The k-variable property is stronger than H-dimension k

Ian Hodkinson* András Simon†

Abstract

We study the notion of H-dimension and the formally stronger k-variable property, as considered by Gabbay and Immerman & Kozen. We exhibit a class of flows of time that has H-dimension 3, and admits a finite expressively complete set of one-dimensional temporal connectives, but does not have the k-variable property for any finite k.

1 Introduction

This paper is concerned with the relationship between the k-variable property and finite H- or Henkin-dimension. These are notions of 'bounded resources' in the monadic first-order logic over a class of structures. (Monadic first-order logic is the expansion of first-order logic by infinitely many unary predicate variables. These are interpreted as arbitrary subsets of the domains of structures in the class.) A class $\mathcal{K}$ of structures has H-dimension (at most) k if, up to equivalence over $\mathcal{K}$, no monadic sentence requires more than k variables; $\mathcal{K}$ has the k-variable property if every monadic formula with at most k free variables is equivalent over $\mathcal{K}$ to one with only k variables altogether. Probably the most important example is the class of all linear orders, which has H-dimension 3, and the 3-variable property. For instance, the monadic sentence

$$\exists x_1 x_2 x_3 x_4 \Big((\bigwedge_{i<j\leq 4} x_i \neq x_j) \wedge \bigwedge_{i \leq 4} P(x_i) \Big),$$

stating that at least four elements satisfy the monadic predicate variable P, is equivalent over linear order to the sentence

$$\exists x(P(x) \wedge \exists y(y > x \wedge P(y) \wedge \exists x(x > y \wedge P(x) \wedge \exists y(y > x \wedge P(y))))),$$

which uses only two variables. More complicated statements will require three variables, but, over linear order, three are always enough.

*Supported by SERC Advanced Fellowship B/ITF/266.

†Supported by National Scientific Grant nos. 1911 and 2258.

Let us review some background. In temporal logic, the notion of finite H-dimension for a class $\mathcal{K}$ of *flows of time* (i.e., irreflexive posets) attracted attention because of the result of Gabbay [G], showing that it is equivalent to the existence of a finite expressively complete set of (possibly many-dimensional) temporal connectives for $\mathcal{K}$. Roughly, 'expressive completeness' means that the temporal logic generated by these connectives is as expressive, over $\mathcal{K}$, as the full monadic first-order logic. Thus, the connectives provide a finite basis for *all* first-order definable temporal connectives over $\mathcal{K}$, just as $\wedge$ and $\neg$ form a basis for the propositional connectives. The theme of expressive completeness had been prominent in temporal logic since Kamp's doctoral thesis, [K], which showed that the connectives Until and Since were expressively complete over the class of Dedekind complete linear flows of time. Gabbay's theorem now gave a handy test for the availability of a finite set of temporal connectives able to express all of monadic first-order logic over any given class. If the class has finite H-dimension, the theorem explicitly provides a finite set of connectives that will be expressively complete. If it doesn't, there can be no such set. This allowed, for the first time, a proof that some classes of flows of time admit no finite expressively complete set of connectives. The class of *all* flows of time is an example: see [G].

In the light of Gabbay's theorem, it was natural to examine H-dimension itself. In [IK], Immerman and Kozen studied the k-variable property in this connection (but, as we will discuss in §2.1 below, they used the term in two different senses). They gave an elegant model-theoretic characterisation of the k-variable property, using 'k-pebble' Ehrenfeucht–Fraïssé games, and used it to unify and simplify previous results on linear flows of time, and extend them to various trees.

Gabbay's theorem had not quite exhausted the topic of expressive completeness, however, since it only showed the equivalence of finite H-dimension and the existence of a finite expressively complete set of *many-dimensional* temporal connectives (both for a class of flows of time). Kamp's connectives Until and Since were *one-dimensional,* in that they are evaluated at a single point of time. The connectives constructed by Gabbay for a class with H-dimension d were $(d+1)$-dimensional: they are evaluated at $d+1$ points of time, the extra d points being termed 'reference points'. One might hope that the distinction was not important, and, indeed, Gabbay conjectured that finite H-dimension would allow the construction of expressively complete one-dimensional connectives. Unfortunately, the conjecture is false: in [H] there was given a class $\mathcal{C}$ of flows of time, with H-dimension 3, that admits no finite expressively complete set of one-dimensional connectives.

One might be tempted to think that the one/many-dimensional distinction for temporal connectives is mirrored by the difference between finite H-dimension and the k-variable property. Finite H-dimension talks about sentences, the k-variable property about formulas with free variables (reference points?). However, it is the k-variable property that is formally stronger than the property of having H-dimension at most k, rather than vice versa. This is proved in proposition 2.3 below. Thus, for any class $\mathcal{K}$ of flows of time, the situation is as follows.

$$\begin{array}{ccc}
\mathcal{K} \text{ has finite H-dimension} & \iff & \begin{array}{c}\mathcal{K}\text{ admits a finite expressively}\\ \text{complete set of (possibly many-}\\ \text{dimensional) temporal connectives}\end{array}\\
\Uparrow^1 & & \Uparrow^2\\
\begin{array}{c}\mathcal{K}\text{ has the}\\ k\text{-variable property}\\ \text{for some finite } k\end{array} & & \begin{array}{c}\mathcal{K}\text{ admits a finite expressively}\\ \text{complete set of one-dimensional}\\ \text{temporal connectives}\end{array}
\end{array}$$

Now [H] showed that the converse of the arrow (2) fails: finite H-dimension is not enough to guarantee a finite expressively complete set of one-dimensional connectives. Could the k-variable property be enough? The answer is 'no', because the class $\mathcal{C}$ given in [H] actually has the 3-variable property.

It is the purpose of the current paper to show that the converse of the arrow (1) also fails. We will prove:

Theorem 1.1 *There is a class of flows of time with H-dimension 3 — and admitting a finite expressively complete set of one-dimensional temporal connectives — but not having the k-variable property for any finite k.*

So no more arrows can be added to the picture above.

The significance of this negative result is firstly that, taken with [H], it proves that the conditions 'k-variable property' and 'there is a finite set of expressively complete one-dimensional temporal connectives' are independent: neither implies the other. This suggests that the two conditions, while both stronger than 'finite H-dimension', have little to do with each other otherwise. On this evidence, Gabbay's H-dimension, and not the k-variable property, is the condition appropriate for temporal logic. The theorem is also of technical interest. Our example codes k-ary relations, essentially using a pairing function. A corresponding example in the case of non-monadic first-order logic, suggested by Ziegler, was given by Flum in [F, remark 2g]; it has infinite H-dimension in our (monadic) sense. (It is interesting to compare the two constructions.) Finally, our result and the one in [F] show that one must carefully distinguish between the two conditions 'k-variable property' and 'H-dimension at most k'; they have sometimes been conflated in the literature.

We note that early authors in the field of expressibility and provability issues for finite variable fragments of first-order logic were Tarski and Henkin (see, e.g., [He]). For example, Tarski proved that the expressive power of the finite variable fragments of first-order logic remain weaker than that of first-order logic even in the presence of finitely many arbitrary new logical connectives. Nevertheless, it follows from the results in Tarski–Givant [TG] that set theory can be built up in the 4-variable fragment of first-order logic; the latter result was improved to 3 in Németi [N].

