

Jean Goubault-Larrecq

On a Generalization of a Result by  
Valk and Jantzen

Research Report LSV-09-09

May 2009

Laboratoire  
Spécification  
et  
Vérification



CENTRE NATIONAL  
DE LA RECHERCHE  
SCIENTIFIQUE

Ecole Normale Supérieure de Cachan  
61, avenue du Président Wilson  
94235 Cachan Cedex France



# On a Generalization of a Result by Valk and Jantzen

Jean Goubault-Larrecq<sup>1,2</sup>

<sup>1</sup> LSV, ENS Cachan, CNRS, France    goubault@lsv.ens-cachan.fr  
<sup>2</sup> INRIA Saclay, France

**Abstract.** We show that, under mild assumptions on the effective, well quasi-ordered set  $X$ , one can compute a finite basis of an upward-closed subset  $U$  of  $X$  if and only if one can decide whether  $U \cap \downarrow z$  is empty for every  $z \in \widehat{X}$ . Here  $\widehat{X}$  is the completion of  $X$  as defined in Finkel and Goubault-Larrecq, *Forward Analysis for WSTS, Part I: Completions*, STACS'09, pages 433-444, 2009. This generalizes a useful result proved by Valk and Jantzen in 1985, which is the case  $X = \mathbb{N}^k$ .

## 1 Introduction

Let  $X$  be a well-quasi-ordered space. We write  $\leq$  its quasi-ordering,  $\downarrow A$  is the downward-closure of the subset  $A$ ,  $\downarrow x = \downarrow \{x\}$ ,  $\uparrow A$  is the upward-closure of  $A$ ,  $\uparrow x = \uparrow \{x\}$ . Since  $X$  is well-quasi-ordered, every upward-closed subset  $U$  of  $X$  has a finite *basis*, i.e., a finite set  $A$  such that  $U = \uparrow A$ . One can even require the basis to be minimal, equivalently to consist of the minimal elements of  $U$ .

A useful theorem by Valk and Jantzen [8, Theorem 2.14] states that, when  $X = \mathbb{N}^k$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap \downarrow z \neq \emptyset$  is empty, for every  $z \in \mathbb{N}_\omega^k$ . Here  $\mathbb{N}_\omega$  is  $\mathbb{N}$  plus a new top element  $\omega$ , and  $U \cap \downarrow z \neq \emptyset$  reduces to checking that  $U$  contains a  $k$ -tuple  $(i_1, \dots, i_k) \in \mathbb{N}^k$  satisfying some constraints  $i_j \leq c_j$ , for some constants  $c_j \in \mathbb{N}$ , and for certain values of  $j$ ,  $1 \leq j \leq k$ . (Namely, for those values of  $j$  such that the  $j$  component of  $z$  is not  $\omega$ ; then  $c_j$  is this  $j$  component.) E.g., with  $k = 5$ , checking  $U \cap \downarrow (1, \omega, 2, 3, \omega) \neq \emptyset$  means checking that there is a tuple  $(i_1, i_2, i_3, i_4, i_5)$  in  $U$  with  $i_1 \leq 1$ ,  $i_3 \leq 2$ , and  $i_4 \leq 3$ .

Our purpose is to show that the Valk-Jantzen generalizes to many spaces  $X$ . Abstractly,  $X$  needs only be an *effective* well poset satisfying the so-called *effective complement property* (Definition 5). We argue that most well-quasi-ordered spaces  $X$  used in verification satisfy these requirements. In fact, all the spaces obtained from any of the data types of [4, Section 5] satisfy these requirements.

The Valk-Jantzen result has a number of applications. Valk and Jantzen combine it with the construction of a coverability graph of a Petri net to show that it is decidable whether, starting from a given marking  $m_0$ , one can fire an sequence of transitions where each transition in a given set  $\hat{T}$  is fired infinitely often (the  $\hat{T}$ -continuity problem); whether each reachable marking from  $m_0$  allows at least one transition from  $\hat{T}$  to be fired (the  $\hat{T}$ -nonblockedness problem); whether there is at least one infinite sequence of transitions out of  $m_0$  (the *liveness* problem); whether there are infinitely many markings reachable from  $m_0$  (the *unboundedness* problem); whether a signal net is prompt, and a variety of other problems. Finkel *et al.* [5, Theorem 4.2] use it to show

that coverability of  $\omega$ -well-structured nets (extending Petri nets strictly) is decidable. The Valk-Jantzen result has also recently been used by Abdulla and Mayr [3] in their study of priced timed Petri nets, where it is used in deriving fine decidability results on ordinary Petri nets constructed from priced timed Petri nets.

It is not our purpose here to give any application of our generalized Valk-Jantzen result. We hope it will be of practical use, and have some reason to. In fact, Theorem 2 is of use in some work by Dimino and Schnoebelen on lossy channel systems, which should be published in the near future.

## 2 Outline and Capsule

Our main theorem is Theorem 1, which reads:

**Theorem 1.** *Let  $X$  be a strongly effective well poset with the effective complement property. Then there are Turing machines that convert from (minimal) finite basis representations to strong oracle representations of upward-closed subsets and back.*

While this does not look anywhere near the announced theorem by Valk-Jantzen, we shall argue in Section 3 that this is really what the generalized Valk-Jantzen Theorem should state. All missing notions are also introduced and explained in Section 3.

Theorem 1 implies a number of special cases, which we list in Section 4. These are obtained by showing that the assumptions of Theorem 1 are satisfied whenever  $X$  is given by any of the data types of [4, Section 5]: this will be Theorem 3 below. However, more concretely, we obtain in particular the following new result:

**Theorem 2.** *Let  $\Sigma$  be a finite alphabet, and order  $\Sigma^*$  by the subword ordering, i.e.,  $w \leq_{\Sigma^*} w'$  iff  $w$  is of the form  $a_1 a_2 \dots a_m$ , and  $w' = w_0 a_1 w_1 a_2 w_2 \dots w_{m-1} a_m w_m$  for some words  $w_0, w_1, w_2, \dots, w_{m-1}, w_m \in \Sigma^*$ .*

*For any upward-closed subset  $U$  of  $\Sigma^*$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap \llbracket P \rrbracket_{\Sigma^*}$  is empty, for every \*-product  $P$ .*

The \*-Products  $P$  are special regular expressions, obtained as concatenations of atomic expressions  $a^?$  ( $a \in \Sigma$ ) and  $A^*$  ( $A$  a non-empty subset of  $\Sigma^*$ ). If  $P$  is a concatenation  $e_1 e_2 \dots e_n$  of atomic expressions  $e_1, e_2, \dots, e_n$ , the language  $\llbracket P \rrbracket_{\Sigma^*}$  of  $P$  is given as the set of words  $w_1 w_2 \dots w_n$  with  $w_i \in \llbracket e_i \rrbracket_{\Sigma^*}$ ,  $1 \leq i \leq n$ ; and  $\llbracket a^? \rrbracket_{\Sigma^*}$  is the set  $\{\epsilon, a\}$ , while  $\llbracket A^* \rrbracket_{\Sigma^*}$  is the set of words whose letters are in  $A$ .

So Theorem 2 reduces the question of constructing a finite basis of an upward-closed subset of words in the subword ordering to the question of whether it meets certain regular languages.

Other similar results are presented in Theorem 4 for the multiset language generators, and in Theorem 5 for the word language generators of [2]. These are bricks in the study of timed Petri nets.

We shall also state the following theorem, which deals with the case of finite sequences of  $k$ -tuples of natural numbers, as used in data nets [7]:

**Theorem 6.** *For any upward-closed subset  $U$  of  $(\mathbb{N}^k)^*$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap \llbracket P \rrbracket_{(\mathbb{N}^k)^*}$  is empty, for every data net product  $P$ .*

Data net products  $P$  are concatenations of atomic expressions  $\mathbf{a}^?$  where  $\mathbf{a} \in \mathbb{N}_\omega^k$ , and  $A^*$ , where  $A$  is any finite subset of  $\mathbb{N}_\omega^k$ . When  $P$  is the concatenation  $e_1 e_2 \dots e_n$ ,

the language of  $P$ ,  $\llbracket P \rrbracket_{(\mathbb{N}^k)^*}$ , is the set of sequences of the form  $w_1 w_2 \dots w_n$  with  $w_1 \in \llbracket e_1 \rrbracket_{(\mathbb{N}^k)^*}$ ,  $w_2 \in \llbracket e_2 \rrbracket_{(\mathbb{N}^k)^*}$ ,  $\dots$ ,  $w_n \in \llbracket e_n \rrbracket_{(\mathbb{N}^k)^*}$ . In turn,  $\llbracket \mathbf{a}^? \rrbracket_{(\mathbb{N}^k)^*}$  is the set consisting of the empty sequence, plus every sequence consisting of just one  $k$ -tuple  $\mathbf{x} \in \mathbb{N}^k$  with  $\mathbf{x} \leq \mathbf{a}$ ; and  $\llbracket A^* \rrbracket_{(\mathbb{N}^k)^*}$  is the set of sequences of  $k$ -tuples  $\mathbf{x}$  from  $\mathbb{N}^k$  such that  $\mathbf{x} \leq \mathbf{a}$  for some  $\mathbf{a} \in A$ .

We finally state the following theorem, which is of use in the context of multiset vector addition systems with states (MVASS). The latter will be the object of a forthcoming study.

**Theorem 7.** *Let  $Q$  be a finite state space, with the trivial ordering (i.e., equality). For any upward-closed  $U$  of  $\llbracket (Q \times \mathbb{N}^k)^{\otimes} \rrbracket$ , one can effectively construct a basis of  $U$  iff one can decide whether  $U \cap \llbracket A^{\otimes} \odot (q_1, \mathbf{a}_1)^{\otimes} \odot \dots \odot (q_n, \mathbf{a}_n)^{\otimes} \rrbracket_{(\mathbb{N}^k)^{\otimes}}$  is empty, for every finite subset  $A$  of  $Q \times \mathbb{N}^k$ , and every elements  $q_1, \dots, q_n \in Q$ ,  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{N}^k$ .  $\llbracket A^{\otimes} \odot (q_1, \mathbf{a}_1)^{\otimes} \odot \dots \odot (q_n, \mathbf{a}_n)^{\otimes} \rrbracket_{\Sigma^{\otimes}}$  is a set of multisets of pairs  $(q, \mathbf{x})$  with  $q \in Q$  and  $\mathbf{x}$  is a  $k$ -tuple of natural numbers. These multisets are those that contain an arbitrary number of  $k$ -tuples of the form  $(q, \mathbf{x})$  for some  $(q, \mathbf{a}) \in A$  such that  $\mathbf{x} \leq \mathbf{a}$ , plus at most one element of the form  $(q_1, \mathbf{x}_1)$  with  $\mathbf{x}_1 \leq \mathbf{a}_1$ , plus  $\dots$  plus at most one element of the form  $(q_n, \mathbf{x}_n)$  with  $\mathbf{x}_n \leq \mathbf{a}_n$ .*

### 3 The Generalized Valk-Jantzen Theorem

*Preliminaries.* Our result will be based on the paper [4], which uses results from [6]. The reader is referred to these papers for basic notions, notations, and theorems. We merely mention that a *well-quasi-ordering* on a set  $X$  is a quasi-ordering  $\leq$  (a reflexive, transitive binary relation) such that any infinite sequence  $(x_n)_{n \in \mathbb{N}}$  is such that there are two indices  $i, j \in \mathbb{N}$  with  $i < j$  and  $x_i \leq x_j$ . Equivalently, a quasi-ordering is well iff it is well-founded (no infinite descending chain) and has no infinite anti-chain (set of incomparable elements). Equivalently again, iff every upward-closed subset of  $X$  has a finite basis, as mentioned in the introduction.

