrak{I} \models_I [FP_{\bar{x},X}\ \varphi]\bar{u} .$$

Hence **IFP** $= (IFP, \models_P) \subseteq$ **PFP**.

So in fact we could do without the relation $\models_I$, and most of the time we will. Therefore we write $\models$ instead of $\models_P$.
However, using $\models_I$ gives us a logic with a recursive syntax.

We close this paragraph with a technical definition that we will need several times:

1.4 Definition : The *quantifier–rank* $\mathrm{qr}(\chi)$ of a fixed–point formula χ is defined inductively in the standard way adding the clause

$$qr(\ [FP_{\overset{k}{x},X}\ \varphi]\ \overset{k}{u}) = qr(\varphi) + k\ .$$

The *fixed–point rank* $fr(\chi)$ is the maximal number of nested fixed–point operators in χ. □

§1.2 Defining the Stages

Each fixed–point formula $\chi = [FP_{\overset{k}{x},X} \varphi(\overset{k}{x}, X)]\ \overset{k}{u}$ associates, for each $i \geq 1$, the relation $X_i^{\mathfrak{I}} \subseteq A^k$ with the interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$. Hence we can consider the *ith stage of* χ as a query.
For each $i \geq 1$ this query is definable as follows:
We let $\chi_0(\overset{k}{x}) := (x_1 \neq x_1)$ and $\chi_{i+1}(\overset{k}{x}) = \varphi(\overset{k}{x}, \chi_i(\_\_))$.[5] Then χ_i defines the ith stage of χ (for each $i \geq 1$).

Note that we have

$$\mathfrak{I} \models \chi \quad \Longleftrightarrow \quad \exists i \geq 0 :\ \mathfrak{I} \models \chi_i(\overset{k}{u})\ \wedge \forall \overset{k}{x}\ (\chi_i(\overset{k}{x}) \leftrightarrow \chi_{i+1}(\overset{k}{x})).$$

For inductive χ this can be simplified to

$$\mathfrak{I} \models \chi \quad \Longleftrightarrow \quad \exists i \geq 0 :\ \mathfrak{I} \models \chi_i(\overset{k}{u}).$$

We can define the *rank* of a fixed–point formula inductively as follows:

- $\mathrm{rank}(\varphi) = 0$ if φ is atomic
- $\mathrm{rank}(\exists x \varphi) = \mathrm{rank}(\forall x \varphi) = \mathrm{rank}(\neg\varphi) = \mathrm{rank}(\varphi) + 1$

[5]This is a notation we are going to use very often: For two formulae $\psi(\overset{l}{x})$ and $\xi(Y)$ (where Y is an l–ary relation variable) we let $\xi(\psi(\_\_))$ be the formula obtained from ξ by replacing each subformula of the form $Y\ \overset{l}{t}$ by $\psi(\overset{l}{t})$. We have to be a bit careful here with the free variables of ψ because they may be quantified in ξ. If this causes trouble we have to rename some of the variables, but clearly this is no problem.

- $\mathrm{rank}(\varphi \vee \psi) = \mathrm{rank}(\varphi \wedge \psi) = \max\{\mathrm{rank}(\varphi), \mathrm{rank}(\psi)\} + 1$
- $\mathrm{rank}(\underbrace{[FP_{\bar{x},X}\varphi]\bar{u}}_{=:\chi}) = \sup\{\mathrm{rank}(\chi_i) + 1 \mid i \geq 0\}$.

Note that the rank of a formula is well–defined. In general it is an infinite ordinal.

§ 1.3 Least Fixed–point Logic

A formula φ is called *positive* in a relation variable X if X only occurs within the scope of an even number of negation symbols. It can be easily verified that if $\varphi(X)$ is positive in X it is *monotone* in X, i.e. for all interpretations $\mathfrak{I} = (\mathfrak{A}, \alpha)$ for φ and $B, C \subseteq A$ we have:

$$B \subseteq C \quad \Longrightarrow \quad (\mathfrak{I} \models \varphi[B] \;\Rightarrow\; \mathfrak{I} \models \varphi[C])$$

Note that if φ is monotone in X then the formula $[FP_{\bar{x},X}\varphi]\bar{u}$ is inductive.

1.5 Definition : The class LFP of *least fixed–point formulae* is the subclass of FP where the fixed–point operator $[FP_{\bar{x},X} \ldots]$ is only applied to formulae that are positive in X.
Least fixed–point logic is the logic $\mathbf{LFP} := (LFP, \models)$. □

Since every formula which is positive in X is also monotone we have $LFP \subseteq IFP$. Gurevich and Shelah proved that in fact:

1.6 Theorem ([GS86]) : $\mathbf{LFP} = \mathbf{IFP}$

This theorem is part of the stronger Theorem 2.5 which we will prove in Paragraph 2.2.

§ 1.4 Existential Fragments

1.7 Definition : A formula is called *existential* if it contains no universal quantifiers and negation symbols only occur in front of atomic subformulae.

If $\mathbf{L} = (L, \models_L)$ is a logic, $\exists L$ denotes the class of the existential formulae in L; we call $\exists\mathbf{L} = (\exists L, \models_L)$ the *existential fragment of* $\mathbf{L}$. □

Hence we can consider the existential fragments $\exists\mathbf{PFP}$, $\exists\mathbf{IFP}$, and $\exists\mathbf{LFP}$ of our fixed–point logics.

However, only $\exists\mathbf{LFP}$ is of interest because universal quantifiers are definable by existential *IFP*–formulae. To see this consider the formula

$$[FP_{x,X}\; x = c \vee (x = d \wedge \exists z \neg Pz) \vee (x = e \wedge Xc \wedge \neg Xd)]e$$

(where c, d, and e are constants and P is a unary relation) which is an existential **IFP**–formula. An easy induction shows that it is equivalent to $\forall y Py$ in all structures in which c, d, e are interpreted by distinct elements.
This gives us an idea how to eliminate universal quantifiers from fixed–point formulae and prove that $\exists$**PFP** = **PFP** and $\exists$**IFP** = **IFP**. However, this method requires a new fixed–point operator for each universal quantifier.
We will strengthen the argument in the next section to prove an existential normal form for **PFP** and **IFP** where only one fixed–point operator occurs.

On the other hand the next Lemma implies **FO** $\not\subseteq$ $\exists$**LFP** thus $\exists$**LFP** $\subsetneq$ **LFP**.

1.8 Lemma : *$\exists LFP$–formulae are preserved under extensions.*

Proof: The proof is a simple induction over the rank of an $\exists LFP$–formula χ.
$\exists LFP$–formulae of rank 0 are atomic formulae, hence preserved under extensions.
The only interesting case of the induction step is $\chi = [FP_{\bar{x},X}\varphi]\bar{u}$. Let $\mathfrak{I} \subseteq \mathfrak{J}$ be interpretations for χ such that $\mathfrak{I} \models \chi$. Then there exists an $i \geq 0$ such that $\mathfrak{I} \models \chi_i(\bar{u})$. Since χ is positive in X, it can be easily seen that χ_i is still existential. Thus by the induction hypothesis we have $\mathfrak{J} \models \chi_i(\bar{u})$, hence $\mathfrak{J} \models \chi$. ■

1.9 Remark : Similarly it can be seen that universal least fixed–point formulae (i.e. formulae that do not contain any existential quantifiers and in which negation symbols only occur in front of atomic subformulae) are preserved under substructures. □

A logic that still has some of the nice properties of $\exists$**LFP**, but on the other hand contains first–order logic, is *stratified fixed–point logic* **SFP**.
The class *SFP* of *stratified fixed–point formulae* is the subclass of *LFP* where the fixed–point operator $[FP_{\bar{x},X} \ldots]$ is only applied to formulae where X does not occur within the scope of any negation symbols, universal quantifiers, or fixed–point operators. We let **SFP** $:= (SFP, \models)$.
Since **FO** $\subseteq$ **SFP** we have $\exists$**LFP** $\subsetneq$ **SFP**; and a result of Dahlhaus [Dah87] and Kolaitis [Kol91] shows that **SFP** $\subsetneq$ **LFP**.

§1.5 Simultaneous Inductions

1.10 Definition : The class *s–FP* of *simultaneous fixed–point formulae* is given by means of the calculus consisting of the first–order clauses and the clause

$$(s\text{–}FP) \quad \frac{\varphi_1, \ldots, \varphi_m}{[FP_{\bar{x}_1,X_1,\ldots,\bar{x}_m,X_m}\ \varphi_1, \ldots, \varphi_m]\bar{u}}$$

where $m \geq 1$ and, for each $i \leq m$, X_i is a relation variable whose arity matches the length of $\bar{x}_i$, and $\bar{u}$ is a tuple of terms of the same length as $\bar{x}_1$. □

It should be remarked here that tuples $\bar{x}_i$ and $\bar{x}_j$ $(i, j \leq m)$ of variables in the definition above may be the same. On the other hand the relation variables X_i and X_j have to be different.

Note that *FP* ⊂*s–FP*.

In order to emphasize that we are dealing with a simultaneous fixed–point formula we usually write

$$[s\text{–}FP_{\bar{x}_1, X_1, \ldots, \bar{x}_m, X_m} \varphi_1, \ldots, \varphi_m]\bar{u}$$

instead of $[FP_{\bar{x}_1, X_1, \ldots, \bar{x}_m, X_m}\ \varphi_1, \ldots, \varphi_m]\bar{u}$.

To extend the semantics to the new formulae, for each simultaneous fixed–point formula

$$[s\text{–}FP_{\bar{x}_1, X_1, \ldots, \bar{x}_m, X_m} \varphi_1, \ldots, \varphi_m]\bar{u}$$

and interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$, we define sequences $(X^{\mathfrak{I}}_{ji})_{i \geq 0}$ $(1 \leq j \leq m)$ of relations on A by

$$\begin{aligned}
X^{\mathfrak{I}}_{j0} &:= \emptyset \\
X^{\mathfrak{I}}_{j(i+1)} &:= \{\bar{a} \in A \mid \mathfrak{I} \models \varphi_j[\bar{a}, X^{\mathfrak{I}}_{1i}, \ldots, X^{\mathfrak{I}}_{mi}]\} \\
X^{\mathfrak{I}}_{j\infty} &:= \begin{cases} X^{\mathfrak{I}}_{jl} & \text{where } l = \min\{i \mid \bigwedge_{j'=1}^{m} X^{\mathfrak{I}}_{j'i} = X^{\mathfrak{I}}_{j'(i+1)}\} \\ & \text{if such an } l \text{ exists} \\ \emptyset & \text{otherwise} \end{cases}
\end{aligned} .$$

We let

$$\mathfrak{I} \models [s\text{–}FP_{\bar{x}_1, X_1, \ldots, \bar{x}_m, X_m} \varphi_1, \ldots, \varphi_m]\bar{u} \quad :\Longleftrightarrow \quad \alpha(\bar{u}) \in X^{\mathfrak{I}}_{1\infty}$$

and extend the satisfaction–relation $\models$ inductively to all *s–FP*–formulae.

Similarly we can extend the relation $\models_I$ to simultaneous fixed–point formulae. It is also clear how to define the class *s–IFP* of *inductive* simultaneous fixed–point formulae.

The well–known Lemma on Simultaneous Induction ([Mos74]) shows that **PFP** $= (s\text{–}FP, \models)$ and **IFP** $= (s\text{–}FP, \models_I) = (s\text{–}IFP, \models)$, so the new formulae do not increase the expressive power of the logics.
Since we not only need this statement, but also some information on the quantifier structure of a formula that simulates a simultaneous induction, we give the following precise formulation of the lemma.

1.11 Definition : The *equality type* of an r–tuple $\overset{r}{a}$ is the formula

$$\varepsilon(\overset{r}{x}) = \bigwedge\{x_i = x_j \mid a_i = a_j\} \wedge \bigwedge\{x_i \neq x_j \mid a_i \neq a_j\} \; .$$

Let, for each $r \geq 1$, $\varepsilon^r_1, \ldots, \varepsilon^r_s$ be an enumeration of all equality types of r–tuples $\overset{r}{a}$ consisting of at most two distinct elements (i.e. $|\{a_1, \ldots, a_r\}| \leq 2$) in a way that $\varepsilon^r_1 = \bigwedge_{i,j \leq r} x_i = x_j$.
Moreover, we define a function let : $\mathbb{N} \longrightarrow \mathbb{N}$ by

$$\text{let}(s) = \min\{r \mid \text{There exist} \geq s \text{ distinct equality types of } r\text{–tuples consisting of at most two elements.}\}.$$

□

Then for each structure[6] $\mathfrak{A}$ (of an arbitrary signature) and each $i \leq r$ there exists a tuple $\overset{r}{a} \in A$ such that $\mathfrak{A} \models \varepsilon^r_i[\overset{r}{a}]$. Moreover, for tuples $\overset{r}{a}, \overset{r}{b} \in A$ such that

$$\mathfrak{A} \models \varepsilon^r_i[\overset{r}{a}] \quad \text{and} \quad \mathfrak{A} \models \varepsilon^r_j[\overset{r}{b}]$$

for some $i \neq j \leq r$ we have $\overset{r}{a} \neq \overset{r}{b}$.

Suggestively we write $\overset{r}{x} = \boldsymbol{i}$ instead of $\varepsilon^r_i(\overset{r}{x})$. If the context allows it we use other similar notations, for example we abbreviate $\exists \bar{x}(\bar{x} = \boldsymbol{i} \wedge R\bar{x}\bar{t})$ by $R\boldsymbol{i}\bar{t}$.[7]

1.12 Lemma on Simultaneous Induction ([Mos74]) : *The following formulae are equivalent:*

$$[s\text{–}FP_{\overset{k_1}{x_1}, X_1, \ldots, \overset{k_m}{x_m}, X_m} \; \varphi_1(\overset{k_1}{x_1}, X_1 \ldots, X_m), \ldots, \varphi_m(\overset{k_m}{x_m}, X_1 \ldots, X_m)] \; \overset{k_1}{u}$$

and

$$\begin{aligned}
\chi' := [FP_{\overset{k}{x}\overset{r}{y}, Y} \quad & \overset{r}{y} = \mathbf{1} \quad \wedge \quad \varphi_1(\overset{k_1}{x}, Y_\mathbf{1}, \ldots, Y_\boldsymbol{m}) \\
\vee \; & \overset{r}{y} = \mathbf{2} \quad \wedge \quad \varphi_2(\overset{k_2}{x}, Y_\mathbf{1}, \ldots, Y_\boldsymbol{m}) \\
& \qquad\qquad \vdots \\
\vee \; & \overset{r}{y} = \boldsymbol{m} \quad \wedge \quad \varphi_m(\overset{k_m}{x}, Y_\mathbf{1}, \ldots, Y_\boldsymbol{m})] \; \overset{k_1}{u} \; \underbrace{u_1 \ldots u_1}_{(k-k_1)\ times} \; \underbrace{u_1 \ldots u_1}_{r\ times}
\end{aligned}$$

[6] Recall that a structure contains at least 2 elements.

[7] The reader may note that we are a bit careless with the variables here. We should mention that the variables in $\bar{x}$ are distinct, and that they are distinct from the variables that occur in $\bar{t}$.

However, we will not hesitate to be careless in the same way in similar situations throughout the whole thesis.

where $r = \mathrm{let}(m)$, $k = \max\{k_1, \ldots, k_m\}$, *and* $\varphi_i(\overset{k_1}{x}, Y\_\boldsymbol{1}, \ldots, Y\_\boldsymbol{m})$ $(i \leq m)$ *is the formula obtained from* φ_i *by replacing each subformula of the form* $X_j \overset{k_j}{t}$ *by* $Y \overset{k_j}{t} \underbrace{t_1 \ldots t_1}_{(k-k_j)\ times} \boldsymbol{j}$ *(for each* $j \leq m$*).*

1.13 Remark : We decided that $Y\_\boldsymbol{i}$ is an abbreviation for $\exists \bar{x}(\bar{x} = \boldsymbol{i} \wedge Y\_\bar{x})$. But note that the formula χ' above would not change its meaning if we would replace it by $\forall \bar{x}(\bar{x} = \boldsymbol{i} \rightarrow Y\_\bar{x})$.
This is useful if we want to preserve the quantifier–structure of the φ_i. For example, if all φ_i $(i \leq m)$ are existential we can obtain a χ' with an existential first–order part by replacing each positive occurrence of a subformula $X_j\bar{t}$ in a φ_i by $\exists \bar{x}(\bar{x} = \boldsymbol{j} \wedge Y\bar{t}\bar{x})$ and each negative occurrence by $\forall \bar{x}(\bar{x} = \boldsymbol{j} \rightarrow Y\bar{t}\bar{x})$. □

§ 1.6 Datalog

We only give a very brief introduction into the database language datalog (which is strongly related to the fixed–point logics); all definitions are given in a way that fits immediately into our formal frame. For a more detailed presentation of the definitions and results given here we refer the reader to [EF].

1.14 Definition : A *general datalog program* Π is a finite set of *clauses* of the form

$$\gamma \leftarrow \gamma_1, \ldots, \gamma_l$$

where $l \geq 0$ and $\gamma_1, \ldots, \gamma_l$ are atomic or negated atomic formulae. Moreover, γ must be of the form $X\bar{x}$ where X is a relation variable. It is called the *head* of the clause; $\gamma_1, \ldots, \gamma_l$ form the *body* of the clause.
Relation variables that occur in the heads of clauses in Π are called *intensional*; all other relation symbols and relation variables that occur in Π are called *extensional*. The *signature* of a general datalog program is the set of all extensional relation symbols and constant symbols that occur in the program.

One intensional relation variable of Π is distinguished to be the *goal* variable. □

To define the semantics, let Π be a general datalog program, $X_1, \ldots, X_m$ the intensional relation variables of Π, and X_1 the goal variable.
For each $j \leq m$ let $\bar{x}_j$ be a tuple (whose length matches the arity of X_j) of distinct individual variables that do not occur in Π. Consider a clause

$$\kappa = X_j\bar{x} \leftarrow \gamma_1, \ldots, \gamma_l\ .$$

We assume that the variables $\bar{x}$ are distinct. (If they are not, say, if $x_i = x_{i'}$ we replace the i'th place of $\bar{x}$ in the head of the clause by a new variable $x'_{i'}$ and extend the body of our clause by $x_i = x'_{i'}$.) Replacing each occurrence of x_i in κ by x_{ji} we obtain a new clause

$$\kappa' = X_j \bar{x}_j \leftarrow \gamma'_1, \ldots, \gamma'_l \ .$$

Let $\bar{y}$ be a tuple consisting of the variables that occur in κ' but not in $\bar{x}_j$,

$$\varphi_\kappa := \exists \bar{y} \bigwedge_{i=1}^{l} \gamma'_i \ ,$$

and

$$\varphi_j := \bigvee \{\varphi_\kappa \mid \kappa \in \Pi, \ X_j \text{ occurs in the head of } \kappa\}.$$

We define Π to be equivalent to the s–FP–formula

$$\varphi_\Pi := [s\text{–}FP_{\bar{x}_1, X_1, \ldots, \bar{x}_m, X_m} \ \varphi_1, \ldots, \varphi_m]\bar{x} \ ,$$

i.e. for each interpretation $\mathfrak{I}$ we let

$$\mathfrak{I} \models \Pi \ :\Longleftrightarrow \ \mathfrak{I} \models \varphi_\Pi \quad \text{and} \quad \mathfrak{I} \models_I \Pi \ :\Longleftrightarrow \ \mathfrak{I} \models_I \varphi_\Pi \ .$$

To be able to define Boolean queries even by datalog programs not containing any constant symbols it is reasonable to allow 0–ary relation variables in general datalog programs.

Denoting the class of all general datalog programs by D we see that

$$\begin{aligned} \textbf{P–Datalog} \ &:= \ (D, \models) \subseteq \textbf{PFP} \quad \text{and} \\ \textbf{I–Datalog} \ &:= \ (D, \models_I) \subseteq \textbf{IFP} \ . \end{aligned}$$

In fact equality holds, a theorem which is essentially due to Abiteboul and Vianu [AV91a]. In the present form it can be found in [EF].

1.15 Theorem ([AV91a, EF]) : **P–Datalog=PFP** *and* **I–Datalog=IFP**.

For a **proof** see Remark 2.3 below.

1.16 Remark : A *(pure) datalog program* is a general datalog program where no negation symbols occur. It can be seen that a query is definable by a pure datalog program if and only if it is definable by a positive $\exists LFP$–formula, i.e. an $\exists LFP$–formula without any negation symbols. (Although this is not completely trivial since proving the Lemma on Simultaneous Induction for positive **∃LFP** requires some thoughts). By **Datalog** we denote the logic whose formulae are the pure datalog programs.

We say that a *datalog$^\neg$ program* is a general datalog program where negation symbols never occur in front of intensional relation variables; and we denote the corresponding logic by **Datalog**$^\neg$. Then **Datalog**$^\neg$=**∃LFP**.

There is also a natural class of general datalog programs, so called *stratified datalog programs*, such that the corresponding logic **S–Datalog** has the same expressive power as **SFP**.

We call **Datalog**, **Datalog**$^\neg$, **S–Datalog**, **I–Datalog**, and **P–Datalog** the *datalogics.* $\square$

§1.7 Ordered Structures

Structures with a built–in order play an important role in descriptive complexity theory. For our investigations they are not so important, but they will occur from time to time when we relate our results to complexity theory.
It is useful then to have the following terminology and results available.

We let $\mathcal{O}$ be the class of all structures $\mathfrak{A}$ whose signature contains the binary relation symbols $<$ and S and the constant symbols $\underline{0}$ and $\underline{e}$ such that:

- $<^{\mathfrak{A}}$ is a linear order of A.
- $S^{\mathfrak{A}}$ is the successor relation that belongs to $<^{\mathfrak{A}}$.
- $\underline{0}^{\mathfrak{A}}$ is the $<^{\mathfrak{A}}$–minimal and $\underline{e}^{\mathfrak{A}}$ is the $<^{\mathfrak{A}}$–maximal element of A.

For two logics $\mathbf{L}$ and $\tilde{\mathbf{L}}$, or more generally two classes of queries, we say that $\mathbf{L}$ is contained in $\tilde{\mathbf{L}}$ *on ordered structures* (we write $\mathbf{L} \subseteq_{\mathcal{O}} \tilde{\mathbf{L}}$) if every query on $\mathcal{O}$ in $\mathbf{L}$ is also in $\tilde{\mathbf{L}}$. We say that $\mathbf{L}$ equals $\tilde{\mathbf{L}}$ *on ordered structures* ($\mathbf{L} =_{\mathcal{O}} \tilde{\mathbf{L}}$) if $\mathbf{L} \subseteq_{\mathcal{O}} \tilde{\mathbf{L}}$ and $\tilde{\mathbf{L}} \subseteq_{\mathcal{O}} \mathbf{L}$.

We treat complexity classes as classes of queries in a standard way. Then the following well–known results hold:

1.17 Theorem ([BG87, Imm86, Var82]) :

$$\exists\mathbf{LFP} =_{\mathcal{O}} \mathbf{IFP} =_{\mathcal{O}} \mathbf{PTIME}$$
$$\mathbf{PFP} =_{\mathcal{O}} \mathbf{PSPACE}$$

The following theorem of Abiteboul and Vianu gives a link between ordered and arbitrary structures:

1.18 Theorem ([AV91b]) :

$$\mathbf{IFP} = \mathbf{PFP} \quad\Longleftrightarrow\quad \mathbf{IFP} =_{\mathcal{O}} \mathbf{PFP} \quad\Longleftrightarrow\quad \mathbf{PTIME} = \mathbf{PSPACE}$$

2 The Quantifier Structure of Fixed–Point Formulae

The starting point for our considerations is the following well–known normal form theorem. It is due to Immerman [Imm87] for least fixed–point formulae and hence inductive formulae. For partial (and also inductive) fixed–point logic it is an easy consequence of Abiteboul's and Vianu's Theorem 1.15. In Appendix A we sketch a direct proof of the **PFP** and **IFP**–part.

2.1 Normal Form Theorem ([Imm87, AV91a]) : *For every fixed–point formula χ there is an equivalent formula χ'* in normal form*, i.e. of the form*

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\varphi]v\dots v,$$ [8]

where φ is first–order and v does not occur in φ.

If χ is inductive then χ' can also be chosen to be inductive. If $\chi \in LFP$ we can choose a φ that is positive in X. If $\chi \in \exists LFP$ then χ' can be chosen to be in $\exists LFP$.

This section is concerned with the question of how far we can simplify the quantifier–structure of the first order part φ of a fixed–point formula in normal form without loosing expressive power. We will see that we can simplify it considerably:
For **IFP** and **PFP** an existential φ suffices, and for *LFP*-formulae the first–order part φ needs to be Δ_2.[9] Because of its similarity to the Skolem normal form for existential second order logic, Dahlhaus calls this last form a *Skolem normal form* for **LFP**.
We will also see that both the existential normal form for **IFP** and **PFP** and the Skolem normal form for **LFP** are optimal (looked at from the standpoint of quantifier–structure).

The results we give in this section are by no means new, the existential normal form for **IFP** and **PFP** given in the first paragraph is an easy consequence of Abiteboul's and Vianu's Theorem 1.15, and the Skolem Normal Form of the second paragraph is due to Dahlhaus [Dah87].

However, our proofs are new and lead to a very unified treatment of the whole subject. In fact it seems to make things easier if we prove the normal form for **PFP** first and then turn to **P-Datalog** than the other way round, as can be

[8] We write $\overset{\exists}{\forall} x\varphi$ for formulae $\exists x\varphi$ (or $\forall x\varphi$, whatever the reader prefers) if $\models \exists x\varphi \leftrightarrow \forall x\varphi$.

[9] $\Sigma_0 = \Pi_0 = \Delta_0$ is the class of quantifier–free first–order formulae. For each $i \geq 0$ we let Σ_{i+1} (Π_{i+1}) be the class of first–order formulae of the form $\exists\bar{x}\varphi$ where φ is Π_i ($\forall\bar{x}\varphi$ where φ is Σ_i respectively). A Δ_i–formula is a formula that is equivalent to a Σ_i–formula and to a Π_i–formula.

done using Abiteboul's and Vianu's theorem. **IFP** can be treated similarly to **PFP**, although the proofs are a bit more involved. The Skolem normal form for **LFP** and the fact that **IFP** = **LFP** are obtained in one final step.
Our approach has one disadvantage, it does not deliver the normal form for $\exists$**LFP**.

The other principal way to obtain the normal forms for **IFP** and **LFP** is to work with **LFP**. The two essential steps are the Transitivity Lemma for positive **LFP** (i.e. the fragment of **LFP** where negation symbols only occur in front of atomic subformulae) and the proof that positive **LFP** is closed under negations. As a by–product of the latter one gets **LFP** = **IFP**.
This gives a proof of Theorem 2.1 for **LFP**, **IFP**, and $\exists$**LFP**. Dahlhaus starts at this point to get his Skolem Normal Form. **PFP** has to be treated seperately.

See Remark A.6 for a further discussion of the two approaches.

§ 2.1 The Existential Normal Form for IFP and PFP

2.2 Theorem : *Every fixed–point formula χ is equivalent to a formula χ' in* existential normal form, *i.e. a formula in normal form where the first–order part is existential.*

If χ is inductive then χ' can be chosen to be inductive.

Proof: By Theorem 2.1 it suffices to prove that for $\chi = \overset{\exists}{\forall} v[FP_{\bar{x},X}\varphi]v\ldots v$ in normal form where φ is Σ_k there exists an equivalent $\chi' = \overset{\exists}{\forall} v[FP_{\bar{y},Y}\varphi']v\ldots v$ in normal form where φ' is Σ_{k-1}.
To understand this proof it might be helpful to recall the example given in Paragraph 1.4 showing how to eliminate universal quantifiers from *IFP*–formulae.

Consider the formula

$$\chi = \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \exists\bar{y}\forall\bar{z}\psi(\bar{x},\bar{y},\bar{z},X)]v\ldots v$$

where ψ is Σ_{k-2}.
We let

$$\begin{aligned}\varphi_1(\bar{x},Y_1,Y_2,Y_3) &:= \exists\bar{y}\neg Y_2\bar{x}\bar{y} \wedge \exists x Y_3 x\\ \varphi_2(\bar{x}\bar{y},Y_1,Y_2,Y_3) &:= \exists\bar{z}\neg\psi(\bar{x},\bar{y},\bar{z},Y_1)\\ \varphi_3(x,Y_1,Y_2,Y_3) &:= x = x\end{aligned}$$

and

$$\xi := \overset{\exists}{\forall} v[s\text{–}FP_{\bar{x},Y_1,\bar{x}\bar{y},Y_2,x,Y_3}\ \varphi_1,\varphi_2,\varphi_3]v\ldots v\ .$$

Note that φ_1 is Σ_1, φ_2 is Σ_{k-1}, and φ_3 is Σ_0.

Let $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation for χ and $X_i^{\mathfrak{I}}$, $Y_{ji}^{\mathfrak{I}}$ (for $i \geq 0, j \leq 3$) be the stages of the respective fixed–point processes of χ and ξ in $\mathfrak{I}$.

Claim 1: For all $m \geq 0$ and $\bar{a}, \bar{b} \in A$ we have:

$$\begin{array}{lll}
(i)\,(a) & \bar{a} \in Y_{1(2m)}^{\mathfrak{I}} & \iff \bar{a} \in X_m^{\mathfrak{I}} \\
\,(b) & & \iff \bar{a} \in Y_{1(2m+1)}^{\mathfrak{I}} \\
(ii)(a) & \bar{a}\bar{b} \in Y_{2(2m+1)}^{\mathfrak{I}} & \iff \neg\forall \bar{c} \in A : \mathfrak{I} \models \psi[\bar{a}, \bar{b}, \bar{c}, X_m^{\mathfrak{I}}] \\
(b) & & \iff \bar{a}\bar{b} \in Y_{2(2m+2)}^{\mathfrak{I}}
\end{array}$$

Moreover we have $Y_{30}^{\mathfrak{I}} = \emptyset$ and $Y_{3m}^{\mathfrak{I}} = A$ for each $m \geq 1$.

Proof: The statement on $(Y_{3m}^{\mathfrak{I}})_{m \geq 0}$ is obvious. We prove (i) and (ii) by induction on m:

$m = 0:$ $(i)(a)$ is immediate, and $(i)(b)$ holds because $Y_{30} = \emptyset$. Then (ii) follows from the definition of φ_2.

$m \to m+1:$ Since $Y_{3(2m+1)}^{\mathfrak{I}} \neq \emptyset$ we have:

$$\begin{array}{rll}
\bar{a} \in Y_{1(2m+2)} & \iff & \exists \bar{b} : \bar{a}\bar{b} \notin Y_{2(2m+1)}^{\mathfrak{I}} \\
& \iff & \exists \bar{b} \forall \bar{c} : \mathfrak{I} \models \psi[\bar{a}, \bar{b}, \bar{c}, X_m^{\mathfrak{I}}] \\
& & \text{(by induction hypothesis } (ii)(a)) \\
& \iff & \bar{a} \in X_{m+1}^{\mathfrak{I}} \ .
\end{array}$$

Statement $(i)(b)$ follows from the induction hypothesis $(ii)(b)$ similarly. Then $(ii)(a)$ and $(ii)(b)$ are immediate from the definition of φ_2.

Hence Claim 1 is proved.
It clearly implies $\mathfrak{I} \models \chi \iff \mathfrak{I} \models \xi$. To complete the proof for arbitrary fixed–point formulae we apply the Lemma on Simultaneous Induction. Note that by Remark 1.13 the first–order part of the resulting formula can be chosen to be equivalent to a Σ_{k-1}–formula.

Inductive Formulae

The situation is more complicated for inductive χ (if we also want to obtain an inductive χ'). So let us suppose now that χ is inductive.
Considering the formula ξ defined above we see that the sequence $(Y_{1i}^{\mathfrak{I}})_{i \geq 0}$ is inductive (i.e. $Y_{1i}^{\mathfrak{I}} \subseteq Y_{1(i+1)}^{\mathfrak{I}}$ for all $i \geq 0$), but we cannot expect the sequence $(Y_{2i}^{\mathfrak{I}})_{i \geq 0}$ to be inductive.
However, we can use a standard trick (that is often applied when working with **IFP**) to obtain an inductive sequence that contains the same information. Instead of the sequence $(Y_{2i}^{\mathfrak{I}})_{i \geq 0}$ we define a sequence

$$\Big(\bigcup_{j=1}^{i} (Y_{2j}^{\mathfrak{I}} \times (Y_{1j}^{\mathfrak{I}} \setminus Y_{1(j-1)}^{\mathfrak{I}}))\Big)_{i \geq 0},$$

i.e. we "stamp" each tuple in $Y^{\mathfrak{I}}_{2i}$ with the time j it has entered the sequence (by $Y^{\mathfrak{I}}_{1j} \setminus Y^{\mathfrak{I}}_{1(j-1)}$). Since the sequence $(Y^{\mathfrak{I}}_{1i})_{i\geq 0}$ is inductive all sets $Y^{\mathfrak{I}}_{1i} \setminus Y^{\mathfrak{I}}_{1(i-1)}$ are disjoint.
However, trying to do so we have some trouble with the first step. To cope with this we have to include an artificial first step where we include all tuples whose (new) last two components are distinct. The "interesting" tuples, however, are only those whose last two components are equal.
We need four simultaneous processes here defined by formulae $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_4$. The processes defined by $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$ correspond to those defined by φ_1 and φ_2 in the general case. The processes defined by $\tilde{\varphi}_3$ and $\tilde{\varphi}_4$ are needed for technical reasons (to calculate $Y^{\mathfrak{I}}_{1i} \setminus Y^{\mathfrak{I}}_{1(i-1)}$ during the simultaneous induction).

We let

$$\begin{aligned}
\tilde{\varphi}_1(\bar{x}x'x'', \tilde{Y}_1, \ldots, \tilde{Y}_4) &:= \tilde{Y}_1\bar{x}x'x'' \vee x' \neq x'' \\
&\quad \vee \exists \bar{y}, \bar{z}z'z''(\tilde{Y}_3\bar{z}z'z'' \wedge \neg\tilde{Y}_4\bar{z}z'z'' \wedge \neg\tilde{Y}_2\bar{x}\bar{y}\bar{z}z'z'') \\
\tilde{\varphi}_2(\bar{x}\bar{y}\bar{z}z'z'', \tilde{Y}_1, \ldots, \tilde{Y}_4) &:= \tilde{Y}_2\bar{x}\bar{y}\bar{z}z'z'' \\
&\quad \vee(\tilde{Y}_1\bar{z}z'z'' \wedge \neg\tilde{Y}_3\bar{z}z'z'' \wedge \exists\bar{w}\neg\psi(\bar{x},\bar{y},\bar{w},\tilde{Y}_1__x_1x_1)) \\
\tilde{\varphi}_3(\bar{x}x'x'', \tilde{Y}_1, \ldots, \tilde{Y}_4) &:= \tilde{Y}_1\bar{x}x'x'' \\
\tilde{\varphi}_4(\bar{x}x'x'', \tilde{Y}_1, \ldots, \tilde{Y}_4) &:= \tilde{Y}_3\bar{x}x'x''
\end{aligned}$$

and

$$\tilde{\xi} = \overset{\exists}{\forall} v[s\text{–}FP_{\bar{x}x'x'',\tilde{Y}_1,\bar{x}\bar{y}\bar{z}z'z'',\tilde{Y}_2,\bar{x}x'x'',\tilde{Y}_3,\bar{x}x'x'',\tilde{Y}_4}\ \tilde{\varphi}_1, \ldots, \tilde{\varphi}_4]v \ldots v\ .$$

Obviously $\tilde{\xi}$ is inductive.

Again we let $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation for χ and we let $X^{\mathfrak{I}}_i$, $\tilde{Y}^{\mathfrak{I}}_{ji}$ (for $i \geq 0, j \leq 4$) be the stages of the respective fixed–point processes. Here the situation looks as follows:

Claim 2: For all $m \geq 0$ and $\bar{a}, a', a'', \bar{b}, \bar{c}, c', c'' \in A$ we have:

$$\begin{array}{lllll}
(i) & (a) & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{1(2m+1)} & \Longleftrightarrow & a' \neq a'' \vee \bar{a} \in X^{\mathfrak{I}}_m \\
 & (b) & & \Longleftrightarrow & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{1(2m+2)} \\
(ii) & (a) & \bar{a}\bar{b}\bar{c}c'c'' \in \tilde{Y}^{\mathfrak{I}}_{2(2m+2)} & \Longleftrightarrow & \exists l \in \{0, \ldots, m\} : \\
 & & & & \quad (\bar{c}c'c'' \in \tilde{Y}^{\mathfrak{I}}_{1(2l+1)} \setminus \tilde{Y}^{\mathfrak{I}}_{1(2l)} \\
 & & & & \quad \wedge \neg\forall \bar{d} \in A : \mathfrak{I} \models \psi[\bar{a}, \bar{b}, \bar{d}, X^{\mathfrak{I}}_l]) \\
 & (b) & & \Longleftrightarrow & \bar{a}\bar{b}\bar{c}c'c'' \in \tilde{Y}^{\mathfrak{I}}_{2(2m+3)} \\
(iii) & (a) & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{3(2m+2)} & \Longleftrightarrow & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{1(2m+1)} \\
 & (b) & & \Longleftrightarrow & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{3(2m+3)} \\
(iv) & (a) & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{4(2m+3)} & \Longleftrightarrow & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{1(2m+1)} \\
 & (b) & & \Longleftrightarrow & \bar{a}a'a'' \in \tilde{Y}^{\mathfrak{I}}_{4(2m+4)}\ .
\end{array}$$

Proof: Again the proof is by induction over m:

$m = 0$: Since $\tilde{Y}_{30}^{\mathfrak{J}} = \emptyset$ we have $\bar{a}a'a'' \in \tilde{Y}_{11}^{\mathfrak{J}}$ if and only if $a' \neq a''$ thus $(i)(a)$ holds. We also have $\tilde{Y}_{31}^{\mathfrak{J}} = \emptyset$ (because $\tilde{Y}_{10}^{\mathfrak{J}} = \emptyset$) hence $(i)(b)$ follows analogously.

(iii) and (iv) are obvious by the definition of $\tilde{\varphi}_3$ and $\tilde{\varphi}_4$ respectively.

$(ii)(a)$ holds by the definition of $\tilde{\varphi}_2$ and $(i)(a)$, and $(ii)(b)$ follows immediately from the fact that $\tilde{Y}_{12}^{\mathfrak{J}} \setminus \tilde{Y}_{32}^{\mathfrak{J}} = \tilde{Y}_{12}^{\mathfrak{J}} \setminus \tilde{Y}_{11}^{\mathfrak{J}} = \emptyset$.

$m \to m+1$: (iii) and (iv) are obvious.

If $m > 0$ and the sequence $(X_i^{\mathfrak{J}})_{i \geq 0}$ has already reached a fixed–point, i.e. $X_{m-1}^{\mathfrak{J}} = X_m^{\mathfrak{J}}$, then by the induction hypothesis we have

$$\tilde{Y}_{1(2m+2)}^{\mathfrak{J}} = \tilde{Y}_{1(2m+1)}^{\mathfrak{J}} = \tilde{Y}_{1(2m)}^{\mathfrak{J}} = \tilde{Y}_{1(2m-1)}^{\mathfrak{J}}$$

thus

$$\tilde{Y}_{3(2m+2)}^{\mathfrak{J}} = \tilde{Y}_{3(2m+1)}^{\mathfrak{J}} = \tilde{Y}_{3(2m)}^{\mathfrak{J}} \quad \text{and} \quad \tilde{Y}_{4(2m+2)}^{\mathfrak{J}} = \tilde{Y}_{4(2m+1)}^{\mathfrak{J}} \, .$$

Moreover, $\tilde{Y}_{1(2m+1)}^{\mathfrak{J}} \setminus \tilde{Y}_{1(2m)}^{\mathfrak{J}} = \emptyset$ implies $\tilde{Y}_{2(2m+2)}^{\mathfrak{J}} = \tilde{Y}_{2(2m+1)}^{\mathfrak{J}}$. Thus all sequences have already reached their fixed–point and we are finished.

Otherwise we have $\tilde{Y}_{1(2m)}^{\mathfrak{J}} \subsetneq \tilde{Y}_{1(2m+1)}^{\mathfrak{J}}$ because one of the following two cases occurs:

- If $m = 0$ we have $\tilde{Y}_{1(2m+1)}^{\mathfrak{J}} \setminus \tilde{Y}_{1(2m)}^{\mathfrak{J}} = \{\bar{a}a'a'' \in A \mid a' \neq a''\} \neq \emptyset$.
- If $m > 0$ and $X_m^{\mathfrak{J}} \neq X_{m-1}^{\mathfrak{J}}$ we have $X_{m-1}^{\mathfrak{J}} \subsetneq X_m^{\mathfrak{J}}$ (since χ is inductive) and thus by induction hypothesis $\tilde{Y}_{1(2m)}^{\mathfrak{J}} \subsetneq \tilde{Y}_{1(2m+1)}^{\mathfrak{J}}$.

Hence:

$$\begin{aligned}
& \bar{a}a'a'' \in \tilde{Y}_{1(2(m+1)+1)}^{\mathfrak{J}} \\
\iff & \bar{a}a'a'' \in \tilde{Y}_{1(2m+2)}^{\mathfrak{J}} \vee a' \neq a'' \\
& \vee \exists \bar{b}, \bar{c}c'c'' (\bar{c}c'c'' \in \tilde{Y}_{3(2m+2)} \setminus \tilde{Y}_{4(2m+2)} \wedge \bar{a}\bar{b}\bar{c}c'c'' \notin \tilde{Y}_{2(2m+2)}^{\mathfrak{J}}) \\
\iff & \bar{a} \in X_m^{\mathfrak{J}} \vee a' \neq a'' \\
& \vee \exists \bar{b}, \bar{c}c'c'' (\bar{c}c'c'' \in \tilde{Y}_{1(2m+1)} \setminus \tilde{Y}_{1(2m)} \wedge \forall \bar{d} : \mathfrak{J} \models \psi[\bar{a}, \bar{b}, \bar{d}, X_m^{\mathfrak{J}}]) \\
& \text{(by the induction hypothesis and since the sets} \\
& \ \tilde{Y}_{1(2l+1)} \setminus \tilde{Y}_{1(2l)} \ (0 \leq l \leq m) \text{ are all disjoint)} \\
\iff & \bar{a} \in X_m^{\mathfrak{J}} \vee a' \neq a'' \vee \bar{a} \in X_{m+1}^{\mathfrak{J}} \\
& \text{(since } \tilde{Y}_{1(2m)}^{\mathfrak{J}} \subsetneq \tilde{Y}_{1(2m+1)}^{\mathfrak{J}}) \\
\iff & a' \neq a'' \vee \bar{a} \in X_{m+1}^{\mathfrak{J}} \\
& \text{(since } X_m^{\mathfrak{J}} \subseteq X_{m+1}^{\mathfrak{J}}).
\end{aligned}$$

This proves $(i)(a)$. Statement $(i)(b)$ follows from the fact that

$$\tilde{Y}_{3(2m+3)}^{\mathfrak{J}} \setminus \tilde{Y}_{4(2m+3)}^{\mathfrak{J}} = \tilde{Y}_{1(2m+2)}^{\mathfrak{J}} \setminus \tilde{Y}_{1(2m+1)}^{\mathfrak{J}} = \emptyset.$$

$(ii)(a)$ follows from $(i)(a)$ and the definition of $\tilde{\varphi}_2$, and $(ii)(b)$ is an immediate consequence of

$$\tilde{Y}^{\mathfrak{J}}_{1(2m+4)} \setminus \tilde{Y}^{\mathfrak{J}}_{3(2m+4)} = \tilde{Y}^{\mathfrak{J}}_{1(2m+4)} \setminus \tilde{Y}^{\mathfrak{J}}_{1(2m+3)} = \emptyset$$

(by $(i)(b)$).

Thus Claim 2 is proved.

Note that item (i) implies that $\chi \dashv\models \tilde{\xi}$.[10]
An application of the Lemma on Simultaneous Induction considering Remark 1.13 completes the proof. ■

2.3 Remark : It is clear that together with the direct proof of Theorem 2.1 in the Appendix this gives us a proof of Abiteboul's and Vianu's Theorem 1.15.

On the other hand, as we have already pointed out, our theorem follows immediately from Abiteboul's and Vianu's by an application of the Lemma on Simultaneous Induction. □

2.4 Remark : We may ask if we can strengthen the theorem by requiring the first–order part to be Δ_1, i.e. we can ask if each fixed–point formula is equivalent to a formula of the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\theta]v \ldots v$$

where θ is a Δ_1–formula.
To see that we cannot, consider a formula $\chi = [FP_{\bar{x},X}\theta]\bar{u}$ where θ is Δ_1.
Recall (from Paragraph 1.2) the definition of the formulae χ_i $(i \geq 0)$ defining the stages. An easy induction shows that all χ_i $(i \geq 0)$ are Δ_1–formulae.

Let $\mathfrak{I} \subseteq \mathfrak{J}$ be two interpretations for χ such that $\mathfrak{J} \models \chi$. Then there exists an $i \geq 1$ such that

$$\mathfrak{J} \models \chi_i(\bar{u}) \wedge \forall \bar{x}(\chi_i(\bar{x}) \leftrightarrow \chi_{i+1}(\bar{x})) \ .$$

Since χ_i and χ_{i+1} are Δ_1–formulae this formula is equivalent to a Π_1–formula hence preserved under substructures.
Thus

$$\begin{aligned} & \mathfrak{I} \models \chi_i(\bar{u}) \wedge \forall \bar{x}(\chi_i(\bar{x}) \leftrightarrow \chi_{i+1}(\bar{x})) \\ \Longrightarrow \ & \mathfrak{I} \models \chi \ . \end{aligned}$$

Hence χ is preserved under substructures.
Since there exist fixed–point formulae which are not preserved under substructures our statement is proved. □

[10] By $\chi \dashv\models \chi'$ we denote that χ and χ' are equivalent, i.e. for all (finite) interpretations $\mathfrak{J}$ we have $\mathfrak{J} \models \chi \iff \mathfrak{J} \models \chi'$.

§2.2 The Skolem Normal Form for LFP

2.5 Theorem ([GS86, Imm86, Dah87]) : *Every IFP–formula (and hence LFP–formula) is equivalent to a formula in* Skolem normal form, *i.e. of the form* $\overset{\exists}{\forall} v[FP_{\bar{x},X}\varphi]v\ldots v$ *with a* Δ_2*–formula* φ *in which* X *only occurs positively and* v *does not occur.*

Proof: By Theorem 2.2 it suffices to consider an *IFP*–formula

$$\chi = [FP_{\overset{k}{x},X}\varphi(\overset{k}{x},X)]v\ldots v$$

in existential normal form.
Let $\mathfrak{I} = (\mathfrak{A},\alpha)$ be an interpretation. For each $\overset{k}{a}\in A$ we let $|\overset{k}{a}|_\chi$ be the least ordinal i such that $\overset{k}{a}\in X_i^{\mathfrak{I}}$ if such an i exists, otherwise we let $|\overset{k}{a}|_\chi = \infty$.
The so called *stage comparison relations* $\leq_\chi$, $\preccurlyeq_\chi$, and $\prec_\chi$ of χ are defined as follows:
For all $\overset{k}{a},\overset{k}{b}\in A$,

$$\overset{k}{a}\leq_\chi\overset{k}{b} \quad :\Longleftrightarrow \quad |\overset{k}{a}|_\chi \leq |\overset{k}{b}|_\chi < \infty\ ,$$

$$\overset{k}{a}\preccurlyeq_\chi\overset{k}{b} \quad :\Longleftrightarrow \quad |\overset{k}{a}|_\chi \leq |\overset{k}{b}|_\chi \quad \text{and} \quad |\overset{k}{a}|_\chi < \infty\ ,$$

and

$$\overset{k}{a}\prec_\chi\overset{k}{b} \quad :\Longleftrightarrow \quad |\overset{k}{a}|_\chi < |\overset{k}{b}|_\chi\ .$$

The following observations can be easily verified:

(i) $\forall\, \overset{k}{a}\in A:\ \mathfrak{I}\models [FP_{\overset{k}{x},X}\varphi]\,\overset{k}{a} \Longleftrightarrow \exists i\geq 1: \overset{k}{a}\in X_i^{\mathfrak{I}} \Longleftrightarrow \overset{k}{a}\leq_\chi\overset{k}{a}$

(ii) $\forall\, \overset{k}{a}\in X_\infty^{\mathfrak{I}}:\ \{\overset{k}{b}\in A^k \mid \overset{k}{b}\leq_\chi\overset{k}{a}\}^C = \{\overset{k}{b}\in A^k \mid \overset{k}{a}\prec_\chi\overset{k}{b}\}$

(iii) $\forall\, \overset{k}{a}\in X_\infty^{\mathfrak{I}}:\ \{\overset{k}{b}\in A^k \mid \overset{k}{b}\prec_\chi\overset{k}{a}\}^C = \{\overset{k}{b}\in A^k \mid \overset{k}{a}\preccurlyeq_\chi\overset{k}{b}\}$

(iv) $\forall\, \overset{k}{a},\overset{k}{b}\in A: \overset{k}{a}\leq_\chi\overset{k}{b} \Longleftrightarrow \exists i\geq 0: \overset{k}{b}\notin X_i^{\mathfrak{I}} \wedge\ \overset{k}{a},\overset{k}{b}\in X_{i+1}^{\mathfrak{I}}$

(v) $\forall\, \overset{k}{a},\overset{k}{b}\in A: \overset{k}{a}\preccurlyeq_\chi\overset{k}{b} \Longleftrightarrow \exists i\geq 0: \overset{k}{b}\notin X_i^{\mathfrak{I}} \wedge\ \overset{k}{a}\in X_{i+1}^{\mathfrak{I}}$

(vi) $\forall\, \overset{k}{a},\overset{k}{b}\in A: \overset{k}{a}\prec_\chi\overset{k}{b} \Longleftrightarrow \exists i\geq 1: \overset{k}{a}\in X_i^{\mathfrak{I}} \wedge\ \overset{k}{b}\notin X_i^{\mathfrak{I}}$

Recalling that $X_{i+1}^{\mathfrak{I}} = \{\overset{k}{a}\in A \mid \mathfrak{I}\models\varphi[\overset{k}{a},X_i^{\mathfrak{I}}]\}$ for each $i\geq 1$, we get the following:

$$\begin{aligned}
\text{(vii)}\ \forall\, \overset{k}{a},\overset{k}{b}\in A: \overset{k}{a}\leq_\chi\overset{k}{b} \Longleftrightarrow \quad & (\mathfrak{I}\models\varphi[\overset{k}{a},\emptyset]\ \wedge\ \mathfrak{I}\models\varphi[\overset{k}{b},\emptyset]) \\
& \vee \exists\, \overset{k}{c}\, (\overset{k}{c}\leq_\chi\overset{k}{c}\ \wedge\ \overset{k}{c}\prec_\chi\overset{k}{b}\ \wedge \mathfrak{I}\models\varphi[\overset{k}{a},\{\overset{k}{d}|\overset{k}{d}\leq_\chi\overset{k}{c}\}] \\
& \qquad \wedge \mathfrak{I}\models\varphi[\overset{k}{b},\{\overset{k}{d}|\overset{k}{d}\leq_\chi\overset{k}{c}\}])
\end{aligned}$$

(viii) $\forall \overset{k}{a}, \overset{k}{b} \in A : \overset{k}{a} \preccurlyeq_\chi \overset{k}{b} \Longleftrightarrow$

$$\mathfrak{I} \models \varphi[\overset{k}{a}, \emptyset] \vee \exists \overset{k}{c} \,(\overset{k}{c} \leq_\chi \overset{k}{c} \wedge \overset{k}{c} \prec_\chi \overset{k}{b} \wedge \mathfrak{I} \models \varphi[\overset{k}{a}, \{\overset{k}{d} | \overset{k}{d} \leq_\chi \overset{k}{c}\}])$$

(ix) $\forall \overset{k}{a}, \overset{k}{b} \in A : \overset{k}{a} \prec_\chi \overset{k}{b} \Longleftrightarrow \overset{k}{a} \leq_\chi \overset{k}{a} \wedge \mathfrak{I} \not\models \varphi[\overset{k}{b}, \{\overset{k}{c} | \overset{k}{c} \prec_\chi \overset{k}{a}\}]$

These three items can be directly translated into a definition of $\leq_\chi$, $\preccurlyeq_\chi$, and $\prec_\chi$ by simultaneous induction:
Let

$$\begin{aligned}\varphi'_{\leq}(\overset{k}{x}\overset{k}{y}, \leq, \preccurlyeq, \prec) := \quad & (\varphi(\overset{k}{x}, \emptyset) \wedge \varphi(\overset{k}{y}, \emptyset)) \\ & \vee \exists \overset{k}{z} \,(\overset{k}{z} \leq \overset{k}{z} \wedge \overset{k}{z} \prec \overset{k}{y} \wedge \varphi(\overset{k}{x}, \_ \leq \overset{k}{z}) \wedge \varphi(\overset{k}{y}, \_ \leq \overset{k}{z})),\end{aligned}$$

$$\varphi'_{\preccurlyeq}(\overset{k}{x}\overset{k}{y}, \leq, \preccurlyeq, \prec) := \varphi(\overset{k}{x}, \emptyset) \vee \exists \overset{k}{z} \,(\overset{k}{z} \leq \overset{k}{z} \wedge \overset{k}{z} \prec \overset{k}{y} \wedge \varphi(\overset{k}{x}, \_ \leq \overset{k}{z})),$$

and

$$\varphi'_{\prec}(\overset{k}{x}\overset{k}{y}, \leq, \preccurlyeq, \prec) := \overset{k}{x} \leq \overset{k}{x} \wedge \neg\varphi(\overset{k}{y}, \_ \prec \overset{k}{x}) \;.$$

To obtain formulae that are positive in $\leq$, $\preccurlyeq$, and $\prec$ we use the following trick: We let $\varphi^+(\overset{k}{x}, \_ \leq \overset{k}{z})$ be the formula obtained from $\varphi(\overset{k}{x}, X)$ by replacing each subformula of the form $X\,\overset{k}{t}$ that occurs positively (i.e. in the scope of an even number of negation symbols) by $\overset{k}{t} \leq \overset{k}{z}$ and each subformula of the form $X\,\overset{k}{t}$ that occurs negatively by $\neg\, \overset{k}{z} \prec \overset{k}{t}$. Then $\varphi^+(\overset{k}{x}, \_ \leq \overset{k}{z})$ is positive in $\leq$, $\preccurlyeq$, and $\prec$.
Similarly, let $\varphi^-(\overset{k}{y}, \_ \prec \overset{k}{x})$ be the formula obtained from $\varphi(\overset{k}{y}, X)$ by replacing each subformula of the form $X\,\overset{k}{t}$ that occurs positively by $\neg\, \overset{k}{x} \preccurlyeq \overset{k}{t}$ and each subformula of the form $X\,\overset{k}{t}$ that occurs negatively by $\overset{k}{t} \prec \overset{k}{x}$. Then $\leq$, $\preccurlyeq$, and $\prec$ only occur negatively in $\varphi^-(\overset{k}{y}, \_ \prec \overset{k}{z})$.
Now we let

$$\begin{aligned}\varphi_{\leq}(\overset{k}{x}\overset{k}{y}, \leq, \prec) := \quad & (\varphi(\overset{k}{x}, \emptyset) \wedge \varphi(\overset{k}{y}, \emptyset)) \\ & \vee \exists \overset{k}{z} \,(\overset{k}{z} \leq \overset{k}{z} \wedge \overset{k}{z} \prec \overset{k}{y} \wedge \varphi^+(\overset{k}{x}, \_ \leq \overset{k}{z}) \wedge \varphi^+(\overset{k}{y}, \_ \leq \overset{k}{z}))\end{aligned}$$

$$\varphi_{\preccurlyeq}(\overset{k}{x}\overset{k}{y}, \leq, \preccurlyeq, \prec) := \varphi(\overset{k}{x}, \emptyset) \vee \exists \overset{k}{z} \,(\overset{k}{z} \leq \overset{k}{z} \wedge \overset{k}{z} \prec \overset{k}{y} \wedge \varphi^+(\overset{k}{x}, \_ \leq \overset{k}{z})),$$

and

$$\varphi_{\prec}(\overset{k}{x}\overset{k}{y}, \leq, \prec) := \overset{k}{x} \leq \overset{k}{x} \wedge \neg\varphi^-(\overset{k}{y}, \_ \prec \overset{k}{x}) \;.$$

Note that $\leq$ and $\prec$ only occur positively in both $\varphi_\leq$ and $\varphi_\prec$. Moreover, since φ is existential $\varphi_\leq$ is also existential and $\varphi_\prec$ is universal.
Let

$$\chi' = \substack{\exists \\ \forall}\, v[s\text{-}FP_{\overset{k}{x}\overset{k}{y}, \leq, \overset{k}{x}\overset{k}{y}, \preccurlyeq, \overset{k}{x}\overset{k}{y}, \prec}\, \varphi_\leq, \varphi_\preccurlyeq, \varphi_\prec]v \ldots v \;.$$

By the Lemma on Simultaneous Induction and Remark 1.13, χ' is equivalent to an *LFP*–formula $\overset{\exists}{\forall} v[FP_{\bar{x},X}\psi]v\dots v$ with a Δ_2–formula ψ.

To verify that $\chi \mathrel{=\!\!|\!\!\models} \chi'$ we shall prove:

CLAIM: $\leq_\infty^{\mathfrak{I}} = \leq_\chi$, $\preccurlyeq_\infty^{\mathfrak{I}} = \preccurlyeq_\chi$, and $\prec_\infty^{\mathfrak{I}} = \prec_\chi$

For each $i \geq 0$ we define relations $\leq_\chi^i$, $\preccurlyeq_\chi^i$, and $\prec_\chi^i$ by

$$\leq_\chi^i := \{\overset{k}{a}\overset{k}{b} \mid \overset{k}{a} \leq_\chi \overset{k}{b},\ \overset{k}{a}, \overset{k}{b} \in X_i^{\mathfrak{I}}\},$$

$$\preccurlyeq_\chi^i := \{\overset{k}{a}\overset{k}{b} \mid \overset{k}{a} \preccurlyeq_\chi \overset{k}{b},\ \overset{k}{a} \in X_i^{\mathfrak{I}}\},$$

and

$$\prec_\chi^i := \{\overset{k}{a}\overset{k}{b} \mid \overset{k}{a} \prec_\chi \overset{k}{b},\ \overset{k}{a} \in X_i^{\mathfrak{I}}\}$$

and note that $\leq_\chi = \bigcup_{i\geq 0} \leq_\chi^i$, $\preccurlyeq_\chi = \bigcup_{i \geq 0} \preccurlyeq_\chi^i$, and $\prec_\chi = \bigcup_{i\geq 0} \prec_\chi^i$.

By (vii)–(ix) we get for each $i \geq 0$

$$\leq_\chi^{i+1} = \varphi'_{\leq}(\_\_, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i),\text{[11]}$$

$$\preccurlyeq_\chi^{i+1} = \varphi'_{\preccurlyeq}(\_\_, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i),$$

and

$$\prec_\chi^{i+1} = \varphi'_{\prec}(\_\_, \leq_\chi^{i+1}, \preccurlyeq_\chi^{i+1}, \prec_\chi^i).$$

Observe that for all $i \geq 0$ and $\overset{k}{b} \in X_{i+1}^{\mathfrak{I}}$ we have

$$\{\overset{k}{a} \in A \mid \overset{k}{a} \leq_\chi^{i+1} \overset{k}{b}\}^C = \{\overset{k}{a} \in A \mid \overset{k}{b} \prec_\chi^{i+1} \overset{k}{a}\} \quad \text{and} \quad \{\overset{k}{a} \in A \mid \overset{k}{a} \prec_\chi^i \overset{k}{b}\}^C = \{\overset{k}{a} \in A \mid \overset{k}{b} \preccurlyeq_\chi^{i+1} \overset{k}{a}\}.$$

Hence by the definition of φ^+ and φ^- we have

$$\leq_\chi^{i+1} = \varphi_{\leq}(\_\_, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i),$$

$$\preccurlyeq_\chi^{i+1} = \varphi_{\preccurlyeq}(\_\_, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i),$$

and

$$\prec_\chi^{i+1} = \varphi_{\prec}(\_\_, \leq_\chi^{i+1}, \preccurlyeq_\chi^{i+1}, \prec_\chi^i).$$

Using the fact that $\varphi_{\leq}$, $\varphi_{\preccurlyeq}$, and $\varphi_{\prec}$ are positive and hence monotone in $\leq$, $\preccurlyeq$, and $\prec$, an easy induction shows that for each $i \geq 0$

$$\leq_\chi^{\lceil \frac{i}{2} \rceil} \subseteq \leq_i^{\mathfrak{I}} \subseteq \leq_\chi^i \quad , \quad \preccurlyeq_\chi^{\lceil \frac{i}{2} \rceil} \subseteq \preccurlyeq_i^{\mathfrak{I}} \subseteq \preccurlyeq_\chi^i \quad \text{and} \quad \prec_\chi^{\lfloor \frac{i}{2} \rfloor} \subseteq \prec_i^{\mathfrak{I}} \subseteq \prec_\chi^i \quad .$$

---

[11]By $\varphi'_{\leq}(\_\_, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i)$ we denote the set $\{\overset{k}{a}\overset{k}{b} \in A \mid \mathfrak{I} \models \varphi'_{\leq}[\overset{k}{a}\overset{k}{b}, \leq_\chi^i, \preccurlyeq_\chi^i, \prec_\chi^i]\}$, the notation in the other cases being similar.