For a recent example of the use of finite H-dimension and the k-variable property in establishing expressive completeness, see Venema [V, chapter 3 ¶4]. The paper [H] gives fuller information on finite H-dimension, the k-variable property, and how they tie up with expressive completeness, and the book [GHR] has more detail on expressive completeness results such as Kamp's. Finite variable fragments have also come to the fore in *finite model theory.* The reader may refer to [DLW, I, KV], among many recent examples, and might try the introductory survey article [H2].

Let us summarise the layout of our paper. In section 2 we give formal definitions of H-dimension and of the k-variable property, and compare them formally with the notions in the literature. In section 3 we list some simple consequences of the definitions. In section 4 we give a class of structures with H-dimension 3 but not having the k-variable property for any finite k, and in section 5 we show that it admits a finite expressively complete set of one-dimensional connectives, Finally, in section 6, we show how to transform our example into a class of flows of time with the same properties.

2 The basic notions

Definition 2.1 *If L is a first-order relational signature, we define the monadic (first-order) language over L as follows:*

- *There are available* monadic predicates, $P_1, P_2, \ldots,$ *not in L.*
- *the atomic formulas are $x = y$, $R(x_1, \ldots, x_k)$ and $P_n(x)$ for each k-ary relation symbol $R \in L$, any n, and any variables $x, y, x_1, x_2, \ldots$*
- *If φ, ψ are formulas and x a variable, then $\neg\varphi$, $\varphi \wedge \psi$, and $\exists x\varphi$ are formulas.*

A non-monadic *(first-order) formula is a monadic formula that does not involve any monadic predicates.*

We will always assume the signature to be relational (i.e., with no function or constant symbols). Notice that '$=$' is always part of the language, and that we do not restrict the variables that are allowed to appear in a particular place in relation symbols: so, for example, $R(x, x, \ldots, x)$ is a formula.

Unless otherwise stated, when we write $\varphi(x_1, \ldots, x_m, P_1, \ldots, P_n)$ it will be implied that the free variables of the monadic formula φ lie amongst $x_1, \ldots,$ x_m but need not include them all, and similarly for the monadic predicates P_1, $\ldots, P_n$. (This is the usual convention.) However, for brevity, we sometimes omit $x_1, \ldots, x_m$ and/or $P_1, \ldots, P_n$ altogether. We sometimes use $\bar{x}$, $\bar{P}$ to denote tuples of variables or predicates.

Definition 2.2 *Let $\mathcal{K}$ be a class of structures of the same signature, and let $k, m < \omega$.*

1. *Let $n < \omega$ and let $\varphi(\bar{x}, P_1, \ldots, P_n)$, $\psi(\bar{x}, P_1, \ldots, P_n)$ be monadic formulas of the signature of $\mathcal{K}$. We say that φ and ψ are* $\mathcal{K}$-equivalent *if for all $M \in \mathcal{K}$ and all subsets $S_1, \ldots, S_n$ of the domain $dom(M)$ of M, we have*
$$M \models \forall \bar{x}(\varphi(\bar{x}, S_1, \ldots, S_n) \leftrightarrow \psi(\bar{x}, S_1, \ldots, S_n)).$$
Note that this definition does not depend on the choice of $\bar{x}$ (containing the free variables of φ and ψ).

2. *We say that $\mathcal{K}$ has property $BV(m, k)$ if for all $n < \omega$, every monadic formula $\varphi(x_1, \ldots, x_m, P_1, \ldots, P_n)$ is $\mathcal{K}$-equivalent to a monadic formula $\varphi^*(x_1, \ldots, x_m, P_1, \ldots, P_n)$ that uses at most k bound variables. (Here and in 3. below, the actual free variables of φ^* need not be the same as those of φ; cf. our convention with regard to free variables.)*

3. *We say that $\mathcal{K}$ has property $V(m, k)$ if for all $n < \omega$, every monadic formula $\varphi(x_1, \ldots, x_m, P_1, \ldots, P_n)$ is $\mathcal{K}$-equivalent to a monadic formula $\varphi^*(x_1, \ldots, x_m, P_1, \ldots, P_n)$ that uses at most $m + k$ variables.*

4. *$\mathcal{K}$ is said to have* H-dimension at most k *if it has property $BV(0, k)$. That is, any monadic sentence is $\mathcal{K}$-equivalent to a monadic sentence with at most k variables. $\mathcal{K}$ is said to have* finite H-dimension *if such a k exists, and* H-dimension k *if k is least such that $\mathcal{K}$ has H-dimension at most k.*

5. *$\mathcal{K}$ is said to have the* k-variable property *if it has property $V(k, 0)$. That is, any monadic formula with at most k free variables is $\mathcal{K}$-equivalent to one written with at most k variables altogether.*

6. *In each case, we say that a structure M has the cited property if the class $\{M\}$ does.*

Proposition 2.3 *Let $\mathcal{K}$ be a class of structures and let $k < \omega$. If $\mathcal{K}$ has the k-variable property then it has H-dimension at most k.*

Proof. If $\mathcal{K}$ has the k-variable property, any monadic sentence σ is $\mathcal{K}$-equivalent to a monadic formula $\varphi(x_1, \ldots, x_k)$ written with variables $x_1, \ldots, x_k$ only. Then $\forall x_1 \ldots \forall x_k \varphi$ is a sentence that is $\mathcal{K}$-equivalent to σ and whose variables are at most $x_1, \ldots, x_k$. So $\mathcal{K}$ has H-dimension at most k. ■

Examples 2.4 The class of all linear orders has the k-variable property if and only if $k \geq 3$ [IK, theorem 4.1], as does the class of all 'circles' [H].

In [IK] various classes of trees with bounded branching were considered, and the least k was determined for which they have the k-variable property.

The class of all sets does not have finite H-dimension, and nor does the class of all flows of time [G].

2.1 Comparison with standard definitions

According to Gabbay [G, definition 1.3], a flow of time $M = (T, <, =)$ is said to have H-dimension k if k is the smallest number such that every monadic formula $\varphi(x_1, \ldots, x_m)$ written with $x_1, \ldots, x_m$ as free variable letters and any number of bound variable letters can be equivalently rewritten over M using at most k bound variable letters and the same free $x_1, \ldots, x_m$.

For a flow of time M, Gabbay's definition and the one in definition 2.2 above are equivalent. Gabbay's definition clearly implies ours. Conversely, assume that M has property $BV(0, k)$. It will be shown in proposition 3.1(1) below that this implies that M has property $BV(m, k)$ for all m. Let $\varphi(x_1, \ldots, x_m)$ be a monadic formula with free variable letters exactly $x_1, \ldots, x_m$. By $BV(m, k)$, there is an M-equivalent monadic formula $\psi(x_1, \ldots, x_m)$ with at most k bound variables and with free variables taken from $x_1, \ldots, x_m$. Then $\psi \wedge \bigwedge_{i \leq m}(x_i = x_i)$ is also M-equivalent to φ, is written with at most k bound variables, and has exactly the same free variables as φ. So M has H-dimension at most k in Gabbay's sense.