*The Valk-Jantzen Theorem as equivalence of representations.* As we have stated it, the Valk-Jantzen Theorem is formally meaningless, despite the fact that it is perfectly understandable. Indeed, to be able to *effectively* construct a basis of  $U$ , i.e., to *compute* a basis, one first need to *represent*  $U$  as some data structure. We must therefore reformulate the theorem in such a way that the representation of upward-closed sets  $U$  is made explicit.

Finite bases are the most canonical representation for upward-closed subsets of a well-quasi-ordered set. However, representing upward-closed subsets by their finite bases would make the theorem vacuous. One may instead represent  $U$  by an oracle that, given an element  $x \in X$ , answers whether  $x \in U$ . (Call this the *weak oracle representation* of  $U$ .) This is fairly general, but not quite what the Valk-Jantzen Theorem states. Indeed, this would assume that  $U$  is a recursive set, while no such assumption is made by Valk and Jantzen.

Instead, one notes that being able to decide whether  $U \cap \downarrow z$  is empty for all  $z \in \mathbb{N}^k_\omega$ , as in the statement of the Valk-Jantzen Theorem, means that we do have a representation of  $U$ , by means of an oracle that, given  $z \in \mathbb{N}^k_\omega$ , decides whether  $U \cap \downarrow z$  is empty.

Call this the *strong oracle representation* of  $U$ , while a finite basis of  $U$  will be a *finite basis representation* of  $U$ . The Valk-Jantzen Theorem then formally means that each representation is computable from the other one, i.e., that there are two algorithms, each converting from one representation to the other.

*Remark 1.* The adjectives ‘strong’ and ‘weak’ are justified by the following fact: one can compute a weak oracle representation of  $U$  from any strong oracle representation of  $U$ . Indeed, the weak oracle takes  $x \in X$  and checks whether  $x \in U$  by checking whether  $U \cap \downarrow x$ , using the strong oracle.

*Generalizing the Valk-Jantzen Theorem.* Now generalize, and replace  $\mathbb{N}^k$  by any well-quasi-ordered set  $X$ . For the strong oracle representation to actually represent  $U$ , we need the quasi-ordering  $\leq$  to actually be an ordering. We shall assume so in the sequel. We also need to find a suitable equivalent of  $\mathbb{N}_\omega^k$ . This was solved in [4], where a suitable *completion*  $\widehat{X}$  of  $X$  is defined, which can be described either as the *ideal completion*  $\text{Idl}(X)$ , or as the *sobrification*  $\mathcal{S}(X_a)$  of  $X_a$ , the topological space obtained by assigning the Alexandroff topology of  $\leq$  to  $X$ . For now, we should be reassured by the fact that, when  $X = \mathbb{N}^k$ , then  $\widehat{X} = \mathbb{N}_\omega^k$ .

For our purposes, it is easier to describe  $\widehat{X}$  as  $\text{Idl}(X)$  than as  $\mathcal{S}(X_a)$ . A subset  $F$  of  $X$  is *directed* if and only if it is non-empty, and every two elements of  $F$  have a common upper bound in  $F$ . An *ideal* is a downward-closed directed subset of  $X$ . Then  $\text{Idl}(X)$  is the set of all ideals of  $X$ , ordered by inclusion.

Then  $X$  embeds into  $\widehat{X}$  through  $\eta : x \mapsto \downarrow x$ . (Check that  $\downarrow x$  is always an ideal.) By embedding, we mean that  $\eta$  is injective, order-preserving, and order-reflecting. That is,  $\eta$  is an order-isomorphism of  $X$  onto its image  $\text{Im } \eta$ . Up to isomorphism,  $\widehat{X}$  is then just  $X$  plus some new elements added.

We then need to be able to represent *elements* of  $X$  (and also of  $\widehat{X}$ , this will come later.)

**Definition 1.** An *effective poset* is a poset  $X$  together with a surjective map  $r : E \rightarrow X$  (the *representation map*), where  $E$  is an r.e. subset of  $\mathbb{N}$ , so that the binary relation  $\sqsubseteq$  on  $E \times E$ , defined by  $e_1 \sqsubseteq e_2$  iff  $r(e_1) \leq r(e_2)$ , is computable.

An *effective well poset* is an effective poset as above such that  $X$  is well-quasi-ordered by  $\leq$ .

We say that  $\sqsubseteq$  is *computable on*  $E \times E$  iff the map  $e_1, e_2 \mapsto (e_1 \sqsubseteq e_2)$  is partial recursive from  $\mathbb{N}^2$  to the Booleans, and its domain of definition contains  $E \times E$ . More informally, Definition 1 requires  $\leq$  be decidable on  $X$ , and that elements of  $X$  can be enumerated.

If  $r(e)$  is defined (if  $e$  is in  $E$ ) and equals  $x$ , we say that  $e \in \mathbb{N}$  is a *code* for  $x \in X$ . It may however be the case that a given element  $x \in X$  has several codes. We shall usually leave  $r$  and  $E$  implicit, and call  $X$  itself an effective poset.

*Example 1.*  $X = \mathbb{N}^k$  is an effective well poset, with the componentwise ordering  $\leq$ . Well-quasi-ordering is by Dickson’s Lemma. Effectiveness is (at least informally) clear: one can compute whether  $i \leq j$ , for any two tuples  $i, j \in \mathbb{N}^k$ .  $\square$

*Example 2.* Let  $\Sigma$  be a finite alphabet, and  $X = \Sigma^*$  be the set of all finite words over  $\Sigma$ , ordered by the subword ordering. This is the least relation such that  $a_1 a_2 \dots a_n \leq w_0 a_1 w_1 a_2 \dots a_n w_n$ , for all letters  $a_1, a_2, \dots, a_n \in \Sigma$  and words  $w_0, w_1, \dots, w_n \in \Sigma^*$ . This ordering is a well-quasi-ordering by Higman's Lemma. Effectiveness is also clear, although finding a polynomial-time algorithm for  $\leq$  is slightly trickier than in Example 1 (use dynamic programming).  $\square$

**Definition 2 (Finite Basis Representation).** *Let  $X$  be an effective well poset, with representation map  $r : E \rightarrow X$ . A finite basis representation for an upward-closed subset  $U$  of  $X$  is any finite subset  $A$  of  $E$  such that  $U = \uparrow r(A)$ .*

For any finite subset  $A$  of  $E$ ,  $\uparrow r(A)$  is upward-closed. Conversely, every upward-closed subset  $U$  of  $X$  is the upward-closure of some finite subset since  $X$  is well-quasi-ordered, and this finite subset is of the form  $r(A)$ , since  $r$  is surjective. So the sets that have a finite basis representation are exactly the upward-closed subsets of  $X$ .

*Remark 2.* One is often interested in a *minimal* finite basis representation of  $U$ . However, it is easy to compute one from any given finite basis, since  $\sqsubseteq$  is computable on the set  $E$  of codes of elements of  $X$ .

To define a generalization of strong oracle representations, we first require elements of  $\widehat{X}$ , and not just  $X$ , to have codes. Recall that  $X$  embeds into  $\widehat{X}$  through  $\eta : x \mapsto \downarrow x$  [4].

**Definition 3 (Strongly Effective Well Poset).** *A strongly effective well poset is any effective well poset  $X$ , together with two surjective maps  $r : E \rightarrow X$  (the subspace representation map) and  $r' : E' \rightarrow \widehat{X}$  (the complete representation map), such that:*

1.  $E \subseteq E' \subseteq \mathbb{N}$ ;
2.  $r'$  extends  $r$ , in the sense that  $r'(e) = \eta(r(e))$  for all  $e \in E$ ;
3.  $E$  is r.e.;
4. the binary relation  $\sqsubseteq$  on  $E \times E'$  defined by  $e \sqsubseteq e'$  iff  $r(e) \in r'(e')$  is computable, i.e., it is partial recursive and its domain of definition contains  $E \times E'$ .

It is sometimes convenient to equate  $X$  with the subset  $\eta(X)$  of  $\widehat{X}$ , up to isomorphism. Under this so-called *subspace convention*, every element of  $X$  is in  $\widehat{X}$ , and  $\widehat{X}$  contains additional, so-called *limit elements*. Note that in strongly effective well posets, we require all elements, including limit elements, to have codes, although only the non-limit elements, in  $X$ , are required to form an r.e. set. We also require the ability to decide when  $x \leq_{\widehat{X}} z$  for  $x \in X$  and  $z \in \widehat{X}$  (where  $\leq_{\widehat{X}}$  is the ordering on  $\widehat{X}$ , i.e., inclusion; this is the last item: observe that  $e \sqsubseteq e'$  iff  $r'(e) \subseteq r'(e')$ ).

We shall sometimes decide to use the subspace convention, sometimes not. While it has the merit of removing some clutter in concrete cases, it causes some type confusions in statements and proofs of theorems.

All strongly effective well posets  $X$  are in particular effective well posets, with representation map  $r$ , and ordering on codes  $e_1, e_2 \in E$  defined by restricting  $\sqsubseteq$  to  $E \times E$ . Indeed,  $e_1 \sqsubseteq e_2$  iff  $r(e_1) \in r'(e_2)$  iff  $r(e_1) \in \downarrow r(e_2)$  (item 2), iff  $r(e_1) \leq r(e_2)$ , matching Definition 1.

Conversely, one may ask when an effective well poset  $X$  can be turned into a strongly effective well poset. That  $r'$  must be onto requires  $\widehat{X}$  to be countable. Since  $r$  is onto,  $X$  is countable, so we may use the following.