This implies the claim. ■

2.6 Remark : It is not hard to see that the normal form given above is optimal. An existential or universal first–order part does not suffice, since $\exists LFP$–formulae are preserved under extensions and universal LFP–formula are preserved under substructures (cf. Lemma 1.8 and Remark 1.9), hence do not suffice to capture **LFP**. Note that here we use that $\overset{\exists}{\forall}$ can be read as an existential or as a universal quantifier. □

3 Bounded Normal Forms

In this section we want to bound the number of occurrences of the relation variable X in a fixed–point formula $[FP_{\bar{x},X}\varphi]\bar{u}$.
We will see that we can restrict ourselves to formulae with one positive and one negative occurrence of X without loosing expressive power. To capture **LFP** and hence **IFP** one positive occurrence already suffices.
However, this is not too surprising because it is an immediate consequence of a theorem of Moschovakis [Mos74] (Canonical Form for Positive Formulae) which essentially says that each formula φ where X only occurs positively is equivalent to a formula φ' where X only occurs once, still positively.

Our new contribution is that we can bound the number of occurrences of X within the normal forms we obtained in the last section, i.e. essentially without changing the quantifier structure of the formulae.
The main result of this section (in fact of the whole Part I) is a *bounded normal form* for **IFP** and **PFP**. We shall see that this normal form is in a sense optimal. As an application of this result we will obtain *complete problems* for the logics under *quantifier–free reductions* in Section 5.

We derive a similar normal form for $\exists$**LFP** and get some information on the structure of *SFP*–formulae (a normal form in our sense does not exist for **SFP**). In the second paragraph we give a *Bounded Skolem Form* for **LFP**.

§ 3.1 Boundedness

3.1 Definition : Let $k, l \geq 0$ and X be a relation variable. A formula φ is called (k,l)*–bounded in* X if it is of the form

$$\psi_0 \vee \exists \bar{x}_1 \in X \ldots \exists \bar{x}_k \in X \; \exists \bar{y}_1 \notin X \ldots \exists \bar{y}_l \notin X \; \psi^{12}$$

where X does not occur in ψ_0 or ψ.

A fixed–point formula φ is called *(k,l)–bounded* if for every subformula of the form $[FP_{\bar{x},X}\varphi]\bar{u}$ there exist $k' \leq k, l' \leq l$ such that φ is (k', l')–bounded in X. $\square$

The ψ_0 in a formula which is (k,l)–bounded in X is needed to take care of the cases in which $X = \emptyset$ or $X = \mathcal{U}$.[13]
The definition is justified by the following lemma:

3.2 Lemma : *Let φ be a fixed–point formula where negation symbols only occur in front of atomic subformulae, and X a relation variable that does not occur within the scope of any universal quantifiers or fixed–point operators in φ.*

[12] $\exists x \in X\varphi$ abbreviates $\exists x(Xx \wedge \varphi)$ and $\exists x \notin X\varphi$ abbreviates $\exists x(\neg Xx \wedge \varphi)$.
[13] $X = \emptyset$ and $X = \mathcal{U}$ are used as a synonym for $\forall x \neg Xx$ and $\forall x Xx$, respectively.

Let k be the number of positive occurrences of X in φ and l the number of negative occurrences.
Then there exists a formula ψ where X does not occur such that:

$$X \neq \emptyset \wedge X \neq \mathcal{U} \models \varphi \leftrightarrow \exists \bar{x}_1 \in X \ldots \exists \bar{x}_k \in X \; \exists \bar{y}_1 \notin X \ldots \exists \bar{y}_l \notin X \; \psi \; .$$

If φ is existential first–order then so is ψ. If $l = 0$ the condition $X \neq \mathcal{U}$ is not needed.

Proof: The proof is a simple induction over φ, using the fact that for formulae ξ, ξ' and a variable x that does not occur freely in ξ we have:

$$\begin{array}{lcl} X \neq \emptyset & \models & \xi \wedge \exists x(Xx \wedge \xi') \leftrightarrow \exists x(Xx \wedge \xi \wedge \xi') \\ X \neq \emptyset & \models & \xi \vee \exists x(Xx \wedge \xi') \leftrightarrow \exists x(Xx \wedge (\xi \vee \xi')) \\ X \neq \emptyset & \models & \exists y \exists x(Xx \wedge \xi') \leftrightarrow \exists x(Xx \wedge \exists y \xi') \; . \end{array}$$

The analogous statements hold for $X \neq \mathcal{U}$ on the left and $\neg Xx$ on the right side of $\models$. ■

3.3 Corollary : *Every SFP–formula is equivalent to a $(k, 0)$–bounded formula (for some $k \geq 0$).*

Proof: Note that for *SFP*–formulae $\chi = [FP_{\bar{x},X} \; \varphi(\bar{x}, X)]\bar{u}$ we have

$$\chi \mathrel{=\!\!|\!\!\models} [FP_{\bar{x},X} \; \varphi(\bar{x}, \emptyset) \vee \varphi(\bar{x}, X)]\bar{u}$$

(by the definition of *SFP*, using the fact that *SFP*–formulae are inductive).
By Lemma 3.2 there exist $k \geq 0, \psi$ such that

$$X \neq \emptyset \models \varphi \leftrightarrow \exists \bar{y}_1 \in X \ldots \exists \bar{y}_k \in X \; \psi \; .$$

and X does not occur in ψ. Then

$$\chi \mathrel{=\!\!|\!\!\models} [FP_{\bar{x},X} \; \varphi(\bar{x}, \emptyset) \vee \exists \bar{y}_1 \in X \ldots \exists \bar{y}_k \in X \; \psi]\bar{u} \; .$$

■

The next definition, together with the following lemma, is very useful and will be needed several times:

3.4 Definition : A fixed–point formula $\chi = [FP_{\overset{r}{x},X} \varphi] \overset{r}{u}$ is called *modest* if for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ and for all $i \geq 1$ we have $\emptyset \subsetneq X_i^{\mathfrak{I}} \subsetneq A^r$. □

3.5 Lemma : *The formulae* $[FP_{\overset{r}{x},X}\varphi(\overset{r}{x},X)]\overset{r}{u}$ *and*

$$\chi' := [FP_{\overset{r+2}{x},Y}(x_{r+1}\neq x_{r+2}\wedge x_r = x_{r+1})\vee(x_{r+1}=x_{r+2}\wedge\varphi(\overset{r}{x},Y\_x_1x_1)]\overset{r}{u}\,u_1u_1$$

are equivalent; and χ' *is modest.*

The **proof** is an easy induction on the stages of the fixed–point processes. ∎

Now we can state our normal form theorem:

3.6 Theorem : *Every fixed–point formula* χ *is equivalent to a formula* χ' *in* bounded normal form*, i.e. of the form*

$$\overset{\exists}{\forall}v[FP_{\bar{x},X}\ \theta_0\vee\exists\bar{y}\in X\exists\bar{z}\notin X\theta]v\ldots v$$

where θ_0 *and* θ *are quantifier–free formulae, and* X *and* v *do not occur in* θ *or* θ_0*.*
If χ *is inductive then* χ' *can also be chosen to be inductive.*

Proof: STEP 1: REDUCTION TO A BOUNDED FORMULA

By the existential normal form Theorem 2.2 and Lemma 3.5 we can assume without loss of generality that χ is a modest formula of the form

$$\overset{\exists}{\forall}v[FP_{\overset{r}{x},X}(\underbrace{x_{r-1}\neq x_r\wedge x_{r-2}=x_{r-1}}_{=:\psi_0(\overset{r}{x})})\vee(x_{r-1}=x_r\wedge\varphi)]v\ldots v$$

where φ is an existential first–order formula and v does not occur in φ. Applying Lemma 3.2 to $(x_{r-1}=x_r\wedge\varphi)$ we get an existential first–order formula ψ in which X does not occur such that

$$X\neq\emptyset\wedge X\neq\mathcal{U}\models\quad(x_{r-1}=x_r\wedge\varphi)$$
$$\leftrightarrow\exists\overset{r}{y}_1\in X\ldots\exists\overset{r}{y}_k\in X\exists\overset{r}{z}_1\notin X\ldots\exists\overset{r}{z}_l\notin X\ \psi\ .$$

An easy induction (which uses the fact that χ is modest[14]) shows that χ is equivalent to

$$\chi_1 := \overset{\exists}{\forall}v[FP_{\overset{r}{x},X}\psi_0(\overset{r}{x})\vee\exists\overset{r}{y}_1\in X\ldots\exists\overset{r}{y}_k\in X\ \exists\overset{r}{z}_1\notin X\ldots\exists\overset{r}{z}_l\notin X\ \psi)]v\ldots v\ .$$

Note that χ_1 is still modest; in particular this implies that we can increase k and l artificially. So we can assume without loss of generality that $k=l\geq 1$.

[14]If a formula ξ is modest, for convenience we also call the formula $\overset{\exists}{\forall}v\xi$ modest.

Step 2: Reduction to a (1,1)–Bounded Formula

The idea of the proof is to replace X by a $2kr$–ary relation Y that always contains k r–tuples in X and k r–tuples not in X. Then one tuple from Y gives us all the information we need to reach the next stage. The problem is that building the next suitable Y requires not only one tuple from X, but k tuples from X and k tuples that are not in X. We take in these tuples one after the other in $2k$ steps.

Recall Definition 1.11 concerning the equality types of 2 elements. Let $s = \mathrm{let}(2k)$. Also recall that we write $\bar{x} = \boldsymbol{i}$ instead of $\varepsilon_i^s(\bar{x})$ and that $\varepsilon_1^s(\bar{x})$ is the formula $\bigwedge_{i,j} x_i = x_j$.
Let Y be a $(s+2kr)$–ary relation variable. The first s places in Y are needed for the equality types. The next kr places are reserved for the k r–tuples in X, and the last kr places for the k r–tuples not in X. We denote tuples intended to be in Y by $\bar{x}_0\bar{x}_1\ldots\bar{x}_{2k}$ instead of $\overset{s}{x}_0\overset{r}{x}_1\ldots\overset{r}{x}_{2k}$.
Let

$$\chi_2 := \overset{\exists}{\forall} v\Big[FP_{\bar{x}_0\bar{x}_1\ldots\bar{x}_{2k},Y}\ (\bar{x}_0 = \mathbf{1} \wedge \psi_0(\bar{x}_1)) \\ \vee \exists \bar{y}_0\bar{y}_1\ldots\bar{y}_{2k} \in Y \exists \bar{z}_0\bar{z}_1\ldots\bar{z}_{2k} \notin Y \textstyle\bigvee_{i=1}^{2k}(\bar{x}_0 = \boldsymbol{i} \wedge \psi_i)\Big] v \ldots v$$

where

$$\psi_1 = \Big(\bar{y}_0 = \boldsymbol{2k} \ \wedge\ \psi(\bar{x}_1, \bar{y}_1, \ldots, \bar{y}_{2k})\Big),$$

$$\psi_{2i} = \Big(\bar{y}_0 = \boldsymbol{2i{-}1} \ \wedge\ \bar{z}_0 = \bar{y}_0 \ \wedge\ \bigwedge_{j=2}^{2k} \bar{z}_j = \bar{y}_j \ \wedge\ \bar{x}_{k+i} = \bar{z}_1 \ \wedge \bigwedge_{\substack{1\le j\le 2k,\\ j\ne k+i}} \bar{x}_j = \bar{y}_j\Big)$$

for $1 \le i \le k$, and

$$\psi_{2i+1} = \Big(\bar{y}_0 = \boldsymbol{2i} \ \wedge\ \bar{z}_0 = \bar{y}_0 \ \wedge \bigwedge_{\substack{1\le j\le 2k,\\ j\ne k+1}} \bar{z}_j = \bar{y}_j \ \wedge\ \bar{x}_i = \bar{z}_{k+1} \ \wedge \bigwedge_{\substack{1\le j\le 2k,\\ j\ne i}} \bar{x}_j = \bar{y}_j\Big)$$

for $1 \le i \le k-1$.

Let $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation for χ and $(X_m)_{m\ge0}$, $(Y_m)_{m\ge0}$ be the stages of the respective fixed–point processes of χ_1 and χ_2 in $\mathfrak{I}$.[15]
By $Y_i\boldsymbol{j}\bar{a}_1\ldots\bar{a}_{2k}$ we abbreviate

$$\exists \bar{a}_0 \in A:\ \mathfrak{A} \models \varepsilon_j^s[\bar{a}_0] \ \wedge\ Y_i\bar{a}_0\bar{a}_1\ldots\bar{a}_{2k}\ .$$

It can be easily verified from the definition of χ_2 that this is equivalent to

$$\forall \bar{a}_0 \in A:\ \mathfrak{A} \models \varepsilon_j^s[\bar{a}_0] \rightarrow Y_i\bar{a}_0\bar{a}_1\ldots\bar{a}_{2k}\ .$$

Let $Y_i\boldsymbol{j}\_\_$ denote the set $\{\bar{a}_1\ldots\bar{a}_{2k} \in A \mid Y_i\boldsymbol{j}\bar{a}_1\ldots\bar{a}_{2k}\}$.

[15] For convenience we leave out the superscripts $\mathfrak{I}$ from the X_i and Y_i.

CLAIM: For all $m \geq 0$ and $1 \leq i \leq k$ we have:

$$(2i-1) \quad Y_{2km+2i-1}\boldsymbol{2i{-}1}\_ = \underbrace{X_{m+1} \times \ldots \times X_{m+1} \times X_{m+1}}_{i \text{ times}} \times \underbrace{A^r \times \ldots \times A^r}_{(k-i) \text{ times}}$$
$$\times \underbrace{X^C_{m+1} \times \ldots \times X^C_{m+1}}_{(i-1) \text{ times}} \times \underbrace{A^r \times A^r \times \ldots \times A^r}_{(k-i+1) \text{ times}}$$

$$(2i) \quad Y_{2km+2i}\boldsymbol{2i}\_ = \underbrace{X_{m+1} \times \ldots \times X_{m+1}}_{i \text{ times}} \times \underbrace{A^r \times \ldots \times A^r}_{(k-i) \text{ times}}$$
$$\times \underbrace{X^C_{m+1} \times \ldots \times X^C_{m+1}}_{i \text{ times}} \times \underbrace{A^r \times \ldots \times A^r}_{(k-i) \text{ times}}$$

and for all $i \leq 2k, j \leq 2k-1$

$$(*) \quad Y_{2km+i}\boldsymbol{i}\_ = Y_{2km+i+j}\boldsymbol{i}\_ \; .$$

The proof is by induction over m:

$m = 0:$ (1)–($2k$) are proved inductively:

> Since $Y_1 = \{\bar{a}_0\bar{a}_1 \ldots \bar{a}_{2k} \in A \mid \bar{a}_0 = \mathbf{1}, \overbrace{\mathfrak{I} \models \psi_0[\bar{a}_1]}^{\Longleftrightarrow X_1 \bar{a}_1}\}$ statement (1) holds for $m = 0$.
>
> Now suppose that $(2i-1)$ is proved and let $\bar{a}_1 \ldots \bar{a}_{2k} \in Y_{2i}\boldsymbol{2i}\_$. Then by the definition of ψ_{2i} there exists an $\bar{a}$ such that
>
> $$\bar{a}_1 \ldots \bar{a}_{k+i-1}\bar{a}\bar{a}_{k+i+1} \ldots \bar{a}_{2k} \in Y_{2i-1}\boldsymbol{2i{-}1}\_$$
>
> and
> $$\bar{a}_{k+i}\bar{a}_2 \ldots \bar{a}_{k+i-1}\bar{a}\bar{a}_{k+i+1} \ldots \bar{a}_{2k} \notin Y_{2i-1}\boldsymbol{2i{-}1}\_ .$$
> By the induction hypothesis this means $\bar{a}_1, \ldots, \bar{a}_i \in X_1$ and $\bar{a}_{k+1}, \ldots, \bar{a}_{k+i-1} \notin X_1$ hence $\bar{a}_{k+i} \notin X_1$ (otherwise we would have
>
> $$\bar{a}_{k+i}\bar{a}_2 \ldots \bar{a}_{k+i-1}\bar{a}\bar{a}_{k+i+1} \ldots \bar{a}_{2k} \in Y_{2i-1}\boldsymbol{2i{-}1}\_$$
>
> by the induction hypothesis).
> This proves the $\subseteq$–direction of $(2i)$.
>
> For the other direction suppose that
>
> $$\bar{a}_1 \ldots \bar{a}_{2k} \in \underbrace{X_1 \times \ldots \times X_1}_{i \text{ times}} \times \underbrace{A^r \times \ldots \times A^r}_{(k-i) \text{ times}}$$
> $$\times \underbrace{X^C_1 \times \ldots \times X^C_1}_{i \text{ times}} \times \underbrace{A^r \times \ldots \times A^r}_{(k-i) \text{ times}} \subseteq Y_{2i-1}\boldsymbol{2i{-}1}\_ .$$

Then

$$\bar{a}_{k+i}\bar{a}_2 \dots \bar{a}_{2k} \notin \underbrace{X_1 \times \dots \times X_1 \times X_1}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \times \underbrace{X_1^C \times \dots \times X_1^C}_{(i-1)\ times} \times \underbrace{A^r \times A^r \times \dots \times A^r}_{(k-i+1)\ times} = Y_{2i-1}\boldsymbol{2i-1}\_ \,.$$

By the definition of ψ_{2i} this implies

$$\bar{a}_1 \dots \bar{a}_{2k} \in Y_{2i}\boldsymbol{2i}\_$$

hence the $\supseteq$–direction.

We proceed similarly to prove $(2i+1)$ from $(2i)$.

$(*)$ can easily be seen considering that $Y_i\boldsymbol{2k}\_$ is empty until $i = 2k$, so $Y_i\boldsymbol{1}\_$ does not change before the $(2k+1)$th step. Thus $Y_i\boldsymbol{2}\_$ does not change (after its first change in the second step) before the $(2k+2)$th step, and so on.

$m \to m+1$: To prove (1) for $m+1$ instead of m recall that $\bar{a} \in X_{m+2}$ if and only if $\mathfrak{I} \models \psi_0[\bar{a}]$ or

$$\exists \bar{a}_1, \dots, \bar{a}_k \in X_{m+1}, \bar{b}_1, \dots, \bar{b}_k \notin X_{m+1} : \mathfrak{I} \models \psi[\bar{a}, \bar{a}_1, \dots, \bar{a}_k, \bar{b}_1, \dots, \bar{b}_k] \,.$$

By the induction hypothesis we have

$$Y_{2km+2k}\boldsymbol{2k}\_ = \underbrace{X_{m+1} \times \dots \times X_{m+1}}_{k\ times} \times \underbrace{X_{m+1}^C \times \dots \times X_{m+1}^C}_{k\ times} \,.$$

Thus (1) is obvious for $m+1$ from the definition of ψ_1. (2)–$(2k)$ follow by an induction completely analogous to the case $m = 0$.

$(*)$ can also be seen analogously.

So the claim is proved, and as a consequence we have:

$$\mathfrak{I} \models \chi_2 \iff \mathfrak{I} \models \chi_1 \,.$$

To conclude: In this step we obtained a formula χ_2 that is equivalent to χ and of the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},Y}\theta_0(\bar{x}) \vee \exists \bar{y} \in Y, \bar{z} \notin Y \xi(\bar{x}, \bar{y}, \bar{z})]v \dots v$$

where θ_0 is quantifier–free, ξ is existential first–order, and Y and v do not occur in θ_0 or ξ.

STEP 3: ELIMINATION OF THE EXISTENTIAL QUANTIFIERS IN ξ.

Let p be the arity of Y and suppose $\xi = \exists u_1 \dots \exists u_q \theta(\overset{p}{x}, \overset{p}{y}, \overset{p}{z}, u_1, \dots, u_q)$ for a quantifier–free θ. Let Z be a $(p+q)$–ary relation variable and

$$\chi' = \overset{\exists}{\forall} v[FP_{\overset{p+q}{x},Z}\ \theta_0(\overset{p}{x}) \vee \exists \overset{p+q}{y} \in Z \exists \overset{p+q}{z} \notin Z\theta(\overset{p}{x}, \overset{p}{y}, \overset{p}{z}, y_{p+1}, \dots, y_{p+q})]v \dots v\ .$$

An easy induction shows that for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ and for each $i \geq 0$ we have:

$$Z_i^{\mathfrak{I}} = Y_i^{\mathfrak{I}} \times A^q$$

Thus $\chi_2 \mathrel{=\!\mid\!\models} \chi'$.

STEP 4: MODIFICATIONS FOR INDUCTIVE FORMULAE

If χ is inductive we want χ' to be inductive, too. Unfortunately the formula χ_2 is not necessarily inductive even if χ is inductive. But some slight modifications will do in this case:
Note first that for inductive χ the formula χ_1 is inductive, so Step 1 goes through without any changes.
To simulate Step 2 we define a formula

$$\begin{aligned}\tilde{\chi}_2 := \overset{\exists}{\forall} v\Big[FP_{\bar{x}_0\bar{x}_1\dots\bar{x}_{2k},Y}\ \ &(\bar{x}_0 = \mathbf{1} \wedge \psi_0(\bar{x}_1))\\ &\vee \exists \bar{y}_0\bar{y}_1 \dots \bar{y}_{2k} \in Y \exists \bar{z}_0\bar{z}_1 \dots \bar{z}_{2k} \notin Y\\ &\qquad (\textstyle\bigwedge_{i=0}^{2k} \bar{x}_i = \bar{y}_i \ \vee\ \bigvee_{i=1}^{2k}(\bar{x}_0 = \boldsymbol{i} \wedge \psi_i))\Big]v \dots v.\end{aligned}$$

(The ψ_i $(0 \leq i \leq 2k)$ are the same formulae as in the original Step 2).
Clearly $\tilde{\chi}_2$ is inductive. To see that $\tilde{\chi}_2 \mathrel{=\!\mid\!\models} \chi_1$ we modify the claim from Step 2 as follows (using the same notation):

MODIFIED CLAIM: For all $m \geq 0$ and $1 \leq i \leq k$ we have:

$$(2i-1) \quad Y_{2km+2i-1}\boldsymbol{2i-1}\,\_\_ = \bigcup_{l=0}^{m} \underbrace{X_{l+1} \times \dots \times X_{l+1} \times X_{l+1}}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \times \underbrace{X_{l+1}^C \times \dots \times X_{l+1}^C}_{(i-1)\ times} \times \underbrace{A^r \times A^r \times \dots \times A^r}_{(k-i+1)\ times}$$

$$(2i) \quad Y_{2km+2i}\boldsymbol{2i}\,\_\_ = \bigcup_{l=0}^{m} \underbrace{X_{l+1} \times \dots \times X_{l+1}}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \times \underbrace{X_{l+1}^C \times \dots \times X_{l+1}^C}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times}$$

and for all $i \leq 2k, j \leq 2k-1$

$$(*) \quad Y_{2km+i}\boldsymbol{i}\,\_\_ = Y_{2km+i+j}\boldsymbol{i}\,\_\_\ .$$

The proof is again by induction over m, for $m=0$ we are in the same situation as before so there is nothing (new) to prove.
So we suppose the claim is proved for m and verify (1)–($2k$) for (m+1) inductively.
To prove (1) let

$$\bar{a}_1 \dots \bar{a}_{2k} \in Y_{2km+2k}\boldsymbol{2k}\_ = \bigcup_{l=0}^{m} \underbrace{X_{l+1} \times \dots \times X_{l+1}}_{k\ times} \times \underbrace{X_{l+1}^C \times \dots \times X_{l+1}^C}_{k\ times},$$

say

$$\bar{a}_1 \dots \bar{a}_{2k} \in X_{l_0+1} \times \dots \times X_{l_0+1} \times X_{l_0+1}^C \times \dots \times X_{l_0+1}^C,$$

and suppose that $\mathfrak{I} \models \psi_1(\bar{a}, \bar{a}_1, \dots \bar{a}_{2k})$. Then $\bar{a} \in X_{l_0+2}$ thus the $\subseteq$–direction of (1) holds. The $\supseteq$–direction is similarly simple.

To prove ($2i$) (from ($2i-1$)) we take some $\bar{a}_1 \dots \bar{a}_{2k} \in Y_{2k(m+1)+2i-1}\boldsymbol{2i-1}\_$, $\bar{b}_1 \dots \bar{b}_{2k} \notin Y_{2k(m+1)+2i-1}\boldsymbol{2i-1}\_$ such that $\bar{a}_j = \bar{b}_j$ for all $j \geq 2$. By the induction hypothesis there exists an $l \leq m+1$ such that

$$\bar{a}_1, \dots, \bar{a}_i \in X_{l+1} \quad \text{and} \quad \bar{a}_{k+1}, \dots, \bar{a}_{k+i-1} \notin X_{l+1}.$$

Then clearly $\bar{b}_1 \notin X_{l+1}$ and thus

$$\begin{aligned}\bar{a}_1 \dots \bar{a}_{k+i-1}\bar{b}_1\bar{a}_{k+i+1} \dots \bar{a}_{2k} \in \quad & \underbrace{X_{l+1} \times \dots \times X_{l+1}}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \\ & \times \underbrace{X_{l+1}^C \times \dots \times X_{l+1}^C}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times},\end{aligned}$$

which proves the $\subseteq$–direction of ($2i$).
For the $\supseteq$–direction suppose that

$$\begin{aligned}\bar{a}_1 \dots \bar{a}_{2k} \in \quad & \underbrace{X_{l+1} \times \dots \times X_{l+1}}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \\ & \times \underbrace{X_{l+1}^C \times \dots \times X_{l+1}^C}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times}\end{aligned}$$

for some $l \geq 0$. Suppose that l is minimal with respect to this condition.
If $l < m+1$ then $\bar{a}_1 \dots \bar{a}_{2k} \in Y_{2k(m+1)+2i}\boldsymbol{2i}\_$ by induction hypothesis and the clause $\bigwedge_{i=0}^{2k} \bar{x}_i = \bar{y}_i$ of $\tilde{\chi}_2$.
So suppose $l = m+1$, i.e.

$$\begin{aligned}\bar{a}_1 \dots \bar{a}_{2k} \in \quad & \underbrace{X_{m+2} \times \dots \times X_{m+2}}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times} \\ & \times \underbrace{X_{m+2}^C \times \dots \times X_{m+2}^C}_{i\ times} \times \underbrace{A^r \times \dots \times A^r}_{(k-i)\ times}\ .\end{aligned}$$

Then

$$\bar{a}_{k+i}\bar{a}_2\bar{a}_3\ldots\bar{a}_{2k} \notin \underbrace{X_{m+2}\times\ldots\times X_{m+2}\times X_{m+2}}_{i\ times}\times\underbrace{A^r\times\ldots\times A^r}_{(k-i)\ times}$$
$$\times\underbrace{X^C_{m+2}\times\ldots\times X^C_{m+2}}_{(i-1)\ times}\times\underbrace{A^r\times\ldots\times A^r\times A^r}_{(k-i+1)\ times}\ .$$

Since χ is inductive we have $X^C_{m+2}\subseteq X^C_{l+1}$ for each $l\le m+1$, thus

$$\bar{a}_{k+i}\bar{a}_2\bar{a}_3\ldots\bar{a}_{2k} \notin \bigcup_{l=0}^{m+1} \underbrace{X_{l+l}\times\ldots\times X_{l+1}\times X_{l+1}}_{i\ times}\times\underbrace{A^r\times\ldots\times A^r}_{(k-i)\ times}$$
$$\times\underbrace{X^C_{l+1}\times\ldots\times X^C_{l+1}}_{(i-1)\ times}\times\underbrace{A^r\times\ldots\times A^r\times A^r}_{(k-i+1)\ times}\ .$$

So by the induction hypothesis $(2i-1)$ we have

$$\bar{a}_{k+i}\bar{a}_2\bar{a}_3\ldots\bar{a}_{2k}\notin Y_{2k(m+1)+2i-1}\boldsymbol{2i-1}\_\ .$$

On the other hand we have $\bar{a}_1\ldots\bar{a}_{2k}\in Y_{2k(m+1)+2i-1}\boldsymbol{2i-1}\_$ and hence

$$\bar{a}_1\ldots\bar{a}_{2k}\in Y_{2k(m+1)+2i}\boldsymbol{2i}\_\ .$$

This proves the $\supseteq$–direction.

$(2i+1)$ can be proved similarly from $(2i)$, and $(*)$ can be seen as in Step 2. Thus the modified claim is proved.

Finally, we apply Step 3 to $\tilde{\chi}_2$ instead of χ_2. It goes through without any difficulties and keeps the formula inductive. ■

3.7 Remark : Note that the formula χ' we obtained in the proof of the last theorem is modest. We will use this when defining complete problems for **PFP** and **IFP** in Paragraph 5.4. □

3.8 Remark : Our normal form is in a sense optimal since formulae with only positive occurrences of the relation variables in fixed–point operators are *LFP*–formulae, and if we want them to be existential we obtain $\exists$*LFP*–formulae which do not suffice to capture **IFP**.
On the other hand, formulae with only negative occurrences of the relation variables in fixed–point operators are equivalent to formulae with only positive occurrences because if X only occurs negatively in $\varphi(\bar{x},X)$ then it only occurs positively in $\varphi(\bar{x},\varphi(\_,X))$ and we have:

$$[FP_{\bar{x},X}\ \varphi(\bar{x},X)]\bar{u} \mathrel{=\!\!|\!\!\models} \underbrace{[FP_{\bar{x},X}\ \varphi(\bar{x},\varphi(\_,X))]\bar{u}}_{=:\psi(\bar{u})} \wedge\forall\bar{y}(\psi(\bar{y})\leftrightarrow\varphi(\bar{y},\psi(\_)))\ .$$

□

The next Theorem is independently due to Henrik Imhof.

3.9 Theorem : *Every SFP–formula χ is equivalent to a formula χ' that is $(2,0)$–bounded.*

If $\chi \in \exists LFP$ then χ' can be chosen of the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\ \theta_0 \vee \exists \bar{y}, \bar{z} \in X\theta]v \dots v$$

where θ_0 and θ are quantifier–free formulae, and X and v do not occur in θ or θ_0.

Proof: By Corollary 3.3 it suffices to consider a formula

$$\chi = [FP_{\bar{x},X}\ \psi_0(\bar{x}) \vee \exists \bar{y}_1 \in X, \dots, \bar{y}_k \in X\psi]\bar{u}$$

where X does not occur in ψ and $k \geq 2$. This time we let $s = \mathrm{let}(k)$ and Y be an $(s+kr)$–ary relation variable (where r is the arity of X). We use a similar notation as in the proof of the previous theorem.
Let χ' be the formula

$$[FP_{\bar{x}_0\bar{x}_1\dots\bar{x}_k,Y}\ (\bar{x}_0 = \mathbf{1}\ \wedge\ \psi_0(\bar{x}_1)) \vee \exists \bar{y}_0\bar{y}_1 \dots \bar{y}_k \in Y, \bar{z}_0\bar{z}_1 \dots \bar{z}_k \in Y \bigvee_{i=1}^{k} (\bar{x}_0 = \boldsymbol{i}\ \wedge\ \psi_i)\Big]\,\underbrace{u_1 \dots u_1}_{s\ times}\,\bar{u}\ \underbrace{u_1 \dots u_1}_{(k-1)r\ times}\ .$$

where

$$\psi_1 = \Big(\bar{y}_0 = \boldsymbol{k}\ \wedge\ \psi(\bar{x}_1, \bar{y}_1, \dots, \bar{y}_k)\Big),$$

for $2 \leq i \leq k$

$$\psi_i = \Big(\bar{y}_0 = \boldsymbol{i{-}1}\ \wedge\ \bar{z}_0 = \boldsymbol{i{-}1}\ \wedge\ \bigwedge_{j=1}^{i-1} \bar{x}_j = \bar{y}_j\ \wedge\ \bar{x}_i = \bar{z}_1\Big)\ .$$

Recalling the proof of the last theorem it is not hard to see that $\chi' \mathrel{=\joinrel\mid\joinrel\models} \chi$.

To prove the stronger statement for $\exists$*LFP*–formulae we proceed in a manner completely analogously to the proof of the last theorem. ■

In fact we can even prove the following slightly stronger result which we will use later to give a vectorized Lindström–quantifier capturing **SFP**.

3.10 Corollary : *Every SFP–formula χ is equivalent to a formula χ' in which every subformula $[FP \dots] \dots$ is of the form*

$$[FP_{\bar{x},X}\psi_0\ \vee\ (\exists \bar{y} \in X\psi_1(\bar{x}, \bar{y})\ \wedge\ \exists \bar{z} \in X\psi_2(\bar{x}, \bar{z}))]\bar{u}$$

where X does not occur in ψ_0, ψ_1, ψ_2.
Every ∃LFP–formula is equivalent to a formula of the form

$$\substack{\exists\\\forall} v[FP_{\bar{x},X}\theta_0 \vee (\exists\bar{y} \in X\theta_1(\bar{x},\bar{y}) \wedge \exists\bar{z} \in X\theta_2(\bar{x},\bar{z}))]v\ldots v$$

with quantifier–free formulae θ_0, θ_1, and θ_2 in which X and v do not occur.

Proof: By Theorem 3.9 it suffices to consider a (2,0)–bounded formula

$$\chi = [FP_{\bar{x},X}\psi_0(\bar{x}) \vee \exists\bar{y},\bar{z} \in X\psi(\bar{x},\bar{y},\bar{z})]\bar{u}\ .$$

We let

$$\chi' = [FP_{\bar{x}_1\bar{x}_2,Y}(\psi_0(\bar{x}_1) \wedge \psi_0(\bar{x}_2)) \vee \exists\bar{y}_1\bar{y}_2, \bar{z}_1\bar{z}_2 \in Y(\psi(\bar{x}_1,\bar{y}_1,\bar{y}_2) \wedge \psi(\bar{x}_2,\bar{z}_1,\bar{z}_2))]\bar{u}\bar{u}$$

(where $\bar{y}_1, \bar{y}_2, \bar{z}_1, \bar{z}_2$ are tuples of variables that do not occur in χ).
Then for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ for χ and each $i \geq 0$ we have

$$X_i^{\mathfrak{I}} \times X_i^{\mathfrak{I}} = Y_i^{\mathfrak{I}}\ .$$

The proof is an easy induction over i. This shows $\chi \mathrel{=\!\mid\!\models} \chi'$.
Since $\bar{y}_1\bar{y}_2$ does not occur in $\psi(\bar{x},\bar{z}_1,\bar{z}_2)$ and $\bar{z}_1\bar{z}_2$ does not occur in $\psi(\bar{x},\bar{y}_1,\bar{y}_2)$ we see that χ' is equivalent to

$$[FP_{\bar{x}_1\bar{x}_2,Y}(\psi_0(\bar{x}_1) \wedge \psi_0(\bar{x}_2)) \vee (\exists\bar{y}_1\bar{y}_2 \in Y\psi(\bar{x}_1,\bar{y}_1,\bar{y}_2) \wedge \exists\bar{z}_1\bar{z}_2 \in Y\psi(\bar{x}_2,\bar{z}_1,\bar{z}_2))]\bar{u}\bar{u},$$

and the latter formula is of the desired form.
Using the stronger statement of Theorem 3.9 concerning **∃LFP** we see that this also proves the **∃LFP**–part. ■

But the Corollary is the strongest statement we can make about **SFP** in this context. Uwe Bosse [Bos92] showed that one (stratified) fixed–point operator does not suffice to capture **SFP**. In fact he proved a stronger hierarchy theorem concerning the quantifier–structure of *SFP*–formulae. This also follows easily from results of Grädel and McColm [GM92b] concerning the quantifier–structure of infinitary formulae. In particular we have:

3.11 Theorem ([Bos92, GM92b]) : *For each $r \geq 1$ there is an SFP-formula of fixed–point rank r that is not equivalent to any formula of fixed–point rank $< r$.*

Moreover, $(1,0)$–bounded formulae do not suffice to capture **SFP** because they coincide with formulae of *transitive closure logic* (cf. Section 4) which is known to be strictly weaker than **SFP**.

§ 3.2 The Bounded Skolem Normal Form

3.12 Theorem : *Every LFP–formula is equivalent to a formula in* bounded Skolem form, *i.e. of the form*

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\ \theta_0\ \vee\ \forall\bar{y}\exists\bar{z}\in X\ \theta]v\ldots v$$

with quantifier–free formulae θ_0 and θ in which X and v do not occur, and also to a formula of the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\ \theta_0'\ \vee\ \exists\bar{y}\forall\bar{z}\notin X\theta']v\ldots v$$

with quantifier–free formulae θ_0' and θ' in which X and v do not occur.

Proof: By Theorem 2.5 it suffices to consider an *LFP*–formula χ in Skolem normal form.

First Step: The Π_2 part.

We first apply Lemma 3.5 and obtain, after bringing the second disjunct of the first–order part in prenex normal form, a formula

$$\chi_1 = \overset{\exists}{\forall} v[FP_{\bar{x},X}\psi_0 \vee \forall\bar{y}\exists\bar{z}\psi]v\ldots v$$

with the following properties:

- ψ_0 and ψ are quantifier–free
- X and v do not occur in ψ_0, and v does not occur in ψ
- ψ is positive in X
- χ_1 is modest
- $\chi_1 \mathrel{=\!\!|\!\!\models} \chi$.

Next we apply Lemma 3.2 to ψ and obtain a formula of the form

$$\exists\bar{z}_1\in X,\ldots,\bar{z}_k\in X\ \xi(\bar{x},\bar{y},\bar{z},\bar{z}_1,\ldots,\bar{z}_k)$$

where ξ is a quantifier–free formula and X does not occur in ξ. Furthermore, the formula is equivalent to ψ in interpretations in which $X\neq\emptyset$.

Let Y be a kr–ary relation variable (where r is the arity of X) and

$$\chi_2 := \overset{\exists}{\forall} v[FP_{\bar{x}_1\ldots\bar{x}_k,Y}\ \bigwedge_{i=1}^{k}\psi_0(\bar{x}_i)\ \vee\ \forall\bar{w}\forall\bar{y}\exists\bar{z}\exists\bar{z}_1\ldots\bar{z}_k\in Y(\bigvee_{i=1}^{k}\bar{w}=\bar{x}_i \\ \rightarrow(\xi(\bar{w},\bar{y},\bar{z},\bar{z}_1,\ldots,\bar{z}_k)\vee\psi_0(\bar{w}))))]v\ldots v.$$

We claim that $\chi_2 \dashv\models \chi_1$.

To see this let $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation and consider the stages $(X_i)_{i\geq 0}$ and $(Y_i)_{i\geq 0}$ of the fixed–point processes of χ_1 and χ_2 respectively.
Then for each $i \geq 0$ we have

$$Y_i = \underbrace{X_i \times \ldots \times X_i}_{k\ times}.$$

The proof is by induction over i, for $i = 0$ and $i = 1$ it is immediate.
So suppose the statement is true for $i \geq 1$ and consider a tuple $\bar{a}_1 \ldots \bar{a}_k$ in Y_{i+1}.
Then for all $\bar{b} \in A$ we have

$$\mathfrak{I} \models \forall \bar{y} \exists \bar{z} \exists \bar{z}_1 \ldots \bar{z}_k \in Y_i (\bigvee_{j=1}^{k} \bar{b} = \bar{a}_j \rightarrow (\xi(\bar{b}, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k) \vee \psi_0(\bar{b}))).$$

In particular for $\bar{b} = \bar{a}_j$ ($j \leq k$ arbitrary) this implies

$$\mathfrak{I} \models \forall \bar{y} \exists \bar{z} \exists \bar{z}_1 \ldots \bar{z}_k \in Y_i \xi(\bar{a}_j, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k) \quad \text{or} \quad \mathfrak{I} \models \psi_0(\bar{a}_j).$$

By the induction hypothesis this means

$$\mathfrak{I} \models \forall \bar{y} \exists \bar{z} \exists \bar{z}_1 \in X_i, \ldots \bar{z}_k \in X_i \xi(\bar{a}_j, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k) \quad \text{or} \quad \mathfrak{I} \models \psi_0(\bar{a}_j).$$

Hence $\bar{a}_j \in X_{i+1}$. Since $j \leq k$ was arbitrary the $\subseteq$–direction is proved.

On the other hand let $\bar{a}_1 \in X_{i+1}, \ldots, \bar{a}_k \in X_{i+1}$. By the definition of X_{i+1} this means

$$\forall j \leq k : \ \mathfrak{I} \models \forall \bar{y} \exists \bar{z} \exists \bar{z}_1 \in X_i \ldots \bar{z}_k \in X_i \xi(\bar{a}_j, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k) \quad \text{or} \quad \mathfrak{I} \models \psi_0(\bar{a}_j)\ .$$

Now everything goes back in the other direction without any problems.

Finally, we want to eliminate the existential quantification $\exists \bar{z}$. We do this in the same way as in the proof of Theorem 3.6 by increasing the arity of Y by the length of $\bar{z}$.

SECOND STEP: The Σ_2 part.

We proceed similarly to the first step:
Starting with the Σ_2–form of χ and applying Lemma 3.5 as above gives us a formula

$$\chi_3 := \overset{\exists}{\forall} v[FP_{\bar{x},X}\psi_0 \vee \exists \bar{y} \forall \bar{z} \psi'] v \ldots v$$

with the following properties:

- ψ_0 is the same formula as above (i.e. a quantifier–free formula in which X and v do not occur)

- ψ' is quantifier–free, and v does not occur in ψ'
- ψ' is positive in X
- χ_3 is modest
- $\chi_3 \dashv\models \chi$.

This time we apply Lemma 3.2 to $\neg\psi'$ and obtain a formula

$$\eta' = \exists \bar{z}_1 \notin X, \ldots, \bar{z}_k \notin X \xi'(\bar{x}, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k)$$

(where ξ' is quantifier–free and X does not occur in ξ') that is equivalent to $\neg\psi'$ in interpretations in which $X \neq \mathcal{U}$.
Let

$$\eta := \forall \bar{z}_1 \notin X, \ldots, \bar{z}_k \notin X \neg\xi'(\bar{x}, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k)\ .$$

Then obviously we have $\eta =\!\mid\models \neg\eta'$ hence $\eta =\!\mid\models \psi'$ in interpretations where $X \neq \mathcal{U}$.
Analogously to the first step we let Y be a relation variable whose arity is k times the arity of X and

$$\chi_4 := \overset{\exists}{\forall} v[FP_{\bar{x}_1 \ldots \bar{x}_k, Y} \overbrace{\bigvee_{i=1}^{k} \psi_0(\bar{x}_i)}^{=:\theta'_0} \vee \exists \bar{w} \exists \bar{y} \forall \bar{z} \forall \bar{z}_1 \ldots \bar{z}_k \notin Y$$
$$(\bigvee_{i=1}^{k} \bar{w} = \bar{x}_i \wedge \neg\xi'(\bar{w}, \bar{y}, \bar{z}, \bar{z}_1, \ldots, \bar{z}_k))]v \ldots v.$$

An induction analogous to the first step shows that for each $i \geq 0$ we have

$$Y_i^C = \underbrace{X_i^C \times \ldots \times X_i^C}_{k\ times}.$$

(where $(X_i)_{i \geq 0}$ and $(Y_i)_{i \geq 0}$ denote the stages of the respective fixed–point processes of χ_3 and χ_4).
Hence $\chi_4 \dashv\models \chi_3$.

To eliminate the quantifiers $\forall \bar{z}$ we increase the arity of Y by the length of $\bar{z}$ and consider the formula

$$\overset{\exists}{\forall} v[FP_{\bar{x}_0 \bar{x}_1 \ldots \bar{x}_k, Y} \theta'_0 \vee \exists \bar{w} \exists \bar{y} \forall \bar{z}_0 \bar{z}_1 \ldots \bar{z}_k \notin Y$$
$$(\bigvee_{i=1}^{k} \bar{w} = \bar{x}_i \wedge \neg\xi'(\bar{w}, \bar{y}, \bar{z}_0, \bar{z}_1, \ldots, \bar{z}_k))]v \ldots v$$

which can easily be seen to be equivalent to χ_4 and thus χ. ∎

4 Transitive Closure Logic

4.1 Definition : The class TC of *transitive closure formulae* is given by the first–order clauses and the new clause:

$$(TC) \quad \frac{\varphi}{[TC_{\overset{k}{x},\overset{k}{y}}\,\varphi]\;\overset{k}{u},\overset{k}{v}} \quad \text{where } k \geq 1 \text{ and } \overset{k}{u} \text{ and } \overset{k}{v} \text{ are tuples of terms.}$$

$[TC_{\overset{k}{x},\overset{k}{y}} \ldots]$ is called a *transitive closure operator.*

The free variables of a transitive closure formula are defined in the standard way, adding the clause

$$\mathrm{free}([TC_{\overset{k}{x},\overset{k}{y}}\,\varphi(\overset{k}{x},\overset{k}{y})]\;\overset{k}{u},\overset{k}{v}) = (\mathrm{free}(\varphi)\backslash\{x_1,\ldots,x_k,y_1,\ldots,y_k\}) \cup \mathrm{free}(\overset{k}{u},\overset{k}{v}).$$

$\square$

4.2 Definition : Let $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation. Then

$$\mathfrak{I} \models [TC_{\overset{k}{x},\overset{k}{y}}\,\varphi]\;\overset{k}{u},\overset{k}{v}$$

if and only if there exist $n \geq 2, \overset{k}{a}_1, \ldots, \overset{k}{a}_n \in A$ such that $\alpha(\overset{k}{u}) = \overset{k}{a}_1, \alpha(\overset{k}{v}) = \overset{k}{a}_n$ and $\mathfrak{I} \models \varphi[\overset{k}{a}_i, \overset{k}{a}_{i+1}]$ for each $i < n$.

Now the semantics of the logic $\mathbf{TC} = (TC, \models)$ can be defined inductively with respect to the calculus given above. $\square$

4.3 Remark : Unlike the standard definition we do not require the transitive closure to be reflexive. There are two reasons in favor of our definition:

- In the next paragraph we are going to introduce the fragment deterministic transitive closure logic. If we would require the formulae to be automatically reflexive the paths would not be deterministic at their starting point (cf. below).

- With our definition we can give a normal form for the existential fragment of **TC** which is very similar to the normal form for fixed–point logic (cf. Theorem 4.4).

Note that the expressive power of both variants is the same anyway. $\square$

To see the relation between transitive closure logic and fixed–point logic we observe that

$$[TC_{\bar{x},\bar{y}}\varphi(\bar{x},\bar{y})]\bar{u},\bar{v} \mathrel{=\!\mid\!\models} [FP_{\bar{y},Y}\varphi(\bar{u},\bar{y}) \vee \exists\bar{x} \in Y\varphi(\bar{x},\bar{y})]\bar{v}.$$

On the other hand, for each $(1,0)$–bounded formula

$$\chi = [FP_{\bar{x},X}\psi_0(\bar{x}) \vee \exists \bar{y} \in X\psi]\bar{v}$$

we have

$$\chi \mathrel{=\!\!|\!\models} \psi_0(\bar{v}) \vee \exists \bar{u}(\psi_0(\bar{u}) \wedge [TC_{\bar{y},\bar{x}}\psi]\bar{u},\bar{v}) \ .$$

Thus **TC** =(1,0)–bounded–**PFP**$\subseteq$ **SFP**.[16] A result of Immerman [Imm81] shows that the inclusion is strict.

The logics **TC** and **SFP** behave very similarly, e.g. Grädel [Grä91] proved that there is no normal form for **TC** with only one *TC*–operator (cf. Theorem 3.11 on **SFP**).

§4.1 Deterministic Transitive Closure Logic

The class *DTC* of *deterministic transitive closure formulae* is the subclass of *TC* built by the first–order clauses and the following variant of the (TC)–clause:

$$(DTC) \quad \frac{\varphi(\overset{k}{x},\overset{k}{y})}{[TC_{\overset{k}{x},\overset{k}{y}}\ \varphi(\overset{k}{x},\overset{k}{y}) \wedge \forall\ \overset{k}{z}\ (\varphi(\overset{k}{x},\overset{k}{z}) \rightarrow \overset{k}{z}=\overset{k}{y})]\ \overset{k}{u},\overset{k}{v}}$$

where $k \geq 1$ and $\overset{k}{u}$, $\overset{k}{v}$ are tuples of terms.
We let **DTC** $= (DTC, \models)$.
For brevity we write $[DTC_{\overset{k}{x},\overset{k}{y}}\ \varphi(\overset{k}{x},\overset{k}{y})]\ \overset{k}{u},\overset{k}{v}$ instead of

$$[TC_{\overset{k}{x},\overset{k}{y}}\ \varphi(\overset{k}{x},\overset{k}{y}) \wedge \forall\ \overset{k}{z}\ (\varphi(\overset{k}{x},\overset{k}{z}) \rightarrow \overset{k}{z}=\overset{k}{y})]\ \overset{k}{u},\overset{k}{v} \ .$$

Grädel and McColm ([GM92a]) showed that **DTC** $\subsetneq$ **TC**.

§4.2 Existential Transitive Closure Logic

Recall from Definition 1.7 that the *existential fragment* $\exists TC$ *of transitive closure logic* is the restriction of **TC** to formulae where no universal quantifiers occur and where negation symbols only occur in front of atomic subformulae.

In analogy with ∃**LFP** (as a fragment of **SFP**), ∃**TC** has a nice normal form.

4.4 Theorem ([Imm87]) : *Every* ∃**TC** *formula is equivalent to a formula of the form*

$$\overset{\exists}{\forall} v[TC_{\bar{x},\bar{y}}\theta]v\ldots v, v\ldots v$$

[16] As the reader might guess, by (k,l)–bounded–**PFP** we denote the sublogic of **PFP** whose formulae are the (k,l)—bounded *FP*–formulae.

where θ is a quantifier–free formula and v does not occur in θ.

Immerman's original theorem is not stated for $\exists$**TC** but for **TC** restricted to ordered structures. As has already been observed by Grädel [Grä91], the order is only needed to eliminate universal quantifiers and negation symbols, so the statement survives for $\exists$**TC** on all structures. However, a few changes are necessary if no constants are present.
Moreover, here the transitive closure is not reflexive; this allows the nice formulation of the theorem which differs a bit from the usual form.
So it seems worthwhile to sketch a **proof** of the theorem:
We first observe that for all $\varphi, \bar{t}, \bar{u}$ we have

$$[TC_{\bar{x},\bar{y}}\varphi]\bar{t},\bar{u} \quad \dashv\models \quad \overset{\exists}{\forall} v[TC_{\bar{x}_0\bar{x},\bar{y}_0\bar{y}}\psi]v\ldots v, v\ldots v,$$

where $\bar{x}_0, \bar{y}_0$ are tuples of new variables of length let(2) (recall Definition 1.11), and

$$\begin{aligned}\psi = \quad & (\bar{x}_0 = \mathbf{1} \wedge \bar{y}_0 = \mathbf{2} \wedge \varphi(\bar{t}, \bar{y}))\\ & \vee(\bar{x}_0 = \mathbf{2} \wedge \bar{y}_0 = \mathbf{2} \wedge \varphi(\bar{x}, \bar{y}))\\ & \vee(\bar{x}_0 = \mathbf{2} \wedge \bar{x} = \bar{u} \wedge \bar{y}_0 = \mathbf{1} \wedge \Delta(\bar{y}))\ .\end{aligned}$$

By $\Delta(\bar{y})$ we abbreviate $\bigwedge_{i,j} y_i = y_j$. (To see the equivalence recall that $\bar{x} = \mathbf{1}$ also abbreviates $\bigwedge_{i,j} x_i = x_j$.)
In the sequel we will only consider formulae of the form $\overset{\exists}{\forall} v[TC_{\bar{x},\bar{y}}\varphi]v\ldots v, v\ldots v$ and abbreviate them by $[TC_{\bar{x},\bar{y}}\varphi]$.

We are only going to give the formulae needed in an inductive proof. It should not be too hard to verify that the stated equivalences are true. For a more detailed proof see [Imm87].

$$\exists u[TC_{\bar{x},\bar{y}}\varphi(u,\bar{x},\bar{y})]$$

$$\dashv\models[TC_{\bar{z}u\bar{x},\bar{z}'u'\bar{y}} \quad (\bar{z} = \mathbf{1} \wedge \bar{z}' = \mathbf{2} \wedge \varphi(u', \bar{x}, \bar{y}))$$

In the beginning the indicator $\bar{z}$ is set to $\mathbf{2}$, a u is chosen and a first step is taken.

$$\vee(\bar{z} = \mathbf{2} \wedge \bar{z}' = \mathbf{2} \wedge u' = u \wedge \varphi(u', \bar{x}, \bar{y}))$$

Then we walk along a φ–path and keep the u fixed.

$$\vee(\bar{z} = \mathbf{2} \wedge \Delta(\bar{x}) \wedge \bar{z}' = \mathbf{1} \wedge u' = y_1 \wedge \Delta(\bar{y}))]$$

If we arrive at Δ we reset the indicator to $\mathbf{1}$ and also everything else to Δ.

$$
\begin{array}{lll}
 & [TC_{\bar{x},\bar{y}}\varphi(\bar{x},\bar{y})] \wedge [TC_{\bar{u},\bar{v}}\psi(\bar{u},\bar{v})] & \\
\dashv\models & [TC_{\bar{z}\bar{x}\bar{u},\bar{z}'\bar{y}\bar{v}} & (\bar{z}=\mathbf{1}\wedge\bar{z}'=\mathbf{2}\wedge\varphi(\bar{x},\bar{y})) \\
 & & \text{We set the indicator to } \mathbf{2} \text{ and start walking along} \\
 & & \text{a } \varphi\text{–path.} \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\bar{z}'=\mathbf{2}\wedge\varphi(\bar{x},\bar{y})) \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\Delta(\bar{x})\wedge\Delta(\bar{u})\wedge\bar{z}'=\mathbf{3}\wedge\psi(\bar{u},\bar{v})) \\
 & & \text{If we arrive at } \Delta \text{ we set the indicator to } \mathbf{3} \text{ and start} \\
 & & \text{a } \psi\text{–path from } \Delta, \\
 & & \vee(\bar{z}=\mathbf{3}\wedge\bar{z}'=\mathbf{3}\wedge\psi(\bar{u},\bar{v})) \\
 & & \vee(\bar{z}=\mathbf{3}\wedge\Delta(\bar{u})\wedge\bar{z}'=\mathbf{1}\wedge\Delta(\bar{y})\wedge\Delta(\bar{v}))] \\
 & & \text{until we arrive at } \Delta. \text{ Then everything is reset to} \\
 & & \Delta.
\end{array}
$$

$$
\begin{array}{lll}
 & [TC_{\bar{x},\bar{y}}\varphi(\bar{x},\bar{y})] \vee [TC_{\bar{u},\bar{v}}\psi(\bar{u},\bar{v})] & \\
\dashv\models & [TC_{\bar{z}\bar{x}\bar{u},\bar{z}'\bar{y}\bar{v}} & (\bar{z}=\mathbf{1}\wedge\bar{z}'=\mathbf{2}\wedge\varphi(\bar{x},\bar{y})) \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\bar{z}'=\mathbf{2}\wedge\varphi(\bar{x},\bar{y})) \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\Delta(\bar{x})\wedge\bar{z}'=\mathbf{1}\wedge\Delta(\bar{y})\wedge\Delta(\bar{v}))] \\
 & & \text{We try a } \varphi\text{–path via } \mathbf{2} \\
 & & \vee(\bar{z}=\mathbf{1}\wedge\bar{z}'=\mathbf{3}\wedge\psi(\bar{u},\bar{v})) \\
 & & \vee(\bar{z}=\mathbf{3}\wedge\bar{z}'=\mathbf{3}\wedge\psi(\bar{u},\bar{v})) \\
 & & \vee(\bar{z}=\mathbf{3}\wedge\Delta(\bar{u})\wedge\bar{z}'=\mathbf{1}\wedge\Delta(\bar{y})\wedge\Delta(\bar{v}))] \\
 & & \text{or a } \psi\text{–path via } \mathbf{3}.
\end{array}
$$

$$
\begin{array}{lll}
 & [TC_{\bar{u},\bar{v}}[TC_{\bar{x},\bar{y}}\varphi(\bar{u},\bar{v},\bar{x},\bar{y})]] & \\
\dashv\models & [TC_{\bar{z}\bar{u}\bar{v}\bar{x},\bar{z}'\bar{u}'\bar{v}'\bar{y}} & (\bar{z}=\mathbf{1}\wedge\bar{z}'=\mathbf{2}\wedge\bar{u}'=\bar{u}\wedge\varphi(\bar{u}',\bar{v}',\bar{x},\bar{y})) \\
 & & \text{We keep } \bar{u} \text{ fixed, try a } \bar{v}, \text{ and start a } \varphi\text{–path.} \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\bar{z}'=\mathbf{2}\wedge\bar{u}'=\bar{u}\wedge\bar{v}'=\bar{v}\wedge\varphi(\bar{u}',\bar{v}',\bar{x},\bar{y})) \\
 & & \text{Walking along the path we keep } \bar{u} \text{ and } \bar{v} \text{ fixed.} \\
 & & \vee(\bar{z}=\mathbf{2}\wedge\Delta(\bar{x})\wedge\bar{z}'=\mathbf{1}\wedge\bar{u}'=\bar{v}\wedge\Delta(\bar{y}))] \\
 & & \text{If we arrive at } \Delta \text{ we let } \bar{u} \text{ be the old } \bar{v} \text{ and go back} \\
 & & \text{to the start — unless } \Delta(\bar{v}), \text{ then we are lucky and} \\
 & & \text{stop.}
\end{array}
$$

Note moreover that for every formula φ in which x, y, v do not occur freely we have:

$$\varphi \quad \dashv\models \quad [TC_{x,y}\varphi]$$

so the cases considered above suffice. ■

4.5 Corollary : *An $\exists TC$–formula whose signature contains a constant symbol c is equivalent to a formula of the form*

$$[TC_{\bar{x},\bar{y}}\theta]c\ldots c, c\ldots c$$

with a quantifier–free θ.

4.6 Corollary : *Every TC–formula χ is equivalent to a fixed–point formula χ' that is $(1,0)$–bounded.*

If $\chi \in \exists TC$ then χ' can be chosen in the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\ \theta_0 \vee \exists \bar{y} \in X\theta]v\ldots v$$

where θ_0 and θ are quantifier–free formulae, and X and v do not occur in θ or θ_0.

Proof: We have already seen the statement for **TC** (cf. page 46); and the stronger statement for **∃TC** is an immediate consequence of the previous theorem since $\overset{\exists}{\forall} v[TC_{\bar{x},\bar{y}}\theta]v\ldots v, v\ldots v$ is equivalent to

$$\overset{\exists}{\forall} v[FP_{\bar{x}\bar{y},X}\theta(\bar{x},\bar{y}) \vee \exists \bar{z}\bar{w} \in X(\bar{x} = \bar{z}\ \wedge\ \theta(\bar{w},\bar{y}))]v\ldots vv\ldots v.$$

■

The following is a generalization of an argument used by Gurevich [Gur84] to prove that the substructure preservation theorem for first–order logic fails in the finite. In particular it shows that a substructure preservation theorem for transitive closure logic does not hold in the finite.

4.7 Theorem : $\mathbf{\exists TC} \subsetneq \mathbf{\exists LFP} \cap \mathbf{TC}$

Proof: Let $\sigma = \{E, L, c\}$ where E is a binary relation symbol, L is a unary relation symbol, and c is a constant symbol.

Recall the definition of forests and trees from Paragraph 0.6. We say that a forest is of width $= w$ ($\leq w$) if all nodes are either leaves or have exactly (respectively, at most) w successors.