Immerman and Kozen's definition in [IK] of the k-variable property agrees with Definition 2.2(5), except in two respects. First, their definition is in terms of a first-order theory Σ in some signature L. Second, and more important, they do not mention 'monadic formulas', monadic predicates being regarded as part of the signature. Thus, in definition 2.2 of [IK], 'a first-order theory Σ is said to satisfy the k-variable property if for all formulas $\varphi \in L$ [L is a first-order language] with free variables among $\{x_1, \ldots, x_k\}$, there exists a $\psi \in L_k$ [where L_k is the sublanguage of L consisting of all formulas containing only variables $\{x_1, \ldots, x_k\}$] such that $\Sigma \models \varphi \leftrightarrow \psi$.' This notion, which we may call the 'non-monadic k-variable property', is used for the technical results of [IK]. However, in §1 of [IK], and also in the abstract, the k-variable property is defined informally and differently as follows: 'A first-order theory Σ satisfies the k-variable property if every first-order formula is equivalent under Σ to a formula with at most k bound variables (possibly reused).' This is clearly a non-monadic version of Gabbay's 'H-dimension at most k'.

The 'non-monadic' approach is continued in Flum's [F], where it is shown that the two definitions in [IK] are not equivalent. Flum gives an example, suggested by Ziegler, of a theory Σ without the non-monadic k-variable property for any k, but such that any formula φ is equivalent over Σ to a formula with at most one bound variable. As we said in the introduction, this is a 'non-monadic' counterpart to our theorem 1.1. The example does not replace ours, however, as the class of models of Σ has infinite H-dimension in our monadic sense.

Flum also notes that, in the non-monadic context, the notion of k-variable property is not 'stable': for example, the theory of discrete linear orderings without endpoints has the non-monadic 1-variable property (as defined in the preceding paragraph), but not the non-monadic 2-variable property. (The 'monadic' k-variable property is stable, by proposition 3.1(4) below.) Flum then defines the notion of a *k-bounded* theory. Roughly, a theory Σ is k-bounded if every formula φ is equivalent over Σ to a formula ψ such that no more than k distinct variables are quantified in any branch of the formation tree of ψ. Any k-bounded theory is $(k+1)$-bounded, and k-boundedness implies the non-monadic k-variable property. Using the techniques of section 3.2 below, it can be shown that if L is relational and all its relations have arity less than k, then Σ is k-bounded iff Σ has the non-monadic n-variable property for all $n \geq k$, iff every L-formula is equivalent over Σ to a boolean combination of L-formulas written with only k variables.

3 Properties of V and BV.

Let us derive some simple facts about the properties V, BV.

3.1 Elementary facts

Proposition 3.1 *Let $k, m < \omega$ be arbitrary, and let $\mathcal{K}$ be a class of similar structures. The following statements refer implicitly to $\mathcal{K}$.*

1. *BV is monotonic in each argument: $BV(m,k) \Rightarrow BV(m+1,k) \wedge BV(m,k+1)$.*
2. *$BV(m,k) \Rightarrow BV(1,k)$ if $m \geq 1$, and $BV(1,k) \Rightarrow BV(0,k+1)$.*
3. *V is also monotonic in each argument.*
4. *$V(m+1,k) \Rightarrow V(m,k+1)$*
5. *$BV(m,k) \Rightarrow V(m,k) \Rightarrow BV(m,k+m)$.*
6. *$V(0,k) \iff BV(0,k)$.*

Proof.

1. Clearly, $BV(m,k)$ is monotonic in k. Assume $BV(m,k)$ holds, and let $\varphi(x_1,\ldots,x_{m+1},P_1,\ldots,P_n)$ be a monadic formula. Introduce a new monadic predicate $Q(x)$ and consider the formula

$$\psi(x_1,\ldots,x_m,Q,P_1,\ldots,P_n) \stackrel{\text{def}}{=} \exists x_{m+1}(Q(x_{m+1}) \wedge \varphi).$$

By $BV(m,k)$, there is a monadic formula $\psi^*(x_1,\ldots,x_m,Q,P_1,\ldots,P_n)$, $\mathcal{K}$-equivalent to ψ and using at most k bound variables. We can assume that these bound variables do not include x_{m+1}. Let $\varphi^*(x_1,\ldots,x_{m+1},P_1,\ldots,P_n)$ be the result of replacing each atomic subformula of

ψ^* of the form $Q(t)$ (for some variable t) by $t = x_{m+1}$. This substitution causes no clashes as x_{m+1} does not occur in ψ^*. Clearly, φ^* uses at most k bound variables. We show that φ^* is $\mathcal{K}$-equivalent to φ. For all $M \in \mathcal{K}$, all $a_1, \ldots, a_{m+1} \in dom(M)$ and all $S_1, \ldots, S_n \subseteq dom(M)$, we have:

$$\begin{aligned} & M \models \varphi(a_1, \ldots, a_{m+1}, S_1, \ldots, S_n) \\ \iff & M \models \psi(a_1, \ldots, a_m, \{a_{m+1}\}, S_1, \ldots, S_n) \\ \iff & M \models \psi^*(a_1, \ldots, a_m, \{a_{m+1}\}, S_1, \ldots, S_n) \\ \iff & M \models \varphi^*(a_1, \ldots, a_{m+1}, S_1, \ldots, S_n). \end{aligned}$$

Hence φ^* is $\mathcal{K}$-equivalent to φ. So $BV(m+1, k)$ holds.

2. Assume that $BV(m, k)$ holds for some $m \geq 1$, and take a monadic formula $\varphi(x_1)$ with at most one free variable. Then φ is $\mathcal{K}$-equivalent to some formula $\psi(x_1, \ldots, x_m)$ with at most k bound variables. We can assume that these bound variables are different from $x_1, \ldots, x_m$. Thus we can substitute x_1 for the variables $x_2, \ldots, x_m$ in ψ to gain a formula with free variables at most x_1 and at most k bound variables that is $\mathcal{K}$-equivalent to φ. Thus $BV(1, k)$ holds.

   Now let σ be a monadic sentence. If $BV(1, k)$ holds, then there is a monadic formula $\varphi(x)$, with at most x free and at most k bound variables, that is $\mathcal{K}$-equivalent to σ. Thus the sentence $\forall x \varphi$ is $\mathcal{K}$-equivalent to σ and uses at most $k+1$ bound variables. Hence $BV(0, k+1)$ holds.

3. This is similar to (1).

4. If $V(m+1, k)$ holds then any monadic formula $\varphi(x_1, \ldots, x_m)$ is $\mathcal{K}$-equivalent to a formula $\varphi^*(x_1, \ldots, x_{m+1})$ using only $m+1+k$ variables. Clearly, $\forall x_{m+1} \varphi^*$ is $\mathcal{K}$-equivalent to φ and involves free variables at most $x_1, \ldots, x_m$ and at most $k+1$ more variables altogether. Hence $\mathcal{K}$ has $V(m, k+1)$.

5. If $BV(m, k)$ holds, then over $\mathcal{K}$, any monadic formula $\varphi(x_1, \ldots, x_m)$ can be equivalently rewritten with free variables from $x_1, \ldots, x_m$ and at most k bound variables. Clearly, the resulting formula will use at most $m + k$ variables altogether. Hence $V(m, k)$ holds. Similarly, if any $\varphi(x_1, \ldots, x_m)$ can be rewritten using at most $m + k$ variables then obviously it can be rewritten using at most $m + k$ bound variables, so $V(m, k) \Rightarrow BV(m, k + m)$.

6. This is a special case of (5).

■

That proposition 3.1(2) cannot be strengthened is shown by the following observation.

Proposition 3.2 *There is a structure M with property $BV(1,2)$ but not $BV(0,2)$.*

Proof. Let M be the seven-element structure shown in figure 3.1 below. The signature consists of '$<$' only. For $a, b \in dom(M)$, we let $M \models a < b$ if and only if the shortest path from a to b is anticlockwise (by following the arrows). The numbers 0–6 are labels, and are not part of the structure.