**Lemma 1.** *For any countable well poset  $X$ ,  $\widehat{X}$  is countable.*

*Proof.*  $X$  has countably many upward-closed subsets, since all of them can be described by a finite basis, as  $X$  is well. Since  $X$  is countable, there are only countably many such finite bases. It follows that  $X$  only has countably many downward-closed subsets, since the latter are exactly the complements of upward-closed subsets. Now notice that elements of  $\widehat{X}$  are certain downward-closed subsets of  $X$ .  $\square$

The same ideas shows that *every* effective well poset can be turned into a strongly effective well poset:

**Proposition 1.** *Let  $X$  be an effective well poset, with representation map  $r_0 : E_0 \rightarrow X$ . Then one can equip  $X$  with subspace representation maps  $r : E \rightarrow X$  and complete representation maps  $r' : E' \rightarrow \widehat{X}$ , turning  $X$  into a strongly effective well poset.*

*Proof.* Fix an injective map from the set of finite subsets of  $\mathbb{N}$  to  $\mathbb{N}$ , say  $A \mapsto \ulcorner A \urcorner = \sum_{n \in A} 2^n$ . Call  $\ulcorner A \urcorner$  the *code* of  $A$ . For any finite subset  $A$  of  $E_0$ , let  $r'_1(\ulcorner A \urcorner) = X \setminus \uparrow r_0(A)$ . Define a *limit code* as any integer of the form  $\ulcorner A \urcorner$  such that  $r'_1(\ulcorner A \urcorner)$  is in  $\widehat{X}$  (that is, is not just downward-closed, but also directed), and not of the form  $\downarrow x$  for any  $x \in X$ . Note that every limit element of  $\widehat{X}$ , i.e., every element outside the range of  $\eta$ , is of the form  $r'_1(n)$  for some limit code  $n$ .

Now define  $E$  as  $\{2n \mid n \in E_0\}$ ,  $r(2n) = r_0(n)$ ,  $E'$  as the (disjoint) union of  $E$  with  $\{2n+1 \mid n \in E'_1\}$ ,  $r'(2n) = \downarrow r_0(n)$  for all  $n \in E_0$ ,  $r'(2n+1) = r'_1(n)$  for all  $n \in E'_1$ . Items 1 through 3 of Definition 3 are clear. (Note that  $E$  is r.e. since  $E_0$  is; however, there is no reason why  $E'_1$ , hence  $E'$ , should be r.e.) As far as item 5 is concerned, for every  $e, e' \in E'$ , define  $e \sqsubseteq e'$  iff  $r(e) \in r'(e')$ ; also, define  $\sqsubseteq_0$  on  $E_0 \times E_0$  by  $m \sqsubseteq_0 n$  iff  $r_0(m) \leq r_0(n)$ : this is computable by assumption. For every  $e = 2m \in E$  (i.e.,  $m \in E_0$ ) and  $e' \in E'$ ,  $e \sqsubseteq e'$  iff either  $e' = 2n$  for some  $n \in E_0$  and  $m \sqsubseteq_0 n$ , or  $e' = 2\ulcorner A \urcorner + 1$  for some finite subset  $A$  of  $E_0$ , and for every  $n \in A$ ,  $m \not\sqsubseteq_0 n$ . So  $\sqsubseteq$  is computable on  $E \times E'$ .  $\square$

Proposition 1 may lead one to think that the notion of *strongly* effective well poset is useless. However, an effective well poset may be strongly effective in more than one way, i.e., for different pairs of representation maps  $r, r'$ , some of which being more practical than others. The crucial effective complement property below (Definition 5) will depend on both  $r$  and  $r'$ , and is therefore more easily defined on strongly effective well posets than on effective well posets.

*Example 3.*  $X = \mathbb{N}^k$  is strongly effective, as it is effective. However, instead of representing limit elements (in  $\widehat{X}$ , outside of  $X$ ) as certain complements of subsets of the form  $\uparrow A$ , where  $A$  are certain finite subsets of  $\mathbb{N}^k$ , it is more convenient to realize that, as we have mentioned earlier,  $\widehat{X} = \mathbb{N}_\omega^k$ , up to order-isomorphism, with the componentwise ordering. So the limit elements will be  $k$ -tuples  $(j_1, \dots, j_k)$  in  $\mathbb{N}_\omega^k$ , where at least one component  $j_p$ ,  $1 \leq p \leq k$ , equals  $\omega$ . Then, the ordering on  $\mathbb{N}_\omega^k$  is decidable,



hence also its restriction to  $X \times \widehat{X}$  (encoded through  $\sqsubseteq$  in Definition 3), defined by  $(i_1, \dots, i_k) \leq (j_1, \dots, j_k)$  iff, for every  $p$ ,  $1 \leq p \leq k$ , either  $j_p = \omega$  or  $i_p \leq j_p$ . Moreover, one can enumerate the elements of  $X$  in an effective way, so  $\mathbb{N}^k$  is strongly effective.  $\square$

*Example 4.* When  $\Sigma$  is a finite alphabet,  $X = \Sigma^*$ , with the subword ordering, is also strongly effective, with representation maps that are again more natural than the canonical ones obtained in Proposition 1. We rely on [4, Proposition 4.1], where elements of  $\widehat{X}$  are obtained as certain regular expressions that were called products there, and which we call  $*$ -products here. Recall from [1] that an *atomic expression* is any regular expression of the form  $a^?$ , with  $a \in \Sigma$  (this denotes the set  $\{\epsilon, a\}$ ), or  $A^*$ , where  $A$  is a non-empty subset of  $\Sigma$  (this denotes the set of all finite words whose letters are all taken from  $A$ ). A  *$*$ -product* is any regular expression of the form  $e_1 e_2 \dots e_n$  ( $n \in \mathbb{N}$ ), where each  $e_i$  is an atomic expression. Such a  $*$ -product denotes the set of all words  $w_1 w_2 \dots w_n$ , where  $w_1$  is taken from  $e_1$ ,  $w_2$  from  $e_2$ ,  $\dots$ ,  $w_n$  from  $e_n$ . A *simple regular expression*, or *SRE*, is a sum, either  $\emptyset$  or  $P_1 + \dots + P_k$ , where  $P_1, \dots, P_k$  are  $*$ -products. Sum is interpreted as union.

Proposition 4.1 of [4] states that  $\widehat{X}$  is exactly the set of (denotations of)  $*$ -products. The ordering on  $\widehat{X}$  (inclusion of denotations of  $*$ -products) can be decided in quadratic time, as shown by Abdulla *et al.* [1]. Since this is decidable and one can effectively enumerate all elements of  $\Sigma^*$ ,  $\Sigma^*$  is strongly effective.

By the way, this is one case where the subspace convention hinders readability. The embedding  $\eta$  maps each word  $w = a_1 a_2 \dots a_n$  to the  $*$ -product  $\eta(w) = a_1^? a_2^? \dots a_n^?$ . Using the subspace convention would force us to equate  $a_1 a_2 \dots a_n$  with  $a_1^? a_2^? \dots a_n^?$ , which is awkward at best.  $\square$

**Definition 4 (Strong Oracle Representation).** *Let  $X$  be a strongly effective well poset, with complete representation map  $r' : E' \rightarrow \widehat{X}$ . A strong oracle representation for an upward-closed subset  $U$  of  $X$  is any code of a Turing machine  $\mathcal{M}$  that, on input  $e' \in E'$ , accepts if  $U \cap r'(e') \neq \emptyset$ , and rejects otherwise.*

So  $\mathcal{M}$  is a Turing machine that is required to terminate on all inputs  $e'$  from  $E'$ . In effect, we require to be able to decide whether  $U \cap r'(e')$  is empty or not, for all codes  $e' \in E'$  of elements of  $\widehat{X}$ .

Note that  $r'(e')$  is in  $\widehat{X}$ , and elements of  $\widehat{X}$  are certain subsets of  $X$ , so  $U \cap r'(e')$  makes sense. However, this statement is more readable if one uses the subspace convention, whereby the *set*  $r'(e')$  really is the set of elements of  $X$  that are below the *element*  $r'(e')$  in the larger space  $\widehat{X}$ . That is, under the subspace convention, we require to be able to decide whether  $U \cap \downarrow_{\widehat{X}}(r'(e'))$  is empty or not. (We write  $\downarrow_{\widehat{X}}$  to make it clear that the downward-closure is taken in  $\widehat{X}$ .) Or, in slightly less formal terms, whether  $U \cap \downarrow_{\widehat{X}} z$  is empty or not for all  $z \in \widehat{X}$ .

To state our generalization of the Valk-Jantzen Theorem, we shall require a final assumption on  $X$ . To appreciate it, one has to realize the following first. Given any upward-closed subset  $U$  of  $X$ , its complement  $X \setminus U$  is downward-closed in  $X$ . But (under the subspace convention) any downward-closed subset  $F$  of  $X$  has a finite *dual basis*  $\{z_1, \dots, z_n\} \subseteq \widehat{X}$ , i.e.,  $F = X \cap \downarrow_{\widehat{X}}\{z_1, \dots, z_n\}$ . This is a reformulation of [4,

Proposition 4.2] using the subspace convention. In other words, we may either represent  $U$  through one of its finite bases ( $\{x_1, \dots, x_m\} \subseteq X$ , so that  $U = \uparrow_X \{x_1, \dots, x_m\}$ ), or through one of the finite dual bases of its complement. However, there is no reason to think that one could *compute* each representation from the other one.

So we shall require the following *effective complement property*. Under the subspace convention, and ignoring representation maps, this states that we can always compute a finite dual basis of the complement of an upward-closed subset  $U$ , represented through one of its finite bases.

**Definition 5 (Effective Complement Property).** *Let  $X$  be a strongly effective well poset, with subspace representation map  $r : E \rightarrow X$ , and complete representation map  $r' : E' \rightarrow \widehat{X}$ .*

*We say that  $X$  has the effective complement property iff there is a Turing machine that, on input  $\{e_1, \dots, e_m\} \subseteq E$ , computes a finite set  $\{e'_1, \dots, e'_n\} \subseteq E'$  such that  $\{r'(e'_1), \dots, r'(e'_n)\}$  is a finite dual basis of the complement of  $\uparrow\{r(e_1), \dots, r(e_m)\}$ , i.e., such that:*

$$X \setminus \uparrow\{r(e_1), \dots, r(e_m)\} = r'(e'_1) \cup \dots \cup r'(e'_n)$$

*Remark 3.* The canonical representation maps built in Proposition 1 may lead one to erroneously think that it is trivial to compute such a representation, because both go through complements. More precisely, let  $U = \uparrow\{x_1, \dots, x_m\}$ , then we would like to represent  $U$  through the finite dual base with just one element  $z_1 = X \setminus U$ . However,  $z_1$  is not in general even an element of  $\widehat{X}$ ; while it is downward-closed, it is in general not directed, so this fails. We shall make this more explicit in Example 6 below.