Consider the σ–query ν which says

> *If E is the edge relation of a forest of width ≤ 3 and L only contains leaves then the subtree with root c is a tree of width $= 3$ whose leaves are all in L.*

It can be defined by the formula

$$\left.\begin{array}{ll}\Big(& \forall x \neg Exx \\ \wedge & \forall x \forall y \forall z \neg (y \neq z \wedge Eyx \wedge Ezx) \\ \wedge & \forall x \forall y (Exy \to \neg [FP_{z,X} z = y \vee \exists z' \in X\, Ez'z]x) \\ \wedge & \forall x \forall y_1, \ldots, y_4 \neg \bigwedge_{1 \leq i < j \leq 4} (y_i \neq y_j \wedge Exy_i) \\ \wedge & \forall x \forall y (Lx \to \neg Exy) \qquad\qquad\qquad\qquad \Big) \end{array}\right\} =: \varphi$$

$$\longrightarrow \quad [FP_{x,X} Lx \vee \exists y_1, y_2, y_3 \bigwedge_{1 \leq i < j \leq 3} (y_i \neq y_j \wedge Xy_i \wedge Exy_i)]c\ .$$

which is obviously equivalent to an $\exists LFP$–formula.
Note that φ is $(1,0)$–bounded and hence equivalent to a TC–formula φ'. Thus we can also define ν by the TC–formula

$$\varphi' \longrightarrow \forall x \Big([TC_{v,w} Evw]c, x \to (Lx \vee \exists y_1, y_2, y_3 \bigwedge_{1 \leq i < j \leq 3} (y_i \neq y_j \wedge Exy_i)) \Big).$$

Claim: ν is not definable in $\exists \mathbf{TC}$.

By the last corollary we can restrict ourselves to sentences χ of the form

$$[TC_{\overset{k}{x},\overset{k}{y}} \theta] c \ldots c, c \ldots c$$

where θ is quantifier–free.

We say that a pair $(\mathfrak{A}, \mathfrak{B})$ of σ–structures has the *k–path–property* if for all $\overset{k}{a}_1, \ldots, \overset{k}{a}_m \in A$ such that $\overset{k}{a}_1 = \overset{k}{a}_m = c^{\mathfrak{A}} \ldots c^{\mathfrak{A}}$ there exist $\overset{k}{b}_1, \ldots, \overset{k}{b}_m \in B$ such that for each $i < m$ the same quantifier–free formulae hold for the $2k$–tuples $\overset{k}{a}_i \overset{k}{a}_{i+1}$ and $\overset{k}{b}_i \overset{k}{b}_{i+1}$.
Certainly, if $(\mathfrak{A}, \mathfrak{B})$ has the k–path–property and χ is an $\exists TC[\sigma]$–sentence of the form above we have

$$\mathfrak{A} \models \chi \quad \Longrightarrow \quad \mathfrak{B} \models \chi\ .$$

Thus it suffices to prove:

For each $k \geq 1$ there exist σ–structures $\mathfrak{A}_k$ and $\mathfrak{B}_k$ such that:

(1) $\mathfrak{A}_k \models \varphi$ *and* $\mathfrak{B}_k \models \varphi$.

(2) $\mathfrak{A}_k$ *is a tree of width* $= 3$ *with root* c *and all leaves in* L.

(3) $\mathfrak{B}_k$ *is not a tree of width* $= 3$ *with root* c *and all leaves in* L.

(4) $(\mathfrak{A}_k, \mathfrak{B}_k)$ *has the* k*–path–property.*

The proof is by induction on k.

$k = 1$: Consider the following structures:

Figure 4.8

$L^{\mathfrak{A}_1}$ and $L^{\mathfrak{B}_1}$ are defined to hold for all leaves in the respective trees (i.e. $L^{\mathfrak{A}_1} = \{b, c, d\}$ and $L^{\mathfrak{B}_1} = \{b', c'\}$) and $c^{\mathfrak{A}_1} = a$, $c^{\mathfrak{B}_1} = a'$.

Clearly (1)–(3) hold for $\mathfrak{A}_1$ and $\mathfrak{B}_1$.

Let $a = a_1, \ldots, a_m = a$ be a path in A_1. Then we can define a path $b_1, \ldots, b_m \in B_1$ as follows:

- If $a_i = a$ then $b_i := a'$.
- If $a_i \in \{b, c, d\}$ then we select $b_i \in \{b', c'\}$ in a way that

$$b_i = b_{i-1} \iff a_i = a_{i-1} \ .$$

Thus $(\mathfrak{A}_1, \mathfrak{B}_1)$ has the 1–path–property.

$k \to k + \mathbf{1}$: The structures $\mathfrak{A}_{k+1}$ and $\mathfrak{B}_{k+1}$ are defined as follows:

Figure 4.9

where $\mathfrak{B}, \mathfrak{C}, \mathfrak{D}, \mathfrak{B}'$, and $\mathfrak{C}'$ are isomorphic copies of $\mathfrak{A}_k{\restriction}_{\{E,L\}}$, and $\mathfrak{D}'$ is an isomorphic copy of $\mathfrak{B}_k{\restriction}_{\{E,L\}}$. Moreover, we let $c^{\mathfrak{A}_{k+1}} := a_{k+1}$, $c^{\mathfrak{B}_{k+1}} := a'_{k+1}$.

Again (1)–(3) are immediate.

Let $\overset{k+1}{a_1}, \overset{k+1}{a_2}, \ldots, \overset{k+1}{a_m} \in A_{k+1}$ with $\overset{k+1}{a_1} = \overset{k+1}{a_m} = a_{k+1} \ldots a_{k+1}$.

To define a path $\overset{k+1}{b_1}, \overset{k+1}{b_2}, \ldots, \overset{k+1}{b_m} \in B_{k+1}$ we first let, for $i \leq m, j \leq k+1$, $b_{ij} := a'_{k+1}$ whenever $a_{ij} = a_{k+1}$.

Next we consider the subsequence $\overset{k+1}{a_{i_1}}, \ldots, \overset{k+1}{a_{i_n}}$ of those tuples that are completely contained in one of the subtrees $\mathfrak{B}$, $\mathfrak{C}$, or $\mathfrak{D}$ of $\mathfrak{A}_{k+1}$. Similarly to the case $k = 1$ we can copy them into the subtrees $\mathfrak{B}'$ and $\mathfrak{C}'$ of $\mathfrak{B}_{k+1}$ (recall that all these trees are isomorphic) such that whenever two succeeding tuples $\overset{k+1}{a_{i_j}}$, $\overset{k+1}{a_{i_{j+1}}}$ are in distinct subtrees so are the corresponding tuples $\overset{k+1}{b_{i_j}}$, $\overset{k+1}{b_{i_{j+1}}}$ (and vice versa).

The remaining a_{ij} form paths of at most k–tuples in the subtrees $\mathfrak{B}$, $\mathfrak{C}$, and $\mathfrak{D}$. We can treat all subtrees independently with the help of the induction hypothesis and thus "fill up" the paths. ■

4.10 Remark : The relation between transitive closure logic and the datalogics has also been investigated (cf. [Grä91]). It is easy to see that a query is definable in **∃TC** if and only if it is definable by a *linear* datalog$^{\neg}$ program, i.e. a datalog$^{\neg}$ program where the body of each clause contains at most one literal $X\bar{t}$ with an intensional X.

Similarly, it can be seen that the queries definable by linear (pure) datalog programs are exactly those definable by positive $\exists TC$–formulae, i.e. $\exists TC$–formulae where no negation symbols occur. Let us denote this class of formulae by $pos\exists TC$ and let **pos∃TC** $:= (pos\exists TC, \models)$.

It also is known that there is a class of linear stratified datalog programs that corresponds to the class of all TC–formulae. □

An interesting question to be mentioned here is:

4.11 Open Problem : *Is* $\mathbf{TC} \cap \mathbf{Datalog} = \mathbf{pos\exists TC}$ *?*

It should be seen in connection with Ajtai's and Gurevich's theorem [AG89] that

$$\mathbf{FO} \cap \mathbf{Datalog} = \mathbf{pos\exists FO}.$$

Note moreover that by Theorem 4.7 and 4.10 we have

$$\mathbf{TC} \cap \mathbf{Datalog}^{\neg} \supsetneq \mathbf{\exists TC}.$$

To complete this section let us mention the following results from descriptive complexity theory:

4.12 Theorem ([Imm87]) :

$$\mathbf{DTC} =_{\mathcal{O}} \mathbf{LOGSPACE}$$
$$\mathbf{\exists TC} =_{\mathcal{O}} \mathbf{TC} =_{\mathcal{O}} \mathbf{NLOGSPACE}$$

5 Lindström–Quantifiers and Complete Problems

Let us assume in this section that classes of structures are always closed under isomorphism.
By $\overset{r\times k}{x}$ we denote the r–tuple of k–tuples

$$\overset{k}{x}_1 \ldots \overset{k}{x}_r = x_{11} \ldots x_{1k} x_{21} \ldots x_{r1} \ldots x_{rk}.$$ [17]

§ 5.1 Lindström Logics

In the following we let $\mathbf{L} = (L, \models)$ be a logic, $\sigma = \{R_1, \ldots, R_l\}$ a signature consisting of r_i–ary relation symbols R_i, and $\mathcal{C}$ a class of σ–structures.

5.1 Definition : (1) The class $L(Q_{\mathcal{C}}^{<\omega})$ of formulae is the closure of L under the following clauses

$$(Q_{\mathcal{C}}^k) \qquad \frac{\varphi_1, \ldots, \varphi_l}{[Q_{\mathcal{C}}^k \overset{r_1\times k}{x_1}, \ldots, \overset{r_l\times k}{x_l} \;\; \varphi_1, \ldots, \varphi_l]}$$

(where $k \geq 1$) and the clauses used to build L.[18]
(The variables in the different tuples $\overset{r_i\times k}{x_i}$ need not necessarily be distinct.)

(2) $L(Q_{\mathcal{C}})$ is the subclass of $L(Q_{\mathcal{C}}^{<\omega})$ built using only the clause $(Q_{\mathcal{C}}^1) =: (Q_{\mathcal{C}})$ and the clauses used to build L.
$Q_{\mathcal{C}} := Q_{\mathcal{C}}^1$ is called the *Lindström–quantifier* associated with $\mathcal{C}$.

(3) The free variables of the formula

$$[Q_{\mathcal{C}}^k \overset{r_1\times k}{x_1}, \ldots, \overset{r_l\times k}{x_l} \;\; \varphi_1, \ldots, \varphi_l]$$

are

$$\bigcup_{h=1}^{l} (\mathrm{free}(\varphi_h) \setminus \{x_{hij} \mid i \leq r_h,\ j \leq k\}) \; .$$

$\square$

To define the semantics, for a formula $\varphi(\overset{r}{x})$ and an interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ we let

$$\varphi(\overset{r}{x})^{\mathfrak{I}} := \{\overset{r}{a} \in A^r \mid \mathfrak{I} \models \varphi[\overset{r}{a}]\} \; .$$

[17] Note that our notation is ambiguous here (this is the price we have to pay for the convenience of writing tuples with no brackets). So let us emphasize that by $\overset{r\times k}{x}$ we mean the r–tuple of k–tuples and not the $r \cdot k$–tuple that we also denote by $x_{11} \ldots x_{1k} x_{21} \ldots x_{r1} \ldots x_{rk}$.
[18] To make this precise we would need an exact definition of the notion "formula". However, it is clear what is meant here if **L** is **QF**, **FO**, or one of the fixed–point logics. This suffices for our purpose.

So $\varphi(\overset{r}{x})$ defines an r–ary relation on A. Similarly a formula $\psi(\overset{r\times k}{x})$ defines an r–ary relation on A^k:

$$\varphi(\overset{r\times k}{x})^{\mathfrak{I}} := \{\overset{r\times k}{a} \in (A^k)^r \mid \mathfrak{I} \models \varphi[\overset{r\times k}{a}]\} .$$

5.2 Definition : Let $k \geq 1$ and $\mathfrak{I} = (\mathfrak{A}, \alpha)$ be an interpretation. Then

$$\mathfrak{I} \models [Q^k_{\mathcal{C}} \overset{r_1\times k}{x_1}, \ldots, \overset{r_l\times k}{x_l} \ \ \varphi_1, \ldots, \varphi_l]$$

if and only if the σ–structure

$$(A^k, \varphi_1(\overset{r_1\times k}{x_1})^{\mathfrak{I}}, \ldots, \varphi_m(\overset{r_l\times k}{x_l})^{\mathfrak{I}})$$

belongs to $\mathcal{C}$.

Now the semantics of the logics $\mathbf{L}(Q_{\mathcal{C}}) = (L(Q_{\mathcal{C}}), \models)$ and $\mathbf{L}(Q^{<\omega}_{\mathcal{C}}) = (L(Q^{<\omega}_{\mathcal{C}}), \models)$ can be defined inductively in the usual way. $\square$

5.3 Example : The existential quantifier can be defined as the Lindström–quantifier associated with the class $\mathcal{EX} = \{(A, P^{\mathfrak{A}}) \mid P^{\mathfrak{A}} \neq \emptyset\}$ (where P is a unary relation symbol).
Similarly the universal quantifier can be defined as the Lindström–quantifier associated with $\mathcal{FA} = \{(A, P^{\mathfrak{A}}) \mid P^{\mathfrak{A}} = A\}$. $\square$

5.4 Example : Let $\mathcal{EVEN} = \{(A, P^{\mathfrak{A}}) \mid |P^{\mathfrak{A}}| \text{ is even}\}$. The associated Lindström–quantifier $Q_{\mathcal{EVEN}}$ is a simple example of a non–trivial (i.e not first–order definable) Lindström–quantifier whose vectorization does not increase the expressive power, i.e.

$$\mathbf{FO}(Q_{\mathcal{EVEN}}) = \mathbf{FO}(Q^{<\omega}_{\mathcal{EVEN}}).$$

To see this it suffices to prove that for each $k \geq 2$ the class

$$\{(A, P_k^{\mathfrak{A}}) \mid |P_k| \text{ is even}\}$$

(where P_k is a k–ary relation symbol) is definable by a formula that only uses the quantifiers $Q^l_{\mathcal{EVEN}}$ for $l < k$.
Note that for each $\{P_k\}$–structure $(A, P_k^{\mathfrak{A}})$ we have

$$|P_k^{\mathfrak{A}}| = \sum_{a\in A} |P_k^{\mathfrak{A}} a\_\_|^{19} = \sum_{a\in A, |P_k^{\mathfrak{A}} a\_\_| \text{ is even}} |P_k^{\mathfrak{A}} a\_\_| + \sum_{a\in A, |P_k^{\mathfrak{A}} a\_\_| \text{ is odd}} |P_k^{\mathfrak{A}} a\_\_|$$

[19] Recall that $P^{\mathfrak{A}} a\_\_ = \{\overset{k-1}{b} \in A \mid P_k^{\mathfrak{A}} a \overset{k-1}{b}\}$.

thus

$$\begin{aligned}
& |P_k^{\mathfrak{A}}| \text{ is even} \\
\iff\ & |\{a \in A \mid |P_k^{\mathfrak{A}} a __| \text{ is odd}\}| \text{ is even} \\
\iff\ & (A, P_k^{\mathfrak{A}}) \models \left[Q_{\mathcal{EVEN}} x \ \neg [Q^{k-1}_{\mathcal{EVEN}} \overset{k-1}{y} \ P_k x \overset{k-1}{y}] \right].
\end{aligned}$$

□

5.5 Definition : We call a logic $\mathbf{L}$ a *vectorized Lindström logic* if there exists a relational signature σ and a class $\mathcal{C}$ of σ–structures such that $\mathbf{L} = \mathbf{FO}(Q_{\mathcal{C}}^{<\omega})$. Then $Q_{\mathcal{C}}^{<\omega}$ is said to *capture* $\mathbf{L}$.

We call $\mathbf{L}$ a *Lindström logic in the strong sense* if there exists a relational signature σ and a class $\mathcal{C}$ of σ–structures such that $\mathbf{L} = \mathbf{FO}(Q_{\mathcal{C}})$. Analogously $Q_{\mathcal{C}}$ is said to *capture* $\mathbf{L}$ (*in the strong sense*). □

5.6 Example : Our canonical example of a vectorized Lindström logic is $\mathbf{TC}$. Let $\sigma_T = \{E, L, R\}$ (E binary, L, R unary) and

$$\mathcal{TC} := \{\mathfrak{A} \in \mathcal{F}[\sigma_T] \mid \exists a \in L^{\mathfrak{A}}, b \in R^{\mathfrak{A}} \exists m \geq 2, a_1 = a, a_2, \ldots, a_m = b$$
$$\forall i < m : E^{\mathfrak{A}} a_i a_{i+1}\}.$$

(Thus $\mathcal{TC}$ is the class of σ_T–structures that have an E–path from the Left to the Right.)
Obviously $\mathcal{TC}$ is definable in $\mathbf{TC}$. Thus $\mathbf{FO}(Q_{\mathcal{TC}}^{<\omega}) \subseteq \mathbf{TC}$. On the other hand for each formula φ we have

$$[TC_{\overset{k}{x},\overset{k}{y}} \varphi] \ \overset{k}{u}, \overset{k}{v} \mathrel{=\!|\!\models} [Q^k_{\mathcal{TC}} \ \overset{k}{x}\overset{k}{y}, \overset{k}{x}, \overset{k}{y} \ \varphi, \overset{k}{x}=\overset{k}{u}, \overset{k}{y}=\overset{k}{v}].$$ [20]

Thus $Q_{\mathcal{TC}}^{<\omega}$ captures $\mathbf{TC}$.

Similarly we can define a class $\mathcal{DTC}$ such that $Q_{\mathcal{DTC}}^{<\omega}$ captures $\mathbf{DTC}$. □

5.7 Example : Another useful example is the *Hamiltonian path quantifier.* $\mathcal{HP}$ is the class of σ_T–structures that have a Hamiltonian path from the Left to the Right, i.e. a path in which every element of the structure occurs exactly once.

[20] We are a bit sloppy here and will be in similar situations to come. More precisely, we should say: For each TC–formula φ which is equivalent to a formula $\varphi' \in FO(Q_{\mathcal{TC}}^{<\omega})$ we have

$$[TC_{\overset{k}{x},\overset{k}{y}} \varphi] \ \overset{k}{u}, \overset{k}{v} \mathrel{=\!|\!\models} [Q^k_{\mathcal{TC}} \ \overset{k}{x}\overset{k}{y}, \overset{k}{x}, \overset{k}{y} \ \varphi', \overset{k}{x}=\overset{k}{u}, \overset{k}{y}=\overset{k}{v}].$$

Then we can prove inductively that $Q_{\mathcal{TC}}^{<\omega}$ captures $\mathbf{TC}$.

Stewart [Ste92] showed that on ordered structures every **NPTIME**–computable query is definable by a $Q_{\mathcal{HP}}$–formula of the form

$$[Q^k_{\mathcal{HP}}\ \overset{k}{x}\overset{k}{y}, \overset{k}{u}, \overset{k}{v}\ \ \theta, \bigwedge_{i=1}^{k} u_i = \underline{0}, \bigwedge_{i=1}^{k} v_i = \underline{e}]$$

where θ is quantifier–free. On the other hand it is easy to see that every query defined by a formula of this form is definable in **NPTIME**.

Similar results hold for other common complexity classes and the quantifiers we are going to define in the following (cf. Remark 5.21). $\square$

We can extend the concept by augmenting a logic by a set of Lindström–quantifiers:

5.8 Definition : Let $\mathbf{L} = (L, \models)$ be a logic and $\mathbf{Q}$ be a set of Lindström–quantifiers. The class $L(\mathbf{Q})$ of formulae is the closure of L under the clauses $(Q_\mathcal{C})$ for $Q_\mathcal{C} \in \mathbf{Q}$ and the rules used to build L. We let $\mathbf{L}(\mathbf{Q}) := (L(\mathbf{Q}), \models)$. $\square$

It is not hard to see that if $\mathbf{Q}$ is a finite class of Lindström–quantifiers then $\mathbf{FO}(\mathbf{Q})$ is a Lindström logic in the strong sense. In general this is not true for infinite $\mathbf{Q}$. For example consider

$$\mathbf{Q} = \{Q_\mathcal{C} \mid \sigma \text{ Signature, } \mathcal{C} \text{ is a class of } \sigma\text{–structures}\}.$$

Then every query is definable in $\mathbf{FO}(\mathbf{Q})$ (hence there are uncountably many queries definable in $\mathbf{FO}(\mathbf{Q})$). On the other hand a Lindström logic in the strong sense (even a vectorized Lindström logic) has only countably many formulae.

5.9 Definition : We call a logic $\mathbf{L}$ a *Lindström logic in the weak sense* if there exists a class $\mathbf{Q}$ of Lindström–quantifiers such that $\mathbf{L} = \mathbf{FO}(\mathbf{Q})$. $\square$

5.10 Example : We have already seen non–trivial examples of Lindström logics in the weak sense, namely vectorized Lindström–logics:
For a relational signature σ let $\sigma_k := \{R_k \mid R \in \sigma\}$ where, for an r–ary R, R_k is a $r \cdot k$–ary relation symbol.
For each class $\mathcal{C}$ of σ–structures we let $\mathcal{C}_k$ be the following class of σ_k–structures: A σ_k–structure $\mathfrak{B}$ belongs to $\mathcal{C}_k$ if and only if the structure $\mathfrak{A}$ defined by $A = B^k$ and, for each r–ary $R \in \sigma$,

$$R^\mathfrak{A} = \{\overset{r\times k}{b} \mid R_k^\mathfrak{B} b_{11} \dots b_{1k}\ \dots\ b_{r1} \dots b_{rk}\}$$

is in $\mathcal{C}$.

Then obviously for all formulae $\varphi_1, \ldots$ we have

$$[Q^k_{\mathcal{C}} \overset{r\times k}{x_1} \ldots \varphi_1 \ldots] \dashv\models [Q_{\mathcal{C}_k} \overset{r\times k}{x_1} \ldots \varphi_1 \ldots]$$

hence $\mathbf{FO}(Q^{<\omega}_{\mathcal{C}}) = \mathbf{FO}(\{Q_{\mathcal{C}_k} \mid k \geq 1\})$. □

In fact every regular logic **L** (cf. [Ebb85]) in which each formula contains only finitely many symbols is a Lindström–logic in the weak sense — just let

$$\mathbf{Q} := \{Q_{\mathcal{C}} \mid \exists\sigma : \mathcal{C} \in \mathcal{F}[\sigma] \text{ is definable in } L[\sigma]\} \,.$$

§ 5.2 A Vectorized Lindström–Quantifier that Captures SFP

Recall the definition of a network from Paragraph 0.6. We are going to define a class that gives rise to a Lindström–quantifier capturing **SFP** which is very similar to the class of networks.
Let $\sigma_S := \{E_1, E_2, L, R\}$ with binary relation symbols E_1 and E_2 and unary relation symbols L, R, and let c be a constant symbol.
A *stratwork* is an $\{E_1, E_2, L, c\}$–structure $\mathfrak{B}$ such that

- $(B, E_1^{\mathfrak{B}} \cup E_2^{\mathfrak{B}})$ is a connected acyclic digraph.
- Each node has either at least one E_1 and one E_2–successor (which may be the same) or no successor at all.
- $c^{\mathfrak{B}}$ is the unique root of $\mathfrak{B}$.
- $L^{\mathfrak{B}}$ is the the set of leaves of $\mathfrak{B}$.

Let $\mathcal{SFP}$ be the class of σ_S–structures $\mathfrak{A}$ such that there exists a stratwork $\mathfrak{B}$ with $B \subseteq A$, $c^{\mathfrak{B}} \in R^{\mathfrak{A}}$, $L^{\mathfrak{B}} \subseteq L^{\mathfrak{A}}$, $E_1^{\mathfrak{B}} \subseteq E_1^{\mathfrak{A}}$, and $E_2^{\mathfrak{B}} \subseteq E_2^{\mathfrak{A}}$.

5.11 Theorem : *The vectorized quantifier $Q^{<\omega}_{\mathcal{SFP}}$ captures* **SFP**.

Proof: $\mathcal{SFP}$ is the class of models of the **SFP**–formula

$$\chi := \exists u \in R\, [FP_{x,X}\; Lx \vee (\exists y \in X E_1xy \,\wedge\, \exists z \in X E_2xz)]u \,.$$

To see this note first that every structure in $\mathcal{SFP}$ is a model of χ.
On the other hand suppose that $\mathfrak{A} \models \chi$. In the proof of Theorem 2.5 we defined the stage comparison relation $\prec_\chi$ as follows:

$$a \prec_\chi b \quad\Longleftrightarrow\quad \underbrace{\min(\{i \in \mathbb{N} \mid a \in X_i^{\mathfrak{A}}\} \cup \{\infty\})}_{=|a|_\chi} < \min(\{i \in \mathbb{N} \mid b \in X_i^{\mathfrak{A}}\} \cup \{\infty\}).$$

We are going to define inductively a stratwork $\mathfrak{B}$ in $\mathfrak{A}$. Select a $\prec_\chi$–minimal $a \in R^{\mathfrak{A}}$ and let $c^{\mathfrak{B}} := a$, $B_1 := \{c^{\mathfrak{B}}\}$, and $E_1^1 := E_2^1 := \emptyset$. Note that since $\mathfrak{A} \models \chi$ we have $|c^{\mathfrak{B}}|_\chi < \infty$.
Having defined B_i, E_1^i, and E_2^i (such that for each $b \in B_i$ we have $|b|_\chi < \infty$), for each $b \in B_i \setminus L^{\mathfrak{A}}$ we select b_1, b_2 such that $E_1^{\mathfrak{A}} bb_1$, $E_2^{\mathfrak{A}} bb_2$, and $b_1 \prec_\chi b$, $b_2 \prec_\chi b$. Such b_1, b_2 exist since $1 < |b|_\chi < \infty$. We let

$$B_{i+1} := \{b_1, b_2 \mid b \in B_i \setminus L^{\mathfrak{A}}\} \quad \text{and} \quad E_j^{i+1} := \{bb_j \mid b \in B_i \setminus L^{\mathfrak{A}}\} \quad (j = 1, 2).$$

We proceed like this until we reach a k such that $B_k \subseteq L^{\mathfrak{A}}$. Letting $B := \bigcup_{i=1}^k B_i$, $E_j^{\mathfrak{B}} := \bigcup_{i=1}^k E_j^i$ $(j = 1, 2)$ we obtain a structure $\mathfrak{B}$ which is clearly a stratwork. (The only thing that is not obvious is that $(B, E_1^{\mathfrak{B}} \cup E_2^{\mathfrak{B}})$ is acyclic. But it is, since $E_j^{\mathfrak{B}} ab$ implies $b \prec_\chi a$ and $\prec_\chi$ is a partial order.)
Using a similar argument it is easy to see that the formulae

$$[FP_{\overset{k}{x}, X} \psi_0 \vee (\exists \overset{k}{y} \in X \psi_1 \wedge \exists \overset{k}{z} \in X \psi_2)] \overset{k}{u}$$

(where X does not occur in ψ_0, ψ_1, ψ_2).
and

$$[Q^k_{\mathcal{SFP}} \overset{k}{x}\overset{k}{y}, \overset{k}{x}\overset{k}{z}, \overset{k}{x}, \overset{k}{x} \;\; \psi_1, \psi_2, \psi_0, \overset{k}{x} = \overset{k}{u}]$$

are equivalent.
By Corollary 3.10 each *SFP*–formula is equivalent to a formula of the form above so this suffices. ■

5.12 Corollary : *Every $\exists LFP$–formula is equivalent to an $L(Q^{<\omega}_{\mathcal{SFP}})$–formula of the form*

$$[Q^k_{\mathcal{SFP}} \overset{k}{x}\overset{k}{y}, \overset{k}{x}\overset{k}{z}, \overset{k}{x}, \overset{k}{x} \;\; \theta_1, \theta_2, \theta_0, \bigwedge_{i,j \leq k} x_i = x_j]$$

where $\theta_0, \theta_1, \theta_2$ are quantifier–free.

Proof: This follows immediately from the proof of the last theorem and the stronger statement Corollary 3.10 makes about $\exists$**LFP**. ■

§ 5.3 Complete Problems

Again let $\mathbf{L} = (L, \models_L)$ be a logic and τ and $\sigma = \{R_1, \ldots, R_m\}$ two signatures (where R_i is r_i-ary).

5.13 Definition : (1) An $\mathbf{L}$–*translation* of σ in τ is a sequence of $L[\tau]$–formulae

$$\left(\varphi_1(\overset{r_1 \times k}{x_1}), \ldots, \varphi_m(\overset{r_m \times k}{x_m})\right)$$

(where the free variables of φ_h are in $\{x_{hij} \mid i \leq r_h, j \leq k\}$) for some $k \geq 1$.

(2) Each **L**–translation $\Phi = (\varphi_1(\overset{r_1 \times k}{x_1}), \ldots, \varphi_m(\overset{r_m \times k}{x_m}))$ associates a σ–structure

$$\mathfrak{A}^\Phi = \left(A^k, \varphi_1(\overset{r_1 \times k}{x_1})^{\mathfrak{A}}, \ldots, \varphi_m(\overset{r_m \times k}{x_m})^{\mathfrak{A}}\right)$$

with each τ–structure $\mathfrak{A}$.
We say that Φ is an **L**–reduction from a class $\mathcal{D}$ of τ–structures to a class $\mathcal{C}$ of σ–structures if for all τ–structures $\mathfrak{A}$ we have

$$\mathfrak{A} \in \mathcal{D} \iff \mathfrak{A}^\Phi \in \mathcal{C} \, .$$

(3) We say that a class $\mathcal{C}$ of σ–structures is *hard for* a logic $\tilde{\mathbf{L}} = (\tilde{L}, \models_{\tilde{L}})$ *under* **L***–reductions* if for every signature τ and every sentence $\varphi \in \tilde{L}[\tau]$ there exists an **L**–reduction from the class of models of φ to $\mathcal{C}$.
$\mathcal{C}$ is *complete for* $\tilde{\mathbf{L}}$ *under* **L***–reductions* if it is definable in $\tilde{\mathbf{L}}$ and hard for $\tilde{\mathbf{L}}$ under **L**–reductions.
Occasionally we say that $\mathcal{C}$ is a *complete problem* for $\tilde{\mathbf{L}}$ (under **L**–reductions).

In particular, if $\mathbf{L} = \mathbf{QF}$ we say that $\mathcal{C}$ is complete (hard, a complete problem) for $\tilde{\mathbf{L}}$ under *quantifier–free reductions.* □

5.14 Example : $\Phi = (Ex_2x_1, Ex_2x_1, Lx, Rx)$ is a quantifier–free reduction from $\mathcal{TC}$ to $\mathcal{SFP}$. □

5.15 Remark : Obviously $\mathcal{C}$ is hard for $\tilde{\mathbf{L}}$ under **L**–reductions if and only if for every sentence $\varphi \in \tilde{L}$ there is an equivalent $L(Q_{\mathcal{C}}^{<\omega})$–sentence *in normal form*, i.e. of the form

$$[Q_{\mathcal{C}}^k \overset{r_1 \times k}{x}, \ldots, \overset{r_m \times k}{x} \quad \varphi_1, \ldots, \varphi_m]$$

with $\varphi_1, \ldots, \varphi_m$ in **L**.
Thus $\mathcal{C}$ is complete for $\mathbf{L}(Q_{\mathcal{C}}^{<\omega})$ under **L**–reductions if and only if every $L(Q_{\mathcal{C}})$–sentence is equivalent to a sentence in normal form. □

§5.4 Complete Problems for Fixed–Point Logics

Let $\sigma_F = \{T, L, R\}$ with ternary T and unary L, R.
Each σ_F–structure $\mathfrak{A}$ gives rise to a sequence $(P_i^{\mathfrak{A}})_{i \geq 1}$ defined by $P_1^{\mathfrak{A}} = L^{\mathfrak{A}}$ and

$$P_{i+1}^{\mathfrak{A}} = \{a \in A \mid \exists b \in P_i^{\mathfrak{A}}, c \notin P_i^{\mathfrak{A}} : T^{\mathfrak{A}}abc\}$$

for each $i \geq 1$. We let

$$P_\infty^{\mathfrak{A}} = \begin{cases} P_m^{\mathfrak{A}} & \text{where } m = \min\{i \mid P_i^{\mathfrak{A}} = P_{i+1}^{\mathfrak{A}}\} \text{ if it exists} \\ \emptyset & \text{otherwise} \end{cases} \quad .$$

Note that if there exists an $i \geq 1$ such that $P_i^{\mathfrak{A}} = \emptyset$ or $P_i^{\mathfrak{A}} = A$ then $P_{i+1}^{\mathfrak{A}} = \emptyset$ and thus $P_\infty^{\mathfrak{A}} = \emptyset$.

Let $\mathcal{PFP}$ be the class consisting of all σ_F–structures $\mathfrak{A}$ with $R^{\mathfrak{A}} \subseteq P_\infty^{\mathfrak{A}}$.

The class $\mathcal{PFP}$ mirrors what happens in the fixed–point process of a fixed–point sentence in bounded normal form. We have:

5.16 Theorem : *The class $\mathcal{PFP}$ is complete for* **PFP** *under quantifier–free reductions.*

Proof: Consider a fixed–point sentence

$$\chi := \overset{\exists}{\forall} v[FP_{\overset{k}{x},X}\ \theta_0(\overset{k}{x}) \vee \exists \overset{k}{y} \in X \ \exists \overset{k}{z} \notin X \ \theta(\overset{k}{x},\overset{k}{y},\overset{k}{z})]v\dots v$$

in bounded normal form. By Remark 3.7 we can also assume that χ is modest (cf. Definition 3.4).
Let Φ be the quantifier–free translation

$$\Big(\theta(\overset{k}{x},\overset{k}{y},\overset{k}{z}) \vee \theta_0(\overset{k}{x}), \theta_0(\overset{k}{x}), \bigwedge_{i,j\leq k} x_i = x_j\Big)$$

of σ_F in σ (where σ is the signature of χ).

We let $\mathfrak{A}$ be a σ–structure and note that $\mathfrak{A} \models \chi$ if and only if

$$\{\overset{k}{a} \in A \mid \forall i,j \leq k : a_i = a_j\} \subseteq X_\infty^{\mathfrak{A}} \ .$$

An easy induction (using that χ is modest) shows that, moreover, for each $i \geq 1$ we have $P_i^{\mathfrak{A}^\Phi} = X_i^{\mathfrak{A}}$. Thus

$$\mathfrak{A} \models \chi \quad \Longleftrightarrow \quad \mathfrak{A}^\Phi \in \mathcal{PFP} \ .$$

By our bounded normal form Theorem 3.6 this suffices to see that $\mathcal{PFP}$ is hard for **PFP** under quantifier–free reductions.

Furthermore, $\mathcal{PFP}$ is the class of models of the fixed–point sentence

$$\xi = \forall u(Ru \longrightarrow [FP_{x,X}\ (Lx \wedge X = \emptyset) \vee \exists y \in X, \exists z \notin X \ Txyz]u)$$

hence complete for **PFP**.
To see this let $\mathfrak{A}$ be a τ–structure and consider the sequences $(P_i^{\mathfrak{A}})_{i\geq 1}$ and $(X_i^{\mathfrak{A}})_{i\geq 0}$ (the fixed–point process inside of ξ). If there exists no $i \geq 1$ such that $P_i^{\mathfrak{A}} = \emptyset$ or $P_i^{\mathfrak{A}} = A$ then for each $i \geq 1$ we have $P_i^{\mathfrak{A}} = X_i^{\mathfrak{A}}$. Otherwise let i_0 be the first such i. We still have $P_i^{\mathfrak{A}} = X_i^{\mathfrak{A}}$ for each $i \leq i_0$. We have seen above that $P_\infty^{\mathfrak{A}} = \emptyset$ in this case; and we have $X_{i_0+1}^{\mathfrak{A}} = \emptyset$ (if $X_{i_0}^{\mathfrak{A}} = P_{i_0}^{\mathfrak{A}} = A$) or $X_{i_0+1}^{\mathfrak{A}} = L^{\mathfrak{A}}$ (if $X_{i_0}^{\mathfrak{A}} = \emptyset$), thus the sequence $(X_i^{\mathfrak{A}})_{i\geq 0}$ becomes periodic or remains empty.

In any case we have $P_\infty^{\mathfrak{A}} = X_\infty^{\mathfrak{A}}$. Thus $\mathfrak{A} \models \xi$ if and only if $\mathfrak{A} \in \mathcal{PFP}$. ∎

To capture **IFP** we let $\mathcal{IFP}$ be the subclass of $\mathcal{PFP}$ consisting of those structures $\mathfrak{A} \in \mathcal{PFP}$ whose sequence $(P_i^{\mathfrak{A}})_{i\geq 1}$ is inductive (i.e. $P_i^{\mathfrak{A}} \subseteq P_{i+1}^{\mathfrak{A}}$ for each $i \geq 1$).

5.17 Theorem : *The class $\mathcal{IFP}$ is complete for* **IFP** *under quantifier–free reductions.*

Proof: It can be seen analogously to the hardness–part of the previous proof that $\mathcal{IFP}$ is hard for **IFP** under quantifier–free reductions.

However, it is not completely trivial to see that it is definable in **IFP**. First, let us consider the *IFP*–formula

$$\varphi(x) := [FP_{x,X} Lx \vee Xx \vee (\exists y \in X \exists z \notin X\, Txyz \wedge \forall x' \in X \exists y \in X \exists z \notin X\, Tx'yz)]x.$$

Let $\mathfrak{A}$ be a σ_F–structure and $m := \min(\{i \mid P_i^{\mathfrak{A}} \not\subseteq P_{i+1}^{\mathfrak{A}}\} \cup \{\infty\})$.
Then for each $i \leq m$ we have $X_i^{\mathfrak{A}} = P_i^{\mathfrak{A}}$. Furthermore $X_i^{\mathfrak{A}} = X_m^{\mathfrak{A}}$ for each $i > m$.
Thus $X_\infty^{\mathfrak{A}} = P_m^{\mathfrak{A}}$.
The following *IFP*–sentence says that the sequence $(P_i^{\mathfrak{A}})_{i\geq 1}$ is inductive:

$$\psi := \forall x(\varphi(x) \leftrightarrow \exists y, z(\varphi(y) \wedge \neg\varphi(z) \wedge Txyz))$$

(since it holds in a structure $\mathfrak{A}$ if and only if $P_m^{\mathfrak{A}}$ is a fixed–point (of the operation used to define the sequence $(P_i^{\mathfrak{A}})_{i\geq 1}$) hence equals $P_\infty^{\mathfrak{A}}$).
So $\mathfrak{I}$ is the class of models of the sentence

$$\psi \wedge \forall x(Rx \to \varphi(x))\ .$$

∎

The following corollary is interesting in so far as it gives a combinatorial characterization (that does not need any notion of computation and very few logic) of the question of whether **PTIME** equals **PSPACE**.

5.18 Corollary : **PTIME=PSPACE** *if and only if there exists a quantifier–free reduction from $\mathcal{PFP}$ to $\mathcal{IFP}$.*

Proof: Recall Abiteboul's and Vianu's Theorem 1.2 which says that **PTIME** equals **PSPACE** if and only if **IFP** equals **PFP**. ∎

We have seen that the question of whether there exists a complete problem for a vectorized Lindström logic is connected with the question of whether there exists a normal form. Since **TC** and **SFP** are vectorized Lindström–logics that do not have a normal form it is no surprise that:

5.19 Theorem : *There are no complete problems for* **SFP** *or* **TC** *under first–order reductions.*

Proof: Suppose a class $\mathcal{C}$ of σ–structures is complete for **SFP** under first–order reductions. Then it is definable by a sentence $\varphi \in SFP$ of fixed–point rank, say, r. Hence each class $\mathcal{C}_k$ (cf. Example 5.10) is definable by a sentence $\varphi_k \in SFP$ of fixed–point rank r.
But then every $\psi \in SFP$ is equivalent to a sentence of fixed–point rank r (that is obtained by substituting atomic subformulae of a φ_k of the form $R_k\bar{x}$ $(R \in \sigma)$ by first–order formulae).
This contradicts Theorem 3.11.

A similar argument holds for **TC**. ■

We have seen that the situation (concerning normal forms) changes for the existential fragments of these logics. And indeed:

5.20 Corollary : *The classes* $\mathcal{TC}$ *and* $\mathcal{SFP}$ *are complete for* ∃**TC** *and* ∃**LFP** *respectively under quantifier–free reductions.*

Proof: The ∃**LFP**–part is essentially Corollary 5.12 (recalling Remark 5.15), and the ∃**TC**–part follows from Theorem 4.4. ■

5.21 Remark : The connection between logical and complexity–theoretical reducibilities has been extensively studied (cf. [LG77, Imm87, MP93]). For example, it is known that on ordered structures a problem is complete for one of the usual complexity–classes under **LOGSPACE**–reductions if and only if it is complete for the class under **DTC**–reductions (which is not very surprising since **DTC** $=_{\mathcal{O}}$ **LOGSPACE** by Immerman's Theorem 4.12). Thus quantifier–free reductions are quite interesting under a complexity theoretic point of view as a stronger form of reductions than **LOGSPACE**–reductions.

The previous results, together with Theorems 1.17 and 4.12 show that on ordered structures we have:

- $\mathcal{PFP}$ is complete for **PSPACE** under quantifier–free reductions.
- $\mathcal{IFP}$ and $\mathcal{SFP}$ are complete for **PTIME** under quantifier–free reductions.
- $\mathcal{TC}$ is complete for **NLOGSPACE** under quantifier–free reductions.

Similarly it can be seen that

- $\mathcal{DTC}$ is complete for **LOGSPACE** under quantifier–free reductions.

To complete the picture, Stewart's result mentioned in Example 5.7 shows:

- $\mathcal{HP}$ is complete for **NPTIME** under quantifier–free reductions.

□

Part II
Arity Hierarchies

6 k–Ary Fragments

In the first part we studied the influence of two syntactical parameters on the expressive power of fixed–point formulae: the quantifier–structure and the number of occurrences of the relation variables that belong to the fixed–point operators. We saw that the influence of both these parameters is unexpectedly small. The situation is drastically different for the third natural parameter, the *arity* of the fixed–point operators.

6.1 Definition : The *arity* of a fixed–point operator $[FP_{\bar{x},X}\ldots]$ is defined to be the arity of X. The arity of a simultaneous fixed–point operator $[s\text{–}FP_{\bar{x}_1,X_1,\ldots,\bar{x}_m,X_m}\ldots]$ is defined to be the maximum of the arities of $X_1,\ldots,X_m$. A (simultaneous) fixed–point formula is *k–ary* if it contains at most k–ary (simultaneous) fixed–point operators. (It is 0–ary if it does not contain any fixed–point operator.) $\Box$

Since we are going to consider arities in several different logics it is convenient to give the following abstract definition:

6.2 Definition : Let $\mathbf{L} = (L, \models)$ be a logic and suppose we have assigned an *arity*, i.e. a non–negative integer, to each formula in L. Then L^k denotes the set of all formulae in L whose arity is $\leq k$.
The *k–ary fragment* $\mathbf{L}^k$ of $\mathbf{L}$ is defined to be $(L^k, \models)$.
We call the sequence $(\mathbf{L}^k)_{k\geq 0}$ the *arity hierarchy* of $\mathbf{L}$. We say that the arity hierarchy of $\mathbf{L}$ is strict if for each $k \geq 1$ there exists a query that is definable in $\mathbf{L}^k$ but not in $\mathbf{L}^{k-1}$. $\Box$

For example, the k–ary fragment of $\exists\mathbf{LFP}$ is $\exists\mathbf{LFP}^k = (\overbrace{\exists LFP \cap FP^k}^{=\exists LFP^k}, \models)$.

Arities of transitive closure formulae are defined as follows:

6.3 Definition : The arity of a transitive closure operator $[TC_{\bar{x},\bar{y}}\ldots]$ is defined to be the length of $\bar{x}$.
A transitive closure formula is k–ary if it contains at most k–ary transitive closure operators. (It is 0–ary if it does not contain any transitive closure operator.) $\Box$

The key result is the following main lemma which implies that the arity hierarchies are strict for all logics considered above and, furthermore, that the queries used to separate the k–ary from the $(k-1)$–ary fragment are rather simple.

6.4 Main Lemma : *For each $k \geq 2$ there exists a graph query which is definable by a k–ary transitive closure formula of the form*

$$[TC_{\overset{k}{x},\overset{k}{y}}\textsc{Edge}_k]\,\overset{k}{u},\overset{k}{v}$$

(where EDGE_k *is a conjunction of atomic formulae),*
but that is not definable in $\mathbf{s\text{–}PFP}^{k-1}$.

6.5 Corollary : *The arity hierarchies are strict for the transitive closure logics* $\exists$**TC**, **TC**, *for the fixed–point logics* $\exists$**LFP**, **SFP**, **LFP**, **IFP**, **PFP**, *and for their simultaneous variants* **s–**$\exists$**LFP**, ... ,**s–PFP**. *This holds even on the class of graphs.*

To prove the main lemma we proceed as follows:
We first construct certain families of structures which are somehow "homogeneous in the $(k-1)$–tuples" (thus $(k-1)$–ary fixed–point formulae become ineffective). To achieve this we use a method that was introduced by Hrushovski [Hru92] (in a completely different context) to extend partial isomorphisms of finite graphs. Then we establish the non–expressibility of our query by $(k-1)$–ary formulae on these structures employing Ehrenfeucht-Fraïssé games.
Therefore we introduce a new game to deal with fragments of bounded–arity of our fixed–point logics. It is based on the well–known pebble game for infinitary logic with only finitely many variables.

We can also define arity hierarchies for the datalogics (the variants of the database language datalog introduced in Paragraph 1.6).

6.6 Definition : The arity of a general datalog program is the maximal arity of its intensional relation variables. $\square$

Since we cannot define a k–ary query by a $(k-1)$–ary datalog program (because the goal variable is at most $(k-1)$–ary) it is reasonable to modify the notion of the arity–hierarchy being strict slightly for the datalogics. We say that the arity hierarchy of a datalogic is strict if, for each $k \geq 1$, there is a Boolean query which is definable in the k–ary fragment, but not in the $(k-1)$–ary fragment.[21]

Thus, to be able to apply the main lemma to datalogics we have to modify our query. We augment our signature by constants $\underline{c}_1, \dots, \underline{c}_k, \underline{d}_1, \dots, \underline{d}_k$ and consider the Boolean graph query

[21] To circumvent this problem Afrati and Cosmadakis [AC89] consider the *right hand side–width (rhs–width)* of a datalog–program which is defined to be the maximal arity of the intensional variables that occur in the body of a clause instead of the arity. The advantage is that it is possible to define queries of arbitrary arity by programs of bounded rhs–width.

Note that every datalog program of rhs–width $\leq k$ is equivalent to a program where all intensional variables except the goal–variable have an arity $\leq k$. Thus a datalog program that defines a Boolean query is of arity $\leq k$ if and only if it is of rhs–width $\leq k$. Since we are going to work with Boolean queries our results also hold for "rhs–width" instead of "arity".

(We prefer our notion of arity because it is immediate that k–ary datalog programs are equivalent to k–ary simultaneous fixed–point formulae.)

There exists an EDGE_k*–path from* $\overset{k}{\underline{c}}$ *to* $\overset{k}{\underline{d}}$.

which can easily be defined by a k–ary (pure) datalog program (but still not in **s–PFP**$^{k-1}$ by the main lemma).
The reader who is bothering with the constants and prefers a pure graph query will be satisfied by Paragraph 8.4 in which we show that we can modify the structures we are going to use in the proof in a way that the constants are definable, even by a pure datalog–program.

6.7 Corollary : *The arity hierarchies are strict (even on the class of graphs) for* **Datalog**, **Datalog**$^{\neg}$, **S–Datalog**, **I–Datalog**, *and* **P–Datalog**.

The reader may miss **DTC** in the list above. The arity hierarchy is also strict for **DTC**, but we cannot strengthen our main lemma in a way that this follows since we have **DTC** $\subseteq$**s–PFP**^{1}. We will study questions concerning the arities of transitive closure operators, in particular deterministic ones, in the last section.

In Sections 9 and 10 we are going to consider variants of our main lemma that provide hierarchy results for implicitly definable queries and Lindström–quantifiers.

7 Tools

§ 7.1 An Ehrenfeucht-Fraïssé Game

7.1 Definition : Let $r, k \geq 0$. The *k–ary r–pebble game* is played by two players, the challenger and the duplicator, on a pair $\mathfrak{A}, \mathfrak{B}$ of structures of the same signature with $2r$ pebbles $P_1, Q_1, \ldots, P_r, Q_r$.
We say that a pair (P_i, Q_i) of pebbles is *on the board* in a situation of the game if it is placed on the structures in that situation. The other pebbles are called *free*. Each situation of the game and each pebble on the board have a depth ≥ 0. The game starts in the situation with depth 0 and all pebbles free.
In each situation of the game the challenger selects one of the following moves:

∃-move : The challenger places a free pebble P_i on an element $a_i \in A$. The duplicator places the corresponding pebble Q_i on a $b_i \in B$. The depth of P_i and Q_i is defined to be the current depth of the game.

∀-move : The challenger places a free pebble Q_i on $b_i \in B$. The duplicator places P_i on $a_i \in A$. The depth of P_i and Q_i is defined to be the current depth of the game.

I-move : The depth of the game is increased by 1.

R-move : This move is only allowed if the current depth of the game is ≥ 1. The challenger reduces the depth of the game to an arbitrary $d \geq 1$ less than or equal to the current depth. Then he[22] selects $l \leq k$ pairs (P_i, Q_i) of depth $\geq d$ to be left on the board. Their depth is reduced to d. All other pebbles of depth $\geq d$ are removed from the structures.

In each situation of the game the pairs of pebbled elements (a_i, b_i) and the pairs $(c^{\mathfrak{A}}, c^{\mathfrak{B}})$ of constants are called *pairs of corresponding elements*.
The duplicator wins the game if in each situation the pairs of corresponding elements form a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$. □

We can straightforwardly extend our definition of the quantifier–rank to simultaneous fixed–point formulae adding the clause

$$qr([s\text{–}FP_{x_1,X_1,\ldots,x_m,X_m}^{l_1 \qquad l_m} \varphi_1, \ldots, \varphi_m] \overset{l_1}{u}) = \max\{qr(\varphi_i) + l_i \mid i \leq m\} .$$

7.2 Theorem : *Let $r, k \geq 0$ and $\mathfrak{A}$ and $\mathfrak{B}$ be two structures such that the duplicator has a winning strategy for the k–ary r–pebble game on $\mathfrak{A}, \mathfrak{B}$.
Then the same s–FP^k–sentences of quantifier–rank $\leq r$ hold in $\mathfrak{A}$ and $\mathfrak{B}$.*

[22] We denote the challenger by male and the duplicator by female pronouns.

7.3 Remark : Let me mention that the analogous theorem (10.6) for the l–bijective k–ary r–pebble game we are going to introduce in Section 10 has nearly the same proof. So the reader who does not want to do the same work twice may read Paragraph 10.1 first and then check both proofs parallel. □

Proof (of the theorem): Let χ be an s–FP^k–sentence of quantifier rank $\leq r$ such that $\mathfrak{A} \models \chi$. We shall prove that $\mathfrak{B} \models \chi$.

There is a canonical way to construct for each s–FP–formula χ an equivalent formula where negation symbols only occur in front of atomic subformulae or in front of s–FP–operators. We denote it by $\overline{\chi}$.
Without loss of generality we assume that $\chi = \overline{\chi}$.

Step 1: Defining the stages

In this step we generalize Paragraph 1.2 to simultaneous fixed–point formulae. In addition, we are a bit more careful with the variables here than we were there. We consider a subformula

$$\psi = [s\text{–}FP_{\overset{l_1}{x}_1,X_1,\ldots,\overset{l_m}{x}_m,X_m}\ \varphi_1,\ldots,\varphi_m]\ \overset{l_1}{u}$$

of χ.

We may assume without loss of generality that none of the free variables of ψ is quantified in a φ_i. For simplicity we also assume that for each subformula of ψ of the form $X_i\ \overset{l_i}{t}$ we have $\overset{l_i}{t} = \overset{l_i}{x}_i$.
(There are two ways to deal with the fact that the original formula ψ might have a subformula $X_i\ \overset{l_i}{t}$ where $\overset{l_i}{t} \neq \overset{l_i}{x}_i$:
The first is to rename variables in the formulae ψ_{ij} (defined below) in a suitable way.
The second way is to replace $X_i\ \overset{l_i}{t}$ by

$$\exists\ \overset{l_i}{z}\ (\ \overset{l_i}{t} = \overset{l_i}{z}\ \wedge \exists\ \overset{l_i}{x}_i\ (\overset{l_i}{z} = \overset{l_i}{x}_i\ \wedge X\ \overset{l_i}{x}_i))$$

where $z_1, \ldots, z_{l_i}$ are variables that do not occur in free$(\overset{l_i}{t}) \cup \{x_{i1}, \ldots, x_{il_i}\}$. Note that this may increase the quantifier–rank of the formula so we have to reformulate the theorem a bit.)

Now, we let $\psi_{i0} := x_{i1} \neq x_{i1}$ and for each $j \geq 0$ we let $\psi_{i(j+1)}$ be the formula obtained from φ_i by replacing, for each $i' \leq m$, each subformula of the form $\neg X_{i'}\ \overset{l_{i'}}{x}_{i'}$ by $\neg\psi_{i'j}$ and each (remaining) subformula of the form $X_{i'}\ \overset{l_{i'}}{x}_{i'}$ by $\psi_{i'j}$.

Then clearly ψ_{ij} defines the set $X_{ij}^{\mathfrak{I}}$ for each interpretation $\mathfrak{I}$ (note that the restriction that none of the free variables of ψ is quantified in a φ_i is needed here) and we have

$$\mathfrak{I} \models \psi \quad \Longleftrightarrow \quad \exists j \geq 1 : \mathfrak{I} \models \underbrace{\exists \overset{l_1}{x_1} \, (\overset{l_1}{x_1} = \overset{l_1}{u} \wedge \psi_{1j}) \wedge \bigwedge_{i=1}^{m} \forall \overset{l_i}{x_i} \, (\psi_{i(j)} \leftrightarrow \psi_{i(j+1)})}_{=:\psi_j^*(\overset{l_1}{u})} \; .$$

We call subformulae ψ_{ij} and $\neg\psi_{ij}$ of the $\psi_{i'j'}$ $(i, i' \leq m,\ 0 \leq j < j')$ which are obtained by replacing subformulae $X \overset{l_i}{x_i}$ or $\neg X \overset{l_i}{x_i}$ of $\varphi_{i'}$ *R(eplacement)–formulae of* ψ.

Step 2: Proof–Sequences for χ

Definition: A *proof–situation* s consists of a formula ψ and a mapping ι from the free variables of ψ to $\{1, \ldots, r\}$. In addition, it may contain a move of the k–ary r–pebble game.
A *proof–sequence* S is a sequence of proof–situations such that

- The sequence of moves in the k–ary r–pebble game that belongs to S is according to the rules of the game and the duplicator's moves are according to her winning strategy.
- Hence each proof–situation s in S corresponds to a situation in the k–ary r–pebble game (obtained by the moves of the previous proof–situations (including s) in S). If ι is the mapping that belongs to s and i is in the range of ι then the pebbles P_i and Q_i are on the board in this situation.
- Thus the mapping ι induces two (partial) assignments α and β (for the free variables of the formula ψ) defined by $\alpha(x) = a$ and $\beta(x) = b$ if $\iota(x) = i$ and P_i and Q_i are placed on $a \in A$ and $b \in B$ respectively.

  We always have $(\mathfrak{A}, \alpha) \models \psi$.

If S and T are two sequences then $S^\frown T$ denotes their concatenation. □

We are going to define a set $\mathcal{S}$ of proof–sequences. The formulae that occur in these proof–sequences are essentially subformulae of χ or formulae $\psi_j^*(\overset{l}{u})$.
Depending on the form of such formulae ψ, we inductively define the set $\mathcal{S}$. The induction will be "from the outside to the inside", i.e. we start with χ and continue with its subformulae and their subformulae and so on.

Let s_0 be the proof–situation with formula χ, no move in the k–ary r–pebble game and the empty mapping ι (χ is a sentence !).
We start our induction with $\mathcal{S} = \{s_0\}$ (here we do not distinguish between the proof–situation s_0 and the proof–sequence consisting only of s_0). Note that s_0 is indeed a proof–sequence since $\mathfrak{A} \models \chi$.

Now, for each proof–sequence $S \in \mathcal{S}$ we consider the formula ψ of its last proof–situation. Let ι be the mapping belonging to this situation and α the assignment for the free variables of ψ in $\mathfrak{A}$ it induces. Since S is a proof–sequence we have $(\mathfrak{A}, \alpha) \models \psi$.
If ψ is not atomic or negated atomic and not an R–formula we proceed as follows depending on the form of ψ (see below for the case that ψ is an R–formula):

$\boldsymbol{\psi = \varphi_1 \lor \varphi_2}$: Since $(\mathfrak{A}, \alpha) \models \psi$ we have $(\mathfrak{A}, \alpha) \models \varphi_1$ or $(\mathfrak{A}, \alpha) \models \varphi_2$; without loss of generality we assume the first. Let s_1 be the proof–situation consisting of no move, the formula φ_1, and the restriction of ι to free(φ_1) as its mapping. $(\mathfrak{A}, \alpha) \models \varphi_1$ guarantees that $S^\frown s_1$ is a proof–sequence.

We replace S in $\mathcal{S}$ by $S^\frown s_1$.

$\boldsymbol{\psi = \varphi_1 \land \varphi_2}$: Let s_i $(i = 1, 2)$ be the proof–situation consisting of no move, the formula φ_i, and the restriction of ι to free(φ_i). Then $S^\frown s_i$ is a proof–sequence.

We remove S from $\mathcal{S}$ and put in $S^\frown s_1$ and $S^\frown s_2$.

$\boldsymbol{\psi = \exists x \varphi(x)}$: Select an $a \in A$ such that $(\mathfrak{A}, \alpha) \models \varphi[a]$.

Let s_a be the proof–situation whose move is the following $\exists$–move: The challenger places a free pebble, say P_i, on a. The duplicator places Q_i on $\mathfrak{B}$ according to her winning strategy. (Our treatment of R–formulae described below guarantees that there are always enough free pebbles.)

Let φ be the formula of s_a and $(\iota \cup \{(x, i)\})\restriction_{\text{free}(\varphi)}$ its mapping.

We replace S in $\mathcal{S}$ by $S^\frown s_a$.

$\boldsymbol{\psi = \forall x \varphi(x)}$: For each $b \in B$ let s_b be the proof–situation whose move is the following $\forall$–move: The challenger places a free pebble, say Q_i, on b. The duplicator places P_i on $\mathfrak{A}$ according to her winning strategy.

Let φ be the formula of s_b and $(\iota \cup \{(x, i)\})\restriction_{\text{free}(\varphi)}$ its mapping.

We replace S in $\mathcal{S}$ by all sequences $S^\frown s_b$ $(b \in B)$.

$\boldsymbol{\psi = [s\text{–}FP_{\overset{l_1}{x_1}, X_1, \ldots, \overset{l_m}{x_m}, X_m} \varphi_1, \ldots, \varphi_m] \, \overset{l_1}{u}}$: Select $j \geq 1$ such that $(\mathfrak{A}, \alpha) \models \psi_j^*(\overset{l_1}{u})$.

Let s_j be the proof–situation consisting of an I–move, the formula $\overline{\psi_j^*(\overset{l_1}{u})}$ and the mapping ι restricted to free$(\psi_j^*(\overset{l_1}{u}))(\subseteq$ free$(\psi))$.

We replace S in $\mathcal{S}$ by $S^\frown s_j$.

$\boldsymbol{\psi = \neg[s\text{–}FP_{\overset{l_1}{x_1}, X_1, \ldots, \overset{l_m}{x_m}, X_m} \varphi_1, \ldots, \varphi_m] \, \overset{l_1}{u}}$: Note that for each $j \geq 1$ we have

$$(\mathfrak{A}, \alpha) \models \neg\psi_j^*(\overset{l_1}{u}).$$

Let s_j be the proof situation consisting of an I–move, the formula $\overline{\neg\psi_j^*(\overset{l_1}{u})}$ and the mapping ι restricted to free$(\psi_j^*(\overset{l_1}{u}))$.

We replace S in $\mathcal{S}$ by the sequences $S^\frown s_j$ $(j \geq 1)$.

However, since the quantifier–rank of formulae of the form $\psi_j^*(\overset{l}{u})$ is not necessarily $\leq r$ it may happen that we run out of free pebbles. To prevent this, whenever ψ is an R–formula (cf. Step 1) we proceed as follows:

ψ R–formula: In this case we have to extend S twice, the first time to include an R–move into our sequence that sets some pebbles free, and the second time in the usual way (described above) to take care of ψ. (Being an R–formula is only an additional feature and does not say anything about the form of ψ.)

FIRST EXTENSION: Suppose ψ is an R–formula $\overline{\xi_{ij}}$ or $\overline{\neg\xi_{ij}}$[23] for a formula

$$\xi = (\neg)[s\text{–}FP_{\overset{l_1}{x_1},X_1,\ldots,\overset{l_m}{x_m},X_m}\ \varphi_1,\ldots,\varphi_m]\ \overset{l_1}{u}\ .$$

Then there exists a proof–situation s in S with formula ξ. The move of the situation following s must have been an I–move which increased the depth of the game to, say, d.