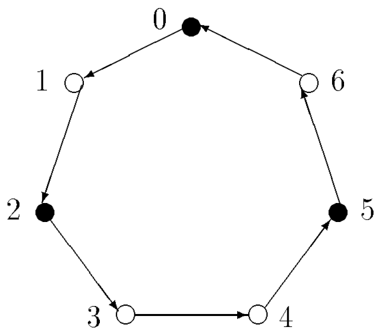

Figure 3.1

It is not hard to see that, over M, any monadic formula $\varphi(x)$ can be equivalently rewritten with only two bound variables. For let $M^* = (M, S_1, \ldots, S_n)$ be any monadic expansion of M (i.e., let $S_1, \ldots, S_n \subseteq M$). We have $M^* \models P_j(a) \iff a \in S_j$, for all j and all $a \in dom(M)$. We can write a formula $\chi^{M^*}(x)$ of the form
$\tau_0(x) \wedge \exists y(y > x \wedge \tau_1(y) \wedge \exists z(z > y \wedge z > x \wedge \tau_2(z)$
$\wedge \exists y(y > x \wedge y > z \wedge \tau_3(y) \wedge \exists z(z < x \wedge z > y \wedge \tau_4(z)$
$\wedge \exists y(y < x \wedge y > z \wedge \tau_5(y) \wedge \exists z(z < x \wedge z > y \wedge \tau_6(z))) \cdots)$
where the formulas $\tau_i(x)$ are the conjunction of those monadic predicates P_j that hold at the i^{th} point of M and the negations of those P_j that do not. It is clear that χ^{M^*} describes M^* up to isomorphism (modulo M): i.e., for all $a \in dom(M)$ and $S'_1, \ldots, S'_n \subseteq dom(M)$, we have $M \models \chi^{M^*}(a, S'_1, \ldots, S'_n)$ if and only if there is an isomorphism $\theta : M^* \to (M, S'_1, \ldots, S'_n)$ with $\theta(0) = a$. Hence any monadic formula $\varphi(x, P_1, \ldots, P_n)$ will be M-equivalent to the finite disjunction

$$\bigvee \{\chi^{(M,\bar{S})}(x) : \bar{S} = S_1, \ldots, S_n \subseteq dom(M), (M, \bar{S}) \models \varphi(0)\},$$

which has only two bound variables. So $BV(1,2)$ holds.

However, the sentence $\exists_{=3} x P(x)$, saying that exactly three elements satisfy P, is not expressible by a *sentence* using only two variables. This can be seen as follows. For $S \subseteq dom(M)$, let us say that (M, S) is *smooth* if for all $x \in dom(M)$ there are $y, z > x$ with $y \in S$ and $z \notin S$, and similarly for '$<$'. It is easily seen (cf. [H] or [IK, theorem 4.1]) that if M^1, M^2 are any two smooth expansions of M then $M^1 \models \sigma \iff M^2 \models \sigma$ for any monadic sentence $\sigma(P)$ using only two variables. Let S be the set of black points (0, 2, 5) in Figure 3.1. If $\exists_{=3}xP(x)$ were expressible by such a sentence $\sigma(P)$, then σ would have different values in (M, S) and $(M, dom(M) \setminus S)$, which is impossible as they are both smooth. So $BV(0, 2)$ fails.

We saw in examples 2.4 that for all m, the class of all linear orders has $BV(m, k)$ if $k \geq 3$. (And in fact 'only if', because of the following. Using the rationals and two copies of the integers with a single new point in between, an argument similar to that of [IK, theorem 4.1] shows that the class of all linear orders does not have the property $BV(1, 2)$. Now the 'only if' direction follows from proposition 3.1.) Hence the phenomenon found in M above — that there is a k such that M has $BV(1, k)$ but not $BV(0, k)$ — does not occur for all classes. ∎

3.2 The strong k-variable property

Here we show that, under restrictions on k, a superficially stronger condition than the k-variable property is in fact equivalent to it. The following preliminary lemma is well-known.

Lemma 3.3 *Let $k \geq 3$ be greater than the arity of any relation in L. Then every monadic formula φ that uses only k variables is a boolean combination of formulas ψ each with at most $k - 1$ free variables.*

Proof. By induction on φ. If φ is atomic, it is $x = y$, $P(x)$ (P a monadic predicate) or $R(x_1, \ldots, x_n)$ for some $n < k$ — so the result is clear in this case. The boolean connectives clearly preserve the given property. Finally, let $\varphi(x_1, \ldots, x_k)$ be a monadic formula written with variables among $x_1, \ldots, x_k$. Then if $1 \leq i \leq k$, $\exists x_i \varphi$ clearly has at most $k - 1$ free variables. ∎

Note that this lemma fails if L has relation symbols of arity k, or function symbols of arity 2 or more. In the case $k = 4$, for example, $R(x_1, \ldots, x_4)$, and, if f is a 2-ary function symbol, the formula $f(f(x_1, x_2), x_3) = x_4$, are not boolean combinations of formulas with at most 3 free variables.

Corollary 3.4 *Let $\mathcal{K}$ be a class of structures and let $k \geq 3$ be greater than the maximum arity of relation symbols in the signature of $\mathcal{K}$. Then $\mathcal{K}$ has the k-variable property if and only if for all n, any monadic formula $\varphi(x_1, \ldots, x_n)$ is $\mathcal{K}$-equivalent to a boolean combination of formulas ψ, each ψ being written with at most k variables. (This latter condition could be called the strong k-variable property.)*

Proof. We only need to prove left-to-right, which we do by induction on n. If $n \leq k$ there is nothing to prove. Assume the result for $n-1 \geq k$ and let the monadic formula $\varphi(x_1, \ldots, x_n)$ be given. By proposition 3.1(3), $\mathcal{K}$ has the n-variable property, so we can assume that φ is written with only n variables.

By lemma 3.3, φ is a boolean combination of monadic formulas ψ with at most $n-1$ free variables. Hence it is enough to deal with each ψ. But we can do this immediately, using the inductive hypothesis. This completes the proof. Observe that by lemma 3.3, each formula of the boolean combination for φ has at most $k-1$ free variables. ■

3.3 V and BV as functions

The main aim of this paper is to refute the converse of proposition 2.3. We can frame our work more broadly, as follows. Fix a class $\mathcal{K}$ of structures and define a partial function $V : \omega \to \omega$ by: $V(m)$ is the least k such that $\mathcal{K}$ has property $V(m,k)$, if such a k exists; otherwise, $V(m)$ is undefined. Define a partial function BV similarly. By proposition 3.1(1–5), either V and BV are both total functions, or they are both undefined everywhere. The following are also immediate consequences of proposition 3.1.

Corollary 3.5 *For any class $\mathcal{K}$ of structures such that V and BV are total functions:*

1. *$\mathcal{K}$ has property $BV(m,k)$ if and only if $k \geq BV(m)$. BV is a non-increasing function.*
2. *$BV(m) = BV(1)$ for all $m \geq 1$, and $BV(0) \leq BV(1)+1$.*
3. *The analogue of (1) for V holds.*
4. *$V(m+1) \geq V(m) - 1$ for all m.*
5. *$BV(m) \geq V(m) \geq BV(m) - m$ for all m.*
6. *$V(0) = BV(0)$.*

There is a structure M with $BV(0) = BV(1)+1$, and a class of structures for which BV is constant (proposition 3.2). $\mathcal{K}$ has the k-variable property if and only if $V(k) = 0$, and its H-dimension is $BV(0)$ (if defined).