*Example 5.*  $X = \mathbb{N}^k$  is easily seen to have the effective complement property, using the strongly effective presentation of  $\widehat{X}$  as  $\mathbb{N}_\omega^k$  (see Example 3). For example, the complement of  $\uparrow(1, 3, 2)$  (if  $k = 3$ ) is the intersection of  $X = \mathbb{N}^3$  with  $\downarrow_{\widehat{X}}\{(0, \omega, \omega), (\omega, 2, \omega), (\omega, \omega, 1)\}$ : not being above  $(1, 3, 2)$  means having at least one component that is too low, i.e., having a first component less than or equal to 0, a second component less than or equal to 2, or a third component less than or equal to 1. In general, the complement of  $\uparrow(i_1, i_2, \dots, i_k)$  in  $X = \mathbb{N}^k$  is equal to  $X \cap \downarrow_{\widehat{X}}\{(\omega, \dots, \omega, i_j - 1, \omega, \dots, \omega) \mid 1 \leq j \leq k, i_j \geq 1\}$ . The complement of  $\uparrow\{x_1, \dots, x_m\}$  can then be computed as the intersection of the complements of  $\uparrow x_1, \dots, \uparrow x_m$ . In turn, one can compute intersections  $\downarrow_{\widehat{X}}\{\mathbf{y}_1, \dots, \mathbf{y}_m\} \cap \downarrow_{\widehat{X}}\{\mathbf{z}_1, \dots, \mathbf{z}_p\}$  as  $\downarrow_{\widehat{X}}\{\min(\mathbf{y}_i, \mathbf{z}_j) \mid 1 \leq i \leq m, 1 \leq j \leq p\}$ , where minima  $\min(\mathbf{y}_i, \mathbf{z}_j)$  are computed componentwise, e.g.,  $\min((1, \omega, 3, \omega, 2), (3, 5, 0, \omega, \omega)) = (1, 5, 0, \omega, 2)$ .  $\square$

*Example 6.* Let us try to compute a finite dual basis of  $U = \uparrow\{(1, 3, 2)\}$  in  $X = \mathbb{N}^3$ , using the canonical representation maps of Proposition 1 instead of the representation of  $\widehat{X}$  as  $\mathbb{N}_\omega^3$ . The minimal finite dual basis is  $\{z_1, z_2, z_3\}$ , where  $z_1$  is the complement of  $\uparrow\{(1, 0, 0)\}$ ,  $z_2$  is the complement of  $\uparrow\{(0, 3, 0)\}$ , and  $z_3$  is the complement of  $\uparrow\{(0, 0, 2)\}$ . Check indeed that  $z_1, z_2, z_3$  are not only downward-closed but also directed. These elements are easily seen to be the set of tuples that are respectively below  $(0, \omega, \omega), (\omega, 2, \omega), (\omega, \omega, 1)$ ; i.e., this is in one-to-one correspondence with the finite dual basis given in Example 5. We hope the reader will share our opinion that the

representation of  $\widehat{X}$  as elements of  $\mathbb{N}_\omega^k$ , as in Example 5, makes the finite dual basis representation clearer.

*Example 7.*  $X = \Sigma^*$  also has the effective complement property, using the description of  $\widehat{X}$  through  $*$ -products (Example 4). For any word  $w \in \Sigma^*$ , say  $a_1 a_2 \dots a_n$ ,  $\uparrow w$  is a regular language, namely  $\Sigma^* a_1 \Sigma^* a_2 \Sigma^* \dots \Sigma^* a_n \Sigma^*$ . From any finite basis  $\{w_1, \dots, w_m\}$  of an upward-closed subset  $U$ , we can therefore compute a regular expression whose language is exactly  $U$ . One can compute its complement as a regular expression, using standard automata-theoretic techniques. Then we must show that, given any regular expression  $E$  whose language is downward-closed, we can effectively convert it into an SRE. This is well-known, although the construction is usually presented as a way of showing that any downward-closed language is the language of some SRE, without any particular stress on computability.

First, we can always assume that  $E$  is a regular expression from the following grammar:

$$\begin{array}{l}
 E ::= a^? \\
 \quad | \quad \epsilon \\
 \quad | \quad \emptyset \\
 \quad | \quad E + E \\
 \quad | \quad EE \\
 \quad | \quad E^*
 \end{array}$$

That is, we can assume that no letter occurs in  $E$  except under a question mark. Otherwise, convert each letter  $a$  not under a question mark to  $a^?$ . Since  $E$  denotes a downward-closed language, it is easy to see that this transformation preserves languages. Moreover, every sub-expression of  $E$  now denotes a downward-closed subset as well.

We now recurse on the syntax of  $E$  to convert it to an SRE. It is easy to see that  $a^?$ ,  $\epsilon$ ,  $\emptyset$  are already SREs, that sums of SREs are SREs, that products of SREs can be converted to SREs by distributing product over sum. It only remains to show how to convert  $L^*$  to an SRE, where  $L$  is any SRE. Define the following rewrite relation, where  $P$  denotes any  $*$ -product,  $L$  denotes any sum of an SRE and of isolated letters:

$$\begin{array}{l}
 \emptyset^* \rightarrow \emptyset \quad (a^? P + L)^* \rightarrow (P + a + L)^* \\
 (\{a_1, \dots, a_n\}^* P + L)^* \rightarrow (P + a_1 + \dots + a_n + L)^*
 \end{array}$$

Under the above assumptions on  $P$  and  $L$ , which imply that the languages of  $P$  and  $L$  are downward-closed, these rules preserve languages. Moreover, they terminate, as each of the last two rules decreases the sum of the lengths of  $*$ -products under the star sign. Its only normal forms are  $\emptyset$  and expressions of the form  $L^*$ , where  $L$  is a sum of isolated letters  $a_1 + \dots + a_n$ ; we then convert the latter to the  $*$ -product  $A^*$ , where  $A = \{a_1, \dots, a_n\}$ .  $\square$

Dual bases provide a third way of representing upward-closed subsets, by representing their complements. This deserves a definition:

**Definition 6 (Finite Dual Basis Representation).** *Let  $X$  be a strongly effective well poset, with complete representation map  $r' : E' \rightarrow \widehat{X}$ . A finite dual basis representation*

for an upward-closed subset  $U$  of  $X$  is any finite subset  $A = \{e'_1, \dots, e'_n\}$  of  $E'$  such that  $X \setminus U = r'(e'_1) \cup \dots \cup r'(e'_n)$ .

Proposition 4.2 of [4] (already cited above) states that every downward-closed subset of  $\widehat{X}$  is a finite union  $C_1 \cup \dots \cup C_n$  of elements  $C_1, \dots, C_n$  of  $\widehat{X}$ , as soon as  $X$  is well-quasi-ordered. The fact that  $r' : E' \rightarrow \widehat{X}$  is surjective entails that we can write each  $C_i$  as  $r'(e'_i)$  for some  $e'_i \in E'$ . So every upward-closed subset  $U$  of  $X$  has a finite dual basis representation (and, of course, every set that has a finite dual basis representation is upward-closed).

The effective complement property states that we can *compute* a finite dual basis representation from any finite basis representation of an upward-closed set.

We now claim:

**Theorem 1 (Valk-Jantzen, Generalized).** *Let  $X$  be a strongly effective well poset with the effective complement property. Then there are Turing machines that convert from (minimal) finite basis representations to strong oracle representations of upward-closed subsets and back.*

More informally, this states the exact analogue of the Valk-Jantzen Theorem on strongly effective well posets with the effective complement property: for any upward-closed subset  $U$  of  $X$ , we can compute a (minimal) finite basis  $\{x_1, \dots, x_m\}$  if and only if, for any  $z \in \widehat{X}$ , we can decide whether  $U \cap \downarrow_{\widehat{X}} z$  is empty.

*Proof of Theorem 1.* For any representations  $A, B$  of upward-closed subsets, write  $A \Rightarrow B$  for the property: “there is a Turing machine that, on input any  $A$ -representation of an upward-closed subset  $U$  of  $X$ , computes a  $B$ -representation of  $U$ .” The effective complement property is Finite Basis  $\Rightarrow$  Finite Dual Basis.

Assuming  $X$  to be a strongly effective well poset, we shall prove:

- Finite Basis  $\Rightarrow$  Strong Oracle.
- if Finite Basis  $\Rightarrow$  Finite Dual Basis, then Strong Oracle  $\Rightarrow$  Finite Basis.

**Proposition 2.** *Let  $X$  be a strongly effective well poset. Then Finite Basis  $\Rightarrow$  Strong Oracle.*

*Proof.* Intuitively, given  $U = \uparrow\{x_1, \dots, x_n\} \subseteq X$ , the oracle checks whether its input  $x \in \widehat{X}$  is such that  $x_i \leq x$  for some  $i, 1 \leq i \leq n$ .

More formally, let  $r : E \rightarrow X$  be the subspace representation map,  $r' : E' \rightarrow \widehat{X}$  the complete representation map. Let also  $\sqsubseteq$  be the relation such that, for all  $e \in E, e' \in E', e \sqsubseteq e'$  iff  $r(e) \in r'(e')$ .

From a finite basis representation  $\{e_1, \dots, e_n\} \subseteq E$  we compute the code of the following Turing machine  $\mathcal{M}$ . On input  $e' \in E', \mathcal{M}$  must check whether  $\uparrow\{r(e_1), \dots, r(e_n)\} \cap r'(e') \neq \emptyset$ . We claim that this is simply done by checking whether  $e_i \sqsubseteq e'$  for some  $i, 1 \leq i \leq n$ . Indeed,  $\uparrow\{r(e_1), \dots, r(e_n)\} \cap r'(e') \neq \emptyset$  iff  $r(e_i) \in r'(e')$  for some  $i$ , since  $r'(e')$ , as an ideal, is downward-closed.  $\square$

We now show that Strong Oracle  $\Rightarrow$  Finite Basis under the effective complement property. We first make Remark 1 formal.

**Lemma 2.** *Let  $X$  be a strongly effective well poset. From any strong oracle representation, one can compute a weak oracle representation of the same upward-closed subset.*

*Proof.* Let  $r : E \rightarrow X$  be the subspace representation map,  $r' : E' \rightarrow \widehat{X}$  the complete representation map. Given any strong oracle representation  $\mathcal{M}$  for an upward-closed subset  $U$ , we compute a weak oracle as follows. The weak oracle takes  $e \in E$ , and checks whether  $r(e) \in U$ , equivalently, whether  $\downarrow r(e)$  intersects  $U$ . Since  $\downarrow r(e) = \eta(r(e)) = r'(e)$ , this amounts to checking whether  $U \cap r'(e) \neq \emptyset$ : then use  $\mathcal{M}$ .  $\square$

**Proposition 3.** *Let  $X$  be a strongly effective well poset. If  $X$  has the effective complement property, then Strong Oracle  $\Leftrightarrow$  Finite Basis.*

*Proof.* Let  $r : E \rightarrow X$  be the subspace representation map,  $r' : E' \rightarrow \widehat{X}$  the complete representation map. We build a Turing machine  $\mathcal{M}$  that takes a strong oracle representation of an upward-closed subset  $U$  of  $X$ , and outputs a finite basis of  $U$ .