Let s' be the situation consisting of the formula ψ, the mapping ι (of the current situation), and the following R–move:

The challenger reduces the depth to d and removes all pebbles of depth $\geq d$ except those pairs (P_n, Q_n) with $n \in \{\iota(x_{ih}) \mid h \leq l_i\}$ from the board.

Note that, since free$(\xi_{ij}) \subseteq$ free$(\xi) \cup \{x_{i1}, \ldots, x_{il_i}\}$, all pebbles (P_n, Q_n) with $n \in \{\iota(x) \mid x \in \text{free}(\psi)\}$ are still on the board as they are either of depth $< d$ (here we use that none of the free variables of ξ is quantified in a $\varphi_{i'}$) or of the form (P_n, Q_n) with $n \in \{\iota(x_{ih}) \mid h \leq l_i\}$.

Thus $S' = S^\frown s'$ is a proof–sequence (because S is one). We replace S by S'.

SECOND EXTENSION: We now proceed with S' as if ψ (the formula belonging to s') was not an R–formula. (I.e. we apply one of the "standard" operations defined above to S', depending on the form of ψ.)

We continue with this procedure until each proof–sequence in $\mathcal{S}$ ends with an atomic formula. Note that, although $\mathcal{S}$ may be infinite, each sequence $S \in \mathcal{S}$ is of finite length. (Formally, we can prove this by transfinite induction.)

[23] More precisely, ξ_{ij} or $\neg\xi_{ij}$ is an R–formula (instead of $\overline{\xi_{ij}}$ or $\overline{\neg\xi_{ij}}$).

Note also that our treatment of the R–formulae guarantees that we never run out of free pebbles. This can also be verified by induction (on the (finite) length of a proof–sequence).

Step 3: Proving $\mathfrak{B} \models \chi$ with χ's Proof–Sequences

Inductively we prove that for each proof–situation s occurring in a sequence $S \in \mathcal{S}$:

Claim: *If ψ is the formula of s and β the interpretation of the free variables of ψ in $\mathfrak{B}$ induced by the mapping ι of s we have $(\mathfrak{B}, \beta) \models \psi$.*

This implies $\mathfrak{B} \models \chi$.

Proof: This time the induction is "from the inside to the outside":

ψ atomic or negated atomic: Let α be the interpretation of the free variables of ψ in $\mathfrak{A}$ induced by the mapping ι of s.

By Step 2 we have $(\mathfrak{A}, \alpha) \models \psi$. Furthermore, since the duplicator has always played according to her winning strategy in the k–ary r–pebble game the mapping between the corresponding elements is a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

By the definition of α and β this means that the same atomic or negated atomic formulae with variables in free(ψ) hold in $(\mathfrak{A}, \alpha)$ and $(\mathfrak{B}, \beta)$. In particular, we have $(\mathfrak{B}, \beta) \models \psi$.

The cases $\psi = \varphi_1 \vee \varphi_2$, $\psi = \varphi_1 \wedge \varphi_2$, $\psi = \exists x\varphi$, and $\psi = \forall x\varphi$ are all easy. To illustrate this we consider only the last:

$\psi = \forall x\varphi$: By the definition of $\mathcal{S}$ for each $b \in B$ there exists a sequence $S \in \mathcal{S}$ where situation s is followed by a situation s_b with formula φ and induced assignment $\beta\frac{b}{x}$ extending β by $\beta\frac{b}{x}(x) = b$.

By the induction hypothesis we know that $(\mathfrak{B}, \beta\frac{b}{x}) \models \varphi$. Hence $(\mathfrak{B}, \beta) \models \psi$.

The fixed–point cases are also easy, e.g.:

$\psi = [s\text{–}FP_{\overset{l_1}{x_1},X_1,\ldots,\overset{l_m}{x_m},X_m}\ \varphi_1,\ldots,\varphi_m]\,\overset{l_1}{u}$: By the definition of $\mathcal{S}$ there exists a $j \geq 1$ and a sequence $S \in \mathcal{S}$ where situation s is followed by situation s_j with formula $\overline{\psi_j^*(\overset{l_1}{u})}$ and induced assignment $\beta' \subseteq \beta$.

By the induction hypothesis we know that $(\mathfrak{B}, \beta') \models \overline{\psi_j^*(\overset{l_1}{u})}$ and thus $(\mathfrak{B}, \beta) \models \psi$.

The case $\psi = \neg[s\text{–}FP_{\overset{l_1}{x_1},X_1,\ldots,\overset{l_m}{x_m},X_m}\ \varphi_1,\ldots,\varphi_m]\,\overset{l_1}{u}$ can be treated similarly.

Note that if ψ is an R–formula it occurs in two succeeding proof–situations, but that does not affect $(\mathfrak{B}, \beta) \models \psi$. ■

§7.2 A Method to Construct "Homogeneous" Structures

For the proofs of our hierarchy theorems we need structures which have a certain property of homogeneity. It is not too difficult to get infinite homogeneous structures. However, we are working with finite structures, and there it is a real problem. Here we use a method that was introduced by Hrushovski [Hru92] to prove the following:

7.4 Hrushovski's Lemma : *Given a (finite) graph $\mathfrak{G}$ and a set of partial isomorphisms of $\mathfrak{G}$ we can find a (finite) graph $\mathfrak{H} \supset \mathfrak{G}$ where all of these partial isomorphisms can be extended to automorphisms.*

Herwig [Her93] extended this to all relational structures. These results were needed in a completely different context, namely to prove the so called *small index property* for certain theories.

We now formulate the basic idea of the method as we use it here.

We let σ be a relational signature, $\mathfrak{A}$ a σ–structure, and $p_1, \ldots, p_l$ partial bijections of A.
Let $\mathtt{P}$ be the free group generated by l elements $\mathtt{p}_1, \ldots, \mathtt{p}_l$. Each element $\mathtt{p} = (\mathtt{p}_{i_1})^{s_1} \ldots (\mathtt{p}_{i_r})^{s_r}$ $(r \geq 1, i_j \leq l, s_j \in \{1, -1\})$ of $\mathtt{P}$ naturally corresponds to a partial bijection $p = (p_{i_1})^{s_1} \ldots (p_{i_r})^{s_r}$ of A. (We write $(p_{i_1})^{s_1} \ldots (p_{i_r})^{s_r}$ for the concatenation of the partial bijections. By $a^{(p_{i_1})^{s_1} \ldots (p_{i_r})^{s_r}}$ we mean $p_{i_r}^{s_r}(\ldots p_{i_1}^{s_1}(a) \ldots)$. This always includes that a is in the domain of this partial bijection.)
We extend this correspondence by associating the unit of $\mathtt{P}$ (the empty word) with the (totally defined) identity mapping on A.

Note that this correspondence is a mapping from $\mathtt{P}$ into the set of all partial bijections of A (which is a semi–group homomorphism).
We are a bit sloppy with this mapping and never refer to it explicitly, but we always let p be the partial bijection corresponding to $\mathtt{p}$.

Let Γ be an arbitrary group and $\rho : \mathtt{P} \longrightarrow \Gamma$ a homomorphism.
We define an equivalence relation $\sim$ on $A \times \Gamma$ by:

$$(a, \gamma) \sim (b, \delta) :\Longleftrightarrow \exists \mathtt{p} \in \mathtt{P} : \delta = \rho(\mathtt{p})\gamma \wedge a = b^{p}$$

In this case we occasionally say $(a, \gamma) \sim (b, \delta)$ *via* $\mathtt{p}$.
It can easily be seen that $\sim$ is indeed an equivalence relation.

We define a σ-structure $\mathfrak{H} = \mathfrak{H}(\mathfrak{A}, p_1, \ldots, p_l, \Gamma, \rho)$ with universe $H = A \times \Gamma/_\sim$ by

$$R^{\mathfrak{H}} b_1 \ldots b_r \;:\Longleftrightarrow \exists \gamma \in \Gamma\ \exists a_1, \ldots, a_r \in A : (a_i, \gamma)/_\sim = b_i\ (1 \leq i \leq r) \wedge R^{\mathfrak{A}} a_1 \ldots a_r$$

for each r–ary $R \in \sigma$.

A *slice* of $\mathfrak{H}$ is a set $A \times \{\gamma\}/_\sim$. Occasionally we say that an element $(a,\gamma)/_\sim \in H$ *belongs to the slice* γ. But note that each element can belong to several slices.

7.5 Lemma : *For each $\gamma \in \Gamma$ the mapping*

$$\pi_\gamma : \mathfrak{A} \to \mathfrak{H} \text{ defined by } \pi_\gamma(a) := (a,\gamma)/_\sim$$

is a homomorphism. (I.e. the slices of $\mathfrak{H}$ are homomorphic images of $\mathfrak{A}$.)

The **proof** is an immediate consequence of the definition of the relations on $\mathfrak{H}$. ■

7.6 Lemma : *For each $\gamma \in \Gamma$ the mapping*

$$f_\gamma : H \to H \text{ defined by } f_\gamma((a,\delta)/_\sim) := (a,\delta\gamma)/_\sim$$

is an automorphism of $\mathfrak{H}$.

Proof: f_γ is well-defined because

$$\begin{aligned}
& (a,\delta)/_\sim = (b,\eta)/_\sim \\
\iff & \exists \mathtt{p} \in \mathtt{P} : \eta = \rho(\mathtt{p})\delta \wedge a = b^p \\
\iff & \exists \mathtt{p} \in \mathtt{P} : \eta\gamma = \rho(\mathtt{p})\delta\gamma \wedge a = b^p \\
\iff & (a,\delta\gamma) \sim (b,\eta\gamma) \\
\iff & f_\gamma((a,\delta)/_\sim) = f_\gamma((b,\eta)/_\sim)
\end{aligned}$$

f_γ is bijective because $(f_\gamma)^{-1} = f_{\gamma^{-1}}$ is also well–defined.

Finally, for r-ary $R \in \sigma$ and $b_1,\ldots,b_r \in H$ we have

$$\begin{aligned}
& R^{\mathfrak{H}} b_1 \ldots b_r \\
\iff & \exists \delta \in \Gamma\ \exists a_1,\ldots,a_r \in A\ \forall j \le r :\ (a_j,\delta)/_\sim = b_j \ \wedge\ R^{\mathfrak{A}} a_1 \ldots a_r \\
\iff & \exists \eta \in \Gamma\ \exists a_1,\ldots,a_r \in A\ \forall j \le r :\ (a_j,\eta)/_\sim = f_\gamma(b_j) \ \wedge\ R^{\mathfrak{A}} a_1 \ldots a_r \\
& (\text{Since } f_\gamma \text{ is bijective.}) \\
\iff & R^{\mathfrak{H}} f_\gamma(b_1) \ldots f_\gamma(b_r) \ .
\end{aligned}$$

■

7.7 Remark : If $\gamma = \rho(\mathtt{p})$ in a sense we can say that f_γ extends p since

$$\pi_{1^\Gamma}(a^p) = (a^p, 1^\Gamma)/_\sim = (a,\gamma)/_\sim = f_\gamma((a,1^\Gamma)/_\sim) = f_\gamma(\pi_{1^\Gamma}(a)) \ .$$

So it seems that we have nearly proved Hrushovski's Lemma. We just have to prove that π_{1^Γ} is an embedding. Unfortunately in general π_{1^Γ} is not even injective, e.g. let Γ be the group $\{1\}$. In fact we are not as near as it seems. $\square$

However, we will stop here following Hrushovski's ideas and turn into another direction. Although the mappings π_γ are not injective in the general case, they are for all structures we are dealing with because Γ and ρ will always be *suitable* for $p_1, \ldots, p_l$ in the following sense:

7.8 Definition : Let $p_1, \ldots, p_l$, Γ and ρ be as above. We call Γ, ρ *suitable* for $p_1, \ldots, p_l$ if

$$\forall \mathtt{p} \in \mathtt{P} \ : (\rho(\mathtt{p}) = 1^\Gamma \Longrightarrow p = id\restriction_{\mathrm{dom}(p)}) \ .$$

$\square$

7.9 Lemma : *Suppose Γ and ρ are suitable for $p_1, \ldots, p_l$. Then for all $\gamma \in \Gamma$ the mapping π_γ is injective.*

Proof: Suppose $(a, \gamma) \sim (b, \gamma) \quad via \quad \mathtt{p}$.
Then

$$\rho(\mathtt{p})\gamma = \gamma \quad \Longrightarrow \quad \rho(\mathtt{p}) = 1^\Gamma \quad \Longrightarrow \quad p = id\restriction_{\mathrm{dom}(p)}$$

But since $(a, \gamma) \sim (b, \gamma)$ *via* $\mathtt{p}$ also implies $a = b^p$ we have $a = b$. $\blacksquare$

An important difference between our construction and Hrushovski's and Herwig's is that we are always going to work with Abelian groups Γ (which we write additively). This makes the resulting structure very homogeneous. We immediately get the following:

7.10 Lemma : *Suppose Γ is an Abelian group. Let $\mathtt{p} \in \mathtt{P}$, $\gamma = \rho(\mathtt{p})$ and $a \in \mathrm{dom}(p)$. Then for all $\delta \in \Gamma$ we have*

$$\pi_\delta(a^p) = (a^p, \delta)/_\sim = f_\gamma((a, \delta)/_\sim) = f_\gamma(\pi_\delta(a)) \ .$$

Proof: $(a^p, \delta)/_\sim = (a, \gamma + \delta)/_\sim = (a, \delta + \gamma)/_\sim = f_\gamma((a, \delta)/_\sim)$. $\blacksquare$

8 Proof of the Main Lemma

§ 8.1 A Weak Version

Before we prove the lemma in its full strength we are going to formulate a weakening whose proof already contains most of the important ideas, but is a bit more transparent.
We will generalize this weak version to logics augmented by Lindström–quantifiers in Section 10. As we will see (cf. Remark 10.9), the generalization is not possible for the stronger (original) main lemma.

For each $k \geq 1$ we let E_k be a $2k$–ary relation symbol.

8.1 Weak Main Lemma : *The transitive closure of* E_k[24] *is not definable in* $\mathbf{s\text{–}PFP}^{k-1}$ *for any* $k \geq 1$.

The proof given here is intended as a first step of the proof of the main lemma, moreover it allows the generalization mentioned above. An easier proof for the weak main lemma based on a proof given in [Gro93] can be found in Appendix B.

It is well–known that the statement holds for $k = 1$. So let us fix a $k \geq 2$.

We start by defining a simple structure $\mathfrak{A} = \mathfrak{A}(k, n)$ for each $n \geq 1$ whereon we apply the construction described in the previous paragraph.
$\mathfrak{A}$ consists of two disjoint E_k-paths of length n, i.e. we let

$$\begin{aligned} A &:= \{1, \ldots, n\} \times \{-1, \ldots, -k, 1, \ldots, k\} \\ E_k^{\mathfrak{A}} &:= \{(I, -1) \ldots (I, -k)(I+1, -1) \ldots (I+1, -k) \mid 1 \leq I < n\} \\ &\quad \cup \{(I, 1) \ldots (I, k)(I+1, 1) \ldots (I+1, k) \mid 1 \leq I < n\} . \end{aligned}$$

To simplify the notation we let $K := \{-1, \ldots, -k, 1, \ldots, k\}$ and define a mapping $æ : K \longrightarrow K$ by $æ(a) = -a$. We extend $æ$ to A by letting $æ(I, a) := (I, -a)$.
We define two more mappings row and col (*column*) on A by

$$\mathrm{row}((I, a)) = I \quad \text{and} \quad \mathrm{col}((I, a)) = |a| .$$

The structure $\mathfrak{B}$ we aim at will consist of n disjoint rows and E_k-edges between k–tuples of succeeding rows. There will be a natural embedding of $\mathfrak{A}$ into $\mathfrak{B}$ and two connected components arising from the two paths in $\mathfrak{A}$. So each element of B naturally belongs to one component, depending on whether it is derived from an element in $\{-1. \ldots, -k\}$ or $\{1. \ldots, k\}$. But for each $(k-1)$–tuple there is an automorphism of $\mathfrak{B}$ mapping it into the "other component" in its rows, but

[24] I.e. the query defined by $[TC_{\overset{k}{x}, \overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{x}, \overset{k}{y}$

leaving all other rows fixed. (Remember that the components in fact consist of k-tuples and not of single elements !)

To achieve this, for each $I \in \{1, \ldots, n\}$ and $a \in \{1, \ldots, k\}$ we define a partial bijection of A by

$$(J, b) \mapsto \begin{cases} (J, b) & \text{if } J \neq I \\ (J, æ(b)) & \text{if } J = I \text{ and } |b| \neq a \\ (\text{undefined} & \text{if } J = I \text{ and } |b| = a) \end{cases} .$$

Let $p_1, \ldots, p_l$ be an enumeration of all these partial bijections, P the free group generated by $\mathsf{p}_1, \ldots, \mathsf{p}_l$ and Γ the (Abelian) group $(\mathbb{Z}_2^l, +)$.
ρ is the extension of the mapping $\mathsf{p}_i \longmapsto (0, \ldots, 0, 1, 0, \ldots, 0)$ (with the 1 at the i-th position) to a homomorphism from P onto Γ.
Let $\mathfrak{B} = \mathfrak{B}(k, n) = \mathfrak{H}(\mathfrak{A}(k, n), p_1, \ldots, p_l, \Gamma, \rho)$.

Note that Γ and ρ are suitable (cf. Definition 7.8) for $p_1, \ldots, p_l$ because for all $i, j \leq l$ we have $p_i = p_i^{-1}$ and $p_i p_j = p_j p_i$. Thus by Lemma 7.9 for every $\gamma \in \Gamma$ the mapping $\pi_\gamma : A \longrightarrow B$ is injective.

Before we start to prove that $\mathfrak{B}$ has the desired properties we introduce some notational conventions:

- We use the variables I, J for the rows.
- We use the variables $a, b, \ldots$ for the elements of K.
- We use the variables $\gamma, \delta, \zeta, \ldots$ for the elements of Γ.
- We write $(I, a, \gamma)/_\sim$ (instead of $((I, a), \gamma)/_\sim$) for the elements of B or use the variables $\boldsymbol{a}, \boldsymbol{b}, \ldots$.

We extend the mappings row and col to B by

$$\begin{aligned} \mathrm{row}((I, a, \gamma)/_\sim) &:= \mathrm{row}((I, a)) = I \\ \mathrm{col}((I, a, \gamma)/_\sim) &:= \mathrm{col}((I, a)) = |a| . \end{aligned}$$

They are well–defined because $p_1, \ldots, p_l$ leave the row and the column of an element fixed.

8.2 Lemma : *The extension of the mapping æ to B defined by*

$$æ(I, a, \gamma)/_\sim := (I, æ(a), \gamma)/_\sim$$

is an automorphism of $\mathfrak{B}$.

For the **proof** just note that the whole construction is absolutely symmetric in $\{1,\ldots,k\}$ and $\{-1,\ldots,-k\}$. ■

Recall that $\Gamma = (\mathbb{Z}_2^l, +)$ and let for each $\gamma \in \Gamma$

$$\mathrm{supp}(\gamma) := \{i \leq l \mid \text{The } i\text{th component of } \gamma \text{ is } 1\}\ .$$

8.3 Lemma : *Let* $\gamma, \delta \in \Gamma$, $I \leq n$, $a, b \in K$, *and* $\mathrm{supp}(\delta - \gamma) = \{i_1, \ldots, i_r\}$. *Then*

$$(I, a, \gamma) \sim (I, b, \delta) \quad \Longleftrightarrow \quad (I, a, \gamma) \sim (I, b, \delta) \quad \textit{via} \quad \mathtt{p}_{i_1} \ldots \mathtt{p}_{i_r}\ .$$

Proof: The direction from the right to the left is trivial. For the other direction note that

$$\begin{aligned}
& (I, a, \gamma) \sim (I, b, \delta) \quad \textit{via} \quad \mathtt{p}_{j_1}^{s_1} \ldots \mathtt{p}_{j_t}^{s_t} \\
\Longrightarrow\ & \{i_1, \ldots, i_r\} = \mathrm{supp}(\delta - \gamma) \subseteq \{j_1, \ldots, j_t\} \\
\Longrightarrow\ & \mathrm{dom}(p_{j_1}^{s_1} \ldots p_{j_t}^{s_t}) \subseteq \mathrm{dom}(p_{i_1} \ldots p_{i_r})\ .
\end{aligned}$$

(Here we use that $\forall i \leq l : p_i = p_i^{-1}$ and $\forall i, j \leq l : p_i p_j = p_j p_i$.)
But $(I, a, \gamma) \sim (I, b, \delta)$ *via* $\mathtt{p}_{j_1}^{s_1} \ldots \mathtt{p}_{j_t}^{s_t}$ also implies

$$(I, b) \in \mathrm{dom}(p_{j_1}^{s_1} \ldots p_{j_t}^{s_t}) \subseteq \mathrm{dom}(p_{i_1} \ldots p_{i_r}).$$

Thus by the suitability of Γ, ρ we have

$$(I, b)^{p_{i_1} \ldots p_{i_r}} = (I, b)^{p_{j_1}^{s_1} \ldots p_{j_t}^{s_t}} = (I, a)$$

hence

$$(I, a, \gamma) \sim (I, b, \delta) \quad \textit{via} \quad \mathtt{p}_{i_1} \ldots \mathtt{p}_{i_r}\ .$$

■

Next we want to prove that $\mathfrak{B}$ has at least two connected components.

8.4 Definition : We say that a tuple $\overset{k}{\boldsymbol{a}} \in B$ belongs to the *right component* if there exist $I \leq n$ and $\gamma \in \Gamma$ such that

$$\forall i \leq k : \boldsymbol{a}_i = (I, i, \gamma)/_\sim\ .$$

We say that $\overset{k}{\boldsymbol{a}}$ belongs to the *left component* if there exist $I \leq n$ and $\gamma \in \Gamma$ such that

$$\forall i \leq k : \boldsymbol{a}_i = (I, -i, \gamma)/_\sim\ .$$

$\square$

8.5 Lemma : *Let $\gamma, \delta \in \Gamma, I \leq n$. Then there exists an $a \in \{1, \dots, k\}$ such that*

$$(I, a, \gamma) \not\sim (I, \ae(a), \delta) \ .$$

Hence no k–tuple belongs to the left and the right component.

Proof: Suppose there exists a $b \in K$ such that

$$(I, b, \gamma) \sim (I, \ae(b), \delta) \quad \textit{via} \quad \mathtt{p} \ .$$

(If no such b exists there is nothing to prove.)
By the last lemma we can assume that

$$\mathtt{p} = \mathtt{p}_{i_1} \dots \mathtt{p}_{i_r} \quad \text{where} \quad \{i_1, \dots, i_r\} = \text{supp}(\delta - \gamma) \ .$$

Then there exists an $i \in \{i_1, \dots, i_r\}$ such that $p_i((I, b)) = (I, \ae(b))$.

By the definition of p_i there exists an $a \in \{1, \dots, k\}$ such that $(I, a), (I, \ae(a)) \notin \text{dom}(p_i)$. Hence

$$\neg\Big((I, a, \gamma) \sim (I, \ae(a), \delta) \quad \textit{via} \quad \mathtt{p}_{i_1} \dots \mathtt{p}_{i_r}\Big)$$

which implies $(I, a, \gamma) \not\sim (I, \ae(a), \delta)$, again by the last lemma. $\blacksquare$

The following theorem summarizes the properties of our structure $\mathfrak{B}$:

8.6 Theorem : *Let $k, n \geq 2$. There exists an $\{E_k\}$–structure $\mathfrak{B} = \mathfrak{B}(k, n)$ with the following properties:*

(1) There exists a mapping row $: B \longrightarrow \{1, \dots, n\}$ *such that $E_k^{\mathfrak{B}} \overset{k}{\boldsymbol{a}} \overset{k}{\boldsymbol{b}}$ implies*

$$\forall i, j \leq k : \ \text{row}(\boldsymbol{a}_i) = \text{row}(\boldsymbol{b}_j) - 1 \ .$$

(2) There exists an automorphism $\ae$ of $\mathfrak{B}$ that is self–inverse and preserves the rows, i.e.

- $\ae^{-1} = \ae$
- $\forall \boldsymbol{b} \in \ B : \text{row}(\ae(\boldsymbol{b})) = \text{row}(\boldsymbol{b})$.

(3) There exist tuples $\overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \in B$ in the first and last row respectively (i.e. $\forall i \leq k : (\text{row}(\boldsymbol{c}_i) = 1 \ \wedge \ \text{row}(\boldsymbol{d}_i) = n)$) such that

$$\mathfrak{B} \models [TC_{\overset{k}{x}, \overset{k}{y}} E_k \overset{k}{x} \overset{k}{y}] \ \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \quad \textit{and} \quad \mathfrak{B} \not\models [TC_{\overset{k}{x}, \overset{k}{y}} E_k \overset{k}{x} \overset{k}{y}] \ \overset{k}{\boldsymbol{c}}, \ae(\overset{k}{\boldsymbol{d}}) \ .$$

(4) For all $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1} \in B$ *there exists an automorphism* f *of* $\mathfrak{B}$ *that is self-inverse, preserves the rows, and switches the component of* $\overset{k-1}{\boldsymbol{b}}$*, but leaves all elements in rows apart from* $\mathrm{row}(\boldsymbol{b}_1), \ldots, \mathrm{row}(\boldsymbol{b}_{k-1})$ *fixed, i.e.*

- $f^{-1} = f$
- $\forall \boldsymbol{b} \in B : \mathrm{row}(f(\boldsymbol{b})) = \mathrm{row}(\boldsymbol{b})$
- $\forall i \leq k-1 : f(\boldsymbol{b}_i) = \text{\ae}(\boldsymbol{b}_i)$
- *For each* $\boldsymbol{a} \in B$ *with* $\forall i \leq k-1 : \mathrm{row}(\boldsymbol{a}) \neq \mathrm{row}(\boldsymbol{b}_i)$ *we have* $f(\boldsymbol{a}) = \boldsymbol{a}$.

Proof: Of course we take the structure $\mathfrak{B}(k,n)$ we have constructed before. Then (1) is immediate.

(2) follows from Lemma 8.2.

For (3), let $\boldsymbol{c}_i = (1, -i, 0)/_\sim$ and $\boldsymbol{d}_i = (n, -i, 0)/_\sim$.
Then $\mathfrak{B} \models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}}$ follows from the fact that π_0 is a homomorphism of $\mathfrak{A}$ into $\mathfrak{B}$ (cf. Lemma 7.5). Recalling the definition of $E_k^{\mathfrak{B}}$, the other statement $\mathfrak{B} \not\models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\boldsymbol{c}}, \text{\ae}(\overset{k}{\boldsymbol{d}})$ is an easy consequence of Lemma 8.5.

To prove (4) suppose that $\boldsymbol{b}_i = (I_i, b_i, \delta_i)/_\sim$ (for $i \leq k-1$).
By the definition of $p_1, \ldots, p_l$ there exists a sequence $i_1, \ldots, i_r \in \{1, \ldots, l\}$ such that for $p = p_{i_1} \ldots p_{i_r}$ we have $p((I,a)) = (I, \text{\ae}(a))$ if $I \in \{I_1, \ldots, I_{k-1}\}$ and $|a| \in \{|b_1|, \ldots, |b_{k-1}|\}$ and $p((I,a)) = (I,a)$ if $I \notin \{I_1, \ldots, I_{k-1}\}$

Let $\gamma = \rho(\mathsf{p})$; by Lemma 7.6 f_γ is an automorphism of $\mathfrak{B}$. By Lemma 7.10 we have for all $i \leq k-1$

$$f_\gamma(\boldsymbol{b}_i) = f_\gamma((I_i, b_i, \delta_i)/_\sim) = ((I_i, b_i)^p, \delta_i)/_\sim = (I_i, \text{\ae}(b_i), \delta_i)/_\sim = \text{\ae}(\boldsymbol{b}_i)$$

and for all $I \notin \{I_1, \ldots, I_{k-1}\}$, $a \in K$, $\delta \in \Gamma$

$$f_\gamma((I, a, \delta)/_\sim) = ((I,a)^p, \delta)/_\sim = (I, a, \delta)/_\sim \; .$$

Clearly, f_γ is self–inverse and preserves the rows. ■

The following lemmata and corollaries will not be needed for the proof of the weak main lemma, but only for a hierarchy result concerning the Hamiltonian path quantifier (cf. Corollary 10.13) at the end of Paragraph 10.3.
In addition, Corollary 8.8 is of its own interest because it completes the picture of our structure $\mathfrak{B}$ in a nice way. However, the reader only interested in our main results may skip the rest of this paragraph and turn to the next.

Let for each $J \leq n$

$$\mathrm{fix}(J) := \{i \leq l \mid \forall a \in K : (J,a)^{p_i} = (J,a)\}.$$

Then clearly for all $i \in \text{fix}(J)$, $a \in K$, and $\gamma \in \Gamma$ we have

$$(J, a, \gamma) \sim (J, a, \rho(\mathtt{p}_i) + \gamma) \ .$$

8.7 Lemma : *Let $I < n, \gamma, \delta \in \Gamma$. Then*

$$E_k^{\mathfrak{B}}(I, 1, \gamma)/_\sim \dots (I, k, \gamma)/_\sim \ (I+1, 1, \delta)/_\sim \dots (I+1, k, \delta)/_\sim \ .$$

Proof: Suppose $\text{supp}(\delta - \gamma) \cap \text{fix}(I) = \{i_1, \dots, i_r\}$ and let $\gamma' = \rho(\mathtt{p}_{i_1} \dots \mathtt{p}_{i_r}) + \gamma$. Then for each $a \in K$

$$(I, a, \gamma)/_\sim = (I, a, \gamma')/_\sim \ .$$

Suppose next that

$$\Big(\text{supp}(\delta - \gamma) \cap \text{fix}(I+1)\Big) \setminus \{i_1, \dots, i_r\} = \{j_1, \dots, j_s\}$$

and let $\delta' = \rho(\mathtt{p}_{j_1} \dots \mathtt{p}_{j_s}) + \delta$. Then for each $a \in K$

$$(I+1, a, \delta)/_\sim = (I+1, a, \delta')/_\sim \ .$$

By the definition of $p_1, \dots, p_l$ each $i \leq l$ belongs to $\text{fix}(I)$ or $\text{fix}(I+1)$ (maybe to both), thus

$$\text{supp}(\delta - \gamma) = \{i_1, \dots, i_r\} \ \dot\cup \ \{j_1, \dots, j_s\} \ .$$

But this implies $\gamma' = \delta'$ and we are done. ■

8.8 Corollary : *The structure $\mathfrak{B}$ consists of exactly two E_k–connected components.*
More precisely we have:

- *Each k–tuple that occurs in an E_k–edge of $\mathfrak{B}$ either belongs to the right or to the left component.*
- *If $\overset{k}{\boldsymbol{a}}$ and $\overset{k}{\boldsymbol{b}}$ belong to the same component and $\text{row}(\overset{k}{\boldsymbol{a}}) < \text{row}(\overset{k}{\boldsymbol{b}})$ then there exists an E_k–path from $\overset{k}{\boldsymbol{a}}$ to $\overset{k}{\boldsymbol{b}}$.*

For each $J \leq n$ we let

$$\text{move}(J) := \{1, \dots, l\} \setminus \text{fix}(J) = \{i \leq l \mid \exists a \in K : (J, a)^{p_i} = \text{æ}((J, a))\}.$$

Note that $|\text{move}(J)| = k$.

8.9 Lemma : *For each $I \le n$ there exists a bijection between* $\mathrm{Pot}(\mathrm{move}(I))$ *and the set*

$$R_I := \{(I,1,\gamma)/_\sim \dots (I,k,\gamma)/_\sim \mid \gamma \in \Gamma\}$$

of k–tuples in row I that belong to the right component.
The analogous statement holds for the set L_I of k–tuples in row I that belong to the left component.

Proof: We define $f : \mathrm{Pot}(\mathrm{move}(I)) \longrightarrow R_I$ by

$$f(\{i_1,\dots,i_r\}) = (I,1,\rho(\mathtt{p}_{i_1}\dots\mathtt{p}_{i_r}))/_\sim \dots (I,k,\rho(\mathtt{p}_{i_1}\dots\mathtt{p}_{i_r}))/_\sim.$$

(For $r=0$ recall that $1^{\mathtt{P}}$ is the empty word and that $\rho(1^{\mathtt{P}}) = 0^\Gamma$.)
Then f is well–defined since Γ is Abelian.
To see that it is surjective let $\gamma \in \Gamma$ and suppose that

$$\mathrm{supp}(\gamma) \setminus \mathrm{move}(I) = \mathrm{supp}(\gamma) \cap \mathrm{fix}(I) = \{i_1,\dots,i_r\}.$$

Let $\gamma' := \rho(\mathtt{p}_{i_1}\dots\mathtt{p}_{i_r}) + \gamma$ and note that

$$(I,a,\gamma')/_\sim = (I,a,\gamma)/_\sim$$

for each $a \in K$. However, since $\mathrm{supp}(\gamma') \subseteq \mathrm{move}(I)$ the tuple

$$(I,1,\gamma')/_\sim \dots (I,k,\gamma')$$

is in the range of f.
So we shall prove that f is injective. Therefore we observe that for all distinct subsets $\{i_1,\dots,i_r\}$, $\{j_1,\dots,j_s\}$ of $\mathrm{move}(I)$ we have

$$\mathrm{dom}(p_{i_1}\dots p_{i_r}) \ne \mathrm{dom}(p_{j_1}\dots p_{j_s}).$$

Suppose now that, say, $a \in \mathrm{dom}(p_{i_1}\dots p_{i_r}) \setminus \mathrm{dom}(p_{j_1}\dots p_{j_s})$.
Then

$$(I,a,\rho(\mathtt{p}_{i_1}\dots\mathtt{p}_{i_r}))/_\sim = (I,a^{p_{i_1}\cdots p_{i_r}},0^\Gamma)/_\sim.$$

On the other hand we have

$$\begin{aligned}
& a \notin \mathrm{dom}(p_{j_1}\dots p_{j_s}) \\
\Longrightarrow\ & \neg\Big((I,a^{p_{i_1}\cdots p_{i_r}},0^\Gamma) \sim (I,a,\rho(\mathtt{p}_{j_1}\dots\mathtt{p}_{j_s})) \quad \textit{via} \quad \mathtt{p}_{j_1}\dots\mathtt{p}_{j_s}\Big) \\
\Longrightarrow\ & (I,a^{p_{i_1}\cdots p_{i_r}},0^\Gamma)/_\sim \ne (I,a,\rho(\mathtt{p}_{j_1}\dots\mathtt{p}_{j_s}))/_\sim \\
& \text{(By Lemma 8.3).}
\end{aligned}$$

Hence $f(\{i_1,\dots,i_r\}) \ne f(\{j_1,\dots,j_s\})$.

The analogous statement for the left component follows from the fact that $\ae$ is an automorphism of $\mathfrak{B}$ switching the components. ■

Note that the E_k–connected components of $\mathfrak{B}$ are exactly the subsets $L := \bigcup_{I=1}^{n} L_I$ and $R := \bigcup_{I=1}^{n} R_I$ of B^k. Thus we get:

8.10 Corollary : *Both E_k–connected components of $\mathfrak{B}$ can be partitioned into 2^k disjoint paths from the first row to the last row.*

Formally, there exist distinct k–tuples $\overset{k}{\boldsymbol{a}}(I,i)$ $(I \le n, i \le 2^k)$ such that

$$L = \{\overset{k}{\boldsymbol{a}}(I,i) \mid I \le n, i \le 2^k\}$$

and for each $I < n, i \le 2^k$ we have $E_k^{\mathfrak{B}}\ \overset{k}{\boldsymbol{a}}(I,i)\ \overset{k}{\boldsymbol{a}}(I+1,i)$.
The tuples $\ae(\overset{k}{\boldsymbol{a}}(I,i))$ $(I \le n, i \le 2^k)$ form the paths for the right component.

Proof: Recall that $|\text{move}(I)| = k$, hence $|L_I| = |\text{Pot}(\text{move}(I))| = 2^k$. We let $\overset{k}{\boldsymbol{a}}(I,1), \ldots, \overset{k}{\boldsymbol{a}}(I,2^k)$ be an enumeration of L_I. Then $E_k^{\mathfrak{B}}\ \overset{k}{\boldsymbol{a}}(I,i)\ \overset{k}{\boldsymbol{a}}(I+1,i)$ holds by Lemma 8.7. ■

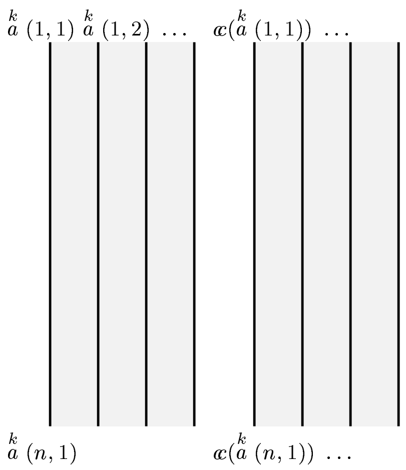

Figure 8.11: The structure $\mathfrak{B}$

§ 8.2 The Game

In this paragraph we are going to prove that the transitive closure of E_k is not $\mathbf{PFP}^{k-1}$ definable on the structures $\mathfrak{B}(k,n)$ we have constructed in the last

paragraph. We do not refer to the exact form of our structures but only to the properties listed in Theorem 8.6.
We keep $k \geq 2$ fixed and augment our signature by constant symbols $\underline{c}_1, \ldots, \underline{c}_k$ and $\underline{d}_1, \ldots, \underline{d}_k$.
For each $n \geq 2$ we expand $\mathfrak{B}(k,n)$ to the structures $\mathfrak{B}_1(k,n) := (\mathfrak{B}(k,n), \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$ and $\mathfrak{B}_2(k,n) := (\mathfrak{B}(k,n), \overset{k}{\boldsymbol{c}}, \text{\ae}(\overset{k}{\boldsymbol{d}}))$ where $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$ are tuples chosen according to 8.6(3).[25]
Then $\mathfrak{B}_1(k,n) \models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}$ and $\mathfrak{B}_2(k,n) \not\models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}$.

On the other hand the classes $\{\mathfrak{B}_1(k,n) \mid n \geq 2\}$ and $\{\mathfrak{B}_2(k,n) \mid n \geq 1\}$ are not distinguishable in $\mathbf{PFP}^{k-1}$.
To see this we let $r \geq 0$, $n = 2^{r+1} + 1$, and $\mathfrak{B} := \mathfrak{B}(k,n)$, $\mathfrak{B}_1 := \mathfrak{B}_1(k,n)$, $\mathfrak{B}_2 := \mathfrak{B}_2(k,n)$.
Clearly we are finished once we have proved that the same PFP^{k-1}–sentences of quantifier rank $\leq r$ hold in $\mathfrak{B}_1$ and $\mathfrak{B}_2$. By Theorem 7.2 it suffices to prove:

8.12 Proposition : *The duplicator has a winning strategy for the* $(k-1)$*–ary* r*–pebble game on* $\mathfrak{B}_1$ *and* $\mathfrak{B}_2$*.*

Proof: Recall that the pairs of corresponding elements in a situation of the game are the pairs of constants and pairs of pebbled elements.

We prove inductively that the duplicator can preserve the following throughout the whole game:
Suppose $r' \leq r$ pairs of pebbles are on the board.
Then there exist $M, N \in \{1, \ldots, n\}$ and an automorphism h of $\mathfrak{B}_1$ such that:

(i) $N - M \geq 2^{r+1-r'}$

(ii) For pairs $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements we have:

$$\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \leq M \quad \text{or} \quad \text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \geq N \ .$$

(iii) h is self–inverse and preserves the rows.

(iv) For each pair $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements the following holds:

- If $\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \leq M$ then $h(\boldsymbol{a}) = \boldsymbol{b}$.
- If $\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \geq N$ then $\text{\ae}(h(\boldsymbol{a})) = \boldsymbol{b}$.

[25] Our notation is explained as follows: $\mathfrak{B}_1(k,n)$ is the structure whose $\{E_k\}$–reduct is $\mathfrak{B}(k,n)$ and where $\underline{c}_i^{\mathfrak{B}_1} = \boldsymbol{c}_i$ and $\underline{d}_i^{\mathfrak{B}_1} = \boldsymbol{d}_i$ for each $i \leq k$. $\mathfrak{B}_2(k,n)$ denotes the same structure, except that $\underline{d}_i^{\mathfrak{B}_2} = \text{\ae}(\boldsymbol{d}_i)$.

Note that this implies that the pairs of corresponding elements form a partial isomorphism from $\mathfrak{B}_1$ to $\mathfrak{B}_2$ since $\ae$ is an automorphism of $\mathfrak{B}$ and we only have edges between succeeding rows of $\mathfrak{B}$.

Clearly (i)–(iv) hold in the beginning with $M = 1, N = n, h = id$ since then the only corresponding elements are the pairs of constants $(\boldsymbol{c}_1, \boldsymbol{c}_1), \ldots, (\boldsymbol{c}_k, \boldsymbol{c}_k)$ in row 1 and $(\boldsymbol{d}_1, \ae(\boldsymbol{d}_1)), \ldots, (\boldsymbol{d}_k, \ae(\boldsymbol{d}_k))$ in row $n = 2^{r+1} + 1$.

So let there be $r' \leq r$ pairs of pebbles on the board and let M, N, h be defined such that (i)–(iv) hold. We are going to define the duplicator's answer to the next move and M', N', h' to preserve (i)–(iv) in the next situation. We have to consider the following moves:

∃-move: The challenger places a pebble on $\boldsymbol{a} \in \mathfrak{B}_1$. Let $I = \mathrm{row}(\boldsymbol{a})$.

Case 1: $I - M \leq N - I$

Let $M' := max\{M, I\}$ and $N' := N$. The duplicator places the corresponding pebble on $\boldsymbol{b} := h(\boldsymbol{a})$.

Case 2: $I - M > N - I$

Let $M' := M$ and $N' := min\{I, N\}$. The duplicator places the corresponding pebble on $\boldsymbol{b} := \ae(h(\boldsymbol{a}))$.

Note that in both cases

$$N' - M' \geq \left\lfloor \frac{N - M}{2} \right\rfloor \geq 2^{r+1-r'-1}$$

so (i) holds. Letting $h' = h$ guarantees (iii) and (iv). Since h preserves the rows we have (ii).

∀-move: Can be treated analogously to the ∃-move.

I-move: Nothing has to be done here. Just note the following:

Say, the depth is increased from $d - 1$ to d. Then all pebbles which have depth $< d$ at the moment cannot be moved until the depth is reduced below d.

R–move: The challenger reduces the depth to $d \geq 1$ and leaves the pairs $(\boldsymbol{a}_1, \boldsymbol{b}_1), \ldots, (\boldsymbol{a}_l, \boldsymbol{b}_l)$ $(l \leq k - 1)$ of depth $\geq d$ pebbled.

All other pebbles that are still on the board (those of depth $< d$) have already been in their position when the last I–move increasing the depth of the game from $d - 1$ to d was made. We look back to that situation; say M'', N'', h'' satisfied (i)–(iv) then. Moreover, we suppose that r'' pairs of pebbles have been on the board then, thus $(r'' + l)$ are on the board now.

By the pigeonhole principle there exist M', N' such that $M'' \leq M' < N' \leq N''$,

$$N' - M' \geq \frac{N'' - M''}{l+1} \geq \frac{2^{r+1-r''}}{l+1} \geq 2^{r+1-(r''+l)},$$

and none of the pairs of corresponding elements that are still on the board are in a row between M' and N'. Note that (i) and (ii) hold for M', N'.

Note that by the induction hypothesis there are also no pairs of corresponding elements in a row between M and N.

But M' and N' have nothing to do with M and N so it may happen that some of the $\boldsymbol{a}_1, \boldsymbol{b}_1, \ldots, \boldsymbol{a}_l, \boldsymbol{b}_l$ are on the "wrong side". In fact the following two cases may cause some trouble:

Case 1: Some of the pairs $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ $(i \leq l)$ are in row between N' and M.

Let $\{\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_m}\}$ be the maximal subset of $\{a_1, \ldots, a_l\}$ with

$$N' \leq \mathrm{row}(\boldsymbol{a}_{i_1}), \ldots, \mathrm{row}(\boldsymbol{a}_{i_m}) \leq M\ .$$

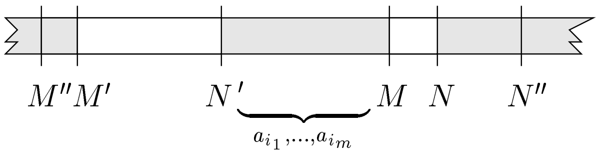

Figure 8.13

Then for each pair of corresponding elements

$$(\boldsymbol{a}, \boldsymbol{b}) \notin \{(\boldsymbol{a}_{i_1}, \boldsymbol{b}_{i_1}), \ldots, (\boldsymbol{a}_{i_m}, \boldsymbol{b}_{i_m})\}$$

we have

$$\forall j \leq m: \ |\mathrm{row}(\boldsymbol{a}) - \mathrm{row}(\boldsymbol{a}_{i_j})| > 1$$

(since $N' - M', M - N \geq 2^1 = 2$).

Since $m \leq l \leq (k-1)$, by Theorem 8.6(4) there exists an automorphism f of $\mathfrak{B}$ such that

- f is self–inverse and preserves the rows.
- $\forall \boldsymbol{a} \in B: \ \mathrm{row}(\boldsymbol{a}) = \mathrm{row}(f(\boldsymbol{a}))$
- $\forall j \leq m: \ f(\boldsymbol{b}_{i_j}) = æ(\boldsymbol{b}_{i_j})$
- If $\boldsymbol{a} \in B$ such that $\forall j \leq m: |\mathrm{row}(\boldsymbol{a}) - \mathrm{row}(\boldsymbol{b}_{i_j})| > 1$ then $f(\boldsymbol{a}) = \boldsymbol{a}$.

(The reader who wonders why we demand $|\text{row}(\boldsymbol{a}) - \text{row}(\boldsymbol{b}_{i_j})| > 1$ instead of $|\text{row}(\boldsymbol{a}) - \text{row}(\boldsymbol{b}_{i_j})| \geq 1$ here is perfectly right. But at least the stronger condition does no wrong (see also Remark 8.14).)

We let $h' := f \circ h$. Since $M'' \geq 1$ and $N'' \leq n$ the automorphism f keeps the first and the last row fixed, in particular the constants $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$, thus h' is in fact an automorphism of $\mathfrak{B}_1$. Moreover, it is clearly self–inverse and preserves the rows hence (iii) holds.

We get (iv) since for each pair of corresponding elements $(\boldsymbol{a}, \boldsymbol{b})$ we have:

- $\text{row}(\boldsymbol{a}) \leq M' \Longrightarrow h'(\boldsymbol{a}) = f(h(\boldsymbol{a})) = h(\boldsymbol{a}) = \boldsymbol{b}$
- $N' \leq \text{row}(\boldsymbol{a}) \leq M \Longrightarrow \exists j \leq m : (\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}_{i_j}, \boldsymbol{b}_{i_j})$
  $\Longrightarrow æ(h'(\boldsymbol{a})) = æ(f(h(\boldsymbol{a}))) = æ(f(\boldsymbol{b})) = æ(æ(\boldsymbol{b})) = \boldsymbol{b}$
- $\text{row}(\boldsymbol{a}) \geq N (\geq N') \Longrightarrow æ(h'(\boldsymbol{a})) = æ(f(h(\boldsymbol{a}))) = æ(h(\boldsymbol{a})) = \boldsymbol{b}$.

Case 2: Some of the $\boldsymbol{a}_i$ $(i \leq l)$ are in row between N and M'.

This case can be treated analogously.

If none of these cases occurs we have no problems anyway; we can just let $h' = h$.

This finishes our proof of Proposition 8.12. ■

8.14 Remark : Note that we did not need Theorem 8.6(4) in its full strength but only the weaker statement:
For each $(k-1)$–tuple $\overset{k-1}{\boldsymbol{a}} \in B$ the structure $\mathfrak{B}$ has an automorphism (which is self–inverse and preserves the rows) mapping $\overset{k-1}{\boldsymbol{a}}$ to $æ(\overset{k-1}{\boldsymbol{a}})$, but leaving rows with distance > 1 from those of $\overset{k-1}{\boldsymbol{a}}$ fixed.
We will need this in the next paragraph to prove the main lemma in its full strength. □

§ 8.3 Reduction to Graphs

We still keep $k \geq 2$ fixed and let E be a binary relation symbol.
To prove the main lemma in its full strength we proceed much like in the proof of the weak main lemma. We are going to construct structures $\mathfrak{C}(k, n)$ and a formula Edge_k that plays the role of E_k ($\text{Edge}_k(\overset{k}{x}, \overset{k}{y})$ simply says that $\overset{k}{x}$ and $\overset{k}{y}$ form a $2k$–clique[26]).

[26] An n–clique in a graph is a subgraph consisting of n vertices and edges between all of them.

Then $\mathfrak{C}(k,n)$ has about the same properties as $\mathfrak{B}(k,n)$, and we can use our game to prove that the transitive closure of Edge_k is not definable on the structures $\mathfrak{C}(k,n)$.

We let

$$\text{Edge}_k(\overset{k}{x},\overset{k}{y}) := \bigwedge_{1\le i,j\le k} Ex_iy_j \wedge \bigwedge_{1\le i\neq j\le k} (Ex_ix_j \wedge Ey_iy_j)\ .$$

For the construction of $\mathfrak{C}(k,n)$ we fix an arbitrary $n \geq 2$.
Instead of the structure $\mathfrak{A}$ defined in Paragraph 8.1 this time we start our construction with a structure $\mathfrak{A}^* = \mathfrak{A}^*(k,n)$ consisting of two disjoint Edge_k–paths of length n.
$\mathfrak{A}^*$ has the same universe $A = \{1,\ldots,n\} \times K$[27] as $\mathfrak{A}$, and $E^{\mathfrak{A}^*}$ is defined by

$$\begin{aligned} E^{\mathfrak{A}^*}(I,a)(J,b) \quad :\Longleftrightarrow \quad & (|I-J| = 1 \,\wedge\, \text{sgn}(a) = \text{sgn}(b)) \\ & \vee (I = J \,\wedge\, \text{sgn}(a) = \text{sgn}(b) \,\wedge\, a \neq b)\ . \end{aligned}$$ [28]

Note that $E^{\mathfrak{A}^*}$ is symmetric and antireflexive.

Unfortunately, if we construct the structure $\mathfrak{C}$ from $\mathfrak{A}^*$ the same way as we constructed $\mathfrak{B}$ from $\mathfrak{A}$ in Paragraph 8.1 we get far too many edges because to have an edge here we do not need $2k$ points (as it was the case for $\mathfrak{B}$) that belong to the same slice of $\mathfrak{C}$, but only 2 points. (Recall that a slice of $\mathfrak{C}$ is a set $A \times \{\gamma\}$.)
So $\mathfrak{C} \models \text{Edge}_k[\overset{k}{\boldsymbol{a}},\overset{k}{\boldsymbol{b}}]$ might hold if all the edges $(\boldsymbol{a}_i,\boldsymbol{b}_j)$ are in different slices and this is not what we want.
To prevent this we use a simple trick:
For each $(k-1)$–tuple $\overset{k-1}{\boldsymbol{a}}$ we force $\mathfrak{C}$ to have an automorphism mapping $\overset{k-1}{\boldsymbol{a}}$ to the other component, but leaving only the rows with distance more than one from those of $\overset{k-1}{\boldsymbol{a}}$ fixed. The proof still works with this weaker condition, but the row inbetween keeps us from having the trouble described above. This is possible as the partial bijections we use for the construction can be chosen to be partial isomorphisms in this case.

Thus here for all $I_1,\ldots,I_{k-1} \in \{1,\ldots,n\}$, $a \in \{1,\ldots,k\}$ consider the partial bijection of $\mathfrak{A}^*$ defined by

$$(J,b) \mapsto \begin{cases} (J,b) & \text{if } \forall i \le k-1 : |J - I_i| > 1 \\ (J,-b) & \text{if } \exists i \le k-1 : J = I_i \text{ and } |b| \neq |a| \\ (\text{undefined} & \text{otherwise}) \end{cases} .$$

Let $p_1,\ldots,p_l$ be an enumeration of all partial bijections defined in this way and P the free group generated by $\mathsf{p}_1,\ldots,\mathsf{p}_l$. Note that the l here is much larger than the respective l in Paragraph 8.1.

[27] Recall that $K = \{-1,\ldots,-k,1,\ldots,k\}$.
[28] $\text{sgn}(a)$ is defined to be 1 if $a \geq 0$ and -1 otherwise.

Note also that for each $\mathtt{p} \in \mathtt{P}$ the corresponding partial bijection p is a partial isomorphism.

Let Γ be the (Abelian) group $(\mathbb{Z}_2^l, +)$ and ρ the extension of the mapping

$$\mathtt{p}_i \longmapsto (0,\ldots,0,1,0,\ldots,0) \quad \text{(with the 1 at the } i\text{th position)}$$

to a homomorphism from P onto Γ. Note that Γ and ρ are suitable for $p_1, \ldots, p_l$. Let $\mathfrak{C} := \mathfrak{C}(k,n) := \mathfrak{H}(\mathfrak{A}^*(k,n), p_1, \ldots, p_l, \Gamma, \rho)$.

We define the mappings row, col and $\mathrm{æ}$ as in Paragraph 8.1 by

$$\begin{aligned} \mathrm{row}((I,a,\gamma)/_\sim) &:= I \\ \mathrm{col}((I,a,\gamma)/_\sim) &:= |a| \\ \mathrm{æ}((I,a,\gamma)/_\sim) &:= (I,-a,\gamma)/_\sim \ . \end{aligned}$$

They are well–defined for the same reason as they were there; also $\mathrm{æ}$ is an automorphism of $\mathfrak{C}$.

To see that $\mathfrak{C}$ is a graph we use the following lemma which implies that $E^{\mathfrak{C}}$ is antireflexive. Clearly the relation $E^{\mathfrak{C}}$ is symmetric because $E^{\mathfrak{A}^*}$ is.

8.15 Lemma : *Let* $\boldsymbol{a}, \boldsymbol{b} \in C$ *such that* $E^{\mathfrak{C}}\boldsymbol{ab}$. *Then* $\mathrm{col}(\boldsymbol{a}) \neq \mathrm{col}(\boldsymbol{b})$ *or* $\mathrm{row}(\boldsymbol{a}) \neq \mathrm{row}(\boldsymbol{b})$.

Proof: Suppose $\mathrm{row}(\boldsymbol{a}) = \mathrm{row}(\boldsymbol{b}) = I$.
$E^{\mathfrak{C}}\boldsymbol{ab}$ implies that there exists a $\zeta \in \Gamma$ and $a', b' \in K$ such that $\boldsymbol{a} = (I, a', \zeta)/_\sim$, $\boldsymbol{b} = (I, b', \zeta)/_\sim$, and $E^{\mathfrak{A}^*}(I,a')(I,b')$. By the definition of $E^{\mathfrak{A}^*}$ this means $|a'| \neq |b'|$, thus $\mathrm{col}(\boldsymbol{a}) \neq \mathrm{col}(\boldsymbol{b})$. ■

For each $\gamma \in \Gamma$ we define $\mathrm{supp}(\gamma)$ in the same way as we did before and obtain the analogue of Lemma 8.3:

8.16 Lemma : *Let* $\gamma, \delta \in \Gamma, I \leq n$ *and* $\mathrm{supp}(\delta - \gamma) = \{i_1, \ldots, i_r\}$. *Then*

$$(I,a,\gamma) \sim (I,b,\delta) \quad \Longleftrightarrow \quad (I,a,\gamma) \sim (I,b,\delta) \quad \textit{via} \quad \mathtt{p}_{i_1} \cdots \mathtt{p}_{i_r} \ .$$

The **proof** of Lemma 8.3 goes through here. ■

8.17 Lemma : *Let* $\gamma, \delta \in \Gamma, I \leq n$. *Then there exists an* $a \in \{1, \ldots, k\}$ *such that*

$$(I,a,\gamma) \not\sim (I, \mathrm{æ}(a), \delta) \ .$$

We just copy the **proof** of Lemma 8.5. ■

We have to define the components slightly differently here:

8.18 Definition : We say that a tuple $\overset{k}{\boldsymbol{a}} \in C$ belongs to the *right (left) component* if there exists a permutation π of $\{1, \dots, k\}$, an $I \leq n$, and a $\gamma \in \Gamma$ such that

$$\forall i \leq k : \ \boldsymbol{a}_{\pi(i)} = (I, (-)i, \gamma)/_{\sim} \ .$$

□

As an immediate consequence of the previous lemma we have:

8.19 Corollary : *No tuple belongs to the right and the left component.*

The difficulty is to prove that there are no Edge_ks between the components. This is done in the following three lemmata.

8.20 Lemma : *Let $I, J \leq n$, $a, b \in K$, $\gamma, \delta \in \Gamma$ such that*

$$E^{\mathfrak{C}}(I, a, \gamma)/_{\sim}(J, b, \delta)/_{\sim} \ .$$

Then there exist $\hat{a}, \hat{b} \in K, \eta \in \Gamma$ such that

(i) $(I, a, \gamma) \sim (I, \hat{a}, \eta)$

(ii) $(J, b, \delta) \sim (J, \hat{b}, \eta)$

(iii) $E^{\mathfrak{A}^*}(I, \hat{a})(J, \hat{b})$

Moreover, we can assume that η is chosen in a way that

(iv) $\text{supp}(\eta - \gamma) \,\dot\cup\, \text{supp}(\eta - \delta) = \text{supp}(\delta - \gamma)$

Proof: We only have to prove that η can be chosen according to (iv), the rest follows immediately from the definition of $E^{\mathfrak{C}}$. Chose $\eta \in \Gamma, \hat{a}, \hat{b} \in K$ such that (i)–(iii) hold.
Note that

$$(\delta - \gamma) = (\eta - \gamma) + (\eta - \delta) \qquad (*)$$

(in the group $\Gamma = (\mathbb{Z}_2^l, +)$). Thus

$$\text{supp}(\delta - \gamma) \subseteq \text{supp}(\eta - \gamma) \cup \text{supp}(\eta - \delta) \ .$$

For each $i \in \operatorname{supp}(\eta-\gamma)\setminus\operatorname{supp}(\delta-\gamma)$ we have, by $(*)$, $i \in \operatorname{supp}(\eta-\delta)\setminus\operatorname{supp}(\delta-\gamma)$. So we suppose that $\operatorname{supp}(\eta-\gamma)\setminus\operatorname{supp}(\delta-\gamma) = \{i_1,\dots,i_r\}$ and let

$$\eta' = \rho(\mathtt{p}_{i_1}) + \ldots + \rho(\mathtt{p}_{i_r}) + \eta\ .$$

We still have

$$(\delta-\gamma) = (\eta'-\gamma) + (\eta'-\delta)$$

and in addition

$$\operatorname{supp}(\delta-\gamma) = \operatorname{supp}(\eta'-\gamma) \cup \operatorname{supp}(\eta'-\delta)\ .$$

Suppose now that there exists an

$$i \in \operatorname{supp}(\eta'-\gamma) \cap \operatorname{supp}(\eta'-\delta).$$

Then

$$i \notin \operatorname{supp}((\eta'-\gamma)+(\eta'-\delta)) = \operatorname{supp}(\delta-\gamma)$$

which is a contradiction. Thus (iv) holds for η' instead of η.

To see (i) (for η') note that $\{i_1,\dots,i_r\} \subseteq \operatorname{supp}(\eta-\gamma)$, thus by (i) (for η) we have $(I,a),(I,-a) \in \operatorname{dom}(p_{i_1}\dots i_{i_r})$. Select $\hat{\hat{a}}$ such that $(I,\hat{\hat{a}}) = (I,\hat{a})^{p_{i_1}\cdots p_{i_r}}$, then (i) holds for η' (instead of η) and $\hat{\hat{a}}$ (instead of $\hat{a}$). Analogously it can be seen that (ii) holds for η' and $\hat{\hat{b}}$ (where $(J,\hat{\hat{b}}) = (J,\hat{b})^{p_{i_1}\cdots p_{i_r}}$).