Figure 3.2 shows (schematically) some possibilities for the graphs of these functions. The example in the next section will show that a graph like the third, in which $V(0) = 3$, $V(1) = 2$ and $V(m) = 1$ for all $m \geq 2$, so that V is defined but is never zero, is a possibility. It remains open to classify all possibilities for the function V. For example, does every non-increasing function $f : \omega \to \omega$ satisfying $f(m+1) \geq f(m) - 1$ for all m occur as V for some class $\mathcal{K}$?

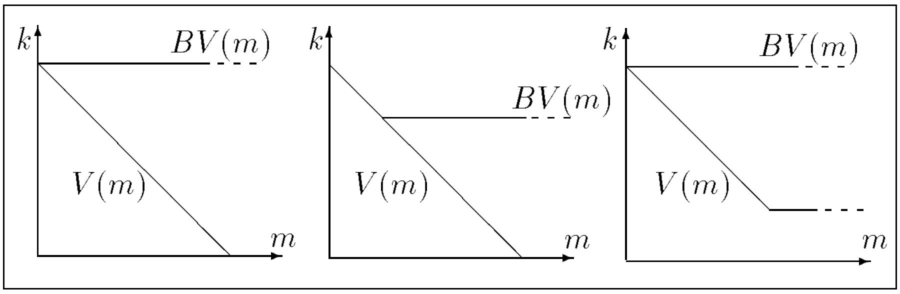

Figure 3.2

4 The example

We will exhibit a class $\mathcal{K}$ of structures such that:

- $\mathcal{K}$ has H-dimension 3;
- $\mathcal{K}$ does not have the k-variable property for any $k < \omega$.

We will use a *pairing function,* expressed as two binary relations.

Definition 4.1 *The signature of $\mathcal{K}$ will consist of two binary relation symbols, l and r. The axioms will be:*

- $\forall x \forall y \exists! z(l(x,z) \wedge r(y,z))$
- $\forall z(\exists x\, l(x,z) \leftrightarrow \exists y\, r(y,z))$
- $\forall xzt(l(x,z) \wedge l(t,z) \rightarrow x = t)$
- $\forall yzt(r(y,z) \wedge r(t,z) \rightarrow y = t)$
- $\forall xy(l(x,y) \vee r(x,y) \rightarrow x \neq y)$

Here, $\exists! z$ means that there exists a unique z. $\mathcal{K}$ will be the class of all models of these axioms.

Example 4.2 If D is any non-empty set, let $D_0 = D$ and $D_{n+1} = D_n \cup (D_n \times D_n)$. Then let $D_\omega = \bigcup_{n<\omega} D_n$. Interpreting l on D_ω as the binary relation $\{(x,(x,y)) : x, y \in D_\omega\}$, and r as $\{(y,(x,y)) : x, y \in D_\omega\}$, so that $l(x,z)$ holds iff $z = (x,y)$ for some $y \in D_\omega$, and similarly for r, we obtain a structure $M[D] \in \mathcal{K}$.

Definition 4.3 *We define formulas $\pi_m(x,y)$ for $m \geq 1$ by induction:*

- $\pi_1(x,y) = l(x,y)$*;*
- $\pi_{m+1}(x,y) = \exists z[r(z,y) \wedge \exists y(y = z \wedge \pi_m(x,y))]$.

Let $M \in \mathcal{K}$, $m \geq 1$ and $a \in dom(M)$. We write $a\lceil m$ for the unique m-tuple $(a_1, \ldots, a_m) \in dom(M)^m$ such that $M \models \pi_i(a_i, a)$ for all $i \leq m$, if such a tuple exists; otherwise, $a\lceil m$ is undefined.

Notice that the formulas π_m use only the variables x, y, z. If $a \in dom(M)$, $a\lceil m$ is defined if and only if $M \models \exists z \pi_m(z, a)$.

Lemma 4.4 *Let $\varphi(v_1, \ldots, v_k, P_1, \ldots, P_n)$ be a monadic formula of the language of $\mathcal{K}$, written only with variables from $v_1, \ldots, v_k$. Then there is a monadic formula $\varphi^*(y, P_1, \ldots, P_n)$ using only variables x, y, z, such that for all $M \in \mathcal{K}$, all $S_1, \ldots, S_n \subseteq dom(M)$ and all $a \in dom(M)$ such that $a\lceil k$ is defined,*

$$M \models \varphi(a\lceil k, S_1, \ldots, S_n) \leftrightarrow \varphi^*(a, S_1, \ldots, S_n).$$

Proof. We define φ^* by induction on φ. (We note that the definition of φ^* depends on the choice of k.) First suppose that $\varphi(v_1, \ldots, v_k)$ is atomic. If it is $R(v_i, v_j)$, where $R \in \{=, l, r\}$ and $i, j \leq k$, then we let $\varphi^*(y)$ be $\exists xz[\pi_i(x, y) \wedge \exists x(x = z \wedge \pi_j(x, y)) \wedge R(x, z)]$. Similarly, if it is a monadic formula $P(v_i)$ we let φ^* be $\exists x(\pi_i(x, y) \wedge P(x))$. We may define $(\neg\varphi)^*$ as $\neg(\varphi^*)$ and $(\varphi \wedge \psi)^*$ as $\varphi^* \wedge \psi^*$. These definitions clearly work.

Finally, if we have defined $\varphi^*(y)$ then for $i \leq k$ we let $(\exists v_i \varphi)^*(y)$ be the formula

$$\exists z \Big(\exists y(y = z \wedge \varphi^*(y) \wedge \exists x \pi_k(x, y)) \wedge \bigwedge_{\substack{j \leq k \\ j \neq i}} \forall x(\exists y(y = z \wedge \pi_j(x, y)) \leftrightarrow \pi_j(x, y))\Big).$$

To confirm that this is what we want, assume the result for φ, written only using variables $v_1, \ldots, v_k$, and let $i \leq k$. Let $M \in \mathcal{K}$, $S_1, \ldots, S_n \subseteq dom(M)$ and $a \in dom(M)$, and suppose that $a\lceil k$ is defined. Then $M \models \exists v_i \varphi(a\lceil k, S_1, \ldots, S_n)$ if and only if there are $b_1, \ldots, b_k \in dom(M)$ such that the k-tuples $(b_1, \ldots, b_k)$ and $a\lceil k$ agree except possibly in their i^{th} entries, and $M \models \varphi(b_1, \ldots, b_k, S_1, \ldots, S_n)$. Since we may always choose $b \in dom(M)$ with $b\lceil k = (b_1, \ldots, b_k)$, by the inductive hypothesis this holds if and only if $M \models \varphi^*(b, S_1, \ldots, S_n)$ for some $b \in dom(M)$ such that $b\lceil k$ is defined and agrees with $a\lceil k$ except possibly in its i^{th} entry. But this is equivalent to $M \models (\exists v_i \varphi)^*(a, S_1, \ldots, S_n)$, as required. ■

Theorem 4.5 *$\mathcal{K}$ has H-dimension at most 3.*

Proof. Let $\sigma(P_1, \ldots, P_n)$ be a monadic sentence. Let k be the number of variables that σ is written with. By the lemma, σ is $\mathcal{K}$-equivalent to the sentence $\forall y(\exists x \pi_k(x, y) \rightarrow \sigma^*(y, P_1, \ldots, P_n))$, which has at most three bound variables. (Of course, σ^* depends on k.) Hence $BV(0, 3)$ holds. ■

A similar argument shows that $\mathcal{K}$ has properties $V(1, 2)$ and $V(m, 1)$ for all $m \geq 2$.

