Varieties of cost functions

Laure Daviaud[1], Denis Kuperberg[2], and Jean-Éric Pin[3]

[1]*LIF, Université d'Aix-Marseille*
[2]*ONERA/DTIM, IRIT, University of Toulouse*
[3]*LIAFA, CNRS and Université Paris Diderot*

Abstract

Regular cost functions were introduced as a quantitative generalisation of regular languages, retaining many of their equivalent characterisations and decidability properties. For instance, stabilisation monoids play the same role for cost functions as monoids do for regular languages. The purpose of this article is to further extend this algebraic approach by generalizing two results on regular languages to cost functions: Eilenberg's varieties theorem and profinite equational characterisations of lattices of regular languages. This opens interesting new perspectives, but the specificities of cost functions introduce difficulties that prevent these generalisations to be straightforward. In contrast, although syntactic algebras can be defined for formal power series over a commutative ring, no such notion is known for series over semirings and in particular over the tropical semiring.

1 Introduction

Quantitative extensions of regular languages have been studied for over 50 years. Most of them rely on the early work of Schützenberger [25, 26, 27], who extended Kleene's theorem to formal power series over a semiring. A very nice presentation of this theory can be found in the book of Berstel and Reutenauer [5]. In this setting, weighted automata play the role of automata and weighted logic was introduced as an attempt to generalise Büchi's characterisation of regular languages in monadic second order logic. See the handbook [12] for an overview and further references.

However, this theory also suffers some weaknesses. For instance, the equality problem for rational series with multiplicities in the tropical semiring is undecidable [15], a major difference with the equality problem for regular languages, which is decidable. To overcome this problem and other related questions, Colcombet introduced the notion of regular cost functions [9], an other quantitative generalisation of regular languages. Cost functions are formally defined as equivalence classes of power series with coefficients in the semiring $\mathbb{N} \cup \{\infty\}$. This equivalence does not retain the exact values of the coefficients of the series but measures boundedness in some precise way. Thus cost functions are less general than power series, but are still more general than languages, which can be viewed as cost functions associated with their characteristic functions.

This approach proved to be very successful. It leads to simplified proofs of several major results related to boundedness (the limitedness problem of distance automata, Kirsten's proof of the star-height problem, etc.). Moreover, the standard equivalences on regular languages

$$\textit{regular languages} \iff \textit{finite automata} \iff \textit{finite monoids} \iff \textit{monadic second order logic}$$

admit the following nontrivial extension

$$\textit{regular cost functions} \iff \textit{cost automata} \iff \textit{stabilisation monoids} \iff \textit{cost monadic logic}$$

Contributions The aim of this paper is to show that the algebraic approach to regular languages also extends to the setting of cost functions. To this end, we change the recognizing object of cost functions from *stabilisation monoid* [9] to a structure with better algebraic properties, called *stabilisation algebra*. This gives us a new way of interpreting cost functions, as particular sets of a free stabilisation algebra $F(A)$ on the alphabet A, generalising the set of words A^*. This allows us to extend the ordered version [19] of Eilenberg's varieties theorem [13], which gives a bijective correspondence between positive varieties of languages and varieties of finite ordered monoids (Theorem 2). We show that the profinite algebra $\widehat{F(A)}$ generalising profinite words is the dual of the lattice of regular cost functions. This leads to an extension of the duality results between profinite words and regular languages. In particular, we extend the equational approach to lattices of regular languages given in [14] (Theorem 7). Our approach not only subsumes the corresponding results on languages but it also gives a nice algebraic framework for the results of [11, 16]. A series of examples is given in Section 8. All statements appearing without proof in the paper are proved in the appendix.

Related work Toruńczyk [28] also established a link between cost functions and profinite words, using a different approach. More precisely, Toruńczyk identifies a regular cost function with the set of profinite words that are limits of infinite sequences of words over which the function is bounded.

It is also interesting to compare these results to similar results on formal power series. Syntactic algebras of formal power series over a commutative ring were introduced by Reutenauer [22, 23], but no such notion is known for semirings. Reutenauer also extended