To see (iii) we just recall that $p_{i_1}\dots p_{i_r}$ is a partial isomorphism. ∎

8.21 Lemma : *Let* $\overset{k}{\boldsymbol{a}},\overset{k}{\boldsymbol{b}} \in C$ *such that* $\mathfrak{C} \models \mathrm{EDGE}_k[\overset{k}{\boldsymbol{a}},\overset{k}{\boldsymbol{b}}]$ *and let* $\overset{k}{\boldsymbol{a}}$ *belong to the right component.*

Then there exist tuples $\overset{k}{\boldsymbol{a}}{}',\overset{k}{\boldsymbol{b}}{}'$ *such that*

- $\mathfrak{C} \models \mathrm{EDGE}_k[\overset{k}{\boldsymbol{a}}{}',\overset{k}{\boldsymbol{b}}{}']$
- $\overset{k}{\boldsymbol{b}}$ *belongs to the right component if and only if* $\overset{k}{\boldsymbol{b}}{}'$ *belongs to the right component.*
- *There exist* $I \le n$, $J \in \{I+1, I-1\}$, $b_i \in \{i,-i\}$ *(for each* $i \le k$*), and* $\gamma,\delta_1,\dots,\delta_k \in \Gamma$ *such that*

  $$\boldsymbol{a}'_i = (I,i,\gamma)/_\sim \quad \textit{and} \quad \boldsymbol{b}'_i = (J,b_i,\delta_i)/_\sim$$

  for all $i \le k$.

  (Note that in particular $\overset{k}{\boldsymbol{a}}{}'$ *also belongs to the right component.)*

A similar statement holds for $\overset{k}{\boldsymbol{a}}$ *belonging to the left component.*

Proof: Select first $\gamma \in \Gamma$, $I \leq n$, and a permutation π of $\{1, \ldots, k\}$ such that $\boldsymbol{a}'_i := \boldsymbol{a}_{\pi(i)} = (I, i, \gamma)/_\sim$ (this is possible since $\overset{k}{\boldsymbol{a}}$ belongs to the right component).

The elements $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k$ form a $2k$–clique (because $\mathfrak{C} \models \mathrm{EDGE}_k[\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}}]$). Thus by Lemma 8.15, the $\boldsymbol{b}_i$ either belong to row $I+1$ or $I-1$ (because otherwise there would be an edge between two elements of the same column of row I). They all belong to the same row J because there are edges between them. Again by Lemma 8.15 this means that they all belong to distinct columns.

So we can select a permutation ρ of $\{1, \ldots, k\}$ and $b_i \in \{i, -i\}$, $\delta_i \in \Gamma$ (for each $i \leq k$) such that $\boldsymbol{b}'_i := \boldsymbol{b}_{\rho(i)} = (J, b_i, \delta_i)/_\sim$.
Clearly $\overset{k}{\boldsymbol{b}}$ belongs to the right component if and only if $\overset{k}{\boldsymbol{b}}{}'$ belongs to the right component, and since

$$\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_k\} = \{\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_k, \boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_k\}$$

the elements $\boldsymbol{a}'_1, \ldots, \boldsymbol{a}'_k, \boldsymbol{b}'_1, \ldots, \boldsymbol{b}'_k$ form a $2k$–clique thus $\mathfrak{C} \models \mathrm{EDGE}_k[\overset{k}{\boldsymbol{a}}{}', \overset{k}{\boldsymbol{b}}{}']$. ■

8.22 Lemma : *Let* $\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}} \in C$ *such that* $\mathfrak{C} \models \mathrm{EDGE}_k[\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}}]$.
Then $\overset{k}{\boldsymbol{a}}$ *belongs to the right component (left component) if and only if* $\overset{k}{\boldsymbol{b}}$ *belongs to the right component (left component respectively).*

Proof: We only prove that if $\overset{k}{\boldsymbol{a}}$ belongs to the right component then $\overset{k}{\boldsymbol{b}}$ also belongs to the right component, the rest follows from the facts that $\mathfrak{x}$ is an automorphism of $\mathfrak{B}$ which switches the components and that EDGE_k is symmetric.

By the last lemma we can assume without loss of generality that $\overset{k}{\boldsymbol{a}}$ and $\overset{k}{\boldsymbol{b}}$ are of the form

$$\boldsymbol{a}_i = (I, i, \gamma)/_\sim \quad \text{and} \quad \boldsymbol{b}_i = (J, b_i, \delta_i)/_\sim \quad (\text{for each } i \leq k)$$

where $I \leq n$, $J \in \{I+1, I-1\}$, $b_i \in \{i, -i\}$, and $\gamma, \delta_1, \ldots, \delta_k \in \Gamma$.

We only consider the case $J = I + 1$, the other possibility $J = I - 1$ can be treated analogously.

During this proof we deviate from our usual notation and use $x, x_1, \ldots, y, y_1, \ldots$ to denote the indices of the $p_1, \ldots, p_l$.
For each $J \leq n$ we let

$$\mathrm{fix}(J) := \{x \leq l \mid \forall a \in K : p_x((J, a)) = (J, a)\}\ .$$

Note that for all $x \in \operatorname{fix}(J)$, $a \in K$ and $\zeta \in \Gamma$ we have

$$(J, a, \zeta) \sim (J, a, \rho(\mathtt{p}_x) + \zeta)\ . \tag{$*$}$$

Furthermore, we let

$$\operatorname{undef}(J) := \{x \leq l \mid \forall a \in K : (J, a) \notin \operatorname{dom}(p_x)\}\ .$$

By the definition of the p_x we have

$$\forall J < n : \operatorname{fix}(J) \setminus \operatorname{fix}(J+1) \subseteq \operatorname{undef}(J+1)\ .$$

In a FIRST STEP we want to show that we can assume without loss of generality that

$$(a) \quad \forall i \leq k : \operatorname{supp}(\delta_i - \gamma) \cap \operatorname{fix}(I+1) = \emptyset$$

and

$$(b) \quad \forall i \leq k : \operatorname{supp}(\delta_i - \gamma) \cap \operatorname{fix}(I) = \emptyset\ .$$

PROOF: To see (a) suppose that for some $i \leq k$

$$\operatorname{supp}(\delta_i - \gamma) \cap \operatorname{fix}(I+1) = \{x_1, \ldots, x_r\}\ .$$

Let $\delta_i' := \rho(\mathtt{p}_{x_1}) + \ldots + \rho(\mathtt{p}_{x_r}) + \delta_i$. Then $\operatorname{supp}(\delta_i' - \gamma) \cap \operatorname{fix}(I+1) = \emptyset$ and $(I+1, b_i, \delta_i) \sim (I+1, b_i, \delta_i')$ by $(*)$. Thus we could have taken δ_i' instead of δ_i to guarantee (a).

To prove (b) note first that there cannot exist $i, j \leq k$ such that

$$\exists x \leq l : x \in \operatorname{supp}(\delta_i - \delta_j) \cap \operatorname{undef}(I+1)\ .$$

Otherwise for each $\delta \in \Gamma$ we would have $x \in \operatorname{supp}(\delta - \delta_i)$ or $x \in \operatorname{supp}(\delta - \delta_j)$. The former would imply (since $x \in \operatorname{undef}(I+1)$) that there exists no $\hat{b}_i$ such that $(I+1, b_i, \delta_i) \sim (I+1, \hat{b}_i, \delta)$, the latter that there exists no $\hat{b}_j$ such that $(I+1, b_j, \delta_j) \sim (I+1, \hat{b}_j, \delta)$. This contradicts $E^{\mathfrak{C}} \boldsymbol{b}_i \boldsymbol{b}_j$.

Together with (a) and since $\operatorname{fix}(I) \setminus \operatorname{fix}(I+1) \subseteq \operatorname{undef}(I+1)$ this implies

$$\forall i, j \leq k : \operatorname{supp}(\delta_i - \gamma) \cap \operatorname{fix}(I) = \operatorname{supp}(\delta_j - \gamma) \cap \operatorname{fix}(I).$$

Suppose $\operatorname{supp}(\delta_i - \gamma) \cap \operatorname{fix}(I) = \{y_1, \ldots, y_s\}$ (for some, hence for all $i \leq k$).
Then, as in the proof of (a), we let $\gamma' := \rho(\mathtt{p}_{y_1}) + \ldots + \rho(\mathtt{p}_{y_s}) + \gamma$ and see that for all $i \leq k$ we have $\operatorname{supp}(\delta_i - \gamma') \cap \operatorname{fix}(I) = \emptyset$ and $(I, i, \gamma) \sim (I, i, \gamma')$. Taking γ' instead of γ proves (b).

SECOND STEP: For each $i \leq k$ there exists an $\hat{\imath} \in \{i, -i\}$ such that $\boldsymbol{b}_i = (I+1, \hat{\imath}, \gamma)/_{\sim}$.

PROOF: Note that for each $x \in \operatorname{supp}(\delta_i - \gamma)$ there exists a $j \in \{1, \ldots, k\}$ such that $(I, j), (I, -j) \notin \operatorname{dom}(p_x)$ (since $x \notin \operatorname{fix}(I)$).
Because of $E^{\mathfrak{C}} \boldsymbol{a}_j \boldsymbol{b}_i$, by Lemma 8.20 there exists an $\eta \in \Gamma$ and $\hat{\jmath} \in \{j, -j\}, \hat{b}_i \in \{i, -i\}$ such that

(i) $(I, j, \gamma) \sim (I, \hat{\jmath}, \eta)$

(ii) $(I+1, b_i, \delta_i) \sim (I+1, \hat{b}_i, \eta)$

(iii) $E^{\mathfrak{A}^*}(I, \hat{\jmath})(I+1, \hat{b}_i)$

(iv) $\operatorname{supp}(\eta - \gamma) \,\dot{\cup}\, \operatorname{supp}(\eta - \delta_i) = \operatorname{supp}(\delta_i - \gamma)$.

Since $(I, j), (I, -j) \notin \operatorname{dom}(p_x)$ we have $x \notin \operatorname{supp}(\eta - \gamma)$ (by (i)) hence $x \in \operatorname{supp}(\eta - \delta_i)$ (because $x \in \operatorname{supp}(\delta_i - \gamma)$). Thus $(I+1, i), (I+1, -i) \in \operatorname{dom}(p_x)$ (by (ii)).

Since $x \in \operatorname{supp}(\delta_i - \gamma)$ was arbitrary this holds for all $x \in \operatorname{supp}(\delta_i - \gamma)$.

Suppose $\operatorname{supp}(\delta_i - \gamma) = \{x_1, \ldots, x_r\}$ and let $p = p_{x_1} \ldots p_{x_r}$. Then $\mathsf{p} \in \rho^{-1}(\delta_i - \gamma)$; and we have just seen that $(I+1, b_i) \in \operatorname{dom}(p)$.
Thus

$$(I+1, b_i, \delta_i) \sim ((I+1, b_i)^p, \gamma)$$

which proves the second step.

We are finished once we have proved:

Third Step: $(I+1, b_i)^p = (I+1, i)$

Let us look back to the beginning of the second step and consider $\eta, \hat{\jmath}, \hat{b}_i$ with properties (i)–(iv) defined there.

Recall that $\operatorname{supp}(\delta_i - \gamma) = \{x_1, \ldots, x_r\}$ and assume (without loss of generality) that $x_1, \ldots, x_r$ are ordered in a way that there exists an $s \leq r$ such that

$$\operatorname{supp}(\eta - \delta_i) = \{x_1, \ldots, x_s\} \quad \text{and} \quad \operatorname{supp}(\eta - \gamma) = \{x_{s+1}, \ldots, x_r\}\,.$$

We have seen in the second step that $p_{x_{s+1}} \ldots p_{x_r}$ is defined for both $(I, \hat{\jmath})$ and $(I+1, \hat{b}_i)$. Since it is a partial isomorphism we have

$$E^{\mathfrak{A}^*}(I, \hat{\jmath})^{p_{x_{s+1}} \cdots p_{x_r}}(I+1, \hat{b}_i)^{p_{x_{s+1}} \cdots p_{x_r}}$$

hence $(I+1, \hat{b}_i)^{p_{x_{s+1}} \cdots p_{x_r}} = (I+1, i)$ (since $(I, \hat{\jmath})^{p_{x_{s+1}} \cdots p_{x_r}} = (I, j)$).
On the other hand we have

$$(I+1, b_i)^{p_{x_1} \cdots p_{x_s}} = (I+1, \hat{b}_i)$$

thus $(I+1, b_i)^p = (I+1, i)$. ■

8.23 Theorem : *Let $k, n \geq 2$. Then there exists a graph $\mathfrak{C} = \mathfrak{C}(k, n)$ and a formula* $\text{Edge}_k(\overset{k}{x}, \overset{k}{y})$ *(which is a conjunction of atomic formulae) such that:*

(1) There exists a mapping row : $C \longrightarrow \{1, \ldots, n\}$ *such that*

$$\forall \boldsymbol{a}, \boldsymbol{b} \in C : (E^{\mathfrak{C}} \boldsymbol{ab} \Longrightarrow \mathrm{row}(\boldsymbol{b}) - \mathrm{row}(\boldsymbol{a}) \leq 1) .$$

(2) There exists an automorphism æ *of* $\mathfrak{C}$ *that is self–inverse and preserves the rows.*

(3) There exist tuples $\overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \in C$ *in the first and last row respectively such that*

$$\mathfrak{C} \models [TC_{\overset{k}{x},\overset{k}{y}} \text{E}\textsc{dge}_k(\overset{k}{x}, \overset{k}{y})]\ \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \quad \textit{and} \quad \mathfrak{C} \not\models [TC_{\overset{k}{x},\overset{k}{y}} \text{E}\textsc{dge}_k(\overset{k}{x}, \overset{k}{y})]\ \overset{k}{\boldsymbol{c}}, \text{æ}(\overset{k}{\boldsymbol{d}}) .$$

(4) For all $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1} \in C$ *there exists an automorphism* f *of* $\mathfrak{C}$ *that is self-inverse, preserves the rows, and maps* $\overset{k-1}{\boldsymbol{b}}$ *to* $\text{æ}(\overset{k-1}{\boldsymbol{b}})$*, but leaves all elements in rows of distance* > 1 *from* $\mathrm{row}(\boldsymbol{b}_1), \ldots, \mathrm{row}(\boldsymbol{b}_{k-1})$ *fixed.*

The **proof** is analogous to the proof of Theorem 8.6 using the lemmata proved above. ■

To finish the proof of the main lemma we have to show that there exists no PFP^{k-1}–formula that is equivalent to

$$\chi = [TC_{\overset{k}{x},\overset{k}{y}} \text{E}\textsc{dge}_k(\overset{k}{x}, \overset{k}{y})]\ \overset{k}{u}, \overset{k}{v}$$

on the class $\{\mathfrak{C}(k, n) \mid n \geq 2\}$. This can be done with our game, proceeding in a manner completely analogous to the proof in the previous paragraph. Note that property (4) of structure $\mathfrak{C}$ is weaker than the respective property of structure $\mathfrak{B}$. But in fact we have already taken this into account, just recall Remark 8.14. ■

§ 8.4 Eliminating the Constants

In the introduction of Part II (Section 6) it was promised to explain how we can extend our structures in order to make the constant–tuples $\overset{k}{\underline{c}}$ and $\overset{k}{\underline{d}}$ definable. This is rather simple in most cases, only for pure datalog there occur some difficulties.

We will first sketch how to make the tuples definable by existential first–order formulae which suffices for all results except from those incorporating pure datalog.
Then we turn to the more involved question how to make them definable by existential first–order formulae without any negation symbols (hence by pure datalog programs).

Consider for example the structure $\mathfrak{C}_1 = (\mathfrak{C}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$. We define a new structure $\mathfrak{C}_1^*$ that still has all the properties we need, but instead of the constants definable elements.

Remember that by Lemma 8.15 in $\mathfrak{C} = \mathfrak{C}(k,n)$ there are no edges between two elements of the same row and column. Since there are only edges between elements of the same row or neighboured rows this implies that $\mathfrak{C}$ contains no cliques of size $\geq 2k+1$.
We augment $\mathfrak{C}$ by cliques of size $(2k+3), (2k+4), \ldots, (4k+2)$ (clique–sizes $(2k+1)$ and $(2k+2)$ are reserved for the next section) and add edges between:

- $\boldsymbol{c}_1$ and a vertex of the $(2k+3)$–clique
- $\boldsymbol{c}_2$ and a vertex of the $(2k+4)$–clique

$\vdots$

- $\boldsymbol{c}_k$ and a vertex of the $(3k+2)$–clique
- $\boldsymbol{d}_1$ and a vertex of the $(3k+3)$–clique

$\vdots$

- $\boldsymbol{d}_k$ and a vertex of the $(4k+2)$–clique.

This suffices to make $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$ definable: $\boldsymbol{d}_k$ is the unique vertex that is connected with a $(4k+2)$–clique in which it is not contained, $\boldsymbol{d}_{k-1}$ is the unique vertex distinct from $\boldsymbol{d}_k$ that is connected with a $(4k+1)$–clique in which it is not contained, and so on.
This can easily be formulated in existential first–order logic.

However, if we have no negation symbols available it is not so easy. For example, we cannot say that $\overset{k}{\boldsymbol{d}}$ is not contained in the $(4k+2)$–clique or that $\boldsymbol{d}_{k-1}$ is distinct from $\boldsymbol{d}_k$ without negation symbols.

For a graph $\mathfrak{G}$ with universe $G = \{g_1, \ldots, g_n\}$ we let

$$\Phi_{\mathfrak{G}} := \bigwedge \{Ex_i x_j \mid E^{\mathfrak{G}} g_i g_j\}.$$

Observe that for each graph $\mathfrak{H}$ we have $\mathfrak{H} \models \exists x_1, \ldots, x_n \Phi_{\mathfrak{G}}$ if and only if there exists a homomorphism from $\mathfrak{G}$ to $\mathfrak{H}$.
The essential step is the following lemma:

8.24 Lemma : *There exist graphs $\mathfrak{G}_1, \ldots, \mathfrak{G}_{2k}$ such that for all $i \neq j \leq 2k$ we have:*

- *There exists no homomorphism from $\mathfrak{G}_i$ to $\mathfrak{G}_j$.*

- $\mathfrak{G}_i$ *is connected.*
- *Each vertex of* $\mathfrak{G}_i$ *is contained in a* $(2k+3)$*–clique.*

Once we have seen this the rest is easy: Let $\tilde{\mathfrak{C}}_1^*$ be the disjoint union of $\mathfrak{C}_1$, $\mathfrak{G}_1$, …, $\mathfrak{G}_{2k}$ with additional edges between

- $\boldsymbol{c}_1$ and a vertex of $\mathfrak{G}_1$
- $\boldsymbol{c}_2$ and a vertex of $\mathfrak{G}_2$

  $\vdots$
- $\boldsymbol{c}_k$ and a vertex of $\mathfrak{G}_k$
- $\boldsymbol{d}_1$ and a vertex of $\mathfrak{G}_{k+1}$

  $\vdots$
- $\boldsymbol{d}_k$ and a vertex of $\mathfrak{G}_{2k}$.

For each $i \leq 2k$ let

$$\Phi_i(x) := \exists x_1, \ldots, x_{n_i}(\Phi_{\mathfrak{G}_i} \wedge \bigvee_{i=1}^{n_i} Exx_i)$$

(where n_i is the size of $\mathfrak{G}_i$). Φ_i is an existential first–order formula that says

x is connected with a homomorphic image of $\mathfrak{G}_i$.

Then obviously we have $\tilde{\mathfrak{C}}_1^* \models \Phi_i[\boldsymbol{c}_i]$ (for each $i \leq k$). Moreover, there is no vertex $\boldsymbol{a} \in C \setminus \{\boldsymbol{c}_i\}$ [29] such that $\tilde{\mathfrak{C}}_1^* \models \Phi_i[\boldsymbol{a}]$ for the following reasons:

- There is no homomorphic image of $\mathfrak{G}_i$ in $\tilde{\mathfrak{C}}_1^*$ that contains vertices of $\mathfrak{C}_1$ because each vertex of $\mathfrak{G}_i$ is part of a $(2k+3)$–clique.
- There is no homomorphic image of $\mathfrak{G}_i$ in $\tilde{\mathfrak{C}}_1^*$ that contains vertices of distinct graphs $\mathfrak{G}_j$, $\mathfrak{G}_{j'}$ since $\mathfrak{G}_i$ is connected.
- There is no homomorphic image of $\mathfrak{G}_i$ in $\tilde{\mathfrak{C}}_1^*$ that is contained in $\mathfrak{G}_j$ for a $j \neq i$ since there is no homomorphism from $\mathfrak{G}_i$ to $\mathfrak{G}_j$.

[29] Recall that C denotes the universe of structure $\mathfrak{C}$ hence of the structure $\mathfrak{C}_1$ we started with.

Hence all homomorphic images of $\mathfrak{G}_i$ in $\tilde{\mathfrak{C}}_1^*$ are subgraphs of $\mathfrak{G}_1$ and $\boldsymbol{c}_i$ is the only vertex in C that is connected with $\mathfrak{G}_i$.
Similarly it can be seen that $\boldsymbol{d}_i$ is the unique vertex of C such that $\tilde{\mathfrak{C}}_1^* \models \Phi_{k+i}[\boldsymbol{d}_i]$ (for each $i \leq k$).

Hence $\overset{k}{\boldsymbol{c}}$ is the unique tuple in $\tilde{\mathfrak{C}}_1^*$ that satisfies

$$\mathrm{C}(\overset{k}{x}) := \bigwedge_{i=1}^{k} \Phi_i(x_i) \wedge \bigwedge_{i \neq j \leq k} E x_i x_j$$

and $\overset{k}{\boldsymbol{d}}$ is the unique tuple in $\tilde{\mathfrak{C}}_1^*$ that satisfies

$$\mathrm{D}(\overset{k}{x}) := \bigwedge_{i=1}^{k} \Phi_{k+i}(x_i) \wedge \bigwedge_{i \neq j \leq k} E x_i x_j \; .$$

Similarly we can define a structure $\tilde{\mathfrak{C}}_2^*$ where $\overset{k}{\boldsymbol{c}}$ and $\alpha(\overset{k}{\boldsymbol{d}})$ are the unique tuples satisfying $\mathrm{C}(\overset{k}{x})$ and $\mathrm{D}(\overset{k}{x})$ respectively.

On the first sight Lemma 8.24 looks rather simple. However, it is not. The reader may only try to find two graphs $\mathfrak{G}$ and $\mathfrak{H}$ such that there exists neither a homomorphism from $\mathfrak{G}$ to $\mathfrak{H}$ nor a homorphism form $\mathfrak{H}$ to $\mathfrak{G}$.

To prove the lemma we make use of a theorem of Erdös [Erd59].
The *girth* of a graph $\mathfrak{G}$ is the size of the smallest circuit. The *chromatic number* $\mathrm{chr}(\mathfrak{G})$ is the minimal number of colours that is needed to colour the vertices of $\mathfrak{G}$ in a way that adjacent vertices have distinct colours.

8.25 Theorem ([Erd59]) : *For each $l, m \geq 1$ there exists a graph of girth $> l$ and chromatic number $> m$.*

A simple proof of this theorem can be found in [AS92] (on page 35).

Observe that if there exists a homomorphism from a graph $\mathfrak{G}$ to a graph $\mathfrak{H}$ we have $\mathrm{chr}(\mathfrak{H}) \geq \mathrm{chr}(\mathfrak{G})$ because each colouring of $\mathfrak{H}$ gives rise to a colouring of $\mathfrak{G}$ with the same colours
In particular this implies that if a graph $\mathfrak{G}$ has a chromatic number $> m$ there exists no homomorphism from $\mathfrak{G}$ to a graph of size $\leq m$.

8.26 Lemma : *For each $r \geq 1$ there exist graphs $\tilde{\mathfrak{G}}_1, \ldots, \tilde{\mathfrak{G}}_r$ such that for all $i \neq j \leq r$ we have:*

- *There exists no homomorphism from $\tilde{\mathfrak{G}}_i$ to $\tilde{\mathfrak{G}}_j$.*
- *$\tilde{\mathfrak{G}}_i$ is connected.*
- $\mathrm{girth}(\tilde{\mathfrak{G}}_i) > 3$.

Proof: We define the graphs inductively.
Select a graph $\tilde{\mathfrak{G}}_1$ with $\mathrm{girth}(\tilde{\mathfrak{G}}_1) > 3$ and $\mathrm{chr}(\tilde{\mathfrak{G}}_1) > 2$ by Theorem 8.25. Without loss of generality we can assume that $\tilde{\mathfrak{G}}_1$ is connected (otherwise there exists a connected subgraph with the same properties). Let n_1 be the size of $\mathfrak{G}_1$.
Suppose now that $\tilde{\mathfrak{G}}_i$ is already defined and that n_i is the size of $\tilde{\mathfrak{G}}_i$. Let $\tilde{\mathfrak{G}}_{i+1}$ be a connected graph with $\mathrm{girth}(\tilde{\mathfrak{G}}_{i+1}) > n_i$ and $\mathrm{chr}(\tilde{\mathfrak{G}}_{i+1}) > n_i$. Note that this implies that the size n_{i+1} of $\tilde{\mathfrak{G}}_{i+1}$ is $> n_i$.

Having defined $\tilde{\mathfrak{G}}_1, \ldots, \tilde{\mathfrak{G}}_r$ like this it is clear from the observation above that there exists no homomorphism from $\tilde{\mathfrak{G}}_i$ to $\tilde{\mathfrak{G}}_j$ if $i > j$ (since $\mathrm{chr}(\tilde{\mathfrak{G}}_i) > n_j$).
Suppose for contradiction that there exists a homomorphism $h : \tilde{\mathfrak{G}}_j \to \tilde{\mathfrak{G}}_i$. Then the size of $h(\tilde{\mathfrak{G}}_j)$ is $\leq n_j$. Since $\mathrm{girth}(\tilde{\mathfrak{G}}_i) > n_{i-1} \geq n_j$ this means that $h(\tilde{\mathfrak{G}}_j)$ is circuit free. It is easy to see that a circuit free graph has a chromatic number ≤ 2.
Thus $\mathrm{chr}(\tilde{\mathfrak{G}}_j) \leq \mathrm{chr}(h(\tilde{\mathfrak{G}}_j)) \leq 2$, a contradiction. ■

Now it is easy to prove Lemma 8.24. Select $\tilde{\mathfrak{G}}_1, \ldots, \tilde{\mathfrak{G}}_{2k}$ by Lemma 8.26 and define for each $i \leq 2k$ a graph $\mathfrak{G}_i$ by

$$G_i := \tilde{G}_i \times \{1, \ldots, 2k+3\}$$

and

$$E^{\mathfrak{G}_i}(a,x)(b,y) \quad \Longleftrightarrow \quad (a = b \wedge x \neq y) \vee E^{\tilde{\mathfrak{G}}_i} ab \ .$$

Then obviously each vertex of $\mathfrak{G}_i$ is contained in a $(2k+3)$–clique and $\mathfrak{G}_i$ is connected.

Suppose for contradiction that there exists a homomorphism $H : \mathfrak{G}_i \to \mathfrak{G}_j$ for some $i \neq j$.
Let $a, b \in \tilde{\mathfrak{G}}_i$ such that $E^{\tilde{\mathfrak{G}}_i} ab$. Then

$$\{(a,z) \mid z \leq 2k+3\} \cup \{(b,z) \mid z \leq 2k+3\}$$

form a $(4k+6)$–clique in $\mathfrak{G}_i$ hence their images under H form a $(4k+6)$–clique in $\mathfrak{G}_j$. Since $\tilde{\mathfrak{G}}_j$ is triangle–free there exist $a', b' \in \tilde{\mathfrak{G}}_j$ such that $E^{\tilde{\mathfrak{G}}_j} a'b'$ and

$$\begin{aligned} & h\Big(\{(a,z) \mid z \leq 2k+3\} \cup \{(b,z) \mid z \leq 2k+3\}\Big) \\ = \ & \{(a',z) \mid z \leq 2k+3\} \cup \{(b',z) \mid z \leq 2k+3\}. \end{aligned}$$

Thus we can "sort" the vertices in the image and define a homomorphism $H' : \mathfrak{G}_i \to \mathfrak{G}_j$ such that for each $a \in \tilde{\mathfrak{G}}_i$ there exists a $a' \in \tilde{\mathfrak{G}}_j$ with

$$h(\{(a,z) \mid z \leq 2k+3\}) = \{(a',z) \mid z \leq 2k+3\}.$$

H' gives rise to a homorphism from $\tilde{\mathfrak{G}}_i$ to $\tilde{\mathfrak{G}}_j$ which is a contradiction.
Hence Lemma 8.24 is proved. ■

Having seen this gives us the reassuring feeling that we can continue using the constants.

9 Implicit Definability

9.1 Definition : Let σ be a signature and $\mathbf{L}$ a logic.
(1) Let $R_1, \ldots, R_m$ be relation symbols not contained in σ. A sentence

$$\varphi \in L[\sigma \cup \{R_1, \ldots, R_m\}]$$

defines $R_1, \ldots, R_m$ *implicitly* if for each σ–structure there exists exactly one expansion to a $(\sigma \cup \{R_1, \ldots, R_m\})$–structure which is a model of φ.[30]

(2) A σ–query ρ of arity ≥ 1 is *implicitly definable in* $\mathbf{L}$ if there exist relation symbols $\{R_1, \ldots, R_m\}$ not contained in σ (where R_1 has the same arity as ρ) and a sentence $\varphi \in L[\sigma \cup \{R_1, \ldots, R_m\}]$ such that

- φ defines $R_1, \ldots, R_m$ implicitly.
- For each $(\sigma \cup \{R_1, \ldots, R_m\})$–structure $\mathfrak{A}$ we have $R_1^{\mathfrak{A}} = \rho(\mathfrak{A})$.

A Boolean query ρ on $\mathcal{C}$ is *implicitly definable in* $\mathbf{L}$ if there exists an implicitly definable query ρ' (of arity ≥ 1) such that $\rho \mathrel{=\!\!|\!\!\models} (\rho' \neq \emptyset)$.[31]

$\mathbf{IMP(L)}$ denotes the class of all queries that are implicitly definable in $\mathbf{L}$.

(3) A query ρ is *k–ary* implicitly definable in $\mathbf{L}$ if the relation symbols $R_1, \ldots, R_m$ used for its implicit definition are all of arity $\leq k$.
$\mathbf{IMP(L)}^k$ denotes the class of all k–ary implicitly definable queries. $\square$

The investigation of implicit definability on finite structures was initiated by Kolaitis [Kol90]. He conjectured that for each $k \geq 1$ there exists a k–ary query in

$$\mathbf{IMP}^{2k}(\mathbf{FO}) \setminus \mathbf{IMP}^{2k-1}(\mathbf{FO})$$

(and proved it for $k = 1$).

In the next section we will define, for each $k \geq 1$, a Boolean graph–query in

$$\mathbf{IMP}^{k}(\mathbf{FO}) \setminus \mathbf{IMP}^{k-1}(\mathbf{PFP}^{k-1}).$$

Obviously this implies Kolaitis' conjecture.

9.2 Remark : Kolaitis' conjecture is also an easy consequence of a theorem of Ajtai [Ajt83] which implies that for a k–ary relation P the query

P contains an even number of elements

[30] Sometimes the notion given here is called "strongly" implicit definition to distinguish it from the other natural way to define (weakly) implicit definitions by replacing "exactly one expansion" by "at most one expansion".

[31] I.e. $\forall \mathfrak{A} : \rho(\mathfrak{A}) = \textsc{True} \iff \rho'(\mathfrak{A}) \neq \emptyset$.

is not definable in $(k-1)$–ary Σ_1^1, even on the class of ordered structures (and hence not $(k-1)$–ary implicitly definable).
But it is easy to see that the query is k–ary implicitly definable in first–order logic on ordered structures.

However, this proof is a bit unsatisfying because it requires a new signature for each k. $\square$

§ 9.1 The Hierarchy Theorem

9.3 Theorem : *For each $k \geq 2$ there is a Boolean graph–query in*

$$\mathbf{IMP}^k(\mathbf{FO}) \setminus \mathbf{IMP}^{k-1}(\mathbf{PFP}^{k-1}).$$

Proof: As usual we extend our signature by the constant–tuples $\overset{k}{\underline{c}}$ and $\overset{k}{\underline{d}}$ and consider the Boolean query

$$\textit{There exists an } \text{EDGE}_k\textit{–path from } \overset{k}{\underline{c}} \textit{ to } \overset{k}{\underline{d}}.$$

(In Paragraph 8.4 we have seen how we can avoid the constants.)

The main problem we have to deal with here is that there is no obvious way to define the query k–ary implicitly in first–order logic. To cope with it we extend our structures to make the order of the rows definable; this makes it easier to define the query.

Let $k, n \geq 2$ and recall that the structure $\mathfrak{C} = \mathfrak{C}(k,n)$ does not contain any cliques of size $\geq 2k+1$.
We are going to define a structure $\mathfrak{D} = \mathfrak{D}(k,n)$ which is an extension of $\mathfrak{C}$ and still has properties (1)–(4) described in Theorem 8.23. But moreover the order of the rows is definable in $\mathfrak{D}$.
Therefore we augment each row of $\mathfrak{C}$ by a $(2k+1)$–clique and by a $(2k+2)$–clique. For each row I we select one vertex a_I that belongs to the $(2k+1)$–clique of this row and one vertex b_I that belongs to the $(2k+2)$–clique. We include an edge between each vertex of row I and a_I and an edge between each vertex of row I and b_I.
Furthermore, for all $I < J \leq n$ we include an edge between a_I and b_J. The resulting structure is $\mathfrak{D}$.
We extend the mapping row to the new vertices in the obvious way, i.e. the vertices of the cliques associated with row I belong to row I. We extend æ to $\mathfrak{D}$ by letting it be the identity on the new vertices.

It is not difficult to define first–order formulae $\text{OLD}_k(x)$, $\prec_k (x,y)$, $\equiv_k (x,y)$ and $\text{SUCC}_k(x,y)$ such that for all $\boldsymbol{a},\boldsymbol{b} \in D$:

$$
\begin{array}{lcl}
\mathfrak{D} \models \text{OLD}_k[\boldsymbol{a}] & \Longleftrightarrow & \boldsymbol{a} \in C \\
\mathfrak{D} \models \boldsymbol{a} \prec_k \boldsymbol{b} & \Longleftrightarrow & \text{row}(\boldsymbol{a}) < \text{row}(\boldsymbol{b}) \\
\mathfrak{D} \models \boldsymbol{a} \equiv_k \boldsymbol{b} & \Longleftrightarrow & \text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \\
\mathfrak{D} \models \text{SUCC}_k[\boldsymbol{a},\boldsymbol{b}] & \Longleftrightarrow & \text{row}(\boldsymbol{a}) + 1 = \text{row}(\boldsymbol{b}).
\end{array}
$$

Finally, we define the formula

$$\text{O–EDGE}_k(\overset{k}{x},\overset{k}{y}) := \text{EDGE}_k(\overset{k}{x},\overset{k}{y}) \wedge \bigwedge_{i=1}^{k}(\text{OLD}_k(x_i) \wedge \text{OLD}_k(y_i)) \wedge \bigwedge_{i,j \le k} \text{SUCC}_k(x_i,y_j)$$

that we consider instead of EDGE_k from now on.

Let Φ_k be the first–order formula saying

- $\equiv_k$ is an equivalence relation.
- $\prec_k$ is a linear order of the $\equiv_k$–classes.
- SUCC_k is the successor–relation that belongs to $\prec_k$.

We need the following properties of our structure $\mathfrak{D}$

(1) $\mathfrak{D} \models \Phi_k$

(2) There exists a mapping row : $D \longrightarrow \{1,\dots,n\}$ such that

$$\text{row}(\boldsymbol{a}) < \text{row}(\boldsymbol{b}) \quad \Longleftrightarrow \quad \mathfrak{D} \models \boldsymbol{a} \prec_k \boldsymbol{b}$$

and $E^{\mathfrak{D}}\boldsymbol{ab}$ implies $|\text{row}(\boldsymbol{b}) - \text{row}(\boldsymbol{a})| \le 1$.

(3) There exists an automorphism $æ$ of $\mathfrak{D}$ that is self inverse and preserves the rows.

(4) There exist tuples $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$ in the first and last row respectively such that

$$\mathfrak{D} \models [TC_{\overset{k}{x},\overset{k}{y}}\text{O–EDGE}_k(\overset{k}{x},\overset{k}{y})]\,\overset{k}{\boldsymbol{c}},\overset{k}{\boldsymbol{d}}$$

but

$$\mathfrak{D} \not\models [TC_{\overset{k}{x},\overset{k}{y}}\text{O–EDGE}_k(\overset{k}{x},\overset{k}{y})]\,\overset{k}{\boldsymbol{c}},æ(\overset{k}{\boldsymbol{d}}).$$

(5) For all $\boldsymbol{b}_1,\dots,\boldsymbol{b}_{k-1} \in D$ there exists an automorphism f of $\mathfrak{D}$ that is self-inverse, preserves the rows, and maps $\overset{k-1}{\boldsymbol{b}}$ to $æ(\overset{k-1}{\boldsymbol{b}})$, but leaves all elements in rows of distance > 1 from $\text{row}(\boldsymbol{b}_1),\dots,\text{row}(\boldsymbol{b}_{k-1})$ fixed.

We let $\mathfrak{D}_1$ and $\mathfrak{D}_2$ be the $\{E, \underline{c}_1, \ldots, \underline{d}_k\}$–structures $(\mathfrak{D}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$ and $(\mathfrak{D}, \overset{k}{\boldsymbol{c}}, \boldsymbol{\alpha}(\overset{k}{\boldsymbol{d}}))$ respectively, where $\overset{k}{\underline{c}}$ and $\overset{k}{\underline{d}}$ are chosen according to property (4) of $\mathfrak{D}$.
Let ρ be the k–ary query that defines the O–$\textsc{Edge}_k$–connected component of $\overset{k}{\underline{c}}$. It is definable by the k–ary *LFP*–formula

$$[FP_{\overset{k}{x},X} \; \overset{k}{x} = \overset{k}{\underline{c}} \vee \exists \overset{k}{y} \in X \, \text{O–}\textsc{Edge}_k(\overset{k}{y}, \overset{k}{x})] \; \overset{k}{x} \, .^{32}$$

Thus by a result of Kolaitis [Kol90] (which says that $\mathbf{LFP}^k \subseteq \mathbf{IMP}^{2k}(\mathbf{FO})$ for each $k \geq 1$) it is definable in $\mathbf{IMP}^{2k}(\mathbf{FO})$.
However, on the class of models of Φ_k it is k–ary implicitly definable by the following first–order formula:

$$\begin{array}{rcl} \varphi(R) & = & R \, \overset{k}{\underline{c}} \\ & \wedge & \forall \overset{k}{x}, \overset{k}{y} \, ((R \, \overset{k}{x} \wedge \text{O–}\textsc{Edge}_k(\overset{k}{x}, \overset{k}{y})) \to R \, \overset{k}{y}) \\ & \wedge & \forall \overset{k}{x} \, (R \, \overset{k}{x} \to (\overset{k}{x} = \overset{k}{\underline{c}} \vee \exists \overset{k}{y} \; (R \, \overset{k}{y} \wedge \text{O–}\textsc{Edge}_k(\overset{k}{y}, \overset{k}{x})))) \, . \end{array}$$

Because let $\mathfrak{A}$ be a graph which is a model of Φ_k and $C \subseteq A^k$ the (uniquely determined) O–$\textsc{Edge}_k$–connected component of $\overset{k}{\underline{c}}$. Then clearly $\mathfrak{A} \models \varphi[C]$.
Let now $B \subseteq A^k$ such that $\mathfrak{A} \models \varphi[B]$. By the first and the second conjunct we have $C \subseteq B$. Moreover, by the third conjunct for each $\overset{k}{a} \in B$ there exists a sequence $\overset{k}{a}_1 = \overset{k}{a}, \overset{k}{a}_2, \ldots$ such that for each $i \geq 1$

$$\mathfrak{D} \models \text{O–}\textsc{Edge}_k[\overset{k}{a}_{i+1}, \overset{k}{a}_i] \quad \text{or} \quad \overset{k}{a}_i = \overset{k}{\underline{c}} \, .$$

We are finished if the latter case occurs. Since the former case implies $\mathfrak{D} \models \overset{k}{a}_{i+1} \prec_k \overset{k}{a}_i$ (by the definition of the formula O–$\textsc{Edge}_k$), and since $\prec_k$ is well–founded, the sequence cannot be infinitely long thus eventually the latter case occurs.
For those readers familiar with Kolaitis' results we remark that the reason this works is that in models of Φ_k the stage–comparison relation of χ is already available — essentially it is $\prec_k$.

We consider the query ρ^* that says:

If Φ_k holds then there exists an O–$\textsc{Edge}_k$*–path from $\overset{k}{\underline{c}}$ to $\overset{k}{\underline{d}}$.*

It can be implicitly defined by the formula

$$\xi(P,R) := \Big(\neg\Phi_k \wedge R = \mathcal{U} \wedge P = \mathcal{U}\Big) \vee \Big(\Phi_k \wedge \varphi(R) \wedge ((R \, \overset{k}{\underline{d}} \wedge P = \mathcal{U}) \vee (\neg R \, \overset{k}{\underline{d}} \wedge P = \emptyset))\Big).$$

[32] Since we are always a bit vague with the notion "connected component" the reader may take this formula as a definition of the query.

Suppose for contradiction that ρ^* is in $\mathbf{IMP}^{k-1}(\mathbf{PFP}^{k-1})$. Then there exists a sentence

$$\psi(R_1,\ldots,R_m) \in FP^{k-1}[\{E, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}, R_1,\ldots,R_m\}]$$

(where $R_1,\ldots,R_m$ are relation symbols of arity $\leq (k-1)$) defining $R_1,\ldots,R_m$ implicitly such that for all $\{E, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}, R_1,\ldots,R_m\}$–structures $\mathfrak{A}$ we have

$$\rho^*(\mathfrak{A}) = \textsc{True} \quad \Longleftrightarrow \quad R_1^{\mathfrak{A}} \neq \emptyset.$$

Without loss of generality we assume that the arity of each R_i $(i \leq m)$ is $(k-1)$.

Let r be the quantifier–rank of ψ and $n = 2^r + 2$. Consider the structures $\mathfrak{D} = \mathfrak{D}(k,n)$, $\mathfrak{D}_1 = \mathfrak{D}_1(k,n)$, and $\mathfrak{D}_2 = \mathfrak{D}_2(k,n)$ defined above. For each $i \leq m$ let $A_i := R_i^{\mathfrak{D}_1}$.
We define $B_1,\ldots,B_m \subseteq D^{k-1}$ as follows:
Let $\overset{k-1}{\boldsymbol{b}} \in D^{k-1}$. There exist (uniquely determined) sets B, B' such that

- $B \cup B' = \{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_{k-1}\}$
- $\forall \boldsymbol{b} \in B, \boldsymbol{b}' \in B' : \mathrm{row}(\boldsymbol{b}) < \mathrm{row}(\boldsymbol{b}') - 1$
- $\forall \boldsymbol{b}' \in B' : \mathrm{row}(\boldsymbol{b}') \geq 3$
- B is chosen minimal with respect to the first three conditions.

For each $i \leq k-1$ we let

$$\boldsymbol{a}_i := \begin{cases} \boldsymbol{b}_i & \text{if } \boldsymbol{b}_i \in B \\ \text{æ}(\boldsymbol{b}_i) & \text{if } \boldsymbol{b}_i \in B' \end{cases},$$

and for each $j \leq m$ we let $\overset{k-1}{\boldsymbol{b}} \in B_j :\Longleftrightarrow \overset{k-1}{\boldsymbol{a}} \in A_j$.

We are going to prove (with the $(k-1)$–ary r–pebble game) that the same $\mathbf{PFP}^{k-1}$–sentences of quantifier–rank r hold in $(\mathfrak{D}_1, A_1,\ldots,A_m)$ and $(\mathfrak{D}_2, B_1,\ldots,B_m)$. In particular $\mathfrak{D}_2 \models \psi(B_1,\ldots,B_m)$, thus for each $i \leq m$ we have $B_i = R_i^{\mathfrak{D}_2}$. But then, since $A_1 \neq \emptyset$ we have $R_1^{\mathfrak{D}_2} = B_1 \neq \emptyset$ thus $\rho^*(\mathfrak{D}_2) = \textsc{True}$ which is a contradiction.

9.4 Lemma : *Let $A, A' \subset D$ such that for all $\boldsymbol{a} \in A, \boldsymbol{a}' \in A'$: $\mathrm{row}(\boldsymbol{a}) < \mathrm{row}(\boldsymbol{a}') - 1$ and $\mathrm{row}(\boldsymbol{a}') \geq 3$. For a tuple $\overset{k-1}{\boldsymbol{a}} \in A \cup A'$ we define $\overset{k-1}{\boldsymbol{b}}$ by*

$$\boldsymbol{b}_i := \begin{cases} \boldsymbol{a}_i & \textit{if } \boldsymbol{a}_i \in A \\ \text{æ}(\boldsymbol{a}_i) & \textit{if } \boldsymbol{a}_i \in A' \end{cases}.$$

Then for each $i \leq m$ we have

$$\overset{k-1}{\boldsymbol{a}} \in A_i \quad \Longleftrightarrow \quad \overset{k-1}{\boldsymbol{b}} \in B_i \; .$$

Proof: Note first that for each automorphism f of $\mathfrak{D}_1$ we have $f(A_i) = A_i$ (for all $i \leq m$).
Without loss of generality we can assume that $A \cup A' = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}\}$ (since the definition of $\overset{k-1}{\boldsymbol{b}}$ only depends on $\overset{k-1}{\boldsymbol{a}}$ and not on the rest of A, A').
Select $\tilde{A}, \tilde{A}' \subset D$ such that

- $\tilde{A} \cup \tilde{A}' = \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}\}$
- $\forall \boldsymbol{a} \in \tilde{A}, \boldsymbol{a}' \in \tilde{A}' : \mathrm{row}(\boldsymbol{a}) < \mathrm{row}(\boldsymbol{a}') - 1$
- $\boldsymbol{a}' \in \tilde{A}' : \mathrm{row}(\boldsymbol{a}') \geq 3$
- $\tilde{A}$ is chosen minimal with respect to the first three conditions.

Thus $\tilde{A} \subseteq A$, and we can divide $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k-1}\}$ into three (possibly empty) parts: $\tilde{A}$, $A \cap \tilde{A}'$, and A'. There is an empty row between each of these parts. Thus by property (5) of structure $\mathfrak{D}$ there exists an automorphism f of $\mathfrak{D}$ that moves each $\boldsymbol{a} \in A \cap \tilde{A}'$ to $\ae(\boldsymbol{a})$ and leaves all elements in rows of distance > 1 from $A \cap \tilde{A}'$ fixed, in particular the elements of $\tilde{A}$ and A' and the tuples $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$.
The latter implies that f is also an automorphism of $\mathfrak{D}_1$ thus

$$f(\overset{k-1}{\boldsymbol{a}}) \in A_i \Longleftrightarrow \overset{k-1}{\boldsymbol{a}} \in A_i$$

for each $i \leq m$.
Let us define sets B and B' by $B := \tilde{A}$ and $B' := \{\boldsymbol{b}_i \mid \boldsymbol{a}_i \in \tilde{A}'\}$. Since $\ae$ preserves the rows we have

- $B \cup B' = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1}\}$
- $\forall \boldsymbol{b} \in B, \boldsymbol{b}' \in B' : \mathrm{row}(\boldsymbol{b}) < \mathrm{row}(\boldsymbol{b}') - 1$
- $\boldsymbol{b}' \in B' : \mathrm{row}(\boldsymbol{b}') \geq 3$
- B is chosen minimal with respect to the first three conditions.

Note that for each $i \leq k-1$ we have

$$f(\boldsymbol{a}_i) = \left\{ \begin{array}{rcll} \boldsymbol{a}_i & = & \boldsymbol{b}_i & \text{if} \quad \boldsymbol{a}_i \in \tilde{A} \\ \ae(\boldsymbol{a}_i) & = & \ae(\boldsymbol{b}_i) & \text{if} \quad \boldsymbol{a}_i \in \tilde{A}' \cap A \\ \boldsymbol{a}_i & = & \ae(\boldsymbol{b}_i) & \text{if} \quad \boldsymbol{a}_i \in A' \end{array} \right.$$

thus

$$f(\boldsymbol{a}_i) = \begin{cases} \boldsymbol{b}_i & \text{if } \boldsymbol{b}_i \in B \\ \text{æ}(\boldsymbol{b}_i) & \text{if } \boldsymbol{b}_i \in B' \end{cases} .$$

Hence by the definition of of B_i we have $f(\overset{k-1}{\boldsymbol{a}}) \in A_i \iff \overset{k-1}{\boldsymbol{b}} \in B_i$.
Thus the statement of the lemma is proved. ∎

The proof of the theorem is now easily finished:
In Section 8 we proved that the duplicator wins the $(k-1)$–ary r–pebble game on $\mathfrak{C}_1$ and $\mathfrak{C}_2$ hence on $\mathfrak{D}_1$ and $\mathfrak{D}_2$ because we only needed properties (1)–(4) of $\mathfrak{C}$, and $\mathfrak{D}$ has got the same properties.
In fact we proved that she can preserve the following throughout the whole game:
Suppose $r' \leq r$ pairs of pebbles are on the board.
Then there exist $M, N \in \{1, \ldots, n\}$ and an automorphism h of $\mathfrak{D}_1$ such that:

(i) $N - M \geq 2^{r+1-r'}$

(ii) For pairs $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements we have:

$$\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \leq M \quad \text{or} \quad \text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \geq N \; .$$

(iii) h is self–inverse and preserves the rows.

(iv) For each pair $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements the following holds:

- If $\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \leq M$ then $h(\boldsymbol{a}) = \boldsymbol{b}$.
- If $\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \geq N$ then $\text{æ}(h(\boldsymbol{a})) = \boldsymbol{b}$.

This implies that the pairs of corresponding elements form a partial isomorphism not only between $\mathfrak{D}_1$ and $\mathfrak{D}_2$, but between $(\mathfrak{D}_1, A_1, \ldots, A_m)$ and $(\mathfrak{D}_2, B_1, \ldots, B_m)$. Because let $A = \{\boldsymbol{a} \in D \mid \text{row}(\boldsymbol{a}) \leq M\}$ and $A' = \{\boldsymbol{a} \in D \mid \text{row}(\boldsymbol{a}) \geq N\}$ and define a partial mapping p on D by

$$p(\boldsymbol{a}) = \begin{cases} h(\boldsymbol{a}) & \text{if } \boldsymbol{a} \in A \\ \text{æ}(h(\boldsymbol{a})) & \text{if } \boldsymbol{a} \in A' \\ \text{undefined} & \text{otherwise} \end{cases}$$

Since $N - M > 2^{r-r'} \geq 1$ and $h, \text{æ}$ are automorphisms of $\mathfrak{D}$, p is a partial isomorphism between $\mathfrak{D}_1$ and $\mathfrak{D}_2$. Lemma 9.4 applied to the sets $h(A)$ and $h(A')$ together with the fact that the A_i are closed under automorphisms (in particular under the automorphism h of $\mathfrak{D}_1$) does the rest.

Thus the duplicator wins the the $(k-1)$–ary r–pebble game on $(\mathfrak{D}_1, A_1, \ldots, A_m)$ and $(\mathfrak{D}_2, B_1, \ldots, B_m)$ and we are finished. ∎

10 Arities of Lindström Quantifiers

In this section we assume again that classes of structures are closed under isomorphism.

10.1 Definition : The *arity* of a Lindström–quantifier $Q_{\mathcal{C}}$ associated with a class $\mathcal{C}$ of σ–structures is defined to be

$$\max(\{r \mid \sigma \text{ contains an } r\text{–ary relation symbol}\} \cup \{0\}) .$$

For each $l \geq 0$ we let

$$\mathbf{Q}^l = \{Q_{\mathcal{C}} \mid Q_{\mathcal{C}} \text{ is an } l\text{–ary Lindström–quantifier.}\}.$$

□

The main result of this section is that the weak main lemma (but not the main lemma) remains true if we augment $\mathbf{PFP}^{k-1}$ by the class $\mathbf{Q}^{2k-1}$.
This does not only improve our weak main lemma, but also a theorem of Hella [Hel92] which says that for each $l \geq 0$ there is a query which is definable in **Datalog** but not in $\mathbf{FO}(\mathbf{Q}^l)$.
Independently, Lauri Hella found a proof of the same result.
Recall that E_k is a $2k$–ary relation symbol.

10.2 Theorem : *The transitive closure of E_k is not definable in* $\mathbf{s}$–$\mathbf{PFP}^{k-1}(\mathbf{Q}^{2k-1})$ *for any $k \geq 1$.*

§ 10.1 The Bijective Pebble Game

Hella [Hel89] introduced the following game that he used to prove his theorem in [Hel92] we mentioned above:

10.3 Definition : Let $r, l \geq 1$. The *l–bijective r–move game* is played by two players, the challenger and the duplicator, on a pair $\mathfrak{A}, \mathfrak{B}$ of structures of the same signature with $2r$ pebbles $P_1, Q_1, \ldots, P_r, Q_r$ as follows:

If A and B have the same cardinality, in each round of the game the duplicator defines a bijection $f : A \to B$. Then the challenger picks up $m \leq l$ yet unused pebbles $P_{i_1}, \ldots, P_{i_m}$ and places them on elements $a_1, \ldots, a_m \in A$ respectively. The corresponding pebbles $Q_{i_1}, \ldots, Q_{i_m}$ are placed on $f(a_1), \ldots, f(a_m) \in B$ respectively.

After each round, the pairs of pebbled elements $(a_i, f(a_i))$ and the pairs $(c^{\mathfrak{A}}, c^{\mathfrak{B}})$ of constants are called *pairs of corresponding elements.*
The duplicator wins the game if, after all pebbles are placed on the structures, the corresponding elements form a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

Otherwise, or if A and B do not have the same cardinality, the challenger wins the game. □

Let $\sigma = \{R_1, \ldots, R_m\}$ where R_i is an r_i–ary relation symbol and let $\mathcal{C}$ be a class of σ–structures. We extend our definition of the quantifier–rank to s–*FP*($Q_\mathcal{C}$)–formulae by

$$\mathrm{qr}([Q_\mathcal{C}\ \overset{r_1}{x_1}, \ldots, \overset{r_m}{x_m}\ \ \varphi_1, \ldots, \varphi_m]) = \max\{\mathrm{qr}(\varphi_1) + r_1, \ldots, \mathrm{qr}(\varphi_m) + r_m\}\ .$$

Using the respective clauses for all quantifiers $Q_\mathcal{C}$ in a set $\mathbf{Q}$ we extend the definition of the quantifier rank to s–*FP*($\mathbf{Q}$)–formulae (for an arbitrary set $\mathbf{Q}$ of Lindström–quantifiers), thus in particular to $FO(\mathbf{Q}^l)$–formulae (for an $l \geq 0$).

10.4 Theorem ([Hel89]) : *Let $r, l \geq 1$ and $\mathfrak{A}$ and $\mathfrak{B}$ be structures of the same signature.*
Then the same $FO(\mathbf{Q}^l)$–sentences of quantifier rank $\leq r$ hold in $\mathfrak{A}$ and $\mathfrak{B}$ if and only if the duplicator has a winning strategy in the l–bijective r–move game on $\mathfrak{A}$ and $\mathfrak{B}$.

To obtain our hierarchy result that incorporates arities of fixed–point operators and Lindström–quantifiers we define the following game:

10.5 Definition : Let $k, l, r \geq 1$. The *k-ary l-bijective r-pebble game* is played by two players, the challenger and the duplicator, on a pair $\mathfrak{A}, \mathfrak{B}$ of structures of the same signature with $2r$ pebbles $P_1, Q_1, \ldots, P_r, Q_r$.
We say that a pair (P_i, Q_i) of pebbles is *on the board* in a situation of the game if it is placed on the structures in that situation. The other pebbles are called *free*.
Moreover, each situation of the game and each pebble on the board have a depth $d \geq 0$. The game starts in the situation with depth 0 and all pebbles free.
If $\mathfrak{A}$ and $\mathfrak{B}$ are of the same cardinality, in each round of the game the challenger selects one of the following moves:

Q-move : The duplicator selects a bijection $f : A \longrightarrow B$.

The challenger places $m \leq l$ free pebbles $P_{i_1}, \ldots, P_{i_m}$ on $a_1, \ldots, a_m \in A$. The corresponding pebbles $Q_{i_1}, \ldots, Q_{i_m}$ are placed on $f(a_1), \ldots, f(a_m) \in B$.

The depth of the P_{i_j} and Q_{i_j} is defined to be the current depth of the game.

I-move : The depth is increased by 1.

R-move : The depth of the game is reduced to a $d \geq 1$ less than or equal to the current depth. The challenger selects $j \leq k$ pairs $(P_{i_1}, Q_{i_1}), \ldots, (P_{i_j}, Q_{i_j})$ of pebbles of depth $\geq d$ to be left on the board. Their depth is reduced to d. All other pebbles of depth $\geq d$ are removed from the structures.

In each situation of the game the pairs of elements which are pebbled by corresponding pebbles P_i, Q_i and the pairs $(c^{\mathfrak{A}}, c^{\mathfrak{B}})$ of constants are called *pairs of corresponding elements.*
The duplicator wins the game if in each situation the corresponding elements form a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$.

Otherwise, or if A and B do not have the same cardinality, the challenger wins the game. $\square$

10.6 Theorem : *Let $k, l, r \geq 1$ and $\mathfrak{A}$ and $\mathfrak{B}$ be two structures of the same signature such that the duplicator has a winning strategy for the k–ary l–bijective r–pebble game on $\mathfrak{A}, \mathfrak{B}$.*
Then the same s–$FP^k(\mathbf{Q}^l)$–sentences of quantifier rank $\leq r$ hold in $\mathfrak{A}$ and $\mathfrak{B}$.

Proof: The proof can be given along the lines of the proof of Theorem 7.2, the analogous statement for the k–ary r–pebble game.

Let χ be a s–$FP^k(\mathbf{Q}^l)$–sentence of quantifier rank $\leq r$ such that $\mathfrak{A} \models \chi$. We shall prove that $\mathfrak{B} \models \chi$. Note that the assumption that the duplicator has a winning strategy implies that $\mathfrak{A}$ and $\mathfrak{B}$ have the same cardinality.

Recall from Example 5.3 that existential and universal quantifiers can be replaced by the unary Lindström–quantifiers $Q_{\mathcal{EX}}$ and $Q_{\mathcal{FA}}$ respectively and note that formulae of the form $\neg[Q_{\mathcal{C}} \ldots]$ are equivalent to $[Q_{\mathcal{C}^C} \ldots]$.[33]
Hence for each $FP^k(\mathbf{Q}^l)$–formula φ there is an equivalent formula $\overline{\overline{\varphi}}$ where no existential or universal quantifiers occur and where negation symbols only occur in front of fixed–point operators or atomic subformulae. We use the formulae $\overline{\overline{\varphi}}$ in the same way that we used $\overline{\varphi}$ in the proof of Theorem 7.2. In particular, we assume without loss of generality that $\chi = \overline{\overline{\chi}}$.
We also assume (without loss of generality) that for each subformula of χ of the form

$$[Q_{\mathcal{C}}\ \overset{r_1}{x}_1, \ldots, \overset{r_m}{x}_m\ \ \varphi_1, \ldots, \varphi_m]$$

the variables $x_{11}, \ldots, x_{1r_1},\ \ldots\ , x_{m1}, \ldots, x_{mr_m}$ are all distinct.

Recall that in the first step of the proof of Theorem 7.2 we constructed, for each s–FP^k–formula of the form

$$\psi = [s\text{–}FP_{\overset{k_1}{x}_1, X_1, \ldots, \overset{k_m}{x}_m, X_m}\ \varphi_1, \ldots, \varphi_m]\ \overset{k_1}{u},$$

a sequence of formulae $(\psi_j^*(\overset{k_1}{u}))_{j \geq 1}$ such that

$$\mathfrak{I} \models \psi \iff \exists j \geq 1 : \mathfrak{I} \models \psi_j^*(\overset{k_1}{u})\ .$$

[33] Here, for a class $\mathcal{C}$ of σ–structures $\mathcal{C}^C$ denotes the class $\mathcal{F}[\sigma] \setminus \mathcal{C}$.

This goes through for s–$FP^k(\mathbf{Q}^l)$–formulae without any problems.

In the second step we introduced the notion of a proof–sequence and inductively defined a set $\mathcal{S}$ of proof–sequences for χ. We proceed analogously here.[34] Note that since no existential or universal quantifiers occur in χ we can omit the cases $\psi = \exists x\varphi(x)$ and $\forall x\varphi(x)$. On the other hand we have to consider the case

$\boldsymbol{\psi} = [\boldsymbol{Q}_{\mathcal{C}}\ \overset{l_1}{\boldsymbol{x}}_1, \ldots, \overset{l_m}{\boldsymbol{x}}_m\ \ \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_m]$**:** For each $i \leq m$ and $a_1, \ldots, a_{l_i}$ let $s_{i,a_1,\ldots,a_{l_i}}$ be the proof–situation whose move is the following Q–move:

The duplicator selects a bijection $f : A \to B$ according to her winning strategy. The challenger places free pebbles $P_{j_1}, \ldots, P_{j_{l_i}}$ on $a_1, \ldots, a_{l_i}$, the corresponding pebbles $Q_{j_1}, \ldots, Q_{j_{l_i}}$ are placed on $f(a_1), \ldots, f(a_{l_i})$.