Theorem 4.6 *For no $k < \omega$ does $\mathcal{K}$ have the k-variable property.*

Proof. Fix $k \geq 3$ and define D to be the disjoint union of k copies E_i ($i \leq k$) of $\{0, 1\}$. Let $M[D] \in \mathcal{K}$ be as in example 4.2 above. Let S be the set

$$\{d \in dom(M[D]) : d\lceil k = (d_1, \ldots, d_k), d_i \in E_i \text{ for } i \leq k, \textstyle\sum_{i\leq k} d_i \text{ is even}\}.$$

Claim: Let $n_1, \ldots, n_{k-1} \leq k$ be any distinct indices, and let a_{n_i}, b_{n_i} be arbitrary elements of E_{n_i}, for each $i < k$. Then for any monadic formula $\psi(x_1, \ldots, x_{k-1}, P)$ we have

$$M[D] \models \psi(a_{n_1}, \ldots, a_{n_{k-1}}, S) \leftrightarrow \psi(b_{n_1}, \ldots, b_{n_{k-1}}, S).$$

Proof of Claim: Let $\{1, \ldots, k\} = \{n_1, \ldots, n_k\}$ for some appropriate $n_k \leq k$. Let $a_{n_k} \in E_{n_k}$ be 0 if $\sum_{i<k} a_{n_i}$ is even, and 1 otherwise. Define $b_{n_k} \in E_{n_k}$ similarly. Define θ to be the permutation of D that, for each $i \leq k$, fixes the two elements of E_i if $a_i = b_i$, and interchanges them otherwise. Then (i) $\theta(a_i) = b_i$ for all i. Also, $|\{i \leq k : \theta(a_i) \neq a_i\}|$ is even, so (ii) if $c_i \in E_i$ for each $i \leq k$ then $\sum_{i\leq k} c_i$ and $\sum_{i\leq k} \theta(c_i)$ have the same parity (odd or even).

Now θ extends inductively to a permutation θ_n of each D_n as defined in example 4.2, as follows. If θ_n is defined, and $x, y \in D_n$, we define $\theta_{n+1}(x) = \theta_n(x)$ and $\theta_{n+1}((x, y)) = (\theta_n(x), \theta_n(y))$. Let θ^* be the union of the permutations θ_n ($n < \omega$). Then θ^* is a well-defined permutation of $dom(M[D])$. Moreover, it preserves l and r, so is an automorphism of $M[D]$. Hence if $d \in dom(M[D])$ and $d\lceil k$ is defined and is $(d_1, \ldots, d_k)$, then $\theta^*(d)\lceil k$ is also defined and is $(\theta^*(d_1), \ldots, \theta^*(d_k))$. So by (ii) we have $d \in S \iff \theta^*(d) \in S$. Hence θ^* is an automorphism of the structure $(M[D], S)$.

It now follows by an easy induction on the structure of ψ that

$$M[D] \models \psi(a_{n_1}, \ldots, a_{n_{k-1}}, S) \leftrightarrow \psi(\theta^*(a_{n_1}), \ldots, \theta^*(a_{n_{k-1}}), S).$$

Since by (i), $\theta^*(a_i) = b_i$ for all i, this proves the claim.

Now suppose for contradiction that $\mathcal{K}$ has the k-variable property. Hence there is a monadic formula $\varphi(v_1, \ldots, v_k, P)$ using only k variables that is $\mathcal{K}$-equivalent to $\exists y(P(y) \wedge \bigwedge_{i\leq k} \pi_i(v_i, y))$. Choose $a_i, b_i \in E_i$ for each $i \leq k$ so that $\sum_{i\leq k} a_i$ is even and $\sum_{i\leq k} b_i$ odd. So clearly, $M[D] \models \varphi(a_1, \ldots, a_k, S) \wedge \neg\varphi(b_1, \ldots, b_k, S)$. But as $k \geq 3$, by lemma 3.3 we see that φ is a boolean combination of formulas $\psi(v_{n_1}, \ldots, v_{n_{k-1}}, P)$ with at most $k - 1$ free variables, where the indices $n_1, \ldots, n_{k-1} \leq k$ are distinct. By the claim, $M[D] \models \psi(a_{n_1}, \ldots, a_{n_{k-1}}, S) \leftrightarrow \psi(b_{n_1}, \ldots, b_{n_{k-1}}, S)$ for all such ψ. Hence $M[D] \models \varphi(a_1, \ldots, a_k, S) \leftrightarrow \varphi(b_1, \ldots, b_k, S)$. This is a contradiction. So $\mathcal{K}$ does not have the k-variable property, and by proposition 3.1(3) it follows that $\mathcal{K}$ does not have the k-variable property for any finite k. ∎

We remark that there is a single structure $M \in \mathcal{K}$ that does not have the k-variable property for any finite k. One such structure is of the form $M[D]$, where D is a structure in a signature with a single n-ary relation R_k for each $k < \omega$. The relations are placed on D 'at random': this can be done using Fraïssé amalgamation (see [C] for details of this). We may code R_k as $S_k = \{b : b\lceil k$ exists and is in D, $D \models R_k(b\lceil k)\}$. Then we can show as in the theorem above, by taking $S = S_k$, that $M[D]$ does not have the k-variable property for any k.

In the proof of theorem 4.6 we used the 'method of big elements' coming from Algebraic Logic. This method, originating with H. Andréka (cf. [A] or [AN]), consists of constructing and using a relation, the information content of which completely vanishes if we forget any one of its coordinates.

5 Temporal connectives

We will prove that the class $\mathcal{K}$ admits a finite set of expressively complete one-dimensional temporal connectives. For more details about the notions below, see e.g., [H].

Definition 5.1 *We define a temporal language using the following set Ξ of one-dimensional temporal connectives.*

- *The boolean connectives* $\wedge, \neg$
- S_i *for* $1 \leq i \leq 3$ *(unary)*
- E_{ij} *for* $1 \leq i, j \leq 3$ *(nullary)*
- L_{ij} *for* $1 \leq i, j \leq 3$ *(nullary)*
- R_{ij} *for* $1 \leq i, j \leq 3$ *(nullary)*
- X_i *for* $1 \leq i \leq 3$ *(unary)*
- $\Diamond$ *(unary)*

The semantics of the connectives are given as follows. Let A be a formula written with these connectives, $M \in \mathcal{K}$, $b \in dom(M)$ and h be an assignment of subsets of $dom(M)$ to the atoms $p_1, p_2, \ldots$, where p_i corresponds to the monadic predicate P_i. We write $\|A\|_b^h = 1$ (respectively, $\|A\|_b^h = 0$) to mean that A is true (respectively, false) at b in M under the assignment h. We also write $(M, h) \models P(b)$ iff $b \in h(p)$. In this way, we can also interpret any monadic formula in (M, h).