Since  $E$  is r.e., enumerate its elements  $e$  (i.e., enumerate the codes of elements of  $X$ ), and test whether  $r(e) \in U$ . This can be done by a weak oracle representation for  $U$ , deduced from the strong oracle (Lemma 2). We can therefore enumerate the codes  $e_1, e_2, \dots$ , of elements of  $U$ . Eventually, all the minimal elements, i.e., the elements of the minimal finite basis, of  $U$  will have been enumerated. So there is an integer  $m$  such that  $U = \uparrow\{r(e_1), \dots, r(e_m)\}$ . To complete the description of  $\mathcal{M}$ , we need to detect when this happens: i.e., for each  $m \in \mathbb{N}$ ,  $\mathcal{M}$  checks whether  $U \subseteq \uparrow\{r(e_1), \dots, r(e_m)\}$ ; if so, it stops and outputs the finite basis  $\{e_1, \dots, e_m\}$ , otherwise it goes on. (This finite basis need not be minimal, however we can convert it to a minimal one, see Remark 2.)

So we need to show that we can decide whether  $U \subseteq \uparrow\{r(e_1), \dots, r(e_m)\}$ . By the effective complement property (Definition 5), we compute a finite set of codes  $e'_1, \dots, e'_n \in E'$  such that  $X \setminus \uparrow\{r(e_1), \dots, r(e_m)\} = r'(e'_1) \cup \dots \cup r'(e'_n)$ . Now  $U \subseteq \uparrow\{r(e_1), \dots, r(e_m)\}$  iff  $U$  does not intersect  $r'(e'_1) \cup \dots \cup r'(e'_n)$ , iff  $U \cap r'(e'_1) = \emptyset$  and  $\dots$  and  $U \cap r'(e'_n) = \emptyset$ . These conditions can be checked by calling the strong oracle on  $e'_1, \dots, e'_n$ .  $\square$

*Other Results.* We explore which other relations  $A \Leftrightarrow B$  between representations do hold.

**Corollary 1.** *Let  $X$  be a strongly effective well poset. If  $X$  has the effective complement property, then Strong Oracle  $\Leftrightarrow$  Finite Dual Basis.*

*Proof.* We obtain Strong Oracle  $\Leftrightarrow$  Finite Basis  $\Leftrightarrow$  Finite Dual Basis: the second  $\Leftrightarrow$  is the effective complement property, and the first one is Proposition 3.  $\square$

**Lemma 3.** *Let  $X$  be a strongly effective well poset. Then Finite Dual Basis  $\Leftrightarrow$  Strong Oracle.*

*Proof.* Intuitively, because  $(X \setminus \downarrow_{\widehat{X}}\{z_1, \dots, z_n\}) \cap \downarrow_{\widehat{X}} z = \emptyset$  iff  $\downarrow_{\widehat{X}} z \subseteq \downarrow_{\widehat{X}}\{z_1, \dots, z_n\}$ , iff  $z \leq_{\widehat{X}} z_i$  for some  $i, 1 \leq i \leq n$ . The details are left to the reader.  $\square$

Combining this with Proposition 3, we obtain:

**Corollary 2.** *Let  $X$  be a strongly effective well poset. If  $X$  has the effective complement property, i.e., if Finite Basis  $\Leftrightarrow$  Dual Finite Basis, then Dual Finite Basis  $\Leftrightarrow$  Finite Basis.*

## 4 Special Cases

$\mathbb{N}^k$ . Examples 1, 3, and 5 show that the assumptions of Theorem 1 are satisfied for  $X = \mathbb{N}^k$ . Then Theorem 1 specializes to the original Valk-Jantzen Theorem.

$\Sigma^*$ . Examples 2, 4, and 7 show that the assumptions of Theorem 1 are satisfied for  $X = \Sigma^*$ . Then Theorem 1 specializes to the following new result, which we phrase in the slightly more informal style of the original Valk-Jantzen Theorem:

**Theorem 2.** *Let  $\Sigma$  be a finite alphabet, and order  $\Sigma^*$  by the subword ordering. For any upward-closed subset  $U$  of  $\Sigma^*$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap \llbracket P \rrbracket_{\Sigma^*}$  is empty, for every  $*$ -product  $P$ .*

Here  $\llbracket P \rrbracket_{\Sigma^*}$  is the language of the product  $P$ . Recall that  $*$ -products are concatenations of atomic expressions  $a^?$  ( $a \in \Sigma$ ) and  $A^*$  ( $A$  a non-empty subset of  $\Sigma^*$ ).

*Other.* We generalize the above constructions to any space  $X$  that can be constructed from the hierarchy of data types of [4, Section 5]. Such data types  $D$  are given by the following grammar:

|                               |                                    |
|-------------------------------|------------------------------------|
| $D ::= \mathbb{N}$            | natural numbers                    |
| $A_{\leq}$                    | finite set $A$ , ordered by $\leq$ |
| $D_1 \times \dots \times D_k$ | finite product                     |
| $D_1 + \dots + D_k$           | finite, disjoint sum               |
| $D^*$                         | finite words                       |
| $D^{\otimes}$                 | finite multisets                   |

Any data type  $D$  denotes a well poset  $\llbracket D \rrbracket$ , with an ordering we call  $\leq_D$ , defined by:  $\llbracket \mathbb{N} \rrbracket$  is the set of natural numbers with its usual ordering,  $\llbracket A_{\leq} \rrbracket$  is the finite set  $A$ , ordered by  $\leq$ . (The data types of [4] only require  $\leq$  to be a quasi-ordering, however we need it to be a partial ordering for  $A$  to have strong oracles at all.) Then,  $\llbracket D_1 \times \dots \times D_k \rrbracket = \llbracket D_1 \rrbracket \times \dots \times \llbracket D_k \rrbracket$ , with the componentwise ordering,  $\llbracket D_1 + \dots + D_k \rrbracket$  is the disjoint sum of  $\llbracket D_1 \rrbracket, \dots, \llbracket D_k \rrbracket$ , with the ordering induced by each summand.

$\llbracket D^* \rrbracket$  is the set of finite words over  $\llbracket D \rrbracket$ . When  $D$  is a finite set  $A_{\leq}$ , with equality as ordering  $\leq$ ,  $\llbracket D^* \rrbracket$  will simply be the set of words over  $A$ , with the subword ordering, which we have dealt with in Theorem 2. In general,  $\llbracket D \rrbracket$  is not a finite alphabet, and is usually equipped with a non-trivial ordering  $\leq_D$ . The partial ordering  $\leq_{D^*}$  on  $\llbracket D^* \rrbracket$  is the *embedding* ordering defined by:  $w \leq_{D^*} w'$  iff, writing  $w$  as the sequence of  $m$  letters  $a_1 a_2 \dots a_m$  from  $\llbracket D \rrbracket$ , one can write  $w'$  as  $w_0 a'_1 w_1 a'_2 w_2 \dots w_{m-1} a'_m w'_m$  with  $a_1 \leq_D a'_1, a_2 \leq_D a'_2, \dots, a_m \leq_D a'_m$ . This is a well-quasi-ordering, as soon as  $\leq_D$  is, by Higman's Lemma.

$\llbracket D^{\otimes} \rrbracket$  is the set of finite multisets  $\{x_1, \dots, x_n\}$  of elements of  $\llbracket D \rrbracket$ , and is partially ordered by  $\leq_{D^{\otimes}}$ , defined as:  $\{x_1, x_2, \dots, x_m\} \leq_{D^{\otimes}} \{y_1, y_2, \dots, y_n\}$  iff there is an injective map  $r : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  such that  $x_i \leq_D y_{r(i)}$  for all  $i, 1 \leq i \leq m$ .

We recall from [4, Theorem 5.3] that  $\widehat{\llbracket D \rrbracket}$  is characterized, computationally (and up to order-isomorphism) for each data type  $D$  as follows. We shall always describe  $\widehat{\llbracket D \rrbracket}$  through a simpler order-isomorphic space, which we shall call  $\widehat{D}$ . We also describe the order-isomorphism  $z \in \widehat{D} \mapsto \llbracket z \rrbracket_D \in \widehat{\llbracket D \rrbracket}$  in each case; note that, in particular,  $\llbracket z \rrbracket_D$  will be an ideal in, hence a set of elements of  $\llbracket D \rrbracket$ .

- $\widehat{\mathbb{N}}$  is  $\mathbb{N}_\omega$ , with the usual ordering; the order-isomorphism is given by  $\llbracket n \rrbracket_{\widehat{\mathbb{N}}} = \downarrow_{\mathbb{N}} n$  for each  $n \in \mathbb{N}$ , and  $\llbracket \omega \rrbracket_{\widehat{\mathbb{N}}} = \mathbb{N}$ .
- $\widehat{A}_{\leq}$ , for finite  $A$ , is  $A$  itself, with  $\leq$  as ordering;  $\llbracket a \rrbracket_{\widehat{A}_{\leq}} = \downarrow_A a$  for all  $a \in A$ .
- When  $D = D_1 \times \dots \times D_k$ ,  $\widehat{D}$  is the product of  $\widehat{D}_1, \dots, \widehat{D}_k$ , with the product ordering; for all  $z_1 \in \widehat{D}_1, \dots, z_k \in \widehat{D}_k$ ,  $\llbracket (z_1, \dots, z_k) \rrbracket_{\widehat{D}_1 \times \dots \times \widehat{D}_k} = \llbracket z_1 \rrbracket_{\widehat{D}_1} \times \dots \times \llbracket z_k \rrbracket_{\widehat{D}_k}$ .
- when  $D = D_1 + \dots + D_k$ ,  $\widehat{D}$  is the disjoint union of  $\widehat{D}_1, \dots, \widehat{D}_k$ , with the induced ordering; for each  $i, 1 \leq i \leq k$ , for every  $z \in \widehat{D}_i$ ,  $\llbracket z \rrbracket_{\widehat{D}_1 + \dots + \widehat{D}_k} = \llbracket z \rrbracket_{\widehat{D}_i}$ .
- $\widehat{D}^*$  is the set of  $*$ -products, defined again as sequences  $P$  of atomic expressions  $a^?$  (with  $a \in \widehat{D}$ ) or  $A^*$  (with  $A$  a non-empty finite subset of  $\widehat{D}$ ). Atomic expressions are ordered by  $a^? \leq_{D^*} a'^?$  iff  $a \leq_D a'$ ,  $a^? \leq_{D^*} A'^*$  iff  $a \leq_D a'$  for some  $a' \in A'$ ,  $A^? \leq_{D^*} A'^?$  iff for every  $a \in A$ , there is an  $a' \in A'$  such that  $a \leq_D a'$ . This is extended to  $*$ -products by:  $\epsilon \leq_{D^*} P'$  always;  $P \leq_{D^*} \epsilon$  iff  $P = \epsilon$ ;  $eP \leq_{D^*} e'P'$  (where  $e, e'$  are atomic expressions,  $P, P'$  are  $*$ -products) iff (1)  $e \not\leq_D e'$  and  $NP \leq_{D^*} P'$ , or (2)  $e = a^?, e' = a'^?, a \leq_D a'$ , and  $P \leq_{D^*} P'$ , or (3)  $e' = A'^*, e \leq_{D^*} A'^*$  and  $P \leq_{D^*} e'P'$ .