The formula of $s_{i,a_1,\ldots,a_{l_i}}$ is $\overline{\overline{\varphi_i}}$ if $(\mathfrak{A}, \alpha) \models \varphi_i[a_1, \ldots, a_{l_i}]$, otherwise it is $\overline{\overline{\neg\varphi_i}}$. Its mapping is

$$(\iota \cup \{(x_1, j_1), \ldots, (x_{l_i}, j_{l_i})\})\restriction_{\text{free}(\varphi_i)} .$$

We replace S in $\mathcal{S}$ by all sequences $S^\frown s_{i,a_1,\ldots,a_{l_i}}$.

Note that the mapping f is the same in each situation $s_{i,a_1,\ldots,a_{l_i}}$ (because the duplicator's winning strategy uniquely determines it). This will be important for the third step.

In the third step of the proof of Theorem 7.2 we checked inductively that for each proof–situation s that occurs in a sequence $S \in \mathcal{S}$ the following holds:

If ψ is the formula of s and β the interpretation of the free variables of ψ in $\mathfrak{B}$ induced by the mapping ι of s we have $(\mathfrak{B}, \beta) \models \psi$.

Again we can omit the cases $\psi = \exists x\varphi(x)$ and $\forall x\varphi(x)$, and instead have to consider:

$\boldsymbol{\psi} = [\boldsymbol{Q}_{\mathcal{C}}\ \overset{l_1}{\boldsymbol{x}}_1, \ldots, \overset{l_m}{\boldsymbol{x}}\ \ \boldsymbol{\varphi}_1, \ldots, \boldsymbol{\varphi}_m]$**:** By the definition of $\mathcal{S}$ there exists a bijection $f : A \to B$ and for each $i \leq m$ and $a_1, \ldots, a_{l_i}$ a proof–sequence where s is followed by a situation $s_{i,a_1,\ldots,a_{l_i}}$ defined as above.

By induction hypothesis and the definition of $s_{i,a_1,\ldots,a_{l_i}}$ this means that

$$(\mathfrak{A}, \alpha) \models \varphi_i[a_1, \ldots, a_{l_i}] \quad \Longleftrightarrow \quad (\mathfrak{B}, \beta) \models \varphi_i[f(a_1), \ldots, f(a_{l_i})]$$

hence

$$\Big(A, \varphi_1(\overset{l_1}{x_1})^{(\mathfrak{A},\alpha)}, \ldots, \varphi_m(\overset{l_m}{x_m})^{(\mathfrak{A},\alpha)}\Big) \cong \Big(B, \varphi_1(\overset{l_1}{x_1})^{(\mathfrak{B},\beta)}, \ldots, \varphi_m(\overset{l_m}{x_m})^{(\mathfrak{B},\beta)}\Big).$$

But since $(\mathfrak{A}, \alpha) \models \psi$ we have

$$\Big(A, \varphi_1(\overset{l_1}{x_1})^{(\mathfrak{A},\alpha)}, \ldots, \varphi_m(\overset{l_m}{x_m})^{(\mathfrak{A},\alpha)}\Big) \in \mathcal{C}$$

[34] using the same notation

hence

$$\left(B, \varphi_1(\overset{l_1}{x_1})^{(\mathfrak{B},\beta)}, \ldots, \varphi_m(\overset{l_m}{x_m})^{(\mathfrak{B},\beta)}\right) \in \mathcal{C}$$

and $(\mathfrak{B}, \beta) \models \psi$.

All the rest can be done as in the old proof. ■

§ 10.2 Proof of the Hierarchy Theorem

Recall that the statement of the theorem is tha the transitive closure of E_k is not definable in $\mathbf{s}\text{–}\mathbf{PFP}^{k-1}(\mathbf{Q}^{2k-1})$ t for any $k \geq 1$.

Before we give a detailed proof of the statement for $k \geq 2$ let us shortly consider the situation for $k = 1$. There the statement reduces to:

The transitive closure of directed graphs is not definable in $\mathbf{FO}(\mathbf{Q}^1)$.

To see this consider a structure $\mathfrak{A}$ that consists of two E_1–paths (a *left* and a *right* path) of length n (for a sufficiently large n). Let $\mathfrak{A}_1$ be the $\{E_1, \underline{c}, \underline{d}\}$–structure where $\underline{c}$ and $\underline{d}$ are the first and last point (respectivly) of the left path and $\mathfrak{A}_2$ be the $\{E_1, \underline{c}, \underline{d}\}$–structure where $\underline{c}$ is the first point of the left and $\underline{d}$ is the last point of the right path.
It can be easily seen that the duplicator has a winning strategy for the 1–bijective r–move game on $\mathfrak{A}_1$ and $\mathfrak{A}_2$ if $n \geq 2^{r+1} + 1$, thus the same $\mathbf{FO}(\mathbf{Q}^1)$–sentences of quantifier–rank $\leq r$ hold in $\mathfrak{A}_1$ and $\mathfrak{A}_2$. This suffices to prove the statement for $k = 1$.
In fact for the Q–move in the following proof we use the same arguments as we would need them here. For this reason we are not going into the details now.

So let us fix a $k \geq 2$.
At the beginning of Paragraph 8.2, for all $n \geq 2$ we have defined $\{E_k, \underline{c}_1, \ldots, \underline{d}_k\}$–structures $\mathfrak{B}_1(k,n) := (\mathfrak{B}(k,n), \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}}\ )$ and $\mathfrak{B}_2(k,n) := (\mathfrak{B}(k,n), \overset{k}{\boldsymbol{c}}, \text{æ}(\overset{k}{\boldsymbol{d}}))$ where $\mathfrak{B}(k,n)$ has been taken from Theorem 8.6 and $\overset{k}{\boldsymbol{c}}$ and $\overset{k}{\boldsymbol{d}}$ have been taken according to item 3 of this theorem. In particular this has been meaning

$$\mathfrak{B}_1(k,n) \models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}} \quad \text{and} \quad \mathfrak{B}_2(k,n) \not\models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}\ .$$

Let $r \geq 0$ and $n := 2^{r+1} + 1$. We are finished once we have proved:

10.7 Proposition : *The duplicator has a winning strategy for the* $(k-1)$*–ary* $(2k-1)$*–bijective* r*–pebble game on* $\mathfrak{B}_1 := \mathfrak{B}_1(k,n)$ *and* $\mathfrak{B}_2 := \mathfrak{B}_2(k,n)$.

Proof: The proof is completely analogous to the proof of Proposition 8.12, we just have to replace the $\exists$ and $\forall$–moves by a Q–move.

As in the old proof it suffices to show inductively that the duplicator can preserve the following throughout the whole game:
Suppose $r' \leq r$ pairs of pebbles are on the board.
Then there exist $M, N \in \{1, \ldots, n\}$ and an automorphism h of $\mathfrak{B}_1$ such that:

(i) $N - M \geq 2^{r+1-r'}$

(ii) For pairs $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements we have:

$$\mathrm{row}(\boldsymbol{a}) = \mathrm{row}(\boldsymbol{b}) \leq M \quad \text{or} \quad \mathrm{row}(\boldsymbol{a}) = \mathrm{row}(\boldsymbol{b}) \geq N \ .$$

(iii) h is self–inverse and preserves the rows.

(iv) For each pair $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements the following holds:

- If $\mathrm{row}(\boldsymbol{a}) = \mathrm{row}(\boldsymbol{b}) \leq M$ then $h(\boldsymbol{a}) = \boldsymbol{b}$.
- If $\mathrm{row}(\boldsymbol{a}) = \mathrm{row}(\boldsymbol{b}) \geq N$ then $\mathit{æ}(h(\boldsymbol{a})) = \boldsymbol{b}$.

In a situation with $r' \leq r$ pairs of pebbles on the board and M, N, h defined such that (i)–(iv) hold we have to define the duplicator's answer to the next move and M', N', h' to preserve (i)–(iv) in the next situation.
The R–move and the I–move can be treated in the same way as in the proof of Proposition 8.12. The only new case is the following:

Q–move: Let $\textsc{Middle} = M + \frac{N-M}{2}$. The duplicator defines a bijection $f : B_1 \to B_2$ by

$$f(\boldsymbol{a}) = \begin{cases} h(\boldsymbol{a}) & \text{if } \mathrm{row}(\boldsymbol{a}) < \textsc{Middle} \\ \mathit{æ}(h(\boldsymbol{a})) & \text{otherwise} \end{cases} \quad .$$

Now, for some $l \leq (2k-1)$, the challenger places free pebbles on $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l \in B_1$. The corresponding pebbles are placed on $\boldsymbol{b}_1 := f(\boldsymbol{a}_1), \ldots, \boldsymbol{b}_l := f(\boldsymbol{a}_l) \in B_2$.

For convenience we assume that we have $l = (2k-1)$, in the case $l < (2k-1)$ we have to be a bit more careful with the appliction of the pigeonhole principle, but the proof still goes through without any problems. (We could also restrict the game to $(2k-1)$–ary moves, and in turn modify the definition of the quantifier rank a bit.)

At most one of the intervals $[M, \textsc{Middle}[$ and $[\textsc{Middle}, N]$ contains $\geq k$ elements of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$.

Case 1: The interval $[M, \textsc{Middle}[$ contains $< k$ elements of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$.

We select M', N' by the pigeonhole principle such that

- $M \leq M' < N' < \textsc{Middle}$
- $$\begin{aligned} N' - M' &\geq \frac{\ulcorner \textsc{Middle} \urcorner - 1 - M}{k} \\ &= \frac{\frac{\ulcorner N-M \urcorner}{2} - 1}{k} \\ &\geq \frac{2^{r+1-r'-1}}{k} - \frac{1}{k} && (\text{since } N - M > 2^{r+1-r'}) \\ &= \frac{2^{r+1-r'}}{l+1} - \frac{1}{k} && (\text{since } l = 2k-1) \\ &\geq 2^{r+1-(r'+l)} && (\text{since } k \geq 2, \text{ hence } l \geq 3) \end{aligned}$$
- None of the pairs $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ $(i \leq l)$ is in a row between M' and N'.

This implies (i) and (ii) for M', N'.

To define h' let $\{\boldsymbol{a}_{i_1}, \ldots, \boldsymbol{a}_{i_m}\}$ be the maximal subset of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$ such that

$$N' \leq \mathrm{row}(\boldsymbol{a}_{i_1}), \ldots, \mathrm{row}(\boldsymbol{a}_{i_m}) < \textsc{Middle}\,.$$

(If $m = 0$ there is nothing to do, we just let $h' := h$.)

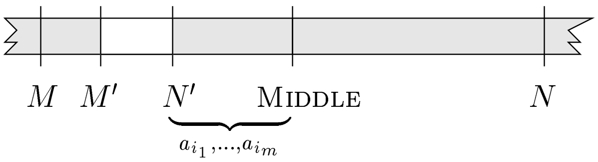

Figure 10.8

By Theorem 8.6(4) there exists an automorphism g of $\mathfrak{B}$ such that

- $\forall \boldsymbol{a} \in B:\ \mathrm{row}(\boldsymbol{a}) = \mathrm{row}(g(\boldsymbol{a}))$
- $\forall j \leq m:\ g(h(\boldsymbol{a}_{i_j})) = æ(h(\boldsymbol{a}_{i_j}))$
- If $\boldsymbol{a} \in B$ such that $\forall j \leq m : \mathrm{row}(\boldsymbol{a}) \neq \mathrm{row}(\boldsymbol{a}_{i_j})$ then $g(\boldsymbol{a}) = \boldsymbol{a}$. $\quad (*)$

Letting $h' := g \circ h$ we get (iii) and (iv) since each old pair of corresponding elements is fixed by g, while for the new pairs note that for each $i \leq l$ we have:

- $\mathrm{row}(\boldsymbol{a}_i) \leq M' \Longrightarrow h'(\boldsymbol{a}_i) = g(h(\boldsymbol{a}_i)) = h(\boldsymbol{a}_i) = f(\boldsymbol{a}_i) = \boldsymbol{b}_i$
- $N' \leq \mathrm{row}(\boldsymbol{a}_i) < \textsc{Middle} \Longrightarrow i \in \{i_1, \ldots, i_m\}$
  $\Longrightarrow æ(h'(\boldsymbol{a}_i)) = æ(g(h(\boldsymbol{a}_i))) = æ(æ(h(\boldsymbol{a}_i))) = h(\boldsymbol{a}_i) = f(\boldsymbol{a}_i) = \boldsymbol{b}_i$
- $\mathrm{row}(\boldsymbol{a}_i) \geq \textsc{Middle}(\geq N') \Longrightarrow æ(h'(\boldsymbol{a}_i)) = æ(h(\boldsymbol{a}_i))) = f(\boldsymbol{a}_i) = \boldsymbol{b}_i$.

Case 2: The interval $[M, \textsc{Middle}[$ contains $\geq k$ elements of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$. Hence the interval $[\textsc{Middle}, N]$ contains $< k$ elements of $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$.

This case can be treated similarly.

■

10.9 Remark : Probably the reader has noticed that in Case 1 of the previous proof we essentially did the same as in Case 1 of the R–move (cf. page 87) of the proof of Proposition 8.12. But in contrast to that case it would not suffice here to demand $\forall j \leq m : |\mathrm{row}(a) - \mathrm{row}(a_{i_j})| > 1$ instead of $\forall j \leq m : \mathrm{row}(a) \neq \mathrm{row}(a_{i_j})$ in line $(*)$ because there may be an a_{i_j} in row ($\textsc{Middle}-1$) and an a in row $\textsc{Middle}$ that have to be treated differently. So we cannot strengthen our result to graphs in the same way as we did with the weak main lemma (cf. also Remark 8.14).

However, this is only reassuring because the theorem is clearly wrong for graphs (i.e. there exists no graph–query in $\mathbf{PFP}^k \setminus \mathbf{PFP}(\mathbf{Q}^{2k-1})$). The reason for this is, as was already pointed out by Hella [Hel92], that obviously *all* graph queries are definable in $\mathbf{FO}(\mathbf{Q}^2)$. □

§10.3 Arity Hierarchies of Vectorized Lindström Logics

Our definition of the arity of a Lindström–quantifier makes sense as long as we consider single Lindström quantifiers. But if we turn to vectorized quantifiers it seems a bit unnatural.
For example, consider the quantifier $Q_{\mathcal{TC}}$ that corresponds to the transitive closure operator. In fact for each $k \geq 1$ we have $\mathbf{FO}(Q^k_{\mathcal{TC}}) = \mathbf{TC}^k$. However, the quantifier $Q_{\mathcal{TC}_k}$ that corresponds to $Q^k_{\mathcal{TC}}$ (cf. Example 5.10) has arity $2k$ which is clearly not what we want.

Having this in mind, the following definition seems to be natural:

10.10 Definition : An $\mathbf{FO}(Q^{<\omega}_{\mathcal{C}})$–formula is *k–ary* if it is built using only the clauses $(Q^l_{\mathcal{C}})$ for $l \leq k$ and the first–order clauses. □

We can now define the k–ary fragment and the arity–hierarchy of a vectorized Lindström logic by Definition 6.2. We usually refer to the arity hierarchy of $\mathbf{L}(Q^{<\omega}_{\mathcal{C}})$ as the arity hierarchy of $Q_{\mathcal{C}}$.
For brevity we are going to denote the k–ary fragment of the logic $L(Q^{<\omega}_{\mathcal{C}})$ by $L(Q^k_{\mathcal{C}})$ instead of $L(Q^{<\omega}_{\mathcal{C}})^k$.
Our results on the arities of fixed–point and transitive closure operators provide some examples of vectorized Lindström logics whose arity hierarchy is strict:

10.11 Corollary : *The arity hierarchies of the quantifiers $Q_{\mathcal{TC}}$, $Q_{\mathcal{SFP}}$, $Q_{\mathcal{IFP}}$, and $Q_{\mathcal{PFP}}$*[35] *are strict.*

Proof: For example, the statement for $Q_{\mathcal{PFP}}$ follows from the fact that the class $\mathcal{PFP}_{k-1}$ (recall that $Q_{\mathcal{PFP}_{k-1}}$ is the quantifier associated with $Q^{k-1}_{\mathcal{PFP}}$) is definable in $\mathbf{PFP}^{k-1}$ on the one hand and that the formula

$$[TC_{\overset{k}{x},\overset{k}{y}}\textsc{Edge}_k(\overset{k}{x},\overset{k}{y})]\ \overset{k}{u},\overset{k}{v}$$

is equivalent to the $FO(Q^k_{\mathcal{PFP}})$–formula

$$\exists\ \overset{k}{w}\ \Big([Q^k_{\mathcal{PFP}}\ \overset{k}{x}\overset{k}{y}\overset{k}{z},\overset{k}{x},\overset{k}{x}\ \ (\overset{k}{y}{=}\overset{k}{x}\ \vee \textsc{Edge}_k(\overset{k}{x},\overset{k}{y}))\wedge\ \overset{k}{y}{\neq}\overset{k}{v},\varphi(\overset{k}{u},\overset{k}{x}),\overset{k}{x}{=}\overset{k}{w})]$$
$$\wedge\,\varphi(\overset{k}{w},\overset{k}{v})\Big)\vee\varphi(\overset{k}{u},\overset{k}{v})$$

on the other hand. (The last formula seems to be awkward, but we have to take care that the process is "modest" (cf. Section 5, in particular the proof of Theorem 5.16).)
The statement for the other quantifiers can be verified similarly. ■

We have also seen an example of a non–trivial Lindström–quantifier whose arity hierarchy collapses: the quantifier $Q_{\mathcal{EVEN}}$ (cf. Example 5.4). In fact it is not hard to find Lindström–quantifiers where the hierarchy collapses on an arbitrary level:

10.12 Example : Let, for each $k \geq 2$,

$$\mathcal{POW}^k = \{(A) \in \mathcal{F}[\emptyset]^{36} \mid \exists \text{ prime number } p : |A| = p^k\}.$$

Since $[Q^l_{\mathcal{POW}^k}] \dashv\models \exists x x \neq x$ for each $l > k$ we have

$$\mathbf{FO}(Q^k_{\mathcal{POW}^k}) = \mathbf{FO}(Q^l_{\mathcal{POW}^k})\ .$$

On the other hand the class of $\emptyset$–structures whose cardinality is a prime number is definable in $\mathbf{FO}(Q^k_{\mathcal{POW}^k})$, but not in $\mathbf{FO}(Q^{k-1}_{\mathcal{POW}^k})$.

To see the latter let p and q be prime numbers, A a set of cardinality p and B a set of cardinality pq. Then for each $l < k$ we have

$$(A) \not\models [Q^l_{\mathcal{POW}^k}] \quad \text{and} \quad (B) \not\models [Q^l_{\mathcal{POW}^k}]\ .$$

Thus the same $FO(Q^l_{\mathcal{POW}^k})$–sentences of quantifier–rank $\leq p$ hold in (A) and (B). □

As a consequence of Theorem 10.2 we obtain:

[35] See Section 5 for the definition of the classes $\mathcal{TC},\mathcal{SFP},\mathcal{IFP},\mathcal{PFP}$.
[36] Recall that $\mathcal{F}[\emptyset]$ denotes the class of finite structures of empty signature.

10.13 Corollary : *For each $k \geq 1$ we have* $\mathbf{FO}(Q^k_{\mathcal{HP}}) \not\subseteq \mathbf{PFP}^{k-1}(\mathbf{Q}^{2k-1})$. *Thus the arity hierarchy is strict for the Hamiltonian path quantifier* $Q_{\mathcal{HP}}$.[37]

Proof: Note that $\mathbf{FO}(Q^{k-1}_{\mathcal{HP}}) = \mathbf{FO}(\{Q_{\mathcal{HP}_1}, \ldots, Q_{\mathcal{HP}_{k-1}}\}) \subseteq \mathbf{FO}(\mathbf{Q}^{2k-1})$ (since the arity of the quantifier $Q_{\mathcal{HP}_l}$ is $2l$). Thus the statement that the arity hierarchy is strict for the Hamiltonian path quantifier follows from $\mathbf{FO}(Q^k_{\mathcal{HP}}) \not\subseteq \mathbf{PFP}^{k-1}(\mathbf{Q}^{2k-1})$ (for each $k \geq 1$).

It is not hard to prove the statement for $k = 1$ (using similar ideas as described at the beginning of the previous paragraph to see that the transitive closure of directed graphs is not definable in $\mathbf{FO}(\mathbf{Q}^1)$). So we fix $k \geq 2$.

In Corollary 8.10 we have seen that the structure $\mathfrak{B}(k, n)$ can be divided into two connected components, and each of these components can be partitioned into 2^k disjoint paths from the first to the last row.

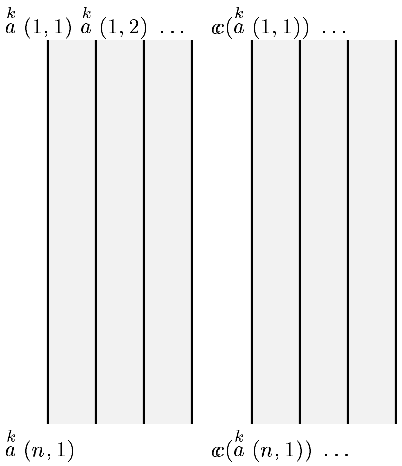

Figure 10.14

We let $\overset{k}{\boldsymbol{c}} = \overset{k}{\boldsymbol{a}}(1, 2^k)$, $\overset{k}{\boldsymbol{d}} = \overset{k}{\boldsymbol{a}}(n, 2^k)$, and, as before, $n = 2^{r+1} + 1$, $\mathfrak{B}_1 := (\mathfrak{B}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$, and $\mathfrak{B}_2 := (\mathfrak{B}, \overset{k}{\boldsymbol{c}}, æ(\overset{k}{\boldsymbol{d}}))$. Then Proposition 10.7 shows that $\mathfrak{B}_1$ and $\mathfrak{B}_2$ cannot be distinguished by a $PFP^{k-1}(\mathbf{Q}^{2k-1})$–sentence. (It does absolutely not matter which tuple $\overset{k}{\boldsymbol{a}}(1, i)$ $(i \leq 2^k)$ we take as $\overset{k}{\boldsymbol{c}}$ and which tuple $\overset{k}{\boldsymbol{a}}(n, i)$ we take as $\overset{k}{\boldsymbol{d}}$ since they all belong to the same component).

[37] For the definition of $Q_{\mathcal{HP}}$ see Example 5.7.

To see that the structures can be distinguished by a $FO(Q^k_{\mathcal{HP}})$–sentence we let

$$\begin{aligned}
\textsc{Path}^1(\overset{k}{x},\overset{k}{y}) := \quad & E_k\,\overset{k}{x}\overset{k}{y} \vee E_k\,\overset{k}{y}\overset{k}{x} \\
& \text{There is an } E_k\text{–edge between } \overset{k}{x} \text{ and } \overset{k}{y}, \\
\vee & \, (\neg\exists\, \overset{k}{z}\,(E_k\,\overset{k}{x}\overset{k}{z} \vee E_k\,\overset{k}{y}\overset{k}{z}) \wedge \exists\,\overset{k}{z}\,(E_k\,\overset{k}{z}\overset{k}{x} \wedge E_k\,\overset{k}{z}\overset{k}{y})) \\
& \text{or } \overset{k}{x} \text{ and } \overset{k}{y} \text{ are both in the last row and both belong} \\
& \text{to the same component,} \\
\vee & \, (\neg\exists\, \overset{k}{z}\,(E_k\,\overset{k}{z}\overset{k}{x} \vee E_k\,\overset{k}{z}\overset{k}{y}) \wedge \exists\,\overset{k}{z}\,(E_k\,\overset{k}{x}\overset{k}{z} \wedge E_k\,\overset{k}{y}\overset{k}{z})) \\
& \text{or } \overset{k}{x} \text{ and } \overset{k}{y} \text{ are both in the first row and both belong} \\
& \text{to the same component.}
\end{aligned}$$

Note that in structure $\mathfrak{B}$ there exists a $\textsc{Path}^1$–path from $\overset{k}{\boldsymbol{a}}(1,1)$ to $\overset{k}{\boldsymbol{a}}(1,2^k)$ that meets each tuple that belongs to the left component exactly once and a path from $\text{\ae}(\overset{k}{\boldsymbol{a}}(n,2^k))$ to $\text{\ae}(\overset{k}{\boldsymbol{a}}(n,1))$ that meets each tuple that belongs to the right component exactly once (see Figure 10.15).

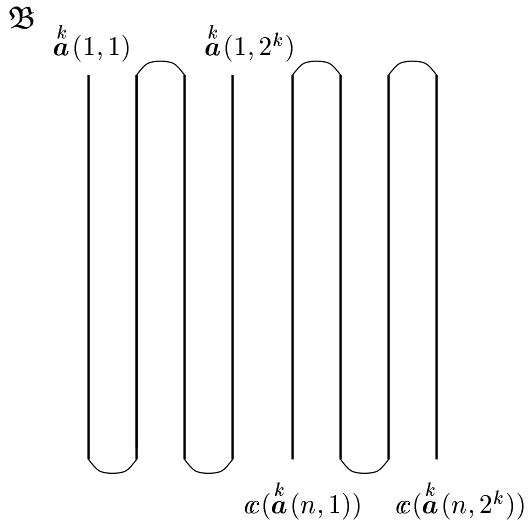

Figure 10.15

We let

$$\textsc{Path}^2(\overset{k}{x},\overset{k}{y}) := \textsc{Path}^1(\overset{k}{x},\overset{k}{y}) \vee (\overset{k}{x}{=}\underline{\overset{k}{c}} \wedge \overset{k}{y}{=}\underline{\overset{k}{d}})$$

and note that in $\mathfrak{B}_2$ there exists a $\textsc{Path}^2$–path that meets each k–tuple that occurs in an E_k–edge exactly once while in $\mathfrak{B}_1$ it does not (see Figure 10.16).

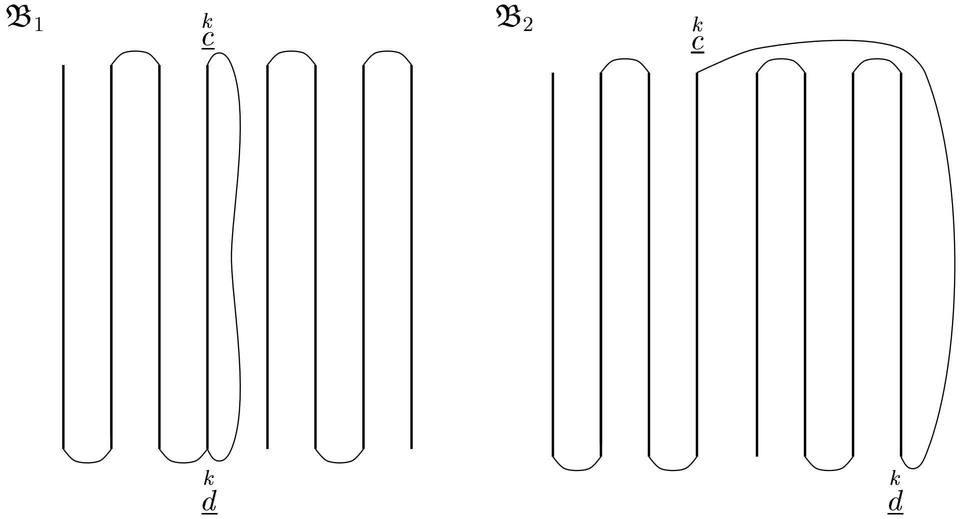

Figure 10.16

Since a Hamiltonian path does not only have to meet those k–tuples that occur in E_k–edges but all k–tuples we let

$$\textsc{Free}(\overset{k}{x}) := \neg\exists\, \overset{k}{y}\ (E_k\ \overset{k}{x}\overset{k}{y}\ \vee E_k\ \overset{k}{y}\overset{k}{x})$$

and

$$\begin{aligned}\textsc{Path}^3(\overset{k}{x},\overset{k}{y}) := \quad & \textsc{Path}^1(\overset{k}{x},\overset{k}{y})\\ & \vee (\textsc{Free}(\overset{k}{x})\ \wedge\ \textsc{Free}(\overset{k}{y}))\\ & \vee (\overset{k}{x}=\underline{\overset{k}{c}}\ \wedge \textsc{Free}(\overset{k}{y}))\\ & \vee (\textsc{Free}(\overset{k}{x}) \wedge\ \overset{k}{y}=\underline{\overset{k}{d}})\ .\end{aligned}$$

It is easy to see now that

$$\mathfrak{B}_1 \not\models [Q^k_{\mathcal{HP}}\ \overset{k}{x}\overset{k}{y}, \overset{k}{x}, \overset{k}{x}\ \ \textsc{Path}^3, x_1 = x_1, x_1 = x_1],$$

but

$$\mathfrak{B}_2 \models [Q^k_{\mathcal{HP}}\ \overset{k}{x}\overset{k}{y}, \overset{k}{x}, \overset{k}{x}\ \ \textsc{Path}^3, x_1 = x_1, x_1 = x_1].$$

■

In the next section we will see that the arity hierarchy of $\mathbf{DTC} = \mathbf{L}(Q^{<\omega}_{\mathcal{DTC}})$ is strict.
But in general the question whether the arity hierarchy of a vectorized Lindström logic is strict remains widely open:

10.17 Open Problem : *Find criteria for the arity hierarchy of a vectorized Lindström logic to be strict.*

11 Arities of Transitive Closure Operators

Our main lemma implies that the arity hierarchies are strict for transitive closure logic and its existential fragment. However, we have not yet solved the arity problem for deterministic transitive closure logic. The main result of this section is:

11.1 Theorem : *For each $k \geq 1$ there exists a graph–query which is definable in* $\mathbf{DTC}^k$*, but not in* $\mathbf{TC}^{k-1}$.

11.2 Corollary : *The arity hierarchy is strict for* **DTC***, even on the class of graphs.*

In the next paragraph we will see why we cannot strengthen this result (and the main lemma) to the statement that there exists a query in $\mathbf{DTC}^k \setminus \mathbf{s\text{–}PFP}^{k-1}$ (for a $k \geq 2$).

§ 11.1 DTC vs Fixed–Point Logics

For a formula $\varphi(\overset{k}{x}, \overset{k}{y})$ we let

$$\varphi_D(\overset{k}{x}, \overset{k}{y}) := \varphi(\overset{k}{x}, \overset{k}{y}) \wedge \forall \overset{k}{z} \, (\varphi(\overset{k}{x}, \overset{k}{z}) \rightarrow \overset{k}{z} = \overset{k}{y}) \; .$$

Recall that by the definition of the DTC–operator the formulae $[DTC_{\overset{k}{x}, \overset{k}{y}} \varphi] \; \overset{k}{u}, \overset{k}{v}$ and $[TC_{\overset{k}{x}, \overset{k}{y}} \varphi_D(\overset{k}{x}, \overset{k}{y})] \; \overset{k}{u}, \overset{k}{v}$ are equivalent.

11.3 Theorem : $\mathbf{DTC} \subseteq \mathbf{s\text{–}PFP}^1$

Proof: The idea is rather simple:
We can code a path $\overset{k}{a}_1, \ldots, \overset{k}{a}_n$ by sets X_{ij} $(1 \leq i \leq n, 1 \leq j \leq k)$ where $X_{ij} = \{a_{ij}\}$. If the path belongs to a deterministic transitive closure formula these sets can be defined to be the stages of a simultaneous fixed–point process. We need the path to be deterministic because otherwise we would get in trouble mixing up different successors of one tuple.

So let us consider the DTC–formula $\chi = [DTC_{\overset{k}{x}, \overset{k}{y}} \varphi] \; \overset{k}{u}, \overset{k}{v}$. For each $i \leq k$ we let

$$\begin{aligned}
\varphi_i := \quad & (X_1 = \emptyset \wedge \exists \overset{k}{y} \, (\varphi_D(\overset{k}{u}, \overset{k}{y}) \wedge x_i = y_i)) \\
& \vee \exists \overset{k}{z}, \overset{k}{y} \, (\textstyle\bigwedge_{j=1}^k X_j z_j \wedge \varphi_D(\overset{k}{z}, \overset{k}{y}) \wedge x_i = y_i) \\
& \vee \textstyle\bigwedge_{j=1}^k X_j v_j \qquad\qquad\qquad\qquad .
\end{aligned}$$

Let

$$\chi' := [s\text{–}FP_{x_1,X_1,\ldots,x_k,X_k}\ \varphi_1,\ldots,\varphi_k]x.$$

An easy induction shows that for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ and path

$$\alpha(\overset{k}{u}) = \overset{k}{a}_0, \overset{k}{a}_1, \ldots$$

such that $\mathfrak{I} \models \varphi_D[\overset{k}{a}_{i-1}, \overset{k}{a}_i]$ (for each $i \geq 1$) we have

$$X_{ij}^{\mathfrak{I}} = \begin{cases} \{a_{ij}\} & \text{if } \forall l < j : \overset{k}{a}_l \neq \alpha(\overset{k}{v}) \\ A & \text{otherwise} \end{cases}$$

(for each $i \geq 1$) Thus $\chi \mathrel{=\!\!|\!\models} \forall x \chi'$. ■

11.4 Lemma : *For each $k \geq 1$ we have* $\mathbf{s\text{–}PFP}^k \subseteq \mathbf{PFP}^{k+1}$.

Proof: We cannot just apply the Lemma on Simultaneous Induction because in our version it increases the arity by more than 1. However, we do not care for the quantifier structure of the formulae here and thus can prove a slightly different version of the Lemma on Simultaneous Induction that only increases the arity by one.
Consider the $s\text{–}FP^k$–formula

$$[s\text{–}FP_{\bar{x}_1,X_1,\ldots,\bar{x}_m,X_m}\ \varphi_1(\bar{x}_1, X_1, \ldots, X_m), \ldots, \varphi_m(\bar{x}_m, X_1, \ldots, X_m)]\bar{u};$$

without loss of generality we assume that the arity of each X_i is k.
It can be easily verified that in structures of size $\geq m$ this formula is equivalent to

$$\exists v_1, \ldots, v_m \Big(\bigwedge_{i<j\leq m} v_i \neq v_j \wedge [FP_{x_0\bar{x},X} \bigvee_{i=1}^{m} (x_0 = v_i \wedge \varphi_i(\bar{x}, Xv_1\_\_, \ldots, Xv_m\_\_))]v_1\bar{u}$$

where $v_1, \ldots, v_m$ are new individual variables and X is a new $(k+1)$–ary relation variable.
Structures of size $< m$ can be treated by a first–order formula. ■

11.5 Corollary : $\mathbf{DTC} \subseteq \mathbf{PFP}^2$

11.6 Open Problem : *Is* $\mathbf{DTC} \subseteq \mathbf{PFP}^1$ *?*

The situation is not so easy for **IFP**, but we can still prove that **DTC** is contained in $\mathbf{IFP}^2$. To see this we take two steps:

The first is to show that **DTC** is contained in $\mathbf{LFP}^2$ (hence $\mathbf{IFP}^2$) on ordered structures. This proof uses about the same ideas as did the proof of Theorem 11.3, but is a bit more complicated.
Analysing it we will see that we only need a definable order of sufficient length to obtain the result. (Which length is "sufficient" depends on the length of the deterministic path we want to simulate.)
Then, in the second step, we are going to show that indeed an order of suffficient length is definable in $\mathbf{IFP}^2$.

11.7 Theorem : $\mathbf{DTC} \subset_{\mathcal{O}} \mathbf{LFP}^2$

Proof: The problem we have here in contrast to the **PFP**–case is that we cannot remove tuples from our relation X in LFP–formulae. If we just let them in we get in trouble because we mix up tuples that belong to different stages. To avoid this we indicate for each tuple at which stage s it came in by the sth element of our order.

If n is the size of our structure we can code deterministic paths of length up to n by this. However, a k–ary deterministic path might have length n^k so we have to iterate our construction which leads to nested fixed–point operators.

For convenience we assume that the universe of an ordered structure is always an initial segment of $\mathbb{N}$. We can code the tuple $\overset{k}{x}$ which is the sth tuple of a deterministic path by the set $\{(s-1,0,x_1),\ldots,(s-1,k-1,x_k)\}$. To make this binary we code the first two places into one by a suitable pairing function:

Let us fix $k \geq 1$. Without loss of generality we only consider structures of size $\geq k$.
We define a function

$$<\cdot,\cdot>: \mathbb{N} \times \{1,\ldots,k\} \longrightarrow \mathbb{N}$$

by $<s,l>= sk+l-1$ and functions

$$L:\mathbb{N}\longrightarrow\mathbb{N} \quad \text{and} \quad R:\mathbb{N}\longrightarrow\{1,\ldots,k\}$$

in a way that $L(<s,l>)=s$ and $R(<s,l>)=l$.

Since it is easy to define addition and multiplication in $\mathbf{LFP}^2$ on ordered structures we can find LFP^2–formulae $\psi_{<\cdot,\cdot>}(x,y,z)$, $\psi_L(x,z)$, and $\psi_R(x,z)$ defining the respective functions on ordered structures.

Let $\varphi(\overset{k}{x},\overset{k}{y})$ be an LFP^2–formula. We shall define an LFP^2–formula $\chi'(\overset{k}{t},\overset{k}{u})$ that is equivalent to

$$\chi(\overset{k}{t},\overset{k}{u}) := [DTC_{\overset{k}{x},\overset{k}{y}}\varphi(\overset{k}{x},\overset{k}{y})]\ \overset{k}{t},\overset{k}{u}\ .$$

Without loss of generality we assume that $\chi(\overset{k}{t},\overset{k}{u}) \models \overset{k}{t} \neq \overset{k}{u}$.

(If this is not the case for the original $\chi(\overset{k}{t},\overset{k}{u})$ we consider the equivalent formula

$$\varphi_D(\overset{k}{t},\overset{k}{u}) \vee \exists \overset{k}{v} \left([DTC_{\overset{k}{x},\overset{k}{y}}\varphi\wedge \overset{k}{t}\neq\overset{k}{v}]\ \overset{k}{t},\overset{k}{v} \wedge[DTC_{\overset{k}{x},\overset{k}{y}}\varphi\wedge \overset{k}{u}\neq\overset{k}{v}]\ \overset{k}{v},\overset{k}{u}\ \right)$$

instead.)

To construct χ' we first define an LFP^2–formula $\varphi^{\frac{n}{\llcorner k \lrcorner}}(\overset{k}{x},\overset{k}{y})$ saying

> *"There exists a deterministic φ–path of length $(\frac{n}{\llcorner k \lrcorner})+1$ from $\overset{k}{x}$ to $\overset{k}{y}$."*

by:

$$\bigwedge_{i=1}^{k}\Big[FP^{2}_{z,Z}\quad \exists \overset{k}{w}\ (\varphi_D(\overset{k}{x},\overset{k}{w})\ \wedge\ \bigvee_{i=1}^{k}(z_1 =< 1,i > \wedge z_2 = w_i)$$

We start the path with the φ_D–successor of $\overset{k}{x}$.

$$\vee \exists \overset{k}{v} \exists \overset{k}{w} \exists t\Big(\bigwedge_{i=1}^{k} Z < t,i > v_i \wedge \varphi_D(\overset{k}{v},\overset{k}{w})$$
$$\wedge \bigvee_{i=1}^{k}(z_1 =< t+1,i > \wedge z_2 = w_i)\Big)\Big] < \tfrac{n}{\llcorner k \lrcorner}, i > \ y_i.$$

Then we walk along the φ_D–path $\frac{n}{\llcorner k \lrcorner}$ steps.

(It is clear how this formula can be turned into a well–defined LFP^2–formula using the formulae $\psi_{<\cdot,\cdot>}(x,y,z)$, $\psi_L(x,z)$, and $\psi_R(x,z)$ and explicit definitions for the numbers $1,\ldots,k$ and $\frac{n}{\llcorner k \lrcorner}$.)

Note that $\varphi^{\frac{n}{\llcorner k \lrcorner}}(\overset{k}{x},\overset{k}{y}) \dashv\models \varphi_D^{\frac{n}{\llcorner k \lrcorner}}(\overset{k}{x},\overset{k}{y})$.

Now we let $\varphi^0 := \varphi$ and $\varphi^{i+1} := (\varphi^i)^{\frac{n}{\llcorner k \lrcorner}}$ for each $i \geq 1$. It is easy to see that φ^i is an LFP^2–formula that says:

> *"There exists a deterministic φ–path of length $((\frac{n}{\llcorner k \lrcorner})^i+1)$ from $\overset{k}{x}$ to $\overset{k}{y}$."*

However, actually we want to define the query

> *"There exists a deterministic φ–path of length $\leq ((\frac{n}{\llcorner k \lrcorner})^i+1)$ from $\overset{k}{t}$ to $\overset{k}{u}$."*

Then we are finished because we can always select an i such that $(\frac{n}{\llcorner k \lrcorner})^i+1 \geq n^k$.

The difficulty is that the formulae $\varphi^i(\overset{k}{x},\overset{k}{y})$ used in the construction need to be deterministic (i.e. $\varphi^i(\overset{k}{x},\overset{k}{y}) \dashv\models \varphi^i_D(\overset{k}{x},\overset{k}{y})$), and obviously a formula saying that there exists a deterministic φ–path of length $\leq ((\frac{n}{\llcorner k \lrcorner})^i+1)$ from $\overset{k}{x}$ to $\overset{k}{y}$ is not deterministic.

To cope with this we start the construction with the formula

$$\tilde{\varphi}(\overset{k}{x},\overset{k}{y}) := \varphi(\overset{k}{x},\overset{k}{y}) \vee (\overset{k}{x}=\overset{k}{u} \wedge \overset{k}{y}=\overset{k}{u})$$

instead of φ.
Then if $\overset{k}{x}\neq\overset{k}{u}$ the formula $\tilde{\varphi}^i(\overset{k}{x},\overset{k}{u})$ in fact says

> *"There exists a deterministic φ–path of length $\leq ((\frac{n}{\llcorner k \lrcorner})^i+1)$ from $\overset{k}{x}$ to $\overset{k}{u}$."*

Thus for an i with $(\frac{n}{\llcorner k \lrcorner})^i + 1 \geq n^k$ we have:

$$\chi(\overset{k}{t},\overset{k}{u}) \dashv\models \underbrace{(\tilde{\varphi}^i(\overset{k}{t},\overset{k}{u}) \wedge \overset{k}{t}\neq\overset{k}{u})}_{=:\chi'(\overset{k}{t},\overset{k}{u})} .$$

■

Consider now a formula $\varphi(\overset{k}{x},\overset{k}{y})$, an interpretation $\mathfrak{I} = (\mathfrak{A},\alpha)$, and a k–tuple $\overset{k}{a}\in A$. We define the *deterministic path-length of* $\varphi(\overset{k}{x},\overset{k}{y})$ *in* $\mathfrak{I}$ *starting at* $\overset{k}{a}$ to be

$$\mathrm{dpl}(\varphi,\mathfrak{I},\overset{k}{a}) := \max\Big\{m \geq 1 \mid \exists\, \overset{k}{a}=\overset{k}{a}_1,\ldots,\overset{k}{a}_m\in A \text{ such that } \forall i < j \leq m : \overset{k}{a}_i\neq\overset{k}{a}_j \text{ and } \forall i < m : \mathfrak{I} \models \varphi_D[\overset{k}{a}_i,\overset{k}{a}_{i+1}] \Big\}.$$

We say that a formula $\psi(x,y)$ *defines an order* if for each interpretation $\mathfrak{I}$ the binary relation $\psi(x,y)^{\mathfrak{I}}$ is a linear order of its support[38].
The *length* $l(\psi,\mathfrak{I})$ *of* ψ *in* $\mathfrak{I}$ is defined to be the cardinality of the support of $\psi(x,y)^{\mathfrak{I}}$.

Similarly to Theorem 11.7 we can prove:

11.8 Lemma : *Consider a formula* $\chi = [DTC_{\overset{k}{x},\overset{k}{y}}\varphi]\ \overset{k}{u},\overset{k}{v}$ *where* $\varphi \in \mathbf{IFP}^2$ *and suppose that there exists an* $\mathbf{IFP}^2$*–formula* $\psi(x,y)$ *defining an order such that for each interpretation* $\mathfrak{I} = (\mathfrak{A},\alpha)$ *we have*

$$l(\psi,\mathfrak{I})^k \geq \mathrm{dpl}(\varphi,\mathfrak{I},\alpha(\overset{k}{u})) .$$

Then there exists an $\mathbf{IFP}^2$*–formula that is equivalent to* χ.

So we shall prove:

[38] The support of a relation R on a set A is the set of all elements of A that occur in a tuple $\bar{a}$ with $R\bar{a}$.

11.9 Lemma : *Let $\varphi(\overset{k}{x}, \overset{k}{y}) \in \mathbf{IFP}^2$ and let $\overset{k}{u}$ be a tuple of terms. Then there exists an $\mathbf{IFP}^2$–formula $\psi(x,y)$ that defines an order such that for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ we have*

$$l(\psi, \mathfrak{I})^k \geq \mathrm{dpl}(\varphi, \mathfrak{I}, \alpha(\overset{k}{u})) \ .$$

Proof: The idea is to define inductively an order $<_{\varphi, \overset{k}{u}}$ as follows:

- We start with the order $u_1 <_{\varphi, \overset{k}{u}} u_2 <_{\varphi, \overset{k}{u}} \ldots <_{\varphi, \overset{k}{u}} u_k$.
- The order defined up to now induces a (lexicographic) order of the k–tuples in its support S. We consider the first tuple $\overset{k}{x} \in S^k$ with respect to this order that has a φ_D–successor $\overset{k}{y} \notin S^k$. We order those elements in $\overset{k}{y}$ that are not contained in S by their indices and extend our order by these elements.

Let X be a binary relation variable and

$$\begin{aligned} \textsc{Supp}_X(x) &:= \exists z (Xxz \lor Xzx) \\ \textsc{Lex}_X(\overset{k}{x}, \overset{k}{y}) &:= \bigvee_{i=1}^{k} (Xx_iy_i \land \bigwedge_{j=i+1}^{k} x_j = y_j) \end{aligned}$$

Consider the formula

$$\begin{aligned} \xi := \quad & Xxy \lor \textstyle\bigvee_{i<j\leq k}(x = u_i \land y = u_j) \\ & \lor \exists \overset{k}{x}, \overset{k}{y} \Big(\textstyle\bigwedge_{i=1}^{k} \textsc{Supp}_X(x_i) \land \bigvee_{i=1}^{k} \neg\textsc{Supp}_X(y_i) \land \varphi_D(\overset{k}{x}, \overset{k}{y}) \\ & \qquad \land \forall \overset{k}{z}, \overset{k}{z}{}' (\textsc{Lex}_X(\overset{k}{z}, \overset{k}{x}) \land \varphi_D(\overset{k}{z}, \overset{k}{z}{}') \to \textstyle\bigwedge_{i=1}^{k} \textsc{Supp}_X(z_i')) \\ & \qquad \land (\quad (\textsc{Supp}_X(x) \land \neg\textsc{Supp}_X(y) \land \textstyle\bigvee_{i=1}^{k} y = y_i) \\ & \qquad\quad \lor (\neg\textsc{Supp}_X(x) \land \neg\textsc{Supp}_X(y) \land \textstyle\bigvee_{i<j\leq k}(x = y_i \land y = y_j)))\Big). \end{aligned}$$

Then $\psi(x,y) := [FP_{x,y,X}\xi]xy \in IFP^2$ defines the order $<_{\varphi, \overset{k}{u}}$ described above. Moreover, we indeed have

$$l(\psi, \mathfrak{I})^k \geq \mathrm{dpl}(\varphi, \mathfrak{I}, \alpha(\overset{k}{u}))$$

for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ since all elements of A that occur in a tuple in the φ_D–path in $\mathfrak{A}$ starting with $\alpha(\overset{k}{u})$ are contained in the support of the order defined by ψ, and at most $l(\psi, \mathfrak{I})^k$ distinct k–tuples can be formed with the $l(\psi, \mathfrak{I})$ elements of the support. ■

An immediate consequence of the last two lemmata is:

11.10 Theorem : $\mathbf{DTC} \subseteq \mathbf{IFP}^2$

11.11 Open Problem : *Is* $\mathbf{DTC} \subseteq \mathbf{IFP}^1$ *?*

Using the usual techniques to translate an *IFP* in to an *LFP*–formula we obtain:

11.12 Corollary : $\mathbf{DTC} \subseteq \mathbf{LFP}^5$

In fact it does not seem to be too hard to show $\mathbf{DTC} \subseteq \mathbf{LFP}^3$ by giving a direct proof of the analogue of Lemma 11.9 for $\mathbf{LFP}^3$. I would conjecture that we even have $\mathbf{DTC} \subseteq \mathbf{LFP}^2$, although I cannot see a proof for that.

11.13 Open Problem : *Is* $\mathbf{DTC} \subseteq \mathbf{LFP}^2$ *?*

On the other hand it is known that $\mathbf{DTC} \not\subseteq \mathbf{LFP}^1$ is an easy consequence of Büchi's theorem. The idea to see this is as follows (as far as I know, it is due to Makowsky):
Büchi's theorem says that on the class $\mathcal{W}$ of word–structures (orders with unary relations; they can be considerd as words) the class of monadic second–order–definable queries coincides with the regular languages. Now the class of words of the form $a^n b^n$ $(n \geq 1)$ can be easily defined by a DTC^2–formula, but it does not form a regular language. Since LFP^1 is contained in monadic second–order logic the statement is proved.

Concerning the relation between **SFP** and **DTC** it seems to be possible to strengthen Theorem 11.1 to the statement that for each $k \geq 1$ there exists a graph–query in $\mathbf{DTC}^k \setminus \mathbf{SFP}^{k-1}$ using similar techniques as in the proof of the original theorem given in the next paragraph. However, we are not going to investigate this strengthening.

Finally, let us briefly consider the situation on ordered structures. Recall Theorem 11.7 saying $\mathbf{DTC} \subseteq_{\mathcal{O}} \mathbf{LFP}^2$.
In [Imh94] Henrik Imhof studied similar questions. Among other things he proved that $\mathbf{DTC} \subseteq_{\mathcal{O}} \mathbf{IFP}^1$ and $\mathbf{TC} \subseteq_{\mathcal{O}} \mathbf{PFP}^1$, and that for each $k \geq 1$:

$$\mathbf{PFP}^k \subsetneq_{\mathcal{O}} \mathbf{PFP}^{k+1}.$$

This also has consequences for the analoguous question for **LFP**:
If there exists a $k \geq 1$ such that $\mathbf{LFP} =_{\mathcal{O}} \mathbf{LFP}^k$ then $\mathbf{LFP} \neq \mathbf{PFP}$ and hence $\mathbf{PTIME} \neq \mathbf{PSPACE}$. On the other hand, if $\mathbf{LFP}^2 \neq_{\mathcal{O}} \mathbf{LFP}$ then $\mathbf{DTC} \neq_{\mathcal{O}} \mathbf{LFP}$ (by Theorem 11.7) hence $\mathbf{LOGSPACE} \neq \mathbf{PTIME}$. (Here we use well–known results from descriptive complexity theory due to Immerman [Imm86, Imm87] and Vardi [Var82] (Theorems 1.17, 4.12).)

Thus the following problem seems to be difficult:

11.14 Open Problem : *Is the arity hierarchy of* **LFP** *strict on ordered structures?*[39]

Except that $\mathbf{DTC}^2 \not\subseteq \mathbf{LFP}^1 \supseteq \mathbf{TC}^1$ on $\mathcal{O}$ (as we have seen above with Büchi's theorem) we do not know anything about:

11.15 Open Problem : *Are the arity hierarchies of* **TC** *and* **DTC** *strict on ordered structures?*

§11.2 The Structures

The rest of this section is devoted to the proof of the hierarchy Theorem 11.1.
The strategy of this proof is similar to the strategy of the proof of the main lemma: We are going to define a class of graphs which has a certain property of homogeneity and show with the help of an Ehrenfeucht-Fraïssé game that the deterministic transitive closure of a certain first–order formula $\text{D–EDGE}_k(\overset{k}{x}, \overset{k}{y})$ is not definable in $\mathbf{TC}^{k-1}$ on this class.
The graphs we need here are very similar to those we used before. But to make a k-ary *DTC*-Operater applicable each k-tuple should have at most one D–EDGE_k-successor and at most one D–EDGE_k-predecessor. The price for this is that we cannot find automorphisms of the structures that move a $(k-1)$–tuple of elements to another component and leave other rows fixed. But it will suffice here to have automorphisms that move some $(k-1)$–tuples and leave $(k-1)$–tuples in different rows fixed.
To deal with nested *TC*–operators we have to iterate the construction.

11.16 Theorem : *Let $k, n \geq 2, t \geq 1$. Then there exists a graph $\mathfrak{G}^t = \mathfrak{G}^t(k,n)$ and a first–order formula $\text{D–EDGE}_k(\overset{k}{x}, \overset{k}{y})$ such that*

(1) There exists a mapping row : $G^t \longrightarrow \{1, \ldots, n\}$ *such that*

$$\forall \boldsymbol{a}, \boldsymbol{b} \in G^t : (E^{\mathfrak{G}^t} \boldsymbol{ab} \Longrightarrow \text{row}(\boldsymbol{b}) - \text{row}(\boldsymbol{a}) \leq 1) \ .$$

(2) There exists an automorphism $æ$ of $\mathfrak{G}^t$ that is self–inverse and preserves the rows.

(3) There exist tuples $\overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \in G^t$ in the first and last row respectively such that

$$\mathfrak{G}^t \models [DTC_{\overset{k}{x}, \overset{k}{y}} \text{D–EDGE}_k(\overset{k}{x}, \overset{k}{y})]\ \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}}$$

$$\textit{and} \quad \mathfrak{G}^t \not\models [DTC_{\overset{k}{x}, \overset{k}{y}} \text{D–EDGE}_k(\overset{k}{x}, \overset{k}{y})]\ \overset{k}{\boldsymbol{c}}, æ(\overset{k}{\boldsymbol{d}}) \ .$$

[39] I.e. is $\mathbf{LFP}^k \subsetneq_{\mathcal{O}} \mathbf{LFP}^{k+1}$ for each $k \geq 1$?

(4) *There exist sets $\mathcal{F}^s$ $(0 \leq s \leq t)$ of automorphisms of $\mathfrak{G}^t$ (which are in fact subgroups of the automorphism–group) such that $\mathcal{F}^0 = \{id, æ\}$ and for each $s \leq t$ the following holds:*

Let $A, B \subset G^t$ such that $\forall \boldsymbol{a} \in A, \boldsymbol{b} \in B : |\mathrm{row}(\boldsymbol{a}) - \mathrm{row}(\boldsymbol{b})| > 1$ and in each row of $\mathfrak{G}^t$ there are at most $(k-1)$ elements of $A \cup B$. Then for all $f_*, f^* \in \mathcal{F}^{s-1}$ there exists an $f \in \mathcal{F}^s$ which is self–inverse and preserves the rows such that

$$\forall \boldsymbol{a} \in A : f(\boldsymbol{a}) = f_*(\boldsymbol{a}) \quad \textit{and} \quad \forall \boldsymbol{b} \in B : f(\boldsymbol{b}) = f^*(\boldsymbol{b}) \ .$$

We fix $k, n \geq 2$ for the construction.
Let $P_1, \ldots, P_k$ be unary relation symbols. It is convenient to construct $\{E, P_1, \ldots, P_k\}$–structures with the desired properties first; they can easily be turned into graphs later. The relation P_i is intended to contain the elements of the ith column.
We let

$$\text{D–E}{\small\text{DGE}}_k(\overset{k}{x}, \overset{k}{y}) := \bigwedge_{1 \leq i,j \leq k} Ex_iy_j \wedge \bigwedge_{1 \leq i \leq k} (P_ix_i \wedge P_iy_i) \ .$$

As the reader might suspect we start our construction with a structure $\mathfrak{G}^0$ consisting of two disjoint D–Edge$_k$–paths. Formally it is defined as follows:
Let Γ^0 be the group $(\mathbb{Z}_2, +)$ and

$$\begin{aligned} G^0 &:= \{1, \ldots, n\} \times \{1, \ldots, k\} \times \Gamma^0 \\ E^{\mathfrak{G}_0} &:= \{((I, a, \gamma), (I+1, b, \gamma)) \mid 1 \leq I < n,\ 1 \leq a, b \leq k,\ \gamma \in \Gamma^0\} \end{aligned}$$

and for each $i \leq k$

$$P_i^{\mathfrak{G}^0} \ := \ \{(I, i, \gamma) \mid 1 \leq I \leq n, \gamma \in \Gamma^0\}.$$

The automorphism $æ$ exchanging the components of $\mathfrak{G}^0$ is defined by

$$æ((I, a, \gamma)) = (I, a, \gamma + 1) \ .$$

We let $\mathcal{F}^0 = \{id, æ\}$.
As usual we let $\mathrm{row}((I, a, \gamma)) := I$ and $\mathrm{col}((I, a, \gamma)) := a$.

We are going to define structures $\mathfrak{G}^t$ for $t \geq 1$ inductively. To define $\mathfrak{G}^t$ we do not only need the previously defined structures $\mathfrak{G}^s$ for $s < t$, but also the groups Γ^s that were used to define $\mathfrak{G}^s$ and groups $\mathcal{F}^s$ of automorphisms of $\mathfrak{G}^s$. Moreover, the mappings row and col need to be extended to $\mathfrak{G}^s$.
So suppose $t \geq 1$ and everything is defined for each $s < t$.

For all $f \in \mathcal{F}^{t-1}$, $R_A, R_B \subseteq \{1, \dots, n\}$ such that $\forall I \in R_A, J \in R_B : |I - J| > 1$, and $a_I \in \{1, \dots, k\}$ (for each $I \in R_A \cup R_B$) we define a partial bijection of $\mathfrak{G}^{t-1}$ by

$$\boldsymbol{a} \mapsto \begin{cases} \boldsymbol{a} & \text{if } I := \mathrm{row}(\boldsymbol{a}) \in R_A \text{ and } \mathrm{col}(\boldsymbol{a}) \neq a_I \\ f(\boldsymbol{a}) & \text{if } J := \mathrm{row}(\boldsymbol{a}) \in R_B \text{ and } \mathrm{col}(\boldsymbol{a}) \neq a_J \\ (\text{undefined} & \text{otherwise}) \end{cases} .$$

We let $p_1^t, \dots, p_{l_t}^t$ be an enumeration of all these partial bijections and $\mathtt{P}^t$ the free group generated by $\mathtt{p}_1^1, \dots, \mathtt{p}_{l_t}^t$.

Note that for each $\mathtt{p} \in \mathtt{P}^t$ the corresponding partial bijection p is a partial isomorphism.
As usual we let $\Gamma^t := (\mathbb{Z}_2^{l_t}, +)$ and define the mapping ρ^t to be the extension of $\mathtt{p}_i^t \longmapsto (0, \dots, 0, 1, 0, \dots, 0)$ (with the 1 at the ith position) to a homomorphism from $\mathtt{P}^t$ to Γ^t.
We let $\mathfrak{G}^t := \mathfrak{H}(\mathfrak{G}^{t-1}, p_1^t, \dots, p_{l_t}^t, \Gamma^t, \rho^t)$.

To simplify the notation we let $\bar{\Gamma} = \bigoplus_{i=0}^t \Gamma^i$ and write $\bar{\gamma}, \bar{\delta}, \dots$ or $(\gamma_0, \dots, \gamma_t)$, $(\delta_0, \dots, \delta_t), \dots$ for elements of $\bar{\Gamma}$. Elements of $\mathfrak{G}^t$ are written in the form $(I, a, \bar{\gamma})/_\sim$ or $(I, a, \gamma_0, \dots, \gamma_t)/_\sim$ instead of

$$((\dots((I, a, \gamma_0), \gamma_1)/_\sim, \dots, \gamma_{t-1})/_\sim, \gamma_t)/_\sim.$$

For each $\bar{\gamma} \in \bar{\Gamma}$ we define a mapping $f_{\bar{\gamma}} : \mathfrak{G}^t \longrightarrow \mathfrak{G}^t$ by

$$f_{\bar{\gamma}}((I, a, \bar{\delta})/_\sim) = (I, a, \bar{\delta} + \bar{\gamma})/_\sim \ ,$$

and we let $\mathcal{F}^t := \{f_{\bar{\gamma}} \mid \bar{\gamma} \in \bar{\Gamma}\}$. We will have to prove (cf. Lemma 11.17) that the $f_{\bar{\gamma}}$ are well–defined. Once we have done this it is obvious that each $f_{\bar{\gamma}}$ is self–inverse, that for $\bar{\gamma}, \bar{\delta} \in \bar{\Gamma}$ we have $f_{\bar{\gamma}} \circ f_{\bar{\delta}} = f_{\bar{\delta}} \circ f_{\bar{\gamma}}$, and that $\mathcal{F}^t$ is closed under concatenation.