- $\|S_i p\|_b^h = 1$ *iff* $(M, h) \models \exists y(\pi_i(y, b) \wedge P(y))$

- $\|E_{ij}\|_b^h = 1$ *iff* $M \models \exists y(\pi_i(y,b) \wedge \pi_j(y,b))$
- $\|L_{ij}\|_b^h = 1$ *iff* $M \models \exists y \exists z(\pi_i(y,b) \wedge \pi_j(z,b) \wedge l(y,z))$
- $\|R_{ij}\|_b^h = 1$ *iff* $M \models \exists y \exists z(\pi_i(y,b) \wedge \pi_j(z,b) \wedge r(y,z))$
- $\|X_i p\|_b^h = 1$ *iff* $(M,h) \models \exists y\Big(P(y) \wedge \exists z \pi_3(z,y) \wedge \bigwedge_{j \in \{1,2,3\} \setminus \{i\}} \forall z(\pi_j(z,b) \leftrightarrow \pi_j(z,y))\Big)$
- $\|\Diamond p\|_b^h = 1$ *iff* $(M,h) \models \exists y(P(y) \wedge \pi_1(b,y) \wedge \exists z \pi_3(z,y))$.

Definition 5.2 *If $\psi(x, P_1, \ldots, P_n)$ is a monadic formula, and $A(p_1, \ldots, p_n)$ a temporal formula, we say that ψ and A are* $\mathcal{K}$-equivalent *if for all $M \in \mathcal{K}$, all $a \in dom(M)$ and all assignments h of subsets of $dom(M)$ to the atoms $p_1, \ldots, p_n$, we have*

$$M \models \psi(a, h(P_1), \ldots, h(P_n)) \iff \|A\|_a^h = 1.$$

Lemma 5.3 *Let $\psi(x_1, P_1, \ldots, P_n)$ be a monadic formula written with variables x_1, x_2, x_3 only. Then there exists a $\mathcal{K}$-equivalent temporal formula $A_\varphi(p_1, \ldots, p_n)$, written with the connectives of Ξ.*

Proof. First let $\psi(x_1, x_2, x_3)$ be any monadic formula using only the variables x_1, x_2 and x_3. We define a temporal formula B_ψ by induction on ψ, as follows:

- $B_{P(x_i)} = S_i p$, where p is the atom corresponding to the monadic predicate P,
- $B_{x_i = x_j} = E_{ij}$,
- $B_{l(x_i, x_j)} = L_{ij}$,
- $B_{r(x_i, x_j)} = R_{ij}$,
- $B_{\neg\psi} = \neg B_\psi$,
- $B_{\psi \wedge \psi'} = B_\psi \wedge B_{\psi'}$,
- $B_{\exists x_i \psi} = X_i(B_\psi)$.

The proof of lemma 4.4 shows that for all $M \in \mathcal{K}$, all assignments h of subsets of $dom(M)$ to the monadic predicates $P_1, \ldots, P_n$ and all $b \in dom(M)$ such that $b\lceil 3$ is defined,

$$M \models \psi(b\lceil 3, h(P_1), \ldots, h(P_n)) \iff \|B_\psi\|_b^h = 1.$$

Now define A_ψ to be $\Diamond B_\psi$. Let M, h, a be as in definition 5.2. Let

$$S = \{b \in dom(M) : b\lceil 3 \text{ is defined and } M \models \pi_1(a,b)\}.$$

Note that $S \neq \emptyset$. We have $M \models \psi(a, h(P_1), \ldots, h(P_n))$
iff $M \models \psi(b\lceil 1, h(P_1), \ldots, h(P_n))$ for some (any) $b \in S$,
iff $\|B_\psi\|_b^h = 1$ for some (any) $b \in S$,
iff $\|A_\psi\|_a^h = 1$, as required. ■

Example 5.4 The formula $\psi(x) = \forall y(l(x, y) \wedge P(y) \to \exists z(r(z, y) \wedge Q(z)))$ is $\mathcal{K}$-equivalent to the temporal formula $\Diamond[\neg X_2 \neg(L_{12} \wedge S_2 p \to X_3(R_{32} \wedge S_3 q)]$.

Definition 5.5 *We say that the connectives* Ξ *are* expressively complete over $\mathcal{K}$ *if for every monadic formula* $\varphi(x, P_1, \ldots, P_n)$ *there is a* $\mathcal{K}$*-equivalent temporal formula* $A(p_1, \ldots, p_n)$.

Theorem 5.6 *The connectives in* Ξ *are expressively complete over* $\mathcal{K}$.

Proof. Let $\varphi(x, P_1, \ldots, P_n)$ be any monadic formula with a single free variable and let k be the number of variables occurring in φ. Let the formula $\varphi^*(y, P_1, \ldots, P_n)$, written with three variables, be as in lemma 4.4. By the argument of theorem 4.5, φ is $\mathcal{K}$-equivalent to the formula

$$\psi(x, P_1, \ldots, P_n) = \exists y(\pi_1(x, y) \wedge \exists x \pi_k(x, y) \wedge \varphi^*(y, P_1, \ldots, P_n)),$$

which uses only three variables. By the lemma above, there is a temporal formula A_ψ, written with the connectives above, that is $\mathcal{K}$-equivalent to ψ. The expressive completeness of the connectives is established. ■

Remarks 5.7 1. Note that the tables of the connectives in definition 5.1 can be written with three variables. So by [G] or [H, theorem 1.5], theorem 5.6 actually implies theorem 4.5.

2. The coding of tuples as single elements is used *twice* in theorem 5.6. First, we use lemma 4.4 to obtain a formula ψ, coding with only three variables the behaviour of the variables in φ, and second, we code the behaviour of these three variables using the temporal connectives. To do this in one step is complicated, involving coding arbitrarily long tuples as single elements, which is easy, and finding a finite set of connectives that can refer to these (arbitrarily many) 'components' of an element, which is not.

6 Flows of time

Let $\mathcal{K}$ be as in section 4. We will describe a new class $\mathcal{K}^*$ of flows of time (i.e., irreflexive posets) obtained from those in $\mathcal{K}$. $\mathcal{K}^*$ will have H-dimension 3, and will admit a finite set of expressively complete one-dimensional connectives, but will not have the k-variable property for any finite k. This slight strengthening of the results of section 4 is of interest because questions about H-dimension etc. were originally framed in the context of flows of time — see [G] or [H]. Moreover, it shows that restricting attention to flows of time need not simplify matters very much, at least as far as H-dimension and the k-variable property are concerned.

Definition 6.1 *Let $M \in \mathcal{K}$. We define a flow of time M^* as follows. $dom(M^*)$ will be $dom(M) \times \{-2, -1, 0, 1, 2\}$. We will write a_i for $(a, i) \in dom(M^*)$. We identify a_0 with $a \in dom(M)$. We define a binary relation $<$ on $dom(M^*)$ as follows. We let $M^* \models a_i < b_j$ iff*

1. *$a = b$ and $i < j$, or*
2. *$M \models l(a, b)$, $i = -2$ and $j \geq 0$, or*
3. *$M \models r(a, b)$, $i \leq 0$ and $j = 2$, or*
4. *$M \models \exists c(l(a, c) \wedge r(c, b))$, $i = -2$ and $j = 2$.*

It is evident that $<$ is irreflexive and transitive. Thus, M^* is a flow of time. Notice that $dom(M)$ is definable in $dom(M^*)$ by the formula $\varphi(x) = \exists y \exists z(z < y < x) \wedge \exists y \exists z(x < y < z)$, which uses only three variables.