The order-isomorphism is given by: for each  $*$ -product  $P = e_1 e_2 \dots e_n$ , where  $e_1, e_2, \dots, e_n$  are atomic expressions,  $\llbracket P \rrbracket_{D^*} = \{w_1 w_2 \dots w_n \mid w_1 \in \llbracket e_1 \rrbracket_{D^*}, \dots, w_n \in \llbracket e_n \rrbracket_{D^*}\}$ , where  $\llbracket e \rrbracket_{D^*}$  is defined for atomic expressions  $e$  by:  $\llbracket a^? \rrbracket_{D^*} = \{\epsilon\} \cup \llbracket a \rrbracket_D$  (where in  $\llbracket a \rrbracket_D$ , we equate letters in  $\llbracket D \rrbracket$  with one-letter words in  $\llbracket D^* \rrbracket$ ), and  $\llbracket A^* \rrbracket_{D^*}$  is the set of words over the alphabet  $\bigcup_{a \in A} \llbracket a \rrbracket_D$ .

- $\widehat{D}^{\otimes}$  is the set of  $\otimes$ -products, defined as expressions of the form  $A^{\otimes} \odot a_1^{\otimes} \odot \dots \odot a_n^{\otimes}$ , where  $A$  is a finite subset of  $\widehat{D}$  (possibly empty),  $n \in \mathbb{N}$ , and  $a_1, \dots, a_n \in \widehat{D}$ . They are ordered by  $P \leq_{D^{\otimes}} P'$ , where  $P = A^{\otimes} \odot a_1^{\otimes} \odot a_2^{\otimes} \odot \dots \odot a_m^{\otimes}$  and  $P' = A'^{\otimes} \odot a_1'^{\otimes} \odot a_2'^{\otimes} \odot \dots \odot a_n'^{\otimes}$ , iff: (1) for every  $a \in A$ , there is an  $a' \in A'$  with  $a \leq_D a'$ , and (2) letting  $I$  be the subset of those indices  $i, 1 \leq i \leq m$ , such that  $a_i \leq_D a'$  for no  $a' \in A'$ , there is an injective map  $r : I \rightarrow \{1, \dots, n\}$  such that  $a_i \leq_D a'_{r(i)}$  for all  $i \in I$ .

The order-isomorphism is given by:  $\llbracket A^{\otimes} \odot a_1^{\otimes} \odot \dots \odot a_n^{\otimes} \rrbracket$  is the set of all multisets that can be written as an arbitrary multiset of elements of  $\bigcup_{a \in A} \llbracket a \rrbracket_D$ , plus at most one element of  $\llbracket a_1 \rrbracket_D$ , at most one element of  $\llbracket a_2 \rrbracket_D$ ,  $\dots$ , and at most one element of  $\llbracket a_n \rrbracket_D$ .

For any data type  $D$ ,  $\llbracket D \rrbracket$  is well-quasi-ordered, and partially ordered by  $\leq_D$ . The definition above makes it clear that  $\widehat{D}$  is a strongly effective well poset. In fact, the set of elements of  $\llbracket D \rrbracket$  and of  $\widehat{D}$  are not just r.e., but recursive.

We proceed to show that  $\llbracket D \rrbracket$  has the effective complement property for all data types  $D$  as above. We first note:

*Remark 4.* For every data type  $D$ ,  $\llbracket D \rrbracket$  has a finite dual basis.

This is because every downward-closed subset of  $\llbracket D \rrbracket$  has one, and  $\llbracket D \rrbracket$  is certainly downward-closed in itself. Concretely, we can build one by induction on  $D$ . A dual basis of  $\llbracket \mathbb{N} \rrbracket$  consists in the single element  $\omega$ , a dual basis of  $\llbracket A_{\leq} \rrbracket$  is given by the maximal elements of  $A$ , ordered by  $\leq$ . A dual basis of  $\llbracket D_1 \times \dots \times D_k \rrbracket$  is given by

all  $k$ -tuples  $(z_1, \dots, z_k)$ , where each  $z_i$  is taken from a dual basis of  $\llbracket D \rrbracket_i$ ,  $1 \leq i \leq k$ . A dual basis of  $\llbracket D_1 + \dots + D_k \rrbracket$  is given by the disjoint union of dual bases of each  $\llbracket D_i \rrbracket$ ,  $1 \leq i \leq k$ . A dual basis of  $\llbracket D^* \rrbracket$  (resp.  $\llbracket D^\circledast \rrbracket$ ) is given by the single  $*$ -product  $A^*$  (resp.,  $A^\circledast$ ), where  $A$  is a finite dual basis of  $\llbracket D \rrbracket$ .

Second, we note that we can compute binary intersections of elements of  $\llbracket D \rrbracket$ , whence we shall be able to compute any finite intersection of such elements. We leave representation functions implicit. Accordingly, the finite set  $\{z_1, \dots, z_n\} \subseteq \widehat{D}$  is a dual basis of a downward-closed subset  $F$  of  $\llbracket D \rrbracket$  iff  $F = \bigcup_{i=1}^n \llbracket z_i \rrbracket_D$  (compare with Definition 6).

**Lemma 4.** *For every data type  $D$ , one can compute a finite dual basis of the intersection  $\llbracket z \rrbracket_D \cap \llbracket z' \rrbracket_D$ , from any  $z, z' \in \widehat{D}$ .*

*Proof.* First,  $\llbracket z \rrbracket_D \cap \llbracket z' \rrbracket_D$  always has such a finite dual basis, since it is downward-closed, as the intersection of two downward-closed subsets of  $\llbracket D \rrbracket$ . That it is computable needs a separate argument. We induct on the structure of  $D$ .

On  $\mathbb{N}$ ,  $\llbracket z \rrbracket_{\mathbb{N}} \cap \llbracket z' \rrbracket_{\mathbb{N}} = \llbracket z' \rrbracket_{\mathbb{N}}$ ,  $\llbracket z \rrbracket_{\mathbb{N}} \cap \llbracket \omega \rrbracket_{\mathbb{N}} = \llbracket z \rrbracket_{\mathbb{N}}$ , and  $\llbracket n \rrbracket_{\mathbb{N}} \cap \llbracket n' \rrbracket_{\mathbb{N}} = \llbracket \min(n, n') \rrbracket_{\mathbb{N}}$  for all  $n, n' \in \mathbb{N}$ .

On  $A_{\leq}$  ( $A$  finite),  $\llbracket z \rrbracket_{A_{\leq}} \cap \llbracket z' \rrbracket_{A_{\leq}}$  is computed by enumerating its elements, and e.g., keeping only the maximal ones. (Moreover, this result can be tabulated in terms of  $z, z'$ .)

If  $D = D_1 \times \dots \times D_k$ , then  $z$  can be written as a tuple  $(z_1, \dots, z_k)$  and  $z'$  as  $(z'_1, \dots, z'_k)$ , so that  $\llbracket z \rrbracket_D \cap \llbracket z' \rrbracket_D = \prod_{j=1}^k (\llbracket z_j \rrbracket_{D_j} \cap \llbracket z'_j \rrbracket_{D_j})$ . Using recursion, compute a finite dual basis  $z''_{ji}$ ,  $1 \leq i \leq n_j$  for each  $\llbracket z_j \rrbracket_{D_j} \cap \llbracket z'_j \rrbracket_{D_j}$ ,  $1 \leq j \leq k$ . The desired dual basis is obtained by distributing unions over products, as the collection of all tuples  $(z''_{1i_1}, \dots, z''_{ki_k})$ , with  $1 \leq i_1 \leq n_1, \dots, 1 \leq i_k \leq n_k$ .

If  $D = D_1 + \dots + D_k$ , then  $z$  is in some  $D_j$  and  $z'$  in some  $D_{j'}$ ,  $1 \leq j, j' \leq k$ . If  $j \neq j'$ , then compute  $\llbracket z \rrbracket_D \cap \llbracket z' \rrbracket_D$  as  $\emptyset$ ; otherwise, this is computed by a recursive call on  $z, z' \in D_j$ .

On the data type  $D^*$ ,  $z$  and  $z'$  must be  $*$ -products. Note that:  $(*)$  for every finite subset  $A$  of  $\llbracket D \rrbracket$ , letting  $\llbracket A \rrbracket_D = \bigcup_{a \in A} \llbracket a \rrbracket_D$ , one can compute a finite dual basis of  $\llbracket A \rrbracket_D \cap \llbracket a' \rrbracket_D$  for any  $a' \in \llbracket D \rrbracket$ , or of  $\llbracket A \rrbracket_D \cap \llbracket A' \rrbracket_D$  for any finite subset  $A'$  of  $\llbracket D \rrbracket$ , by induction hypothesis and by distributing unions over intersections. The desired intersection  $\llbracket z \rrbracket_{D^*} \cap \llbracket z' \rrbracket_{D^*}$  is computed by using the following equations, which can either be proved directly, or derived from [4, Lemma E.14, long version, available on the Web]. These are then used to define a procedure that recurses on the length of  $z, z'$  first, and on  $D$  second.

- $\llbracket \epsilon \rrbracket_{D^*} \cap \llbracket P \rrbracket_{D^*} = \llbracket P \rrbracket_{D^*} \cap \llbracket \epsilon \rrbracket_{D^*} = \llbracket \epsilon \rrbracket_{D^*}$  for every  $*$ -product  $P$ .
- $\llbracket a^? P \rrbracket_{D^*} \cap \llbracket a'^? P' \rrbracket_{D^*}$  is the union of: (1) the union over all elements  $a''$  of a finite dual basis of  $\llbracket a \rrbracket_D \cap \llbracket a' \rrbracket_D$  (recursing on  $D$ ) and over all elements  $P''$  of a finite dual basis of  $\llbracket P \rrbracket_{D^*} \cap \llbracket P' \rrbracket_{D^*}$  of  $\llbracket (a'')^? P'' \rrbracket_{D^*}$ ; (2)  $\llbracket a^? P \rrbracket_{D^*} \cap \llbracket P' \rrbracket_{D^*}$ ; and (3)  $\llbracket P \rrbracket_{D^*} \cap \llbracket (a')^? P' \rrbracket_{D^*}$ .
- $\llbracket a^? P \rrbracket_{D^*} \cap \llbracket A'^* P' \rrbracket_{D^*}$  is the union of: (1) the union over all elements  $a''$  of a finite dual basis of  $\llbracket a \rrbracket_D \cap \llbracket A' \rrbracket_D$  (using  $(*)$ ) and over all elements  $P''$  of a finite dual basis of  $\llbracket P \rrbracket_{D^*} \cap \llbracket A'^* P' \rrbracket_{D^*}$  of  $\llbracket (a'')^? P'' \rrbracket_{D^*}$ ; and (2)  $\llbracket a^? P \rrbracket_{D^*} \cap \llbracket P' \rrbracket_{D^*}$ .