However, having seen that the $f_{(\gamma_0, \dots, \gamma_{t-1})} \in \mathcal{F}^{t-1}$ have these properties it is obvious that Γ^t and ρ^t are suitable for $p_1^t, \dots, p_{l_t}^t$ (since for all $i, j \leq l_t$ we have $(p_i^t)^{-1} = p_i^t$ and $p_i^t p_j^t = p_j^t p_i^t$ by the analogous properties of the $f_{(\gamma_0, \dots, \gamma_{t-1})} \in \mathcal{F}^{t-1}$ and the choice of the domains of the p_i^t).
Note also that the $f_{(\gamma_0, \dots, \gamma_{t-1})} \in \mathcal{F}^{t-1}$ respect rows and columns hence the mappings row and col defined as usual by

$$\mathrm{row}((I, a, \bar{\gamma})/_\sim) := I \quad \text{and} \quad \mathrm{col}((I, a, \bar{\gamma})/_\sim) := a$$

are well–defined.

Before we prove the next lemma we should have a closer look at the partial bijections p for $\mathtt{p} \in \mathtt{P}^t$. By its definition, on each row each p_i^t $(1 \leq i \leq l_t)$ is the restriction of an $f \in \mathcal{F}^{t-1}$. Thus for each fixed row I, on the intersection of this

row with its domain p is the concatenation of mappings in $\mathcal{F}^{t-1}$. Since $\mathcal{F}^{t-1}$ is closed under concatenation it is itself in $\mathcal{F}^{t-1}$. Hence by the definition of $\mathcal{F}^{t-1}$ we have:
For all $I \leq n, \mathtt{p} \in \mathtt{P}^t$ there exist $\zeta_0 \in \Gamma^0, \dots, \zeta_{t-1} \in \Gamma^{t-1}$ such that

$$\forall \boldsymbol{a} \in \mathrm{dom}(p) : \left(\mathrm{row}(\boldsymbol{a}) = I \Longrightarrow \boldsymbol{a}^p = f_{(\zeta_0,\dots,\zeta_{t-1})}(\boldsymbol{a})\right) .$$

11.17 Lemma : *For each $\bar{\gamma} \in \bar{\Gamma}$ the mapping $f_{\bar{\gamma}}$ is an automorphism of $\mathfrak{G}^t$.*

Proof by induction over t:
Clearly the lemma is true for $t = 0$. Let $t \geq 1$ and suppose it is true for $t-1$. Let $(\gamma_0, \dots, \gamma_t) \in \bar{\Gamma}$.
First, we have to prove that $f_{(\gamma_0,\dots,\gamma_t)}$ is well–defined which is a bit tiring:

$$\begin{aligned} & (I, a, \delta_0, \dots, \delta_t)/_\sim = (I, a, \eta_0, \dots, \eta_t)/_\sim \\ \Longleftrightarrow\ & \exists \mathtt{p} \in (\rho^t)^{-1}(\eta_t - \delta_t) : \\ & \qquad (I, a, \delta_0, \dots, \delta_{t-1})/_\sim = ((I, a, \eta_0, \dots, \eta_{t-1})/_\sim)^p. \end{aligned}$$

Choose $\zeta_0 \in \Gamma^0, \dots, \zeta_{t-1} \in \Gamma^{t-1}$ such that

$$\forall \boldsymbol{a} \in \mathrm{dom}(p) : \left(\mathrm{row}(\boldsymbol{a}) = I \Longrightarrow \boldsymbol{a}^p = f_{(\zeta_0,\dots,\zeta_{t-1})}(\boldsymbol{a})\right) .$$

Then

$$\begin{aligned} & (I, a, \delta_0, \dots, \delta_{t-1})/_\sim = ((I, a, \eta_0, \dots, \eta_{t-1})/_\sim)^p \\ \Longleftrightarrow\ & (I, a, \delta_0, \dots, \delta_{t-1})/_\sim = (I, a, \eta_0 + \zeta_0, \dots, \eta_{t-1} + \zeta_{t-1})/_\sim \\ \Longleftrightarrow\ & (I, a, \delta_0 + \gamma_0, \dots, \delta_{t-1} + \gamma_{t-1})/_\sim = \\ & \qquad (I, a, \eta_0 + \zeta_0 + \gamma_0, \dots, \eta_{t-1} + \zeta_{t-1} + \gamma_{t-1})/_\sim \\ & \text{(By the induction hypothesis } f_{(\gamma_0,\dots,\gamma_{t-1})} \in \mathcal{F}^{t-1} \text{ is a bijection.)} \\ \Longleftrightarrow\ & (I, a, \delta_0 + \gamma_0, \dots, \delta_{t-1} + \gamma_{t-1})/_\sim = \\ & \qquad ((I, a, \eta_0 + \gamma_0, \dots, \eta_{t-1} + \gamma_{t-1})/_\sim)^p \\ & \text{(since } \bar{\Gamma} \text{ is Abelian).} \end{aligned}$$

Remembering $\mathtt{p} \in (\rho^t)^{-1}(\eta_t - \delta_t)$ this implies

$$\begin{aligned} & (I, a, \delta_0 + \gamma_0, \dots, \delta_{t-1} + \gamma_{t-1}, \delta_t)/_\sim = \\ & \qquad (I, a, \eta_0 + \gamma_0, \dots, \eta_{t-1} + \gamma_{t-1}, \eta_t)/_\sim \\ \Longleftrightarrow\ & (I, a, \delta_0 + \gamma_0, \dots, \delta_{t-1} + \gamma_{t-1}, \delta_t + \gamma_t)/_\sim = \\ & \qquad (I, a, \eta_0 + \gamma_0, \dots, \eta_{t-1} + \gamma_{t-1}, \eta_t + \gamma_t)/_\sim \\ & \text{(By Lemma 7.6 } f_{\gamma_t} \text{ is an automorphism of } \mathfrak{G}^t, \\ & \text{in particular well–defined and bijective.)} \\ \Longleftrightarrow\ & f_{(\gamma_0,\dots,\gamma_t)}((I, a, \delta_0, \dots, \delta_t)/_\sim) = f_{(\gamma_0,\dots,\gamma_t)}((I, a, \eta_0, \dots, \eta_t)/_\sim) \end{aligned}$$

$f_{(\gamma_0,\dots,\gamma_t)}$ is bijective because

$$f_{(-\gamma_0,\dots,-\gamma_t)} \circ f_{(\gamma_0,\dots,\gamma_t)} = id \,.$$

To see that $f_{(\gamma_0,\dots,\gamma_t)}$ is indeed an automorphism, let $\boldsymbol{a}, \boldsymbol{b} \in G^t$ such that $E^{\mathfrak{G}^t}\boldsymbol{ab}$. Then there exist $I < n$, $a, b \in \{1, \dots, k\}$, $\delta_0, \zeta_0 \in \Gamma^0, \dots, \delta_{t-1}, \zeta_{t-1} \in \Gamma^{t-1}$, and $\eta \in \Gamma^t$ such that

$$\boldsymbol{a} = (I, a, \delta_0, \dots, \delta_{t-1}, \eta)/_\sim, \quad \boldsymbol{b} = (I+1, b, \zeta_0, \dots, \zeta_{t-1}, \eta)/_\sim,$$

and

$$E^{\mathfrak{G}^{t-1}}(I, a, \delta_0, \dots, \delta_{t-1})/_\sim(I+1, b, \zeta_0, \dots, \zeta_{t-1})/_\sim \,.$$

By the induction hypothesis, $f_{(\gamma_0,\dots,\gamma_{t-1})} \in \mathcal{F}^{t-1}$ is an automorphism, hence

$$\begin{aligned}
&E^{\mathfrak{G}^{t-1}}(I, a, \delta_1+\gamma_0, \dots, \delta_{t-1}+\gamma_{t-1})/_\sim(I+1, b, \zeta_1+\gamma_0, \dots, \zeta_{t-1}+\gamma_{t-1})/_\sim \\
\Longrightarrow\ &E^{\mathfrak{G}^t}(I, a, \delta_1+\gamma_0, \dots, \delta_{t-1}+\gamma_{t-1}, \eta)/_\sim(I+1, b, \zeta_1+\gamma_0, \dots, \zeta_{t-1}+\gamma_{t-1}, \eta)/_\sim \\
&\text{(By Lemma 7.5.)} \\
\Longrightarrow\ &E^{\mathfrak{G}^t} f_{(\gamma_0,\dots,\gamma_t)}(\boldsymbol{a}) f_{(\gamma_0,\dots,\gamma_t)}(\boldsymbol{b}) \\
&\text{(By Lemma 7.6).}
\end{aligned}$$

Clearly $f_{(\gamma_0,\dots,\gamma_t)}$ also preserves the relations $P_i^{\mathfrak{G}^t}$ since it preserves the columns. ■

For $s < t$ we can identify $\mathcal{F}^s$ with the subset $\{f_{\bar\gamma} \mid \bar\gamma \in \bigoplus_{i=0}^s \Gamma^i\}$ of $\mathcal{F}^t$. But we must be aware that $f_{(\gamma_0,\dots,\gamma_s)}$ is an automorphism of $\mathfrak{G}^s$ and that $f_{(\gamma_0,\dots,\gamma_s,0,\dots,0)}$ extends it to an automorphism of $\mathfrak{G}^t$.
In particular $f_{(1,0,\dots,0)}$ extends α to an automorphism of $\mathfrak{G}^t$.

Now we can already prove that the structure $\mathfrak{G}^t$ has the property (4) stated in Theorem 11.16:

11.18 Lemma : *Let $A, B \subset G^t$ such that $\forall \boldsymbol{a} \in A, \boldsymbol{b} \in B : |\mathrm{row}(\boldsymbol{a}) - \mathrm{row}(\boldsymbol{b})| > 1$ and in each row of $\mathfrak{G}^t$ there are at most $(k-1)$ elements of $A \cup B$. Then for all $s \le t$, $f_*, f^* \in \mathcal{F}^{s-1}$ there exists an $f \in \mathcal{F}^s$ which is self–inverse and preserves the rows such that*

$$\forall \boldsymbol{a} \in A : f(\boldsymbol{a}) = f_*(\boldsymbol{a}) \quad \textit{and} \quad \forall \boldsymbol{b} \in B : f(\boldsymbol{b}) = f^*(\boldsymbol{b}) \,.$$

Proof: Say, $f_*^{-1} \circ f^*$ is the automorphism $f_{(\gamma_0,\dots,\gamma_{s-1},0,\dots,0)}$ of $\mathfrak{G}^t$.
We can assume without loss of generality that A and B are of the following form (otherwise we extend them):

There exist sets $R_A, R_B \subseteq \{1, \dots, n\}$ such that $\forall I \in R_A, J \in R_B : |I - J| > 1$, and for each $I \in R_A \cup R_B$ there exists an $a_I \in \{1, \dots, k\}$ such that

$$A = \{\boldsymbol{a} \in G^t \mid \mathrm{row}(\boldsymbol{a}) \in R_A, \mathrm{col}(\boldsymbol{a}) \neq a_{\mathrm{row}(\boldsymbol{a})}\}$$

and

$$B = \{\boldsymbol{b} \in G^t \mid \mathrm{row}(\boldsymbol{b}) \in R_B, \mathrm{col}(\boldsymbol{b}) \neq a_{\mathrm{row}(\boldsymbol{b})}\}.$$

For each $r \leq t$ we let

$$A_r = \{\boldsymbol{a} \in G^r \mid \mathrm{row}(\boldsymbol{a}) \in R_A, \mathrm{col}(\boldsymbol{a}) \neq a_{\mathrm{row}(\boldsymbol{a})}\}$$

and

$$B_r = \{\boldsymbol{b} \in G^r \mid \mathrm{row}(\boldsymbol{b}) \in R_B, \mathrm{col}(\boldsymbol{b}) \neq a_{\mathrm{row}(\boldsymbol{b})}\} .$$

Thus $A = A_t$ and $B = B_t$.
There exists an $i \leq l_s$ such that the partial bijection p_i^s of $\mathfrak{G}^{s-1}$ is the identity on A_{s-1} and maps each element of B_{s-1} to its image under $f_{(\gamma_0, \dots, \gamma_{s-1})}$.
Let $\delta_s := \rho^s(\mathrm{p}_i^s)$. By Lemma 7.6 $f_{(0,\dots,0,\delta_s)}$ is an automorphism of $\mathfrak{G}_s$ which, by Lemma 7.10, is the identity on A_s and maps each element of B_s to its image under $f_{(\gamma_0, \dots, \gamma_{s-1}, 0)}$.
Then for each $(I, a, \zeta_0, \dots, \zeta_s)/_\sim \in A_s$ we have

$$\begin{aligned} & (I, a, \zeta_0, \dots, \zeta_{s-1}, \zeta_s + \delta_s)/_\sim \\ = \; & f_{(0,\dots,0,\delta_s)}((I, a, \zeta_0, \dots, \zeta_s)/_\sim) \\ = \; & (I, a, \zeta_0, \dots, \zeta_s)/_\sim \end{aligned}$$

and for each $(I, b, \eta_0, \dots, \eta_s)/_\sim \in B_s$ we have

$$\begin{aligned} & (I, b, \eta_0, \dots, \eta_{s-1}, \eta_s + \delta_s)/_\sim \\ = \; & f_{(0,\dots,0,\delta_s)}((I, b, \eta_0, \dots, \eta_s)/_\sim) \\ = \; & (I, b, \eta_0 + \gamma_0, \dots, \eta_{s-1} + \gamma_{s-1}, \eta_s)/_\sim . \end{aligned}$$

Thus for each $\boldsymbol{a} \in A_t$ we have

$$f_{(0,\dots,0,\delta_s,0,\dots,0)}(\boldsymbol{a}) = \boldsymbol{a}$$

and for each $\boldsymbol{b} \in B_t$

$$f_{(0,\dots,0,\delta_s,0,\dots,0)}(\boldsymbol{b}) = f_{(\gamma_0,\dots,\gamma_{s-1},0,\dots,0)}(\boldsymbol{b}) = f_*^{-1} \circ f^*(\boldsymbol{b}) .$$

We let $f := f_* \circ f_{(0,\dots,0,\delta_s,0,\dots,0)}$ and see that the the Lemma holds. ∎

For each $\gamma_t \in \Gamma^t$ we define $\mathrm{supp}(\gamma_t)$ in the usual way and obtain completely analogously to Lemma 8.3:

11.19 Lemma : *Let* $\bar{\gamma} = (\gamma_1, \ldots, \gamma_t), \bar{\delta} = (\delta_1, \ldots \delta_t) \in \bar{\Gamma}, I \leq n$ *and suppose that* $\mathrm{supp}(\delta_t - \gamma_t) = \{i_1, \ldots, i_r\}$*. Then*

$$(I, a, \bar{\gamma}) \sim (I, b, \bar{\delta}) \quad \Longleftrightarrow \quad (I, a, \bar{\gamma}) \sim (I, b, \bar{\delta}) \quad \textit{via} \quad \mathtt{p}^t_{i_1} \ldots \mathtt{p}^t_{i_r} \ .$$

The next lemma is a strong version of Lemma 8.5:

11.20 Lemma : *Let* $I \leq n, \bar{\gamma} \neq \bar{\delta} \in \bar{\Gamma}$*. Then there exists an* $a \in \{1, \ldots, k\}$ *such that*

$$(I, a, \bar{\gamma}) \not\sim (I, a, \bar{\delta}) \ .$$

Proof: The proof is by induction over t, and again the case $t = 0$ is trivial.
Suppose the lemma is true for $(t - 1)$ and let $I \leq n$ and $(\gamma_0, \ldots, \gamma_t) \neq (\delta_0, \ldots, \delta_t) \in \bar{\Gamma}$. We can assume that there exists a $b \in \{1, \ldots, k\}$ such that

$$(I, b, \gamma_0, \ldots, \gamma_t) \sim (I, b, \delta_0, \ldots, \delta_t) \ .$$

(If no such b exists there is nothing to prove.)
Then

$$(I, b, \gamma_0, \ldots, \gamma_t) \sim (I, b, \delta_0, \ldots, \delta_t) \quad \textit{via} \quad \mathtt{p}^t_{i_1} \ldots \mathtt{p}^t_{i_r}$$

where $\{i_1, \ldots, i_r\} = \mathrm{supp}(\delta_t - \gamma_t)$.
If $r > 0$ we can choose $a \in \{1, \ldots, k\}$ such that $(I, a, \delta_0, \ldots, \delta_{t-1}) \notin \mathrm{dom}(p^t_{i_1})$.

If $r = 0$, i.e. $\gamma_t = \delta_t$, then $(\gamma_0, \ldots, \gamma_{t-1}) \neq (\delta_0, \ldots, \delta_{t-1})$. By the induction hypothesis there exists an $a \in \{1, \ldots, k\}$ such that

$$(I, a, \gamma_0, \ldots, \gamma_{t-1}) \not\sim (I, a, \delta_0, \ldots, \delta_{t-1}) \ .$$

Since π_{γ_t} is injective by Lemma 7.9 this implies

$$(I, a, \gamma_0, \ldots, \gamma_t) \not\sim (I, a, \delta_0, \ldots, \delta_t) \ .$$

■

Finally we want to prove that our paths are really deterministic, which is in a sense the analogon of Lemma 8.22 for this case. The proof is similar to the proof of Lemma 8.22, in particular we need the following preliminary Lemma 11.21 corresponding to Lemma 8.20.

11.21 Lemma : *Let* $a, b \in G^{t-1}$ *and* $\gamma_t, \delta_t \in \Gamma^t$ *such that*

$$E^{\mathfrak{G}^t}(a, \gamma_t)/_\sim(b, \delta_t)/_\sim \ .$$

Then there exist $\hat{a}, \hat{b} \in G^{t-1}$ *and* $\eta_t \in \Gamma_t$ *such that* *(i)* $(a, \gamma_t) \sim (\hat{a}, \eta_t)$

(ii) $(b, \delta_t) \sim (\hat{b}, \eta_t)$

(iii) $E^{\mathfrak{G}^{t-1}} \hat{a}\hat{b}$

(iv) $\operatorname{supp}(\eta_t - \gamma_t) \,\dot{\cup}\, \operatorname{supp}(\eta_t - \delta_t) = \operatorname{supp}(\delta_t - \gamma_t)$

The **proof** can be given analogously to the proof of Lemma 8.20.

11.22 Lemma : *Let* $\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}} \in G^t$ *such that* $\mathfrak{G}^t \models \text{D–EDGE}_k[\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}}]$ *and*

$$\exists \bar{\gamma} \in \bar{\Gamma}, I \leq n \; \forall i \leq k : \boldsymbol{a}_i = (I, i, \bar{\gamma})/_{\sim} \,.$$

Then $\forall i \leq k : \boldsymbol{b}_i = (I+1, i, \bar{\gamma})/_{\sim}$.

As one could expect remembering the proof of Lemma 8.22 the proof of this lemma is not so easy. On top of everything we have to incorporate the additional parameter t here.
However, it becomes a bit simpler because the direction of the paths is determined by the direction of E and the columns are fixed by the P_i. So an analogue of Lemma 8.21 is not necessary here. Moreover, the p_i^t never keep whole rows fixed thus we can leave out the first step of the proof of Lemma 8.22. For this reason we also do not need edges between x_i and x_j $(i, j \leq k)$ or between y_i and y_j $(i, j \leq k)$ in the formula $\text{D–EDGE}_k(\overset{k}{x}, \overset{k}{y})$.
The **proof** is by induction on t:
It can be immediately seen from the definition of $\mathfrak{G}^0$ that the statement of the lemma holds for $t = 0$.

So suppose that $t \geq 1$ and $\bar{\gamma} = (\gamma_0, \ldots, \gamma_t)$.
$\mathfrak{G}^t \models \text{D–EDGE}_k[\overset{k}{\boldsymbol{a}}, \overset{k}{\boldsymbol{b}}]$ means that

$$\forall i, j \leq k : E^{\mathfrak{G}^t} \boldsymbol{a}_i \boldsymbol{b}_j$$

and

$$\forall i \leq k : P_i^{\mathfrak{G}^t} \boldsymbol{a}_i \quad \text{and} \quad P_i^{\mathfrak{G}^t} \boldsymbol{b}_i.$$

Note that for all $i \leq k$ the former implies $\operatorname{row}(\boldsymbol{b}_i) = I + 1$ and the latter implies $\operatorname{col}(\boldsymbol{a}_i) = \operatorname{col}(\boldsymbol{b}_i) = i$.
We shall prove:

CLAIM:

$$\forall h \leq k \; \forall j < t \; \exists \delta_{hj} \in \Gamma^j : \; \boldsymbol{b}_h = (I+1, h, \delta_{h1}, \ldots, \delta_{h(t-1)}, \gamma_t)/_{\sim}$$

and

$$\begin{aligned}\mathfrak{G}^{t-1} \models \text{D–EDGE}_k[\ & (I,1,\gamma_1,\ldots,\gamma_{t-1})/_\sim,\ldots,(I,k,\gamma_1,\ldots,\gamma_{t-1})/_\sim,\\ & (I+1,1,\delta_{11},\ldots,\delta_{1(t-1)})/_\sim,\ldots,(I+1,k,\delta_{k1},\ldots,\delta_{k(t-1)})/_\sim]\end{aligned}$$

Once we have proved this claim we can apply the induction hypothesis and are finished.

PROOF: We fix a $h \le k$ and suppose $\boldsymbol{b}_h = (I+1,h,\delta_0,\ldots,\delta_t)/_\sim$.
By the definition of the partial bijections $p_1^t,\ldots,p_{l_t}^t$ for each $x \in supp(\delta_t - \gamma_t)$ there exists an $i_x \in \{1,\ldots,k\}$ such that $a \notin \mathrm{dom}(p_x^t)$ for each $a \in G^{t-1}$ with $\mathrm{row}(a) = I$ and $\mathrm{col}(a) = i_x$.
By $E^{\mathfrak{G}^t}\boldsymbol{a}_{i_x}\boldsymbol{b}_h$ and the previous Lemma 11.21 there exist $\eta_t \in \Gamma^t$ and $\gamma_i', \delta_i' \in \Gamma^i$ (for all $i \le t-1$) such that

$$\left.\begin{array}{rl} & (I,i_x,\gamma_1,\ldots,\gamma_t) \sim (I,i_x,\gamma_1',\ldots,\gamma_{t-1}',\eta_t)\\ \wedge & (I+1,h,\delta_1,\ldots,\delta_t) \sim (I+1,h,\delta_1',\ldots,\delta_{t-1}',\eta_t)\\ \wedge & E^{\mathfrak{G}^{t-1}}(I,i_x,\gamma_1',\ldots,\gamma_{t-1}')/_\sim(I+1,h,\delta_1',\ldots,\delta_{t-1}')/_\sim\\ \wedge & \mathrm{supp}(\eta_t-\gamma_t) \,\dot\cup\, \mathrm{supp}(\eta_t-\delta_t) = \mathrm{supp}(\delta_t-\gamma_t) \qquad . \end{array}\right\} (*)$$

Since for all $a \in G^{t-1}$ with $\mathrm{row}(a) = I$ and $\mathrm{col}(a) = i_x$ we have $a \notin \mathrm{dom}(p_x^t)$ this implies

$$\begin{array}{rl} & x \notin \mathrm{supp}(\eta_t - \gamma_t)\\ \Longrightarrow & x \in \mathrm{supp}(\eta_t - \delta_t)\\ \Longrightarrow & \forall b \in G^{t-1} \text{ with } \mathrm{row}(b) = I+1 \text{ and } \mathrm{col}(b) = h:\ b \in \mathrm{dom}(p_x^t). \end{array}$$

As x was arbitrary we have proved this for all $x \in \mathrm{supp}(\delta_t - \gamma_t)$. $(**)$

Let $\mathrm{supp}(\delta_t-\gamma_t) = \{x_1,\ldots,x_r\}$ be ordered such that $\mathrm{supp}(\delta_t-\eta_t) = \{x_1,\ldots,x_s\}$ and $\mathrm{supp}(\eta_t-\gamma_t) = \{x_{s+1},\ldots,x_r\}$ for an $s \le r$ (where η_t is taken from $(*)$).
Then

$$\begin{aligned} ((I+1,h,\delta_1,\ldots,\delta_{t-1})/_\sim)^{p^t_{x_1},\ldots,p^t_{x_s}} &= (I+1,h,\delta_1',\ldots,\delta_{t-1}')/_\sim\ ,\\ ((I,i_x,\gamma_1',\ldots,\gamma_{t-1}')/_\sim)^{p^t_{x_{s+1}},\ldots,p^t_{x_r}} &= (I,i_x,\gamma_1,\ldots,\gamma_{t-1})/_\sim\ . \end{aligned}$$

(Since $\forall i,j \le l_t : p_i^t = (p_i^t)^{-1} \wedge p_i^t p_j^t = p_j^t p_i^t$.)
Moreover, by $(**)$ we can choose $\delta_{h1},\ldots,\delta_{h(t-1)}$ such that

$$((I+1,h,\delta_1',\ldots,\delta_{t-1}')/_\sim)^{p^t_{x_{s+1}},\ldots,p^t_{x_r}} = (I+1,h,\delta_{h1},\ldots,\delta_{h(t-1)})/_\sim\ .$$

Hence $\boldsymbol{b}_h = (I+1,h,\delta_{h1},\ldots,\delta_{h(t-1)},\gamma_t)/_\sim$.
Since $p^t_{x_{s+1}},\ldots,p^t_{x_r}$ is a partial isomorphism

$$E^{\mathfrak{G}^{t-1}}(I,i_x,\gamma_1',\ldots,\gamma_{t-1}')/_\sim(I+1,h,\delta_1',\ldots,\delta_{t-1}')/_\sim$$

implies

$$E^{\mathfrak{G}^{t-1}}(I, i_x, \gamma_1, \dots, \gamma_{t-1})/_\sim(I+1, h, \delta_{h1}, \dots, \delta_{h(t-1)})/_\sim .$$

So the claim is proved and we are finished. ■

These lemmata show that $\mathfrak{G}^t$ has all the properties stated in the theorem:
(1) is obvious. (2) follows from Lemma 11.17 with $\text{æ} = f_{(1,0,\dots,0)}$.
To prove (3) let $\boldsymbol{c}_i = (1, i, 0, \dots, 0)/_\sim$ and $\boldsymbol{d}_i = (n, i, 0, \dots, 0)/_\sim$ (for each $i \leq k$).
Then there exists a D–EDGE$_k$–path from $\overset{k}{\boldsymbol{c}}$ to $\overset{k}{\boldsymbol{d}}$, and by Lemma 11.22 it is deterministic. Hence there exists no path from $\overset{k}{\boldsymbol{c}}$ to $\text{æ}(\overset{k}{\boldsymbol{d}})$ because by Lemma 11.20 we have $\overset{k}{\boldsymbol{d}} \neq \text{æ}(\overset{k}{\boldsymbol{d}})$.
We have already proved (4) in Lemma 11.18.

However, $\mathfrak{G}^t$ is a $\{E, P_1, \dots, P_k\}$–structure and not a graph.
Since there are only edges between succeeding rows $\mathfrak{G}^t$ does not contain any cliques of size ≥ 3.
To eliminate the relations $P_1, \dots, P_k$ we augment each row of our structure by k cliques of sizes $3, \dots, (k+2)$. Let a_{Ii} be a vertex of the $(2+i)$–clique of row I. For all $I < n, i \leq k$ we include edges between all elements of the ith column in row I and a_{Ii}. In this new structure the relations $P_1, \dots, P_k$ are clearly definable, so we can erase them.
The structure we obtain is a directed graph. We replace all edges by symmetric edges and obtain a graph, but in this graph the direction of a D–EDGE$_k$–path is no longer recognizable. So for each $I < n$ we include an edge between the a_{I1} and $a_{(I+1)2}$.

Let us call the resulting graph $\mathfrak{G}^t_*$. We define a formula D–EDGE$_k{}^*$ by

$$\text{D–EDGE}_k{}^*(\overset{k}{x}, \overset{k}{y}) := \bigwedge_{i,j \leq k} \mathrm{E}(x_i, y_j) \wedge \bigwedge_{i=1}^{k} (\mathrm{P}_i(x_i) \wedge \mathrm{P}_i(y_i))$$

where

$$\begin{aligned}\mathrm{E}(x,y) := Exy \wedge \exists \overset{3}{x}, \overset{4}{y} \Big(& \textstyle\bigwedge_{i<j\leq 3}(Ex_ix_j \wedge x \neq x_i) \\ & \wedge \textstyle\bigwedge_{i<j\leq 4}(Ey_iy_j \wedge y \neq y_i) \\ & \wedge Exx_1 \wedge Eyy_1 \wedge Ex_1y_1\Big)\end{aligned}$$

and

$$\begin{aligned}\mathrm{P}_i(x) := \quad & \exists \overset{2+i}{x} \, (\textstyle\bigwedge_{i<j<2+i}(Ex_ix_j \wedge x \neq x_i) \wedge Exx_1) \\ \wedge \neg(\exists & \overset{3+i}{x} \, (\textstyle\bigwedge_{i<j<3+i}(Ex_ix_j \wedge x \neq x_i) \wedge Exx_1)\end{aligned}$$

for each $i \leq k$.
We extend the mapping row to $\mathfrak{G}^t_*$ in the obvious way and each $f \in \mathcal{F}^t$ by letting it be the identity on the new vertices.

Clearly, $\mathfrak{G}^t_*$ together with D–$\textsc{Edge}_k{}^*$ still has properties (1)–(4), so Theorem 11.16 is proved. ■

11.23 Remark : If we only wanted to prove that $\mathbf{DTC}^k \not\subseteq \mathbf{DTC}^{k-1}$ we could have already worked with the structure $\mathfrak{G}^1$. The reason for this is that $(k-1)$–tuples cannot define any $(k-1)$–tuples in other rows of $\mathfrak{G}^1$. So there are no deterministic paths of $(k-1)$–tuples crossing longer distances and $\mathbf{DTC}^{k-1}$–formulae are ineffective. □

§11.3 An Ehrenfeucht-Fraïssé Game for Transitive Closure Logic

We want to prove that the deterministic transitive closure of D–$\textsc{Edge}_k$ is not TC^{k-1}–expressible on the class $\{\mathfrak{G}^t(k,n) \mid n \geq 2, t \geq 1\}$ of structures defined in the last paragraph. Again the proof uses an Ehrenfeucht-Fraïssé game.
Let $k, t \geq 0$. We define a game for k–ary transitive closure logic with at most t nested TC–operators; it is only a slight modification of the TC-Ehrenfeucht-Fraïssé game introduced by Grädel [Grä91].

11.24 Definition : The k–ary t–nested r–move TC–game is played by two players, the challenger and the duplicator, on a pair $\mathfrak{A}, \mathfrak{B}$ of structures of the same signature with r pairs of pebbles $P_1, Q_1, \ldots, P_r, Q_r$ following the moves defined below.

We say that two elements $a \in A$ and $b \in B$ correspond in a situation of the game if they are the interpretations of the same constant in $\mathfrak{A}$ and $\mathfrak{B}$ respectively or if a pair P_i, Q_i of pebbles is placed on them.

In each situation of the game the challenger selects one of the following moves, respecting that throughout the whole game at most t TC or ¬TC–moves are allowed.

∃-move: The challenger places a yet unused pebble P_i on (an element of structure) $\mathfrak{A}$, the duplicator answers by placing the corresponding pebble Q_i on $\mathfrak{B}$.

∀-move: The challenger places Q_i on $\mathfrak{B}$, the duplicator answers by placing P_i on $\mathfrak{A}$.

TC-move: The challenger selects an $l \leq k$ such that there are still $2l$ yet unused pairs of pebbles. Then he selects a sequence $\overset{l}{a}_1, \ldots, \overset{l}{a}_m$ of l-tuples in $\mathfrak{A}$ where $\overset{l}{a}_1$ and $\overset{l}{a}_m$ consist of elements already pebbled and constants.

The duplicator selects a sequence $\overset{l}{b}_1, \ \dots\ , \overset{l}{b}_n$ of l-tuples in $\mathfrak{B}$ where b_{1i} and b_{ni} $(1 \le i \le l)$ correspond to a_{1i} and a_{mi} respectively.

Now the challenger selects an $i < n$ and places $2l$ yet unused pebbles on $b_{i1}, \dots, b_{il}, b_{(i+1)1}, \dots, b_{(i+1)l}$ respectively.

The duplicator selects a $j < m$ and places the corresponding pebbles on $a_{j1}, \dots, a_{jl}, a_{(j+1)1}, \dots, a_{(j+1)l}$ respectively.

¬TC-move: Analogous to the TC^k-move with structures $\mathfrak{A}$ and $\mathfrak{B}$ exchanged.

The duplicator wins if after all pebbles are placed on the structures the mapping between the corresponding elements is a partial isomorphism between $\mathfrak{A}$ and $\mathfrak{B}$. □

To relate the game to transitive closure logic we need the following notions of rank for TC–formulae:
We let

$$qr([TC_{\overset{l}{x},\overset{l}{y}}\ \varphi(\overset{l}{x},\overset{l}{y})]\ \overset{l}{u}\overset{l}{v}) = qr(\varphi) + 2l$$

and define the *quantifier-rank* inductively in the natural way.

The *TC–rank* $tr(\varphi)$ of a TC–formula φ is the maximal number of nested TC–operators in φ.

11.25 Theorem ([Grä91]) : *Let $\mathfrak{A}$ and $\mathfrak{B}$ be structures of the same signature. Then for all $k, t, r \ge 0$ the following two statements are equivalent:*

(1) For each TC^k–sentence φ with $qr(\varphi) \le r$ and $tr(\varphi) \le t$ we have

$$\mathfrak{A} \models \varphi \quad \Longleftrightarrow \quad \mathfrak{B} \models \varphi.$$

(2) The duplicator has a winning strategy for the k–ary t–nested r–move TC–game on $\mathfrak{A}, \mathfrak{B}$.

§ 11.4 The Game

To calculate the number of rows of our structures we define a function

$$H_k : \mathbb{N}^2 \longrightarrow \mathbb{N} \quad \text{by} \quad H_k(r,t) := (3^k k!)^t 2^{r+1}.$$

We fix $r, t \ge 0$ which are intended to be the quantifier rank and the TC-rank of a formula and let $n := H_k(r,t) + 1$ and $\mathfrak{G} := \mathfrak{G}^{t(k-1)}(k, n)$ (chosen according to Theorem 11.16).

We select $\overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}} \in G$ satisfying property (3) of the structure and let $\mathfrak{G}_1 := (\mathfrak{G}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$ and $\mathfrak{G}_2 := (\mathfrak{G}, \overset{k}{\boldsymbol{c}}, \text{æ}(\overset{k}{\boldsymbol{d}}))$; hence $\mathfrak{G}_1$ and $\mathfrak{G}_2$ can be distinguished by a DTC^k–sentence that does not depend on r or t.
Thus by Theorem 11.25 we are finished once we have proved:

11.26 Proposition : *The duplicator wins the $(k-1)$–ary t–nested r–move TC–game on $\mathfrak{G}_1$ and $\mathfrak{G}_2$.*

Proof: For all $\boldsymbol{a}, \boldsymbol{b} \in \mathfrak{G}$ we define the distance between $\boldsymbol{a}$ and $\boldsymbol{b}$ by

$$d(\boldsymbol{a}, \boldsymbol{b}) := |\text{row}(\boldsymbol{a}) - \text{row}(\boldsymbol{b})| \; .$$

Since corresponding elements will always be in the same row we can extend this definition to two pairs of corresponding elements.

By a simultaneous induction over the number of pebbles that are already placed on the board and the number of TC–moves that have already been made we prove that the duplicator can preserve the following throughout the whole game:

Suppose that $r' \leq r$ pairs of pebbles have already been placed on the board and $t' \leq t$ is the number of TC-moves that have been made yet.
Then

(i) If $(\boldsymbol{a}, \boldsymbol{b})$ is a pair of corresponding elements we have $\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b})$.

(ii) For each pair $\boldsymbol{p} = (\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements there exists an automorphism $f_{\boldsymbol{p}} \in \mathcal{F}^{t'(k-1)}$ of $\mathfrak{G}$ such that $f_{\boldsymbol{p}}(\boldsymbol{a}) = \boldsymbol{b}$.
If $\boldsymbol{q}$ is another pair of corresponding elements and $d(\boldsymbol{p}, \boldsymbol{q}) < H_k(r - r', t - t')$ then $f_{\boldsymbol{p}} = f_{\boldsymbol{q}}$.

Note that we are finished once we have proved this.

In the beginning we let $f_{(\boldsymbol{c}_i, \boldsymbol{c}_i)} = id$ and $f_{(\boldsymbol{d}_i, \text{æ}(\boldsymbol{d}_i))} = \text{æ}$ (for each $i \leq k$). Then (i) and (ii) hold because $id, \text{æ} \in \mathcal{F}^0$ and $\forall i, j \leq k : \; d(\boldsymbol{c}_i, \boldsymbol{d}_j) = n - 1 = H_k(r, t)$.

∃-move: The challenger places a pebble on $\boldsymbol{a} \in \mathfrak{G}_1$. Let $\boldsymbol{p}$ be a nearest pair of corresponding elements. The duplicator places the corresponding pebble on $\boldsymbol{b} := f_{\boldsymbol{p}}(\boldsymbol{a})$, and we define $f_{(\boldsymbol{a}, \boldsymbol{b})} := f_{\boldsymbol{p}}$.

Then for all pairs $\boldsymbol{q}$ of corresponding elements with $f_{\boldsymbol{q}} \neq f_{\boldsymbol{p}}$ (hence $d(\boldsymbol{q}, \boldsymbol{p}) \geq H_k(r - r', t - t')$ by induction hypothesis) we have

$$d((\boldsymbol{a}, \boldsymbol{b}), \boldsymbol{q}) \geq \left\lfloor \frac{H_k(r - r', t - t')}{2} \right\rfloor = H_k(r - r' - 1, t - t') \; .$$

$\forall$-move: is treated analogously.

The essential part of the duplicator's answer to a TC–move is hidden in Lemma 11.27 below. However, to understand the lemma it is easier to see its application first. Hence we finish our game (and thus the proof of Proposition 11.26) first:

TC-move: The challenger selects a sequence $\overset{l}{\boldsymbol{a}}_1, \dots, \overset{l}{\boldsymbol{a}}_m \in \mathfrak{G}_1$ where $\overset{l}{\boldsymbol{a}}_1$ and $\overset{l}{\boldsymbol{a}}_m$ consist of constants or elements that have already been pebbled in previous moves and $l \leq k-1$ such that $r' + 2l \leq r$.

The duplicator has to define a sequence $\overset{l}{\boldsymbol{b}}_1, \dots, \overset{l}{\boldsymbol{b}}_m \in \mathfrak{G}_2$ such that if the challenger places pebbles on $\overset{l}{\boldsymbol{b}}_i, \overset{l}{\boldsymbol{b}}_{i+1}$ afterwards she can answer by placing the corresponding pebbles on $\overset{l}{\boldsymbol{a}}_i, \overset{l}{\boldsymbol{a}}_{i+1}$. To preserve (ii) we also have to define automorphisms $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})}$ and $f_{(\boldsymbol{a}_{i+1j}, \boldsymbol{b}_{i+1j})}$ in $\mathcal{F}^{(t'+1)(k-1)}$.

Formally, we will define suitable automorphisms first; they will determine the $\boldsymbol{b}_{ij}$. That is, we are going to define automorphisms $f_{ij} \in \mathcal{F}^{t'(k-1)+l} \subseteq \mathcal{F}^{(t'+1)(k-1)}$ for $i \leq m, j \leq l$ such that $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})}$ will be f_{ij} and thus $\boldsymbol{b}_{ij} = f_{ij}(\boldsymbol{a}_{ij})$.

Unfortunately this does not suffice: Suppose we have defined our f_{ij} and hence $\boldsymbol{b}_{ij}$ and the challenger places pebbles on, say, $\overset{l}{\boldsymbol{b}}_i, \overset{l}{\boldsymbol{b}}_{i+1}$ (where $\boldsymbol{b}_{ij} = f_{ij}(\boldsymbol{a}_{ij})$ and $\boldsymbol{b}_{i+1j} = f_{i+1j}(\boldsymbol{a}_{i+1j})$). Then the duplicator answers by placing her pebbles on $\overset{l}{\boldsymbol{a}}_i, \overset{l}{\boldsymbol{a}}_{i+1}$ and we let $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})} := f_{ij}$ and $f_{(\boldsymbol{a}_{i+1j}, \boldsymbol{b}_{i+1j})} := f_{i+1j}$.

Now consider another choice of the challenger, suppose he places his pebbles on $\overset{l}{\boldsymbol{b}}_{i-1}, \overset{l}{\boldsymbol{b}}_i$ (where $\boldsymbol{b}_{i-1j} = f_{i-1j}(\boldsymbol{a}_{i-1j})$). The duplicator answers by placing her pebbles on $\overset{l}{\boldsymbol{a}}_{i-1}, \overset{l}{\boldsymbol{a}}_i$. But there is no reason that we should also let $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})} := f_{ij}$ in this case. (We only need $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})}(\boldsymbol{a}_{ij}) = \boldsymbol{b}_{ij} = f_{ij}(\boldsymbol{a}_{ij})$.) In fact it is necessary to define it in a different way here.

Therefore we also have to define automorphisms $f^+_{ij} \in \mathcal{F}^{t'(k-1)+l}$ $(2 \leq i \leq m, j \leq l)$ such that we can define $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})} = f^+_{ij}$ in the second case. Clearly this requires $\forall j \leq l : f_{ij}(\boldsymbol{a}_{ij}) = f^+_{ij}(\boldsymbol{a}_{ij})$.

These considerations lead us to the following:

Claim: There exist automorphisms $f_{ij} \in \mathcal{F}^{t'(k-1)+l}$ $(1 \leq i \leq m-1, j \leq l)$ and $f^+_{ij} \in \mathcal{F}^{t'(k-1)+l}$ $(2 \leq i \leq m, j \leq l)$ of $\mathfrak{G}$ such that

(1) $f_{ij}(\boldsymbol{a}_{ij}) = f^+_{ij}(\boldsymbol{a}_{ij})$ (for all $2 \leq i \leq m-1, j \leq l$)

(2) If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'}) < H_k(r-r', t-t'-1)$ then $f_{ij} = f_{ij'}$ and $f^+_{ij} = f^+_{ij'}$.
If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{i+1j'}) < H_k(r-r', t-t'-1)$ then $f_{ij} = f^+_{i+1j'}$.
(for all possible i, j, j').[40]
(This item guarantees that (ii) will be preserved for the new pebbled elements among themselves.)

(3) If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}) < H_k(r-r', t-t'-1)$ for an $\boldsymbol{a}$ that belongs to a pair $(\boldsymbol{a}, \boldsymbol{b})$ of (already) corresponding elements then $f_{ij} = f^+_{ij} = f_{(\boldsymbol{a},\boldsymbol{b})}$
(for all possible i, j).
(This item guarantees that the relation between new and old corresponding elements preserves (ii).)

We have seen how to give the duplicator's answer to the TC–move once we have proved this:

She defines her path by $\boldsymbol{b}_{ij} := f_{ij}(\boldsymbol{a}_{ij})$ for each $i < m$, $j \leq l$ and $\boldsymbol{b}_{mj} := f^+_{mj}(\boldsymbol{a}_{mj})$ for each $j \leq l$. If the challenger places pebbles on $\overset{l}{\boldsymbol{b}}_i, \overset{l}{\boldsymbol{b}}_{i+1}$ she answers by placing the corresponding pebbles on $\overset{l}{\boldsymbol{a}}_i, \overset{l}{\boldsymbol{a}}_{i+1}$. We let, for each $j \leq l$, $f_{(\boldsymbol{a}_{ij}, \boldsymbol{b}_{ij})} := f_{ij}$ and $f_{(\boldsymbol{a}_{i+1j}, \boldsymbol{b}_{i+1j})} := f^+_{i+1j}$.

PROOF (of the claim): By the induction hypotheses (i) and (ii) we can find integers $h \geq 1$ and

$$1 = N_0 \leq M_1 < N_1 \leq M_2 < \ldots < N_{h-1} \leq M_h < N_h \leq M_{h+1} = n$$

such that

- $N_i - M_i \geq H_k(r-r', t-t')$ $(1 \leq i \leq h)$.
- For each pair $\boldsymbol{p}$ of corresponding elements there exists an $i \in \{0, \ldots, h\}$ such that $\mathrm{row}(\boldsymbol{p}) \in [N_i, M_{i+1}]$.
- On the other hand for each $i \in \{0, \ldots, h\}$ there exists a $\boldsymbol{p}$ such that $\mathrm{row}(\boldsymbol{p}) \in [N_i, M_{i+1}]$.
- If $\boldsymbol{p}$ and $\boldsymbol{p}'$ are pairs of corresponding elements such that $\boldsymbol{p}, \boldsymbol{p}' \in [N_i, M_{i+1}]$ then $f_{\boldsymbol{p}} = f_{\boldsymbol{p}'} =: f_{[N_i, M_{i+1}]}$.

We divide $\{1, \ldots, n\}$ into the non-empty intervals

$$[N_0, M_1],]M_1, N_1[, \ldots,]M_h, N_h[, [N_h, M_{h+1}].$$

These intervals are treated independently in the following way:

[40] Hopefully it is clear which i, j, j' are "possible" here. For example, the statement "If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'}) < H_k(r-r', t-t'-1)$ then $f_{ij} = f_{ij'}$" is claimed for all $i \leq m-1, j, j' \leq l$ and the analogous statement for f^+_{ij} is claimed for all $2 \leq i \leq m, j, j' \leq l$.

For each $p \leq h$ we consider only those $\boldsymbol{a}_{ij}$ with $N_{p-1} \leq \text{row}(\boldsymbol{a}_{ij}) \leq M_{p+1}$. Hence we obtain a sequence of (at most) l–tuples in each interval $[N_{p-1}, M_{p+1}]$. Without loss of generality we can assume that it is in fact a sequence of l–tuples.

We apply Lemma 11.27 to this sequence with $s = t'(k-1)$, $M = M_p$, $N = N_p$, $f_* = f_{[N_{p-1}, M_p]}$, $f^* = f_{[N_p, M_{p+1}]}$, and $d = H_k(r-r', t-t'-1)$ (Notice the definition of the function h preceding the lemma and the observation that
$$h(H_k(r-r', t-t'-1), k-1) \leq H_k(r-r', t-t').)$$
We obtain functions f_{ij}, f_{ij}^+ with properties (1)–(3) because:

(1) and (2) coincide with the corresponding items in the lemma; and (3) is guaranteed by (3) of the lemma because if $\boldsymbol{p}$ is a pair of (old) corresponding elements in a row between $N_{p-1} - H_k(r-r', t-t'-1)$ and $M_{p+1} + H_k(r-r', t-t'-1)$ (so it might be "near" an $\boldsymbol{a}_{ij}$ in question) then either $\text{row}(\boldsymbol{p}) \in [N_{p-1}, M_p]$ hence $f_{\boldsymbol{p}} = f_{[N_{p-1}, M_p]} = f_*$ or $\text{row}(\boldsymbol{p}) \in [N_p, M_{p+1}]$ hence $f_{\boldsymbol{p}} = f_{[N_p, M_{p+1}]} = f^*$.

We do this for each $p \leq h$. Hence on the intersections of the intervals $[N_{p-1}, M_{p+1}]$ $(1 \leq p \leq h)$ the definitions of those f_{ij} and f_{ij}^+ with $\text{row}(\boldsymbol{a}_{ij})$ in this intersection should coincide. But obviously they do because the intersections are intervals $[N_p, M_{p+1}]$ $(0 \leq p \leq h$ (!)) and for $\boldsymbol{a}_{ij}$ with rows in such intervals f_{ij} and f_{ij}^+ are $f_{[N_p, M_{p+1}]}$ anyway (by (3) of the lemma).

Hence the claim is proved (modulo Lemma 11.27).

$\neg$**TC-move** is treated analogously.

■

Before formulating the missing lemma it is helpful to define a new function $h : \mathbb{N}^2 \longrightarrow \mathbb{N}$ by
$$h(d, 0) := d + 2 \quad \text{and} \quad h(d, l) := 3(l+1)h(d, l-1)\ .$$
An easy induction shows that for all $d, l \geq 0$ we have $h(d, l) = 3^l (l+1)!(d+2)$. Thus
$$\begin{aligned} h(H_k(r-r', t-t'-1), k-1) &= 3^{k-1}k!(H_k(r-r', t-t'-1) + 2) \\ &\leq 3^k k! H_k(r-r', t-t'-1) \\ &= H_k(r-r', t-t')\ . \end{aligned}$$

11.27 Lemma : *Let $d \geq 1$, $0 \leq l \leq k-1$, $s \leq t(k-1)-l$, $f_*, f^* \in \mathcal{F}^s$ and $m \geq 2$.*
Let $\overset{l}{\boldsymbol{a}}_1, \ldots, \overset{l}{\boldsymbol{a}}_m \in \mathfrak{G}$ and $M, N \in \{1, \ldots, n\}$ such that

- $N - M \geq h(d, l)$
- $\forall j \leq l : \mathrm{row}(\boldsymbol{a}_{1j}), \mathrm{row}(\boldsymbol{a}_{mj}) \notin]M, N[$.

Then there exist automorphisms $f_{ij} \in \mathcal{F}^{s+l}$ $(1 \leq i \leq m-1, j \leq l)$ and $f_{ij}^+ \in \mathcal{F}^{s+l}$ $(2 \leq i \leq m, j \leq l)$ of $\mathfrak{G}$ such that

(1) $f_{ij}(\boldsymbol{a}_{ij}) = f_{ij}^+(\boldsymbol{a}_{ij})$ *(for all possible i, j)*

(2) If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'}) \leq d$ *then* $f_{ij} = f_{ij'}$ *and* $f_{ij}^+ = f_{ij'}^+$.
If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{i+1j'}) \leq d$ *then* $f_{ij} = f_{i+1j'}^+$.
(for all possible i, j, j')

(3) If $\mathrm{row}(\boldsymbol{a}_{ij}) - M \leq d$ *then* $f_{ij} = f_{ij}^+ = f_*$.
If $N - \mathrm{row}(\boldsymbol{a}_{ij}) \leq d$ *then* $f_{ij} = f_{ij}^+ = f^*$.
(for all possible i, j)

Proof: The proof is harder than it seems at the first sight, although the idea is relatively simple: If the path $\overset{l}{\boldsymbol{a}}_1, \ldots, \overset{l}{\boldsymbol{a}}_m$ starts in a row below M, then crosses the interval $]M, N[$, and ends up above N, there will be two tuples $\overset{l}{\boldsymbol{a}}_i$ and $\overset{l}{\boldsymbol{a}}_j$ such that $\overset{l}{\boldsymbol{a}}_i$ is near M and $\overset{l}{\boldsymbol{a}}_j$ is near N. By the homogeneity property of structure $\mathfrak{G}$ we find an automorphism f such that $f(\overset{l}{\boldsymbol{a}}_i) = f_*(\overset{l}{\boldsymbol{a}}_i)$ and $f(\overset{l}{\boldsymbol{a}}_j) = f^*(\overset{l}{\boldsymbol{a}}_j)$. We define the $f_{i'j'}$ to be f_* if $i' \leq i$, f if $i \leq i' < j$, and f^* if $i' \geq j$. Then the path switches its component on the way from $\overset{l}{\boldsymbol{a}}_i$ to $\overset{l}{\boldsymbol{a}}_j$.
However, the situation is much more complicated. For example, the path may cross the interval $]M, N[$ more than once. To deal with this, we have to cut the path into short, "simple" pieces.
The most serious problem is that we are dealing with a path of l–tuples, and each component may behave completely differently. Therefore our proof is by induction over l:

For $l = 0$ there is nothing to prove.

Suppose $l \geq 1$ and the lemma is proved for $l - 1$.
The idea of the following induction step is a (more or less) straightforward generalization of a proof of the statement for $l = 1$. So while checking the proof it might be helpful to keep the situation for $l = 1$ in mind. The idea becomes much clearer in this case because we do not need the disturbing applications of

the induction hypothesis (formally we need the induction hypothesis $l = 0$ but it is trivial).

Step 1: Subdividing $[1, m]$ into Simple Intervals

We subdivide $[M, N]$ into three parts having the following idea in mind:
As long as there is at least one element in the middle there can be at most $(l-1)$ elements in the bottom or in the top part and we can apply the induction hypothesis there. If nothing is in the middle there is no danger because bottom and top are not connected.
We let $M' := M + (l+1)h(d, l-1)$ and $N' := N - (l+1)h(d, l-1)$ and note that $N' - M' \geq (l+1)h(d, l-1)$.

In a sense we are mainly interested in the functions

$$R_j : \{1, \ldots, m\} \longrightarrow \{1, \ldots, n\} \quad \text{defined by} \quad R_j(i) := \text{row}(\boldsymbol{a}_{ij}) \ .$$

Figure 11.28 illustrates our situation.

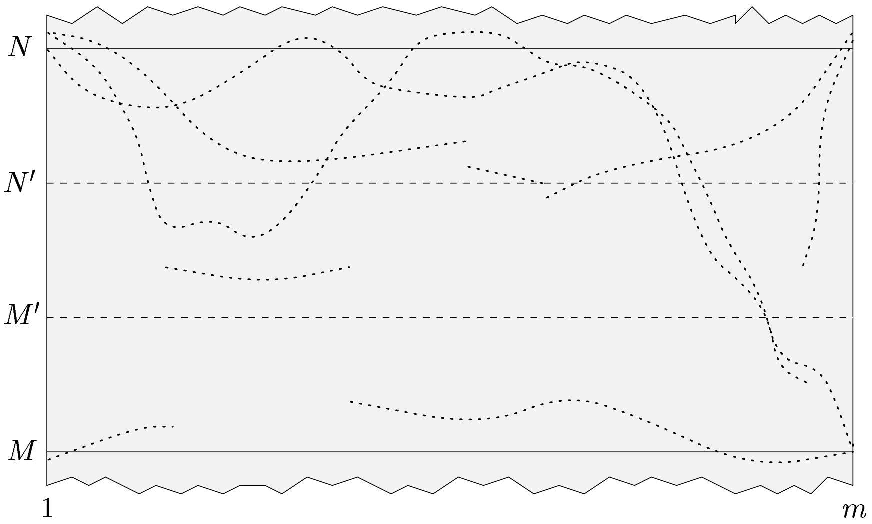

Figure 11.28

Next, we subdivide $[1, m]$ into intervals such that either for all i in such an interval there exists a j with $R_j(i)$ in the middle (i.e. in $]M', N'[$) or for no i in the interval there exists such a j.
We let

$$\begin{aligned} x_1 &:= 1 \\ x_{2h} &:= min\{i \mid x_{2h-1} \leq i < m,\ \forall j \leq l : R_j(i) \notin]M', N'[, \\ &\qquad\qquad\qquad\qquad\qquad\quad \exists j \leq l : R_j(i+1) \in]M', N'[\ \} \end{aligned}$$

$$x_{2h+1} \;:=\; min\{i \mid x_{2h} < i \leq m,\; \forall j \leq l : R_j(i) \notin]M', N'[,\; \exists j \leq l : R_j(i-1) \in]M', N'[\,\}$$

and h_{max} be the maximal number such that $x_{h_{max}}$ is defined. Thus for all $h < \frac{h_{max}}{2}$ we have:

$$\forall i \in [x_{2h-1}, x_{2h}]\; \forall j \leq l:\; R_j(i) \notin]M', N'[$$
$$\forall i \in]x_{2h}, x_{2h+1}[\; \exists j \leq l:\; R_j(i) \in]M', N'[\;.$$

We call $[x_{2h-1}, x_{2h+1}]$ a *simple interval.*

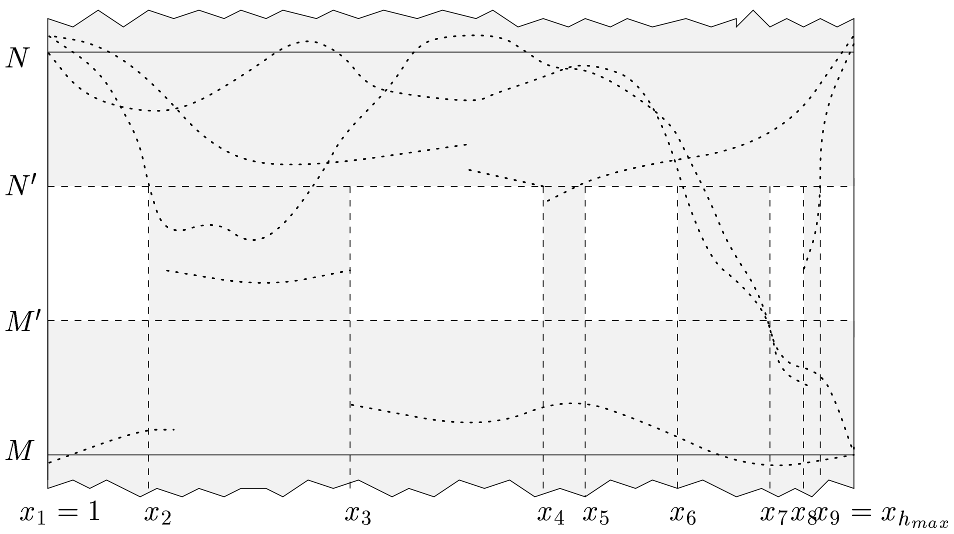

Figure 11.29

Before we consider the simple intervals we have to look at the rightmost interval $[x_{h_{max}}, m]$. Note that h_{max} is odd since $\forall j \leq l : R_j(m) \notin]M', N'[$. Hence for all $i \geq x_{h_{max}}$, $j \leq l$ we have $R_j(i) \notin]M', N'[$.
If $x_{h_{max}} < m$ for each $i \in [x_{h_{max}}, m[$ we let

$$f_{ij} := \begin{cases} f_* & if \quad R_j(i) \leq M' \\ f^* & if \quad R_j(i) \geq N' \end{cases}$$

and

$$f^+_{i+1j} := \begin{cases} f_* & if \quad R_j(i+1) \leq M' \\ f^* & if \quad R_j(i+1) \geq N' \end{cases} \quad .$$

We are finished once we have proved:

THE STATEMENT FOR SIMPLE INTERVALS: For each $h \geq 1$ such that $2h+1 \leq h_{max}$ and there exist automorphisms $f_{ij} \in \mathcal{F}^{t'+l}$ $(x_{2h-1} \leq i < x_{2h+1}, j \leq l)$ and $f^+_{ij} \in \mathcal{F}^{t'+l}$ $(x_{2h-1} < i \leq x_{2h+1}, j \leq l)$ of $\mathfrak{G}$ such that (1)–(3) hold and in addition

(4) If $R_j(x_{2h-1}) \le M'$ then $f_{x_{2h-1}j}(\boldsymbol{a}_{x_{2h-1}j}) = f_*(\boldsymbol{a}_{x_{2h-1}j})$.
If $R_j(x_{2h-1}) \ge N'$ then $f_{x_{2h-1}j}(\boldsymbol{a}_{x_{2h-1}j}) = f^*(\boldsymbol{a}_{x_{2h-1}j})$.

(5) If $R_j(x_{2h+1}) \le M'$ then $f^+_{x_{2h+1}j}(\boldsymbol{a}_{x_{2h+1}j}) = f_*(\boldsymbol{a}_{x_{2h+1}j})$.
If $R_j(x_{2h+1}) \ge N'$ then $f^+_{x_{2h+1}j}(\boldsymbol{a}_{x_{2h+1}j}) = f^*(\boldsymbol{a}_{x_{2h+1}j})$.

(4) and (5) make it possible to paste together the intervals $[x_{2h-1}, x_{2h+1}]$ to $[1, m]$ afterwards because they guarantee that for the endpoints x_{2h+1} $(1 \le h \le \frac{h_{max}}{2})$ of the simple intervals we have $f_{x_{2h+1}j}(\boldsymbol{a}_{x_{2h+1}j}) = f^+_{x_{2h+1}j}(\boldsymbol{a}_{x_{2h+1}j})$ for all $j \le l$. (But note that we do not demand $f_{x_{2h+1}j} = f^+_{x_{2h+1}j}$.)

So we consider an interval $[x_{2h-1}, x_{2h+1}]$. To simplify the notation we let $x := x_{2h-1}$, $y := x_{2h}$ and $z := x_{2h+1}$.

STEP 2: THE SPARSE PART

A place $i \in [x, z]$ is called *dense* if $\forall j \le l : R_j(i) \in]M', N'[$. Otherwise it is called *sparse*. Let w be the minimal dense place in $[x, z]$; if no such place exists we let $w := z$. Note that we always have $w > y$ by the definition of $y = x_{2h}$. (See Figure 11.30)

Recalling that $M' - M,\ N' - M'\ N - N' \ge (l+1)h(d, l-1)$, by the pigeonhole–principle we can find $M_*, N_*, M'', N'', M^*, N^*$ such that

- $M \le M_* < M_* + h(d, l-1) \le N_* \le M'$
- $M' \le M'' < M'' + h(d, l-1) \le N'' \le N'$
- $N' \le M^* < M^* + h(d, l-1) \le N^* \le N$
- $\forall j \le l :\ R_j(w) \notin]M'', N''[\wedge R_j(z) \notin]M_*, N_*[\cup]M^*, N^*[$.