Definition 6.2 *We define the following formulas:*

1. *$\alpha_{x=y}$ is $x = y$,*
2. *$\alpha_{l(x,y)}$ is $x \neq y \wedge \exists z(z < x \wedge \exists x(x < z \wedge x < y))$,*
3. *$\alpha_{r(x,y)}$ is $x \neq y \wedge \exists z(z > y \wedge \exists y(y > z \wedge y > x))$,*
4. *For each k with $-2 \leq k \leq 2$, $\chi_k(x, y)$ is defined so that for any $M \in \mathcal{K}$, $a, b \in dom(M)$ and $i, j \in \{-2, -1, 0, 1, 2\}$,*

   $$M^* \models \chi_k(a_i, b_j) \iff a = b, j = 0 \text{ and } i = k.$$

   We can easily find suitable formulas χ that use only three variables. For example, $\chi_{-1}(x, y)$ is $\varphi(y) \wedge x < y \wedge \exists z(z < x)$.

5. *For $i, j \in \{-2, -1, 0, 1, 2\}$ we let $=_{ij}(x, y)$ be the formula $x \neq x$ if $i \neq j$, and $x = y$ if $i = j$. Than for any $M \in \mathcal{K}$ and any $a, b \in dom(M)$,*

   $$M^* \models a_i = b_j \iff M \models =_{ij}(a, b).$$

6. *For all $i, j \in \{-2, -1, 0, 1, 2\}$ we let $<_{ij}(x, y)$ be a formula of the signature $\{l, r\}$ of $\mathcal{K}$ chosen so that for any $M \in \mathcal{K}$ and any $a, b \in dom(M)$,*

   $$M^* \models a_i < b_j \iff M \models <_{ij}(a, b).$$

   Reading off from definition 6.1, we see that $<_{ij}$ will be the disjunction of:

   - *$x = y$, if $i < j$*
   - *$l(x, y)$, if $i = -2$ and $j \geq 0$*

- $r(x, y)$, *if* $i \leq 0$ *and* $j = 2$
- $\exists z(l(x, z) \wedge r(z, y))$, *if* $i = -2$ *and* $j = 2$.

All these formulas are written with at most three variables.

Now let $\mathcal{K}^* = \{M^* : M \in \mathcal{K}\}$. The argument of theorem 4.6 shows that $\mathcal{K}^*$ does not have the k-variable property for any k. (Notice that automorphisms of M extend canonically to automorphisms of M^*.) But by theorem 4.4 and results in §6 of [H], $\mathcal{K}^*$ has H-dimension at most 3. We sketch the proof of this. Consider a monadic sentence $\sigma(P_1, \ldots, P_n)$ of the signature $\{<\}$ of $\mathcal{K}^*$. Introduce new monadic predicates P_k^i ($k \leq n, i \in \{-2, -1, 0, 1, 2\}$), and the corresponding atoms p_k^i. If $M \in \mathcal{K}$ and h is an assignment into M^*, we will define an assignment h_0 into M by

$$h_0(p_k^i) = \{a \in dom(M) : a_i \in h(p_k)\}.$$

Since every $a_i \in dom(M^*)$ has a 'representative' $a \in dom(M)$, and the relation $a_i < b_j$ on M^* is definable in terms of the representatives a, b, using $<_{ij}$, (M, h_0) 'knows' all about (M^*, h), and we can write a monadic sentence $\sigma_0(P_1^{-2}, \ldots, P_n^2)$ of the signature $\{l, r\}$ such that $(M^*, h) \models \sigma \iff (M, h_0) \models \sigma_0$. By theorem 4.5, σ_0 can be equivalently rewritten over $\mathcal{K}$ to a sentence $\sigma_1(P_1^{-2}, \ldots, P_n^2)$ using only three variables. We then simultaneously relativise the quantifiers of σ_1 to φ, replace all atomic subformulas $l(x, y)$ of σ_1 by $\alpha_{l(x,y)}$, etc., and, for each i, k, replace all monadic subformulas $P_k^i(y)$ of σ_1 by $\exists x(\chi_i(x, y) \wedge P_k(x))$ (see definition 6.2). This can be done without using more variables, since φ and the α- and χ-formulas use only three variables. We obtain a sentence $\sigma_2(P_1, \ldots, P_n)$ that is $\mathcal{K}^*$-equivalent to σ and uses only three variables. Thus we see that $\mathcal{K}^*$ has H-dimension (at most) three, as required.

A similar argument shows that $\mathcal{K}^*$ admits a finite expressively complete set of temporal connectives. See [H, §6] for details. This completes the proof of theorem 1.1.

References

A H. Andréka, *Complexity of Equations Valid in Algebras of Relations,* Dissertation for D.Sc. with the Hung. Acad. of Sci., Budapest, 1992

AN H. Andréka, I. Németi, *Relation algebraic conditions for representability of cylindric and polyadic algebras,* Preprint, Math. Inst. of the Hung. Acad. of Sci., Budapest, 1988. Submitted.

C P.J. Cameron, *Oligomorphic Permutation Groups,* London Mathematical Society Lecture Notes 152, Cambridge University Press, 1990.

DLW A. Dawar, S. Lindell, S. Weinstein, *Infinitary logic and inductive definability over finite structures,* submitted to Information and Computation.

F J. Flum, *On bounded theories,* E. Boerger, G. Jaeger, H. Kleine Buening, M.M. Richter (eds.), Proc. Computer Science Logic 91, Berne, Lecture Notes in Computer Science 626, Springer-Verlag, 111–118.

G D.M. Gabbay, *Expressive functional completeness in tense logic,* in: Aspects of Philosophical Logic, ed. U. Monnich, Reidel, Dordrecht, 1981, 91–117.

GHR D. M. Gabbay, I. M. Hodkinson, M. A. Reynolds, *Temporal logic, Volume 1,* Oxford University Press, 1994.

He L. Henkin, *Logical systems containing only a finite number of symbols,* Les Presses de l'Université de Montréal, Montréal, Canada, 1967.

H I. Hodkinson, *Finite H-dimension does not imply expressive completeness,* J. Philosophical Logic 23 (1994) 535–573.

H2 I. Hodkinson, *Finite variable logics,* Bull. Europ. Assoc. Theor. Comp. Sci. 51 (1993) 111–140.

I N. Immerman, *Upper and lower bounds for first-order expressibility,* J. Comput. System Sci. 25 (1982), 76–98.

IK N. Immerman, D. Kozen, *Definability with bounded number of bound variables,* Proceedings IEEE 1987, 236–244.

K J.A.W. Kamp, *Tense Logic and the Theory of Linear Order,* Ph.D. thesis, University of California, 1968.

KV Ph. Kolaitis, M. Vardi, *Infinitary logics and 0–1 laws,* Information and Computation 98 (1992), 258–294.

N I. Németi, *Free Algebras and Decidability in Algebraic Logic,* Dissertation for D.Sc. with the Hung. Acad. Sci., Budapest, 1986 (in Hungarian).

TG A. Tarski, S. Givant, *A Formalization of Set Theory without Variables,* AMS Colloquium publications, Providence, R.I., Vol 41, 1987.

V Y. Venema, Many-dimensional Modal Logic, Ph.D. Thesis, University of Amsterdam, 1992.

Ian Hodkinson, Department of Computing, Imperial College, 180 Queen's Gate, London SW7 2BZ, U.K.

András Simon, Mathematical Institute of the Hungarian Academy of Sciences, Budapest P. O. B. 127, H-1364 Hungary.