- $\llbracket A^*P \rrbracket_{D^*} \cap \llbracket A'^*P' \rrbracket_{D^*}$  is the union over all elements  $a''$  of a finite dual basis of  $\llbracket A \rrbracket_D \cap \llbracket A' \rrbracket_D$  (using  $(*)$  again) of: (1) the union over all elements  $P''$  of a finite dual basis of  $\llbracket A^*P \rrbracket_{D^*} \cap \llbracket P' \rrbracket_{D^*}$  of  $\llbracket a''^? P'' \rrbracket_{D^*}$ , and (2) the union over all elements  $P''$  of a finite dual basis of  $\llbracket P \rrbracket_{D^*} \cap \llbracket A'^*P' \rrbracket_{D^*}$  of  $\llbracket a''^? P'' \rrbracket_{D^*}$ .

On the data type  $D^\otimes$ ,  $z$  and  $z'$  must be  $\otimes$ -products. Let  $z = A^\otimes \odot a_1^\otimes \odot \dots \odot a_m^\otimes$ ,  $z' = A'^\otimes \odot a'_1 \odot \dots \odot a'_n$ . Call a *partial injection*  $f : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  any bijection from some subset of  $\{1, \dots, m\}$  (the *domain*  $\text{dom } f$ ) to some subset of  $\{1, \dots, n\}$  (the *codomain*  $\text{cod } f$ ). Then a finite dual basis of  $\llbracket z \rrbracket_{D^\otimes} \cap \llbracket z' \rrbracket_{D^\otimes}$  is given by enumerating all partial injections  $f : \{1, \dots, m\} \rightarrow \{1, \dots, n\}$  (let  $k$  below be the cardinality of its domain), and for each one, outputting an expression of the form  $A''^\otimes \odot b_1^\otimes \odot \dots \odot b_{m-k}^\otimes \odot c_1^\otimes \odot \dots \odot c_{n-k}^\otimes \odot d_1^\otimes \odot \dots \odot d_k^\otimes$ , where:

1.  $A''$  is a finite dual basis of  $\llbracket A \rrbracket_D \cap \llbracket A' \rrbracket_D$ ;
2.  $b_1, \dots, b_{m-k}$  are each taken from finite dual bases of the  $m-k$  downward-closed subsets  $\llbracket a_i \rrbracket_D \cap \llbracket A' \rrbracket_D, i \in \{1, \dots, m\} \setminus \text{dom } f$ ;
3.  $c_1, \dots, c_{n-k}$  are each taken from finite dual bases of the  $n-k$  downward-closed subsets  $\llbracket A \rrbracket_D \cap \llbracket a'_j \rrbracket_D, j \in \{1, \dots, n\} \setminus \text{cod } f$ ;
4.  $d_1, \dots, d_k$  are each taken from finite dual bases of the  $k$  downward-closed subsets  $\llbracket a_i \rrbracket_D \cap \llbracket a'_{f(i)} \rrbracket_D, i \in \text{dom } f$ .

Indeed, any multiset in  $\llbracket z \rrbracket_{D^\otimes} \cap \llbracket z' \rrbracket_{D^\otimes}$  can be split as  $m_0 \uplus m_1 \uplus m_2 \uplus m_3$ , where  $m_0$  is a multiset of elements in  $\llbracket A \rrbracket_D \cap \llbracket A' \rrbracket_D$ ,  $m_1$  is a multiset of elements in some  $\llbracket a_i \rrbracket_D \cap \llbracket A' \rrbracket_D, 1 \leq i \leq m$ ,  $m_2$  is a multiset of elements in some  $\llbracket A \rrbracket_D \cap \llbracket a'_j \rrbracket_D, 1 \leq j \leq n$ , and finally  $m_3$  is a multiset of elements in some  $\llbracket a_i \rrbracket_D \cap \llbracket a'_j \rrbracket_D$ , for some  $i, j$  with  $1 \leq i \leq m, 1 \leq j \leq n$ . Pairing the latter  $i$ s with the corresponding  $j$ s, we obtain a partial injection  $f$ , with domain  $\text{dom } f = \{i_1, \dots, i_k\}$ , so that  $m_3$  is a multiset with  $k$  elements taken from  $\llbracket a_{i_1} \rrbracket_D \cap \llbracket a'_{f(i_1)} \rrbracket_D, \dots, \llbracket a_{i_k} \rrbracket_D \cap \llbracket a'_{f(i_k)} \rrbracket_D$  respectively: so  $m_3 \in \llbracket d_1^\otimes \odot \dots \odot d_k^\otimes \rrbracket_{D^\otimes}$  for some  $d_1, \dots, d_k$  as in item 4 above. Similarly, the at most  $m-k$  elements of  $m_1$  are in some  $\llbracket b_1^\otimes \odot \dots \odot b_{m-k}^\otimes \rrbracket_{D^\otimes}$ , for some  $b_1, \dots, b_{m-k}$  as in item 2 above; and  $m_2$  is in  $\llbracket c_1^\otimes \odot \dots \odot c_{n-k}^\otimes \rrbracket_{D^\otimes}$  for some  $c_1, \dots, c_{n-k}$  as in item 3 above. Finally,  $m_0$  is in  $\llbracket A \rrbracket_D \cap \llbracket A' \rrbracket_D$ , hence in  $\llbracket A'' \rrbracket_D$ , matching item 1.

The converse inclusion is easy, and left to the reader.  $\square$

**Lemma 5.** *For every data type  $D$ , one can compute a finite dual basis representation of  $\uparrow x$  for every  $x \in \llbracket D \rrbracket$ .*

*Proof.* On  $\mathbb{N}$ , the complement of  $\uparrow n$  in  $\mathbb{N}$  is computed as  $\mathbb{N} \cap \downarrow_{\mathbb{N}_\omega}(n-1)$  if  $n \neq 0$ , and as  $\emptyset$  if  $n = 0$ .

Given any finite set  $A$ , partially ordered by  $\leq$ , it is trivial to compute a dual basis of any downward-closed subset of  $A$ ; in fact one can just tabulate them all.

If  $D = D_1 \times \dots \times D_k$ , one can compute the complement of  $\uparrow(x_1, \dots, x_k)$  as the union of  $E_i = \llbracket D_1 \rrbracket \times \dots \times \llbracket D_{i-1} \rrbracket \times (\llbracket D_i \rrbracket \setminus \uparrow x_i) \times \llbracket D_{i+1} \rrbracket \times \dots \times \llbracket D_k \rrbracket$  over  $i, 1 \leq i \leq k$ . In turn, each  $\llbracket D_j \rrbracket$  ( $j \neq i$ ) has a finite dual basis by Remark 4, and one can compute a finite dual basis of  $\llbracket D_i \rrbracket \setminus \uparrow x_i$  by induction hypothesis. One

gets a finite dual basis of  $E_i$  by distributing unions over products, i.e., as all tuples  $(z_1, \dots, z_{i-1}, z_i, z_{i+1}, \dots, z_k)$ , where  $z_j$  is taken from a finite dual basis of  $\llbracket D_j \rrbracket$  (for each  $j \neq i$ ) and  $z_i$  is taken from a finite dual basis of  $\llbracket D_i \rrbracket \setminus \uparrow x_i$ .

If  $D = D_1 + \dots + D_k$ , one can compute the complement of  $\uparrow x_i$  (where  $x_i \in \llbracket D_i \rrbracket$ ) as the disjoint union of all  $\llbracket D_j \rrbracket$ ,  $j \neq i$ , with the complement (in  $\llbracket D_i \rrbracket$ ) of  $\uparrow x_i$ . This is done using Remark 4 and the induction hypothesis, and is left to the reader.

On  $D^*$ , one is led to compute the complement of  $\uparrow w$ , for some word  $w = a_1 a_2 \dots a_n$ . When  $n = 0$ , this is  $\emptyset$ . Otherwise, this is the complement of the language  $\llbracket D^* \rrbracket (\uparrow a_1) \llbracket D^* \rrbracket (\uparrow a_2) \llbracket D^* \rrbracket \dots \llbracket D^* \rrbracket (\uparrow a_n) \llbracket D^* \rrbracket$ , and is given by [4, Lemma E.5, long version, available on the Web], as the union of all  $\llbracket A_1^* a_1' A_2^* a_2' \dots a_{n-1}' A_n^* \rrbracket_{D^*}$ , where for each  $i$ ,  $1 \leq i \leq n$ ,  $A_i$  is a finite dual basis of the complement of  $\uparrow a_i$ , obtained by induction, and  $a_1', a_2', \dots, a_{n-1}'$  range over  $(n-1)$ -tuples of elements of a finite dual basis of  $\llbracket D \rrbracket$  itself, using Remark 4.

On  $D^\otimes$ , one must compute the complement of  $\uparrow m$ , for some multiset  $m = \{x_1, \dots, x_n\}$ . We claim that this is equal to the union over all proper subsets  $I$  of  $\{1, \dots, n\}$ , say of cardinality  $k$ , of  $\llbracket A_I^\otimes \odot D^\otimes \odot \dots \odot D^\otimes \rrbracket_{D^\otimes}$ , where there are  $k$  copies of  $D^\otimes$ , and  $A_I$  is a finite dual basis of  $\downarrow \{x_i \mid i \in \{1, \dots, n\} \setminus I\}$ . Therefore, a finite dual basis of the complement of  $\uparrow m$  is given as the union over all such  $I$ , and over  $k$ -tuples  $a_1', \dots, a_k'$  of elements of a finite dual basis of  $\llbracket D \rrbracket$ , of  $\llbracket A_I^\otimes \odot a_1'^\otimes \odot \dots \odot a_k'^\otimes \rrbracket_{D^\otimes}$ .