Then we also have

- $\forall j \le l : R_j(x) \notin]M'', N''[$
  (because $\forall j \le l : R_j(x) \notin]M', N'[$),
- $\forall j \le l : R_j(w) \notin]M_*, N_*[\cup]M^*, N^*[$
  (because $\forall j \le l : R_j(w) \in]M', N'[$), and
- $\forall j \le l : R_j(z) \notin]M'', N''[$
  (because $\forall j \le l : R_j(z) \notin]M', N'[$).

(see Figure 11.30)

Since all $i \in [x, w[$ are sparse there exists a sequence $(j_i)_{x \le i \le w-1}$ such that $R_{j_i}(i) \notin]M', N'[\supseteq]M'', N''[$. Without loss of generality we assume that $j_i = l$ for all $i \in [x, w[$.

Now we apply the induction hypothesis to

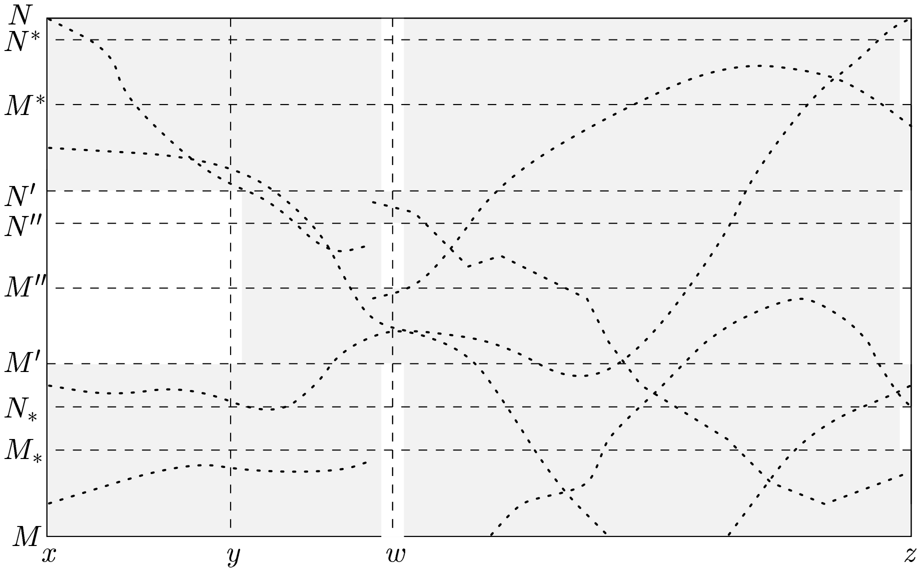

Figure 11.30

- the path $\overset{l-1}{\boldsymbol{a}}_i$ $(i \in [x, w])$ (i.e. we take $\{x, \ldots, w\}$ instead $\{1, \ldots, m\}$ as index–set)
- the interval $[M'', N'']$ instead of $[M, N]$ (Recall that N'', M'' are chosen in a way that $\forall j \leq l : R_j(w), R_j(x) \notin]M'', N''[$ and that $N'' - M'' \geq h(d, l-1)$.)
- $d,\ s,\ f_*, f^*$ (unchanged).

We get automorphisms $f_{ij} \in \mathcal{F}^{s+l-1}$ $(x \leq i \leq w-1, j \leq l-1)$ and $f^+_{ij} \in \mathcal{F}^{s+l-1}$ $(x+1 \leq i \leq w, j \leq l-1)$ of $\mathfrak{G}$ such that

- $f_{ij}(\boldsymbol{a}_{ij}) = f^+_{ij}(\boldsymbol{a}_{ij})$ (for all possible i, j)
- If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'}) \leq d$ then $f_{ij} = f_{ij'}$ and $f^+_{ij} = f^+_{ij'}$.
  If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{i+1j'}) \leq d$ then $f_{ij} = f^+_{i+1j'}$.
  (for all possible i, j, j').
- If $\mathrm{row}(\boldsymbol{a}_{ij}) - M'' \leq d$ then $f_{ij} = f^+_{ij} = f_*$.
  If $N'' - \mathrm{row}(\boldsymbol{a}_{ij}) \leq d$ then $f_{ij} = f^+_{ij} = f^*$.
  (for all possible i, j).

Furthermore, for each $i \in [x, w[$ we define

$$f_{il} := \begin{cases} f_* & if \quad R_l(i) \leq M'' \\ f^* & if \quad R_l(i) \geq N'' \end{cases}$$

and

$$f^+_{i+1l} := \begin{cases} f_* & if \quad R_l(i+1) \leq M'' \\ f^* & if \quad R_l(i+1) \geq N'' \end{cases} .$$

Note that these automorphisms perfectly fit into the items above , so in fact we have functions f_{ij} and f^+_{ij} for each $j \leq l$ such that the statements above hold.

Thus items (1)–(3) are true for the interval $[x, w]$ (with the functions defined above). Recall that for each $j \leq l$ we have $R_j(x) \notin]M', N'[$ thus

$$f_{xj} = \begin{cases} f_* & if \quad R_j(x) \leq M' \\ f^* & if \quad R_j(x) \geq N' \end{cases}$$

and (4) holds.
On the other side we have

$$f^+_{wj} = \begin{cases} f_* & if \quad R_j(w) \leq M'' \\ f^* & if \quad R_j(w) \geq N'' \end{cases}$$

(for each $j \leq l$).
If $w = z$ this implies (5) because there exists no $j \leq l$ such that $R_j(z) \in]M', N'[$; thus we are finished.
So we assume $w < z$ and consider the interval $[w, z]$.

Step 3: The Dense Part

By Theorem 11.16(4) applied to the sets

$$\begin{aligned} A &:= \{\boldsymbol{a}_{ij} \mid i \in \{w, z\}, j \leq l, R_j(i) \in [N_*, M'']\} \\ B &:= \{\boldsymbol{a}_{ij} \mid i \in \{w, z\}, j \leq l, R_j(i) \in [N'', M^*]\} \end{aligned}$$

there exists an $f \in \mathcal{F}^{s+1}$ such that

$$f(\boldsymbol{a}_{wj}) = \begin{cases} f_*(\boldsymbol{a}_{wj}) & if \quad R_j(w) \in]M', M''] \\ f^*(\boldsymbol{a}_{wj}) & if \quad R_j(w) \in [N'', N'[\end{cases}$$

and

$$f(\boldsymbol{a}_{zj}) = \begin{cases} f_*(\boldsymbol{a}_{zj}) & if \quad R_j(z) \in [N_*, M'] \\ f^*(\boldsymbol{a}_{zj}) & if \quad R_j(z) \in [N', M^*] \end{cases} .$$

Here we use that

$$\forall j \leq l : R_j(w) \in]M', N'[\wedge R_j(z) \notin]M', N'[$$

to guarantee that there are at most $l \leq (k-1)$ elements of $A \cup B$ in a row.
Moreover, $N'' - M'' > 1$ so for all $\boldsymbol{a} \in A, \boldsymbol{b} \in B$ we have $\mathrm{row}(\boldsymbol{b}) - \mathrm{row}(\boldsymbol{a}) > 1$.

Recall that for each $i \in]y, z] \supseteq [w, z]$ there exists a $j_i \le l$ such that $R_{j_i}(i) \in]M', N'[$ (this is by the definition of $y = x_{2h}$ and $z = x_{2h+1}$). Without loss of generality we can assume that $j_i = l$.
For each $i \in [w, z[$ we let $f_{il} := f$, and for each $i \in]w, z]$ we let $f^+_{il} := f$.

Step 4: Final Applications of the Induction Hypothesis

The situation we are in now is the most difficult. To cope with it we have to apply the induction hypothesis twice:
Once to

- the path of (at most) $(l-1)$–tuples consisting of the elements $\boldsymbol{a}_{ij}$, $j \le l-1$, $i \in [w, z]$ with $\mathrm{row}(\boldsymbol{a}_{ij}) \le M^*$
- the interval $[M_*, N_*]$ instead of $[M, N]$
- f instead of f^*
- $s+1$ instead of s
- d, f_* (unchanged)

and once to

- the path of (at most) $(l-1)$–tuples consisting of the elements $\boldsymbol{a}_{ij}$, $j \le l-1$, $i \in [w, z]$ with $\mathrm{row}(\boldsymbol{a}_{ij}) \ge N_*$
- the interval $[M^*, N^*]$ instead of $[M, N]$
- f instead of f_*
- $s+1$ instead of s
- d, f^* (unchanged).

What we get are automorphisms $f_{ij} \in \mathcal{F}^{s+1+l-1}$ $(w \le i \le z-1, j \le l-1)$ and $f^+_{ij} \in \mathcal{F}^{s+l}$ $(w+1 \le i \le z, j \le l-1)$ of $\mathfrak{G}$ such that

- $f_{ij}(\boldsymbol{a}_{ij}) = f^+_{ij}(\boldsymbol{a}_{ij})$ (for all possible i, j)
- If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{ij'}) \le d$ then $f_{ij} = f_{ij'}$ and $f^+_{ij} = f^+_{ij'}$.
  If $d(\boldsymbol{a}_{ij}, \boldsymbol{a}_{i+1j'}) \le d$ then $f_{ij} = f^+_{i+1j'}$.
  (for all possible i, j, j').
- If $\mathrm{row}(\boldsymbol{a}_{ij}) - M_* \le d$ then $f_{ij} = f^+_{ij} = f_*$.
  If $N^* - \mathrm{row}(\boldsymbol{a}_{ij}) \le d$ then $f_{ij} = f^+_{ij} = f^*$.
  (for all possible i, j).

- If $N_* - d \leq \mathrm{row}(\boldsymbol{a}_{ij}) \leq M^* + d$ then $f_{ij} = f^+_{ij} = f$
  (for all possible i, j).

Note that by the definition of the $f^{(+)}_{il}$ in fact we have functions f_{ij} and f^+_{ij} for each $j \leq l$ such that the statements above hold.
Clearly these statements guarantee (1)–(3) for the interval $[w, z]$.
Since for each $j \leq l$

$$f^+_{zj} := \begin{cases} f_* & if \quad R_j(z) \leq M_* \\ f & if \quad N_* \leq R_j(z) \leq M^* \\ f^* & if \quad R_j(z) \geq N^* \end{cases}$$

and, by the definition of f,

$$f(\boldsymbol{a}_{zj}) := \begin{cases} f_*(\boldsymbol{a}_{zj}) & if \quad R_l(i+1) \leq M' \\ f^*(\boldsymbol{a}_{zj}) & if \quad R_l(i+1) \geq N' \end{cases}$$

we also obtain (5).

Finally we note that our definitions for the intervals $[x, w]$ and $[w, z]$ match ; i.e. for all $j \leq l$ we have

$$f_{wj}(\boldsymbol{a}_{wj}) = f(\boldsymbol{a}_{wj}) = \left\{ \begin{array}{c} \overbrace{f_*(\boldsymbol{a}_{wj})}^{if\ R_j(w) \leq M''} \\ \underbrace{f^*(\boldsymbol{a}_{wj})}_{if\ R_j(w) \geq N''} \end{array} \right\} = f^+_{wj}(\boldsymbol{a}_{wj}) \; .$$

(The first equality holds because $R_j(w) \in]M', N'[$, for the last cf. Step 2.) ■

Appendix A: Proof of the Normal Form Theorem for IFP and PFP

In this appendix we are going to sketch a direct proof of the normal form theorem for **IFP** and **PFP** our considerations in Section 2 started with:

A.1 Theorem : *For every fixed–point formula χ there is an equivalent formula χ' in normal form.*

If χ is inductive then χ' can also be chosen to be inductive.

The proof is essentially due to [Gro92]. The main tool in the proof is the so called Transitivity Lemma (which is due to Moschovakis for **LFP**).
For the proof of the **PFP**–case it is convenient to pass over to the smaller class of *totally defined fixed–point formulae*:

A.2 Definition : A fixed–point formula χ is *totally defined* if for each subformula of the form $[FP_{\overset{k}{x},X}\ \varphi]\ \overset{k}{u}$ and each interpretation $\mathfrak{I}$ the sequence $(X_i^{\mathfrak{I}})_{i\geq 0}$ reaches a fixed–point.

TFP denotes the class of all totally defined fixed–point formulae. □

A.3 Lemma : *Every fixed–point formula is equivalent to a TFP–formula.*

Proof (sketch): The proof is by induction on fixed–point formulae; the only interesting step is to find a *TFP*–formula that is equivalent to a fixed–point formula of the form

$$\chi = [FP_{\bar{x},X}\varphi(\bar{x},X)]\bar{u}$$

where $\varphi \in$ *TFP*.

We first observe that for each interpretation $\mathfrak{I} = (\mathfrak{A}, \alpha)$ for χ there exists an $l > 0$ such that $X_l^{\mathfrak{I}} = X_{2l}^{\mathfrak{I}}$.
Because, since A is finite the sequence $(X_i^{\mathfrak{I}})_{i\geq 0}$ becomes periodic. Let $p \geq 1$ be the length of the period and m_0 be chosen such that for all $m \geq m_0$ we have $X_{m+p}^{\mathfrak{I}} = X_m^{\mathfrak{I}}$. We select l to be a multiple of p larger than m_0.

We are going to define a simultaneous fixed–point formula χ' one of whose fixed–point processes is the process belonging to χ. Another of χ''s processes is the same process proceeding with twice the velocity. By the observation above these two processes eventually reach the same state. Once they do we stop and look wether a fixed–point is reached or not. A third process is needed to indicate that this moment has come.

Formally, we let Y and Z be relation variables of the same arity as X, I a unary relation variable, and

$$\chi' = [s\text{–}FP_{\bar{y},Y,\bar{z},Z,v,I}\ \varphi_Y, \varphi_Z, \varphi_I]\bar{u}$$

where

$$\begin{aligned}
\varphi_Y &= (I = \emptyset \wedge \varphi(\bar{y}, Y)) \vee (I = \mathcal{U} \wedge Y = \varphi(_, Y) \wedge Y\bar{y})\\
\varphi_Z &= (I = \emptyset \wedge \varphi(\bar{z}, \varphi(_, Z))) \vee (I = \mathcal{U} \wedge Y = \varphi(_, Y) \wedge Y\bar{z})\\
\varphi_I &= (Iv \vee \varphi(_, Y) = \varphi(_, \varphi(_, Z))).
\end{aligned}$$ [41]

Let $\mathfrak{I}$ be an interpretation for χ and consider the fixed–point processes of χ and χ'. We choose $l \geq 1$ minimal such that $X_l^{\mathfrak{I}} = X_{2l}^{\mathfrak{I}}$.
Instead of giving a formal (inductive) proof that $\chi \dashv\models \chi'$[42] and that χ' in fact reaches a fixed–point we illustrate the processes in a table (where, for example, column l in row Y gives $Y_l^{\mathfrak{I}}$) which is much easier to understand.

	1	2	3	...	$l-1$	l	$l+1$	$l+2$	...
X	$X_1^{\mathfrak{I}}$	$X_2^{\mathfrak{I}}$	$X_3^{\mathfrak{I}}$	...	$X_{l-1}^{\mathfrak{I}}$	$X_l^{\mathfrak{I}}$	$X_{l+1}^{\mathfrak{I}}$	$X_{l+2}^{\mathfrak{I}}$	...
Y	$X_1^{\mathfrak{I}}$	$X_2^{\mathfrak{I}}$	$X_3^{\mathfrak{I}}$	...	$X_{l-1}^{\mathfrak{I}}$	$X_l^{\mathfrak{I}}$	$X_\infty^{\mathfrak{I}}$	$X_\infty^{\mathfrak{I}}$	...
Z	$X_2^{\mathfrak{I}}$	$X_4^{\mathfrak{I}}$	$X_6^{\mathfrak{I}}$	...	$X_{2l-2}^{\mathfrak{I}}$	$X_{2l}^{\mathfrak{I}}$	$X_\infty^{\mathfrak{I}}$	$X_\infty^{\mathfrak{I}}$	...
I	$\emptyset$	$\emptyset$	$\emptyset$	...	$\emptyset$	$\mathcal{U}$	$\mathcal{U}$	$\mathcal{U}$	...

We complete the proof with an application of the Lemma on Simultaneous Induction to χ'. ■

A.4 Lemma : *Let $\chi = [FP_{\bar{x},X}\varphi(\bar{x}, X, \bar{v})]\bar{u}$ where $\bar{v}$ is the tuple of all free individual variables of φ distinct from $\bar{x}$.*
Then χ is equivalent to the formula

$$\chi' = [FP_{\bar{x}\bar{y},Y}\varphi(\bar{x}, Y\_\bar{y}, \bar{y})]\bar{u}\bar{v}\ .$$

The **proof** is an easy induction. ■

[41] The notation is (hopefully) self–explanatory. For example, $\varphi(\bar{z}, \varphi(\_, Z))$ denotes the formula obtained from $\varphi(\bar{z}, X)$ by replacing each subformula of the form $X\bar{t}$ by $\varphi(\bar{t}, Z)$; and $Y = \varphi(\_, Y)$ abbreviates $\forall\bar{y}(Y\bar{y} \leftrightarrow \varphi(\bar{y}, Y))$.
$I = \emptyset$ abbreviates $\forall x\neg Ix$ and $I = \mathcal{U}$ stands for $\forall x Ix$.

[42] By $\chi \dashv\models \chi'$ we denote that χ and χ' are equivalent, i.e. for all (finite) interpretations $\mathfrak{I}$ we have $\mathfrak{I} \models \chi \iff \mathfrak{I} \models \chi'$.

A.5 Transitivity Lemma : *Let* $\chi = [FP_{\bar{x},X}\varphi]\bar{u} \in$ *TFP such that at most one fixed–point operator occurs in* φ*. Then there exists a TFP-formula* χ' *which is equivalent to* χ *of the form*

$$[FP_{\bar{z},Z}\varphi']\bar{u}u_1 \ldots u_1$$

where φ' *is first–order.*

If χ *is inductive then* χ' *can also be chosen to be inductive.*

Proof (sketch): Let $\zeta = [FP_{\bar{y},Y}\psi(\bar{y}, X, Y)]\bar{v}$ be a subformula of φ. (If no subformula of this form exists there is nothing to prove.)
By Lemma A.4 we can assume without loss of generality that ψ has no free individual variables besides $\bar{y}$. For convenience we assume without loss of generality that χ has no free relation variables so all free variables of ψ are in $\{\bar{y}, X, Y\}$.
Let

$$\begin{array}{ll} \varphi_X = & (Y = \psi(\_, X, Y) \wedge \varphi^*(\bar{x}, X, Y)) \\ & \quad \text{If } Y \text{ reaches a fixed–point } X \text{ enters the next stage,} \\ & \vee (Y \neq \psi(\_, X, Y) \wedge X\bar{x}) \\ & \quad \text{otherwise it remains unchanged.} \end{array}$$

(where φ^* is the formula obtained from φ by replacing the subformula $[FP_{\bar{y},Y}\psi]\bar{v}$ by $Y\bar{v}$)
and

$$\begin{array}{ll} \varphi_Y = & (Y \neq \psi(\_, X, Y) \wedge \psi(\bar{y}, X, Y)) \\ & \quad \text{If } Y \text{ is not a fixed–point it goes to the next stage,} \\ & \quad \text{if it reaches a fixed–point it is set back to } \emptyset, \\ & \vee (X = \varphi^*(\_, X, Y) \wedge Y = \psi(\_, X, Y) \wedge Y\bar{y}) \\ & \quad \text{unless } X \text{ also reaches a fixed–point, then } Y \text{ remains unchanged.} \end{array}$$

Note that no fixed–point operators occur in φ_X or φ_Y.
It can be verified inductively that χ is equivalent to the formula

$$[s\text{–}FP_{\bar{x},X,\bar{y},Y}\ \varphi_X, \varphi_Y]\bar{u}.$$

The idea is that the fixed–point process of ψ is simulated until Y reaches a fixed–point. At this stage X takes its next step and Y is reset to be $\emptyset$. Since χ is totally defined Y always reaches a fixed–point.
Note that it is important that ψ has no free variables besides $\bar{y}$, X, and Y because otherwise the fixed–point process of Y could depend on variables that are quantified within φ, but we use that the process only depends on X.

An application of the Lemma on Simultaneous Induction (1.12) completes the proof for *TFP*–formulae.

The situation is a bit more complicated for *IFP*–formulae (if we want χ' to be inductive) because there we cannot reset Y to $\emptyset$ every time it reaches a fixed–point. So we have to indicate, within the process that simulates the process of ζ, at which stage of the process that simulates the process of χ we are.
Therefore we note that for inductive χ each tuple enters X at most once, so we can use the new tuples as indicators.
To guarantee that there always exist such tuples until a fixed–point is reached (in particular after the first step) we increase the arity of X by 2 and always include all tuples whose last two components are distinct. The interesting tuples, however, are those whose last two components are equal.

Say, X is r–ary and Y s–ary. Let X' and X^- be new $(r+2)$–ary relation variables and Y' a new $(r+2+s)$–ary relation variable. Again we will simulate χ by a simultaneous induction, this time over three formulae $\varphi_{X'}$, φ_{X^-}, and $\varphi_{Y'}$ (where $\varphi_{X'}$and $\varphi_{Y'}$ have the same function as φ_X and φ_Y had in the general case and φ_{X^-} is needed for technical reasons).
We let

$$\varphi_{X'}(\overset{r+2}{x}, X', X^-, Y') := X' \overset{r+2}{x} \vee \; x_{r+1} \neq x_{r+2}$$

In the first step we include all tuples whose last two components are distinct.

$$\vee (\quad x_{r+1} = x_{r+2} \wedge X' \neq \emptyset$$
$$\wedge Y' \_ = \varphi_{Y'}(\_, X', X^-, Y')$$
$$\wedge \varphi^*(\overset{r}{x}, X' \_ x_1 x_1, \exists \overset{r+2}{z} \in (X' \setminus X^-) \; Y \overset{r+2}{z} \_))$$

Once the first step is taken we take the next step in the process of χ whenever Y' reaches a fixed–point.

(where $\varphi^*(\overset{r}{x}, X, Y)$ is defined as above),

$$\varphi_{X^-}(\overset{r+2}{x}, X', X^-, Y') := X^- \overset{r+2}{x} \vee (Y' \_ = \varphi_{Y'}(\_, X', X^-, Y') \wedge X' \overset{r+2}{x}),$$

Whenever X' goes on one step X^- takes its old value. Otherwise it remains unchanged.

and

$$\varphi_{Y'}(\overset{r+2}{x}\overset{s}{y}, X', X^-, Y') := \quad Y' \overset{r+2}{x}\overset{s}{y}$$

We keep Y inductive.

$$\vee (X' \overset{r+2}{x} \wedge \neg X^- \overset{r+2}{x} \wedge \psi(\overset{s}{y}, X' \_ x_1 x_1, Y' \overset{r+2}{x} \_)).$$

Moreover, we include all those tuples whose $\bar{x}$–component is in $X' \setminus X^-$ and whose $\bar{y}$–component is in the next stage of the process of ζ with X interpreted as the "relevant" part of X'.

Note that $\varphi_{X'}$, φ_{X^-}, and $\varphi_{Y'}$ are first–order. Let

$$\eta := [s\text{–}FP_{\bar{x}, X', \bar{x}, X^-, \bar{x}\bar{y}, Y'} \varphi_{X'}, \varphi_{X^-}, \varphi_{Y'}] \bar{u} u_1 u_1$$

It is immediate from the first disjunct of the definition of $\varphi_{X'}$, φ_{X^-}, and φ'_Y that η is inductive. Once we have seen that it is equivalent to χ we can apply the Lemma on Simultaneous Induction and we are finished.

So let $\mathfrak{I}$ be an interpretation and let X'_i, X^-_i and Y'_i (for $i \geq 0$) denote the stages of respective processes in the simultaneous induction.[43]
Moreover, let X_i (for $i \geq 0$) denote the stages of the fixed–point process of χ and, for each $i \geq 0$, let Y_{ij} (for $j \geq 0$) denote the stages of the fixed–point process of ζ when X is interpreted by X_i. Note that since ψ has no free variables besides $\bar{y}$, X, and Y this fixed–point process only depends on the interpretation of X.
By X_∞ and X'_∞ we denote the fixed–points of the sequences $(X_i)_{i\geq 1}$ and $(X'_i)_{i\geq 1}$ respectively.

Claim: $X'_\infty = \{\overset{r+2}{a} \in A \mid a_{r+1} \neq a_{r+2}\} \cup \{\overset{r+2}{a} \mid a_{r+1} = a_{r+2} \wedge X_\infty \overset{r}{a}\}$.

Clearly this implies $\eta \mathrel{=\!\mid\!\models} \chi$.

Proof: Let $k, l \geq 0$ and consider the following statement $(S_{k,l})$:

- $X'_k = \{\overset{r+2}{a} \in A \mid a_{r+1} \neq a_{r+2}\} \cup \{\overset{r+2}{a} \in A \mid a_{r+1} = a_{r+2} \wedge X_l \overset{r}{a}\}$
- $X^-_k = X'_{k-1} \subsetneq X'_k$
- For all $\overset{r+2}{a} \in X'_k - X^-_k$ we have $Y'_k \overset{r+2}{a} \_ = \emptyset = Y_{l0}$.

It can be easily verified that the statement $(S_{1,0})$ holds.

Let now $k \geq 1, l \geq 0$ such that $(S_{k,l})$ holds.
Let $m = \min\{i \mid Y_{li} = Y_{l(i+1)}\}$. Inductively it can be seen that for each $i \leq m$ we have $Y'_{k+i} \overset{r+2}{a} \_ = Y_{li}$ (for each $\overset{r+2}{a} \in X'_k \setminus X^-_k$), $X'_{k+i} = X'_k$, and $X^-_{k+i} = X^-_k$.
But then we have

$$
\begin{aligned}
X'_{k+m+1} &= X'_k \cup \{\overset{r+2}{a} \mid a_{r+1} = a_{r+2} \wedge \mathfrak{I} \models \varphi^*[\overset{r}{a}, X'_k \_ a_1 a_1, Y_{lm}]\} \\
&= X'_k \cup \{\overset{r+2}{a} \mid a_{r+1} = a_{r+2} \wedge \mathfrak{I} \models \varphi[\overset{r}{a}, X_l]\} \\
&= \{\overset{r+2}{a} \in A \mid a_{r+1} \neq a_{r+2}\} \cup \{\overset{r+2}{a} \mid a_{r+1} = a_{r+2} \wedge X_{l+1} \overset{r}{a}\},
\end{aligned}
$$

$X^-_{k+m+1} = X'_{k+m}$, and $Y'_{k+m+1} = Y'_{k+m}$. Note that $X'_{k+m} = X'_{k+m+1}$ if and only if $X_l = X_{l+1}$. In this case the sequences have reached their fixed–point and our claim is proved. Since inductive formulae always reach a fixed–point this case eventually occurs.
Otherwise, the statement $(S_{k+m+1,l+1})$ holds and we continue with our induction.

■

This was a bit technical, but with the help of the Transitivity Lemma it is rather simple to give an inductive proof of the normal form theorem.

[43] For convenience we suppress the superscript $\mathfrak{I}$ here.

Observe that for each fixed–point formula ψ and v, x, X that do not occur freely in ψ we have

$$\psi \mathrel{=\!\!|\!\models} \overset{\exists}{\forall} v[FP_{x,X}\psi]v\ .$$

We are going to prove inductively over a *TFP*–formula χ that it is equivalent to a *TFP*–formula χ' of the form

$$\overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi]v\ldots v$$

and that χ' can be chosen to be inductive if χ is. This suffices to prove Theorem A.1 because by Lemma A.3 every fixed–point formula is equivalent to a *TFP*–formula.

$\chi \in FO$:
: By the observation above χ is equivalent to $\chi' := \overset{\exists}{\forall} v[FP_{x,X}\chi]v$ (where v, x, X do not occur in χ).

$\chi = \neg \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi]v\ldots v$, $\chi = \exists y \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi]v\ldots v$, or
$\chi = \forall y \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi]v\ldots v$:
: By the observation above χ is equivalent to $\psi := \overset{\exists}{\forall} w[FP_{z,Z}\chi]w$ (where w, z, Z do not occur in χ). We apply the Transitivity Lemma to ψ and obtain a formula χ' of the desired form.

$\chi = \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi_1]v\ldots v\ \wedge\ \underbrace{\overset{\exists}{\forall} w[FP_{\bar{y},Y}\ \varphi_2]w\ldots w}_{=:\chi_2}$:
: Without loss of generality we can assume that the variables $\bar{x}$ and X do not occur freely in χ_2. Then an easy induction shows that χ is equivalent to

$$\psi := \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi_1\ \wedge\ \chi_2]v\ldots v$$

To obtain χ' we apply the Transitivity Lemma to ψ.

$\chi = \overset{\exists}{\forall} v[FP_{\bar{x},X}\ \varphi_1]v\ldots v\ \vee\ \overset{\exists}{\forall} w[FP_{\bar{y},Y}\ \varphi_2]w\ldots w$:
: Can be treated with the $\wedge$ and $\neg$–case.

$\chi = [FP_{\bar{x},X}\ \overset{\exists}{\forall} v[FP_{\bar{y},Y}\ \varphi]v\ldots v]\bar{u}$:
: An application of the Transitivity Lemma to χ gives us a formula ψ equivalent to χ of the form

$$[FP_{\bar{z},Z}\xi]\bar{u}u_1\ldots u_1\ .$$

To put the variables in the right form we apply the Transitivity Lemma again to

$$\overset{\exists}{\forall} w[FP_{z,Z}\psi]w$$

(where w, z, Z do not occur in ψ) to obtain χ'.

■

A.6 Remark : As mentioned in the introduction of Section 2 there is a different way to establish the normal form for **IFP**: One proves the Transitivity Lemma for positive **LFP** (i.e. the fragment of **LFP** where negation symbols only occur in front of atomic subformulae) and then shows that positive **LFP** is closed under negation and equals **IFP**(which can essentially both be done in one step). The proof of the Transitivity Lemma for positive **LFP** is much easier than the proof for **IFP** because one can use the monotonicity of *LFP*–operators. So probably this approach is a bit simpler.
We have chosen our approach for the following three reasons: The first is pure interest in the question if there is a direct prove that does not incorporate **LFP** for a result that only concerns **IFP**. Second we want to stress the similarities between **IFP** and **PFP** which become clearer in this approach.
And third, our proof goes through without any difficulties if we augment **IFP** by a family of (maybe none-monotone) Lindström–quantifier's. So we also obtain a normal–form result for such stronger logics. It is not possible to prove this result the other way because it does not even make sense to augment **LFP** by none-monotene Lindström–quantifiers. □

Appendix B:
A Simple Proof of the Weak Main Lemma

As usual we fix $k \geq 2$ and let E_k be a $2k$–ary relation variable. Recall that the weak main lemma (8.1) says that the transitive closure of E_k is not definable in $\mathbf{PFP}^{k-1}$.
In this appendix we sketch a different proof of the weak main lemma which was essentially given in [Gro93].
It uses a probabilistic argument to construct structures which are, in some sense, homogeneous in the k–tuples, and again the k–ary r–pebble game to prove that the transitive closure of E_k is not $\mathbf{PFP}^{k-1}$–definable on these structures.

§B.1 Extension Axioms for Parametric Classes

In this paragraph we take care of the probabilistic part of our proof. We want to find structures where the k–tuples (of distinct elements) can be divided into two parts, but where each $(k-1)$–tuple can be extended to k–tuples in both parts.
The idea is to take structures with only one k-ary relation P that is chosen randomly on the set of k–tuples of distinct elements. It is plausible that if the structures are big enough the statement above holds.

To realize this we make use of a theorem by Oberschelp [Obe82] concerning 0–1 laws on parametric classes.

Let us fix a signature σ only consisting of finitely many relation symbols.

For each $n \geq 1$ we can define a probability measure l_n on the class of all classes of σ-structures, letting

$$l_n(\mathcal{C}) = \frac{|\{\mathfrak{A} \in \mathcal{C} \mid A = \{1, \ldots, n\}\}|}{|\{\mathfrak{A} \mid \mathfrak{A} \ \sigma\text{-structure}, \ A = \{1, \ldots, n\}\}|}$$

for each class $\mathcal{C}$ of σ-structures.
l_n gives the probability that a randomly chosen σ-structure with universe $\{1, \ldots, n\}$ is in $\mathcal{C}$.

If $\mathcal{D}$ is another class of σ-structures with $l_n(\mathcal{D}) > 0$ we can define the conditional probability of $\mathcal{C}$ with respect to $\mathcal{D}$ by

$$l_n(\mathcal{C}|\mathcal{D}) = \frac{l_n(\mathcal{C} \cap \mathcal{D})}{l_n(\mathcal{D})} \ .$$

If $\mathcal{C} = Mod(\varphi)$ we write $l_n(\varphi)$ instead of $l_n(\mathcal{C})$ and $l_n(\varphi|\mathcal{D})$ instead of $l_n(\mathcal{C}|\mathcal{D})$.

It is a well–known result ([GKLT69, Fag76]) that a *0–1 law* holds for first–order logic, i.e. for each first–order sentence φ the *labeled asymptotic probability*

$$l(\varphi) := \lim_{n \to \infty} l_n(\varphi)$$

exists and is 0 or 1.

Oberschelp extended this result to so called parametric classes:

B.1 Theorem ([Obe82]) : *A first–order 0-1 law holds on parametric classes, i.e. if $\mathcal{P}$ is a parametric class then for each first–order sentence φ the conditional labeled asymptotic probability $l(\varphi|\mathcal{P}) := \lim_{n\to\infty} l_n(\varphi|\mathcal{P})$ exists and is 0 or 1.*

We do not want to introduce parametric classes formally (they are classes that are definable by certain universal first–order sentences) but just point out that for each $k \geq 1$ the class $\mathcal{P}_k$ of $\{P\}$–structures with a k–ary relation P that only contains k–tuples of distinct elements is parametric.
Oberschelp's theorem is proved (cf. [Com88]) using the so called *extension axioms.*

B.2 Definition : (1) The *quantifier–free type of a tuple* $\overset{r}{a}$ *in a structure* $\mathfrak{A}$ is the formula

$$\varphi^0_{\mathfrak{A},\overset{r}{a}} := \bigwedge \Big\{\theta \mid \theta \text{ atomic or negated atomic first–order, } \text{free}(\theta) \subseteq \{x_1,\ldots,x_r\}, \mathfrak{A} \models \theta[\overset{r}{a}]\Big\}.$$

A *quantifier–free type* is a formula $\varphi^0_{\mathfrak{A},\bar{a}}$ for some structure $\mathfrak{A}$ and tuple $\bar{a} \in A$.

(2) Let $\varphi(\overset{r}{x})$, $\psi(\overset{r+1}{x})$ be quantifier–free types such that $\psi(\overset{r+1}{x})$ extends $\varphi(\overset{r}{x})$, i.e. $\models \psi(\overset{r+1}{x}) \to \varphi(\overset{r}{x})$.
Then we call the sentence

$$\forall x_1 \ldots \forall x_r (\varphi(\overset{r}{x}) \to \exists x_{r+1} \psi(\overset{r+1}{x}))$$

an $(r+1)$-*extension axiom.*

Φ_r denotes the set of all s–extension axioms where $s \leq r$.

(3) Let $\mathcal{P}$ be a parametric class. We call an extension axiom

$$\forall x_1 \ldots \forall x_r (\varphi(\overset{r}{x}) \to \exists x_{r+1} \psi(\overset{r+1}{x}))$$

compatible with $\mathcal{P}$ if $\exists x_1 \ldots \exists x_{r+1} \psi$ has a model in $\mathcal{P}$.

$\Phi_r(\mathcal{P})$ denotes the set of all s-extension axioms ($s \leq r$) compatible with $\mathcal{P}$. $\square$

The essential step in the proof of Oberschelp's Theorem is to show that $l(\bigwedge \Phi_r(\mathcal{P})|\mathcal{P}) = 1$ for each parametric class $\mathcal{P}$ and $r \geq 1$, and actually this is what we use.

Turning to the parametric class $\mathcal{P}_k$ defined above $l(\bigwedge \Phi_r(\mathcal{P}_k)|\mathcal{P}_k) = 1$ implies that for sufficiently large m (depending on k and r) we have $l_m(\bigwedge \Phi_r(\mathcal{P}_k)|\mathcal{P}_k) > \frac{1}{2}$.

B.3 Lemma : *Let $k, r \geq 1$, P k-ary. Then we can find an $m(k,r) \geq 1$ and $\{P\}$-structures $\mathfrak{L}(k,r)$ and $\mathfrak{R}(k,r)$ in $\mathcal{P}_k$ with universe $\{1,\ldots,m(k,r)\}$ which are both models of $\Phi_r(\mathcal{P}_k)$ such that for all distinct $a_1,\ldots,a_k \in \{1,\ldots,m(k,r)\}$:*

$$\overset{k}{a} \in P^{\mathfrak{L}(k,r)} \iff \overset{k}{a} \notin P^{\mathfrak{R}(k,r)}.$$

Proof: For each $\mathfrak{A} \in \mathcal{P}_k$ with universe $\{1,\ldots,m\}$ there exists exactly one $\mathfrak{B}_{\mathfrak{A}} \in \mathcal{P}_k$ with universe $\{1,\ldots,m\}$ such that for all distinct $a_1,\ldots,a_k \in \{1,\ldots,m\}$:

$$\overset{k}{a} \in P^{\mathfrak{A}} \iff \overset{k}{a} \notin P^{\mathfrak{B}_{\mathfrak{A}}}.$$

If for all $\mathfrak{A} \in \mathcal{P}_k$ with universe $\{1,\ldots,m\}$ either $\mathfrak{A} \not\models \Phi_r(\mathcal{P}_k)$ or $\mathfrak{B}_{\mathfrak{A}} \not\models \Phi_r(\mathcal{P}_k)$ we have $l_m(\bigwedge \Phi_r(\mathcal{P}_k) \mid \mathcal{P}_k) \leq \frac{1}{2}$. Thus for sufficiently large m the statement holds.
■

§B.2 The Structures

Let $k, n \geq 2$, $r \geq 0$, and $m = m(k,r)$. Let $\mathfrak{L} = \mathfrak{L}(k,r)$ and $\mathfrak{R} = \mathfrak{R}(k,r)$ be the structures obtained from Lemma 2.3.
We define an E_k–structure $\mathfrak{F} = \mathfrak{F}(k,r,n)$ as follows:
We let

$$F = \{1,\ldots,n\} \times \{1,\ldots,m\}\ .$$

Denoting tuples $(I,a_1)\ldots(I,a_k) \in F^k$ by $(I,\overset{k}{a})$ we define

$$\begin{aligned} E_k^{\mathfrak{F}} :=\ & \{(I,\overset{k}{a})(I+1,\overset{k}{b}) \mid I < n \wedge P^{\mathfrak{L}}\overset{k}{a} \wedge P^{\mathfrak{L}}\overset{k}{b})\} \\ & \cup\{(I,\overset{k}{a})(I+1,\overset{k}{b}) \mid I < n \wedge P^{\mathfrak{R}}\overset{k}{a} \wedge P^{\mathfrak{R}}\overset{k}{b}\}. \end{aligned}$$

We define the mapping row on F by $\mathrm{row}((I,a)) = I$.
We say that a k–tuple $\overset{k}{\boldsymbol{a}} \in F$ belongs to the *right component (left component)* if it is of the form $(I,\overset{k}{a})$, where $P^{\mathfrak{R}}\overset{k}{a}$ ($P^{\mathfrak{L}}\overset{k}{a}$ respectively). We immediately see:

B.4 Lemma : *There exists an E_k–path from a tuple $\overset{k}{\boldsymbol{a}} \in F$ to a tuple $\overset{k}{\boldsymbol{b}} \in F$ if and only if* $\mathrm{row}(\boldsymbol{a}_1) < \mathrm{row}(\boldsymbol{b}_1)$ *and either both $\overset{k}{\boldsymbol{a}}$ and $\overset{k}{\boldsymbol{b}}$ belong to the right component or both tuples belong to the left component.*

The following lemma shows that the rows of our structure can be treated independently.

B.5 Lemma : *Let π_1 and π_2 be partial isomorphisms both either from $\mathfrak{L}$ to $\mathfrak{L}$ (case 1) or from $\mathfrak{L}$ to $\mathfrak{R}$ (case 2).*
Then for each $I < n$ the partial mapping π on F defined by

$$\pi((J,a)) = \begin{cases} (J, \pi_1(a)) \text{ if } J = I \\ (J, \pi_2(a)) \text{ if } J = I+1 \end{cases}$$

is a partial isomorphism of $\mathfrak{F}$.

Proof: Let us consider for example case 2:
Clearly π is injective. Let $\overset{k}{a} \in \mathrm{dom}(\pi_1), \overset{k}{b} \in \mathrm{dom}(\pi_2)$. Then

$$\begin{array}{ll} & E^{\mathfrak{F}}(I, \overset{k}{a})(I+1, \overset{k}{b}) \\ \iff & (P^{\mathfrak{L}}\overset{k}{a} \wedge P^{\mathfrak{L}}\overset{k}{b}) \vee (P^{\mathfrak{R}}\overset{k}{a} \wedge P^{\mathfrak{R}}\overset{k}{b}) \\ \iff & (P^{\mathfrak{R}}\pi_1(\overset{k}{a}) \wedge P^{\mathfrak{R}}\pi_2(\overset{k}{b})) \vee (P^{\mathfrak{L}}\pi_1(\overset{k}{a}) \wedge P^{\mathfrak{L}}\pi_2(\overset{k}{b})) \\ \iff & E^{\mathfrak{F}}(I, \pi_1(\overset{k}{a}))(I+1, \pi_2(\overset{k}{b})) \\ \iff & E^{\mathfrak{F}}\pi(I, \overset{k}{a})\pi(I+1, \overset{k}{b}). \end{array}$$

■

§B.3 The Game

Fix $q \geq 0$, let $r := q + k$ and $n := 2^{q+1} + 1$, and consider the structure $\mathfrak{F} := \mathfrak{F}(k, r, n)$ defined in the previous section. Select tuples $\overset{k}{c} \in P^{\mathfrak{L}}$ and $\overset{k}{d}\,' \in P^{\mathfrak{R}}$ (where $\mathfrak{R} := \mathfrak{R}(k, r)$ and $\mathfrak{L} := \mathfrak{L}(k, r)$ are the structures used to define $\mathfrak{F}$), let $\overset{k}{\boldsymbol{c}} := (1, \overset{k}{c})$, $\overset{k}{\boldsymbol{d}} := (n, \overset{k}{c})$, and $\overset{k}{\boldsymbol{d}}{}' = (n, \overset{k}{d}{}')$.
We define $\{E_k, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}\}$–structures $\mathfrak{F}_1 := (\mathfrak{F}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}})$ and $\mathfrak{F}_2 := (\mathfrak{F}, \overset{k}{\boldsymbol{c}}, \overset{k}{\boldsymbol{d}}{}')$. Then by Lemma B.4 we have

$$\mathfrak{F}_1 \models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}} \quad \text{and} \quad \mathfrak{F}_2 \not\models [TC_{\overset{k}{x},\overset{k}{y}} E_k \overset{k}{x}\overset{k}{y}]\, \overset{k}{\underline{c}}, \overset{k}{\underline{d}}\ .$$

We have to prove:

B.6 Proposition : *The duplicator has a winning strategy for the $(k-1)$–ary q–pebble game on $\mathfrak{F}_1$ and $\mathfrak{F}_2$.*

Proof: We are going to show that the duplicator can preserve the following throughout the whole game:
Suppose there are $q' \leq q$ pebbles on the board.
Then there exist $M, N \in \{1, \ldots, n\}$ such that $N - M \geq 2^{q+1-q'}$ and

$(*)$ For pairs $(\boldsymbol{a}, \boldsymbol{b})$ of corresponding elements we have:

$$\text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \leq M \quad \text{or} \quad \text{row}(\boldsymbol{a}) = \text{row}(\boldsymbol{b}) \geq N$$

and for each $I \in \{1, \ldots, n\}$

$(*_I)$ Suppose that $(I, a_1), (I, b_1), \ldots, (I, a_s), (I, b_s)$ are the pairs of corresponding elements in row I.

If $I \leq M$ then $\varphi^0_{\mathfrak{L}, \overset{s}{a}} = \varphi^0_{\mathfrak{L}, \overset{s}{b}}$.

If $I \geq N$ then $\varphi^0_{\mathfrak{L}, \overset{s}{a}} = \varphi^0_{\mathfrak{R}, \overset{s}{b}}$.

Once we have seen this we are finished by Lemma B.5 (making use of the fact that there are only edges between succeeding rows of $\mathfrak{F}$).
Obviously $(*)$ and $(*_I)$ (for all I) hold in the beginning with $M = 1$ and $N = n$.

So let there be $q' \leq q$ pairs of pebbles on the board and let M, N be defined such that $(*)$ and $(*_I)$ (for all I) hold. We are going to define the duplicator's answer to the next move and M', N' to preserve $(*)$ and $(*_I)$ (for all I).

$\exists$-move: The challenger places a pebble on $(I, a) \in \mathfrak{F}_1$.

Suppose that $(I, a_1), (I, b_1), \ldots, (I, a_s), (I, b_s)$ are the pairs of corresponding elements in row I. Note that $s < r$.

Case 1: $I - M \leq N - I$

Let $M' := max\{M, I\}$ and $N' := N$.

By the induction hypothesis we have $\varphi^0_{\mathfrak{L}, \overset{s}{a}} = \varphi^0_{\mathfrak{L}, \overset{s}{b}}$.

Since $s < r$ and $\mathfrak{L} \models \varphi^0_{\mathfrak{L}, \overset{s}{a}a}[\overset{s}{a}, a]$ the extension axiom

$$\forall \overset{s}{v} \, (\varphi^0_{\mathfrak{L}, \overset{s}{b}} \rightarrow \exists v_{s+1} \varphi^0_{\mathfrak{L}, \overset{s}{a}a})$$

is in $\Phi_r(\mathcal{P}_k)$. Select $b \in \{1, \ldots, n\}$ such that $\mathfrak{L} \models \varphi^0_{\mathfrak{L}, \overset{s}{a}a}[\overset{s}{b}, b]$. The duplicator places the corresponding pebble on (I, b). Then clearly $(*_I)$ still holds.

Case 2: $I - M > N - I$

Let $M' := M$ and $N' := min\{I, N\}$.

By the induction hypothesis we have $\varphi^0_{\mathfrak{L}, \overset{s}{a}} = \varphi^0_{\mathfrak{R}, \overset{s}{b}}$.

Since $\mathfrak{L} \models \varphi^0_{\mathfrak{L}, \overset{s}{a}a}[\overset{s}{a}, a]$ the extension axiom

$$\forall \overset{s}{v} \, (\varphi^0_{\mathfrak{R}, \overset{s}{b}} \rightarrow \exists v_{s+1} \varphi^0_{\mathfrak{L}, \overset{s}{a}a})$$

is in $\Phi_r(\mathcal{P}_k)$. Select $b \in \{1, \ldots, n\}$ such that $\mathfrak{R} \models \varphi^0_{\mathfrak{L}, \overset{s}{a}a}[\overset{s}{b}, b]$. The duplicator places the corresponding pebble on (I, b) and $(*_I)$ still holds.

Note that in both cases

$$N' - M' \geq \left\lfloor \frac{N-M}{2} \right\rfloor \geq 2^{q+1-q'-1}$$

and $(*)$ holds in the new situation with M' instead of M and N' instead of N. For all $J \neq I$, $(*_J)$ is not affected by this move so it still holds.

$\forall$-move: Can be treated analogously to the $\exists$-move.

I-move: Nothing has to be done here.

R–move: The challenger reduces the depth to $d \geq 1$ and leaves the elements $\boldsymbol{a}_1, \boldsymbol{b}_1, \ldots, \boldsymbol{a}_l, \boldsymbol{b}_l$ $(l \leq k-1)$ of depth $\geq d$ on the board.

All other pebbles that are still on the board (those of depth $< d$) have already been in their position when the last I–move increasing the depth of the game from $d-1$ to d was made. We look back to that situation; say M'', N'' satisfied $(*)$ and $(*_I)$ (for all I) then. Moreover we suppose that q'' pairs of pebbles have been on the board then, thus $(q''+l)$ are on the board now.

By the pigeonhole principle there exist M', N' such that $M'' \leq M' < N' \leq N''$,

$$N' - M' \geq \frac{N'' - M''}{l+1} \geq \frac{2^{q+1-q''}}{l+1} \geq 2^{q+1-(q''+l)},$$

and none of the pairs of corresponding elements that are still on the board are in a row between M' and N'.

Note that by the induction hypothesis there are also no pairs of corresponding elements in a row between M and N.

Let I be a row. If $I \leq \min\{M', M\}$ or $I \geq \max\{N', N\}$ then $(*_I)$ clearly still holds in the new situation (with M' instead of M and N' instead of N).

If $N' \leq I \leq M$ or $N \leq I \leq M'$ suppose that I contains elements $(I, a_{i_1}), \ldots, (I, a_{i_p}) \in \{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_l\}$.

Let $(I, b_{i_1}), \ldots, (I, b_{i_p})$ be the corresponding elements. Note that $p \leq k-1$ and recall that $P^{\mathfrak{L}}$ and $P^{\mathfrak{R}}$ only contain k–tuples of distinct elements. Let $\bar{a} = a_{i_1} \ldots a_{i_p}$ and $\bar{b} = a_{i_1} \ldots a_{i_p}$. Then we have

$$\varphi^0_{\mathfrak{L},\bar{a}} = \varphi^0_{\mathfrak{L},\bar{b}} = \varphi^0_{\mathfrak{R},\bar{a}} = \varphi^0_{\mathfrak{R},\bar{b}} \,.$$

Hence $(*_I)$ still holds in the new situation.

$\blacksquare$

Appendix C: The Structure of Fixed–Point Logics

Let us finally give a schematic survey of the structure of fixed–point logics as far as it is known by now.

C.1 The Relation between the Logics

$$
\begin{tikzcd}
 & \exists\mathbf{TC} \arrow[r, phantom, "\subsetneq_1"] \arrow[dr, phantom, sloped, "\subsetneq_1"] & \exists\mathbf{LFP} \arrow[dr, phantom, sloped, "\subsetneq_2"] & & & & \\
\exists\mathbf{FO} \arrow[ur, phantom, sloped, "\subsetneq"] & & & \mathbf{SFP} \arrow[r, phantom, "\subsetneq_4"] & \mathbf{LFP} \arrow[r, phantom, "=_5"] & \mathbf{IFP} \arrow[r, phantom, "\subseteq_6"] & \mathbf{PFP} \\
 & \mathbf{FO} \arrow[ul, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \mathbf{TC} \arrow[ur, phantom, sloped, "\subsetneq_3"] & & & &
\end{tikzcd}
$$

[1] Theorem 4.7

[2] Lemma 1.8

[3] [Imm81]

[4] [Dah87, Kol91]

[5] [GS86]

[6] Equality holds if and only if **PTIME=PSPACE** [AV91b]

C.2 Normal Forms

$$
\begin{aligned}
& \mathbf{PFP} && = (1,1)\text{–bounded } \mathbf{PFP} \text{ in normal form}^{7} \\
& \mathbf{IFP} && = (1,1)\text{–bounded } \mathbf{IFP} \text{ in normal form}^{7} \\
= \;& \mathbf{LFP} && = \mathbf{LFP} \text{ in bounded Skolem form}^{8} \\
& \mathbf{SFP} && = (2,0)\text{–bounded } \mathbf{PFP}^{9} \\
& \exists\mathbf{LFP} && = (2,0)\text{–bounded } \mathbf{PFP} \text{ in normal form}^{9} \\
& \mathbf{TC} && = (1,0)\text{–bounded } \mathbf{PFP}^{10} \\
& \exists\mathbf{TC} && = (1,0)\text{–bounded } \mathbf{PFP} \text{ in normal form}^{10}
\end{aligned}
$$

[7] Theorem 3.6

[8] Theorem 3.12

[9] Theorem 3.9

[10] Corollary 4.6

C.3 Arity Hierarchies

$$
\begin{array}{ccccc}
\vdots & & & & \vdots \\
\exists\mathbf{TC}^{k-1} & \subseteq & \cdots & \subseteq & \mathbf{PFP}^{k-1} \\
\rotatebox{-90}{$\not\subseteq$} & & \rotatebox{20}{$\not\subseteq$} & & \rotatebox{-90}{$\not\subseteq$} \\
\exists\mathbf{TC}^{k} & \subseteq & \cdots & \subseteq & \mathbf{PFP}^{k} \\
\vdots & & & & \vdots
\end{array}
$$

(Main Lemma)

$$
\begin{array}{ccc}
\vdots & & \vdots \\
\mathbf{DTC}^{k-1} & \subseteq & \mathbf{TC}^{k-1} \\
\rotatebox{-90}{$\not\subseteq$} & \rotatebox{45}{$\not\subseteq$} & \rotatebox{-90}{$\not\subseteq$} \\
\mathbf{DTC}^{k} & \subseteq & \mathbf{TC}^{k} \\
\vdots & & \vdots \\
\rotatebox{-90}{$\subseteq$} & & \\
\mathbf{DTC} & & \\
\rotatebox{-90}{$\subseteq$} & & \\
\mathbf{IFP}^{2} & &
\end{array}
$$

(Theorem 11.1 and Theorem 11.10)

References

[AC89] F. Afrati and S. Cosmadakis. Expressiveness of restricted recursive queries. In *Proceedings of 21st ACM symposium on Theory of Computing*, pages 113–126, 1989.

[AG89] M. Ajtai and Y. Gurevich. Datalog vs. first–order logic. In *30th Annual Symposium on Foundations of Computer Science*, pages 142–147. IEEE Computer Society, 1989.

[Ajt83] M. Ajtai. Σ_1^1-formulae on finite structures. *Annals of pure and applied logic*, 24:1–48, 1983.

[AS92] N. Alon and J. Spencer. *The Probabilistic Method.* Wiley, 1992.

[AV91a] S. Abiteboul and V. Vianu. Datalog extensions for database queries and updates. *Journal of Computer and System Sciences*, 43:62–124, 1991.

[AV91b] S. Abiteboul and V. Vianu. Generic computation and its complexity. In *Proceedings of 23rd ACM symposium on theory of computing*, pages 209–219, 1991.

[BG87] A. Blass and Y. Gurevich. Existential fixed-point logic. In E. Börger, editor, *Computation Theory and Logic*, volume 270 of *Lecture Notes in Computer Science*, pages 20–36. Springer-Verlag, 1987.

[Bos92] U. Bosse. An "Ehrenfeucht-Fraïssé Game" for fixpoint logic and stratified fixpoint logic. In *Proceedings of the 6th workshop on computer science logic*, volume 702 of *Lecture Notes in Computer Science*, pages 100–114. Springer-Verlag, 1992.

[CH82] A. Chandra and D. Harel. Structure and complexity of relational queries. *Journal of Computer and System Sciences*, 25:99–128, 1982.

[CM92] A. Calo and J.A. Makowsky. The Ehrenfeucht-Fraïssé game for transitive closure. In A. Nerode and M. Taitsin, editors, *Logical Foundation of Computer Science - Tver'92*, volume 620 of *Lecture Notes in Computer Science*, pages 57–68. Springer-Verlag, 1992.

[Com88] K. Compton. 0–1 laws in logic and combinatorics. In I. Reidel, editor, *Proc. NATO Advanced Study Instute on Algorithms and Order*, pages 353–383. Reidel, Dordrecht, 1988.

[Dah87] E. Dahlhaus. Skolem normal forms concerning the least fixed point. In E. Börger, editor, *Computation Theory and Logic*, volume 270 of *Lecture Notes in Computer Science*, pages 20–36. Springer-Verlag, 1987.

[DM89] P. Dublish and S. Maheshwari. Expressibility of bounded arity fixed-point queries. In *Proceedings of 8th ACM symposium on principles of database systems*, pages 324–335, 1989.

[Ebb85] H.-D. Ebbinghaus. Extended logics: The general framework. In J. Barwise and S. Feferman, editors, *Model–Theoretic Logics*, pages 25–76. Springer-Verlag, 1985.

[EF] H.-D. Ebbinghaus and J. Flum. Finite model theory. to appear.

[EFT94] H.-D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic.* Springer-Verlag, 2nd edition, 1994.

[Erd59] P. Erdös. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959.

[Fag76] R. Fagin. Probabilities on finite models. *Journal of Symbolic Logic*, 41:50–58, 1976.

[GKLT69] Y.V. Glebskiĭ, D.I. Kogan, M.I. Liogon'kiĭ, and V.A. Talanov. Range and degree of realizability of formulas in the restricted predicate calculus. *Kibernetika*, 2:17–28, 1969. Englisch translation, Cybernetics 5:142-154,1969.

[GM92a] E. Grädel and G. McColm. Deterministic versus nondeterministic transitive closure logic. In *Proc. 7th IEEE Symp. on Logic in Computer Science*, pages 58–63, 1992.

[GM92b] E. Grädel and G. McColm. Hierarchies in transitive closure logic, stratified datalog and infinitary logic. In *Proc. 33rd IEEE Symp. on Foundations of Computer Science*, pages 167–176, 1992.

[Grä91] E. Grädel. On transitive closure logic. In *Proceedings of the 5th workshop on computer science logic*, volume 626 of *Lecture Notes in Computer Science*, pages 149–163. Springer-Verlag, 1991.

[Gro92] M. Grohe. Fixpunktlogiken in der endlichen Modelltheorie, 1992. Diplomarbeit.

[Gro93] M. Grohe. Bounded arity hierarchies in fixed–point logics. In *Proceedings of the 7th workshop on computer science logic*, volume 832 of *Lecture Notes in Computer Science*, pages 150–164. Springer-Verlag, 1993.

[GS86] Y. Gurevich and S. Shelah. Fixed point extensions of first–order logic. *Annals of pure and applied logic*, 32:265–280, 1986.

[Gur84] Y. Gurevich. Toward logic tailored for computational complexity. In M.M. Richter et al., editor, *Computation and Proof Theory*, volume 1104 of *Lecture Notes in Mathematics*, pages 175–216. Springer-Verlag, 1984.

[Hel89] L. Hella. Definability hierarchies of generalized quantifiers. *Annals of pure and applied logic*, 43:235–271, 1989.

[Hel92] L. Hella. Logical hierarchies in PTIME. In *Proceedings of the 7th IEEE Symposium in Logic and Computer science*, 1992.

[Hel94] L. Hella. Fixpoint logic vs generalized quantifiers, 1994. manuscript.

[Her93] B. Herwig. Extending partial isomorphisms, 1993. to appear.

[Hru92] E. Hrushovski. Extending partial isomorphisms of graphs. *Combinatorica*, 12:411–416, 1992.

[Imh94] H. Imhof. The arity problem for fixed point logic – a complexity theoretical point of view, 1994.

[Imm81] N. Immerman. Number of quantifiers is better than number of tape cells. *Journal of Computer and System Sciences*, 22:384–406, 1981.

[Imm86] N. Immerman. Relational queries computable in polynomial time. *Information and Control*, 68:86–104, 1986.

[Imm87] N. Immerman. Languages that capture complexity classes. *SIAM Journal of Computing*, 16:760–778, 1987.

[Kol90] Ph. G. Kolaitis. Implicit definability on finite structures and unambiguous computations. In *Proceedings of the 5th IEEE Symposium on Logic in Computer science*, 1990.

[Kol91] Ph.G. Kolaitis. The expressive power of stratified logic programs. *Information and Computation*, 90:50–66, 1991.

[LG77] L. Lovász and P. Gács. Some remarks on generalized spectra. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 23:27–144, 1977.

[Mos74] Y.N. Moschovakis. *Elementary Induction on Abstract Structures*. North Holland, 1974.

[MP93] J.A. Makowsky and Y.B. Pnueli. Oracles and quantifiers. In *Proceedings of the 7th workshop on computer science logic*, volume 832 of *Lecture Notes in Computer Science*, pages 150–164. Springer-Verlag, 1993.

[Obe82] W. Oberschelp. Asymptotic laws in combinatorics. In *Combinatorial Theory, Proc. Schloss Rauischholzhausen*, volume 969 of *Lecture Notes in Mathematics*, pages 276–292. Springer-Verlag, 1982.

[Ste92] I. Stewart. Using the hamiltonian path operator to capture NP. *Journal of Computer and System Sciences*, 45:127–151, 1992.

[Var82] M. Vardi. The complexity of relational query languages. In *Proceedings of 14th ACM symposium on the theory of computing*, pages 137–146, 1982.