Indeed, write  $\sqsubseteq$  the ordering on multisets defined by  $\{a_1, \dots, a_k\} \sqsubseteq \{b_1, \dots, b_\ell\}$  iff  $k = \ell$ , and there is a permutation  $\pi$  of  $\{1, \dots, k\}$  such that  $a_i \leq_D b_{\pi(i)}$  for all  $i$ ,  $1 \leq i \leq k$ . So  $m \leq_{D^\otimes} m'$  iff  $m'$  can be split as  $m'_0 \uplus m'_1$  with  $m \sqsubseteq m'_0$ . Any multiset  $m'$  can be split as  $m'_0 \uplus m'_1$ , and correspondingly  $m$  can be split as  $m_0 \uplus m_1$ , in such a way that  $m_0 \sqsubseteq m'_0$ , and the cardinality  $k$  of  $m_0$  (and of  $m'_0$ ) is maximal. Now assume  $m'$  is in the complement of  $\uparrow m$ , and let  $I$  be a subset of those indices  $i \in \{1, \dots, n\}$  chosen so that  $m_0 = \{x_i \mid i \in I\}$ . Since  $m' \not\sqsubseteq \uparrow m$ ,  $I$  is a proper subset of  $\{1, \dots, n\}$ . By maximality, for every  $x' \in m'_1$ , there can be no  $x \in m_1$  with  $x \leq x'$ , so  $x'$  is in the complement of  $\llbracket A_I \rrbracket$ . It follows that  $m'$  is in  $\llbracket A_I^\otimes \odot D^\otimes \odot \dots \odot D^\otimes \rrbracket_{D^\otimes}$ , where there are  $k$  copies of  $D^\otimes$ . The converse inclusion is obvious.  $\square$

**Proposition 4.** *For every data type  $D$ ,  $\llbracket D \rrbracket$  has the effective complement property.*

*Proof.* One can compute the complement of  $\uparrow \{x_1, \dots, x_m\}$  as the intersection of the complements of  $\uparrow x_i$ ,  $1 \leq i \leq m$ . The latter are computed by Lemma 5. Finite intersections are computed by Remark 4 (if  $m = 0$ ) or by iterating Lemma 4 (if  $m \geq 1$ ).  $\square$

From Theorem 1, it follows immediately:

**Theorem 3.** *For every data type  $D$ , and any upward-closed subset  $U$  of  $\llbracket D \rrbracket$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap \llbracket z \rrbracket_D$  is empty, for every  $z \in \widehat{D}$ .*

When  $D$  is  $\mathbb{N}^k$ , i.e., the  $k$ -fold product of  $\mathbb{N}$ , Theorem 3 specializes to the original Valk-Jantzen Theorem. When  $D = \Sigma^*$ , where  $\Sigma$  is any finite set with the trivial ordering (i.e., equality), one retrieves Theorem 2.

Here are a few other special cases. The (normal) *multiset language generators* of [2, Section 5] are elements of  $\llbracket \Sigma^\otimes \rrbracket$ , where  $\Sigma$  is some finite alphabet (with the trivial ordering  $\Rightarrow$ ). These are used in the study of timed Petri nets.

**Theorem 4 (Multiset Language Generators).** *Let  $\Sigma$  be a finite alphabet. For any upward-closed subset  $U$  of  $[[\Sigma^{\otimes}]]$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap [[A^{\otimes} \odot a_1^{\otimes} \odot \dots \odot a_n^{\otimes}]]_{\Sigma^{\otimes}}$  is empty, for every subset  $A$  of  $\Sigma$ , and every letters  $a_1, \dots, a_n \in \Sigma$ .*

The word language generators of [2, Section 6] are elements of the slightly more complex space  $[[\Sigma^{\otimes}]^*]$ . So:

**Theorem 5 (Word Language Generators).** *Let  $\Sigma$  be a finite alphabet. For any upward-closed subset  $U$  of  $[[\Sigma^{\otimes}]^*]$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap [[P]]_{(\Sigma^{\otimes})^*}$  is empty, for every  $*$ -product  $P$ . Here  $*$ -products are products of atomic expressions of the form  $(A^{\otimes} \odot a_1^{\otimes} \odot \dots \odot a_n^{\otimes})^?$ , or  $\{A_1^{\otimes} \odot a_{1n_1}^{\otimes} \odot \dots \odot a_{1n_1}^{\otimes}, \dots, A_k^{\otimes} \odot a_{kn_1}^{\otimes} \odot \dots \odot a_{kn_k}^{\otimes}\}^*$ , where  $a_i, a_{ij} \in \Sigma$ ,  $A, A_i \subseteq \Sigma$  for all  $i, j$ .*

Explicitly,  $[[A^{\otimes} \odot a_1^{\otimes} \odot \dots \odot a_n^{\otimes}]^?]_{(\Sigma^{\otimes})^?}$  is the set of sequences of multisets that are either the empty set, or consist of just one multiset with as many letters from  $A$  as one wishes, plus at most one  $a_1, \dots$ , at most one  $a_n$ . The denotation of  $\{A_1^{\otimes} \odot a_{1n_1}^{\otimes} \odot \dots \odot a_{1n_1}^{\otimes}, \dots, A_k^{\otimes} \odot a_{kn_1}^{\otimes} \odot \dots \odot a_{kn_k}^{\otimes}\}^*$  is the set of sequences of multisets  $m$  such that, for some  $i$ ,  $1 \leq i \leq k$ ,  $m$  consists of an arbitrary number of elements from  $A_i$ , plus at most one  $a_{i1}, \dots$ , at most one  $a_{in_i}$ .

The region generators of [2, Section 7] are elements of  $[[\Sigma^{\otimes} \times (\Sigma^{\otimes})^* \times P^{\otimes}]]$ , where  $P$  is a finite set with the trivial ordering  $=$ , and  $\Sigma = P \times \{0, \dots, max\}$ ,  $max$  being a fixed non-negative integer. Again Theorem 3 applies, although we believe stating the corresponding result would not bring much, and is now an easy exercise.

The data nets of [7] are transition systems over  $(\mathbb{N}^k)^*$ .

**Theorem 6.** *For any upward-closed subset  $U$  of  $(\mathbb{N}^k)^*$ , one can effectively construct a basis of  $U$  if and only if one can decide whether  $U \cap [[P]]_{(\mathbb{N}^k)^*}$  is empty, for every data net product  $P$  in  $\widehat{(\mathbb{N}^k)^*}$ .*

Because of the definition of  $(\mathbb{N}^k)^*$ , data net products  $P$  are therefore concatenations of atomic expressions  $\mathbf{a}^?$  where  $\mathbf{a} \in \mathbb{N}_{\omega}^k$ , and  $A^*$ , where  $A$  is any finite subset of  $\mathbb{N}_{\omega}^k$ . When  $P$  is the concatenation  $e_1 e_2 \dots e_n$ , the language of  $P$ ,  $[[P]]_{(\mathbb{N}^k)^*}$ , is the set of sequences of the form  $w_1 w_2 \dots w_n$  with  $w_1 \in [[e_1]]_{(\mathbb{N}^k)^*}$ ,  $w_2 \in [[e_2]]_{(\mathbb{N}^k)^*}$ ,  $\dots$ ,  $w_n \in [[e_n]]_{(\mathbb{N}^k)^*}$ . In turn,  $[[\mathbf{a}^?]]_{(\mathbb{N}^k)^*}$  is the set consisting of the empty sequence, plus every sequence consisting of just one  $k$ -tuple  $\mathbf{x} \in \mathbb{N}^k$  with  $\mathbf{x} \leq \mathbf{a}$ ; and  $[[A^*]]_{(\mathbb{N}^k)^*}$  is the set of sequences of  $k$ -tuples  $\mathbf{x}$  from  $\mathbb{N}^k$  such that  $\mathbf{x} \leq \mathbf{a}$  for some  $\mathbf{a} \in A$ .

We finish this series of special cases with the case  $D = (Q \times \mathbb{N}^k)^{\otimes}$ , with  $Q$  finite (with the trivial ordering), a data type we shall use in our forthcoming study of MVASS (multiset vector addition systems with states).

**Theorem 7.** *For any upward-closed  $U$  of  $[[Q \times \mathbb{N}^k]^{\otimes}]$ , one can effectively construct a basis of  $U$  iff one can decide whether  $U \cap [[A^{\otimes} \odot (q_1, \mathbf{a}_1)^{\otimes} \odot \dots \odot (q_n, \mathbf{a}_n)^{\otimes}]]_{(\mathbb{N}^k)^{\otimes}}$  is empty, for every finite subset  $A$  of  $Q \times \mathbb{N}_{\omega}^k$ , and every elements  $q_1, \dots, q_n \in Q$ ,  $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{N}_{\omega}^k$ .*

Explicitly,  $\llbracket A^{\otimes} \odot (q_1, \mathbf{a}_1)^{\textcircled{2}} \odot \dots \odot (q_n, \mathbf{a}_n)^{\textcircled{2}} \rrbracket_{\Sigma^{\otimes}}$  is a set of multisets of pairs  $(q, \mathbf{x})$  where  $q \in Q$  and  $\mathbf{x}$  is a  $k$ -tuple of natural numbers. These multisets are those that contain an arbitrary number of  $k$ -tuples of the form  $(q, \mathbf{x})$  for some  $(q, \mathbf{a}) \in A$  such that  $\mathbf{x} \leq \mathbf{a}$ , plus at most one element of the form  $(q_1, \mathbf{x}_1)$  with  $\mathbf{x}_1 \leq \mathbf{a}_1$ , plus  $\dots$  plus at most one element of the form  $(q_n, \mathbf{x}_n)$  with  $\mathbf{x}_n \leq \mathbf{a}_n$ .

## Acknowledgments

Philippe Schnoebelen gets the credit, both for asking me whether I had a proof of Theorem 2, and for suggesting numerous improvements to the paper.

## References

1. P. Abdulla, A. Bouajjani, and B. Jonsson. On-the-fly analysis of systems with unbounded, lossy Fifo channels. In *10th CAV*, pages 305–318. Springer Verlag LNCS 1427, 1998.
2. P. A. Abdulla, J. Deneux, P. Mahata, and A. Nylén. Forward reachability analysis of timed Petri nets. In Y. Lakhnech and S. Yovine, editors, *FORMATS/FTRTFT*, volume 3253 of *Lecture Notes in Computer Science*, pages 343–362. Springer, 2004.
3. P. A. Abdulla and R. Mayr. Minimal cost reachability/coverability in priced timed petri nets. In L. de Alfaro, editor, *Proc. FOSSACS 2009, 12th International Conference on Foundations of Software Science and Computation Structures*, pages 348–363, York, UK, Mar. 2009. Springer Verlag LNCS 5504.
4. A. Finkel and J. Goubault-Larrecq. Forward analysis for WSTS, part I: Completions. In S. Albers and J.-Y. Marion, editors, *Proceedings of the 26th Annual Symposium on Theoretical Aspects of Computer Science (STACS'09)*, pages 433–444, Freiburg, Germany, Feb. 2009.
5. A. Finkel, P. McKenzie, and C. Picaronny. A well-structured framework for analysing Petri net extensions. *Information and Computation*, 195(1-2):1–29, 2004.
6. J. Goubault-Larrecq. On Noetherian spaces. In *22nd LICS*, pages 453–462, Wrocław, Poland, July 2007. IEEE Computer Society Press.
7. R. Lazic, T. C. Newcomb, J. Ouaknine, A. W. Roscoe, and J. Worrell. Nets with tokens which carry data. In *ICATPN*, volume 4546 of *LNCS*, pages 301–320. Springer, 2007.
8. R. Valk and M. Jantzen. The residue of vector sets with applications to decidability problems in Petri nets. *Acta Informatica*, 21:643–674, 1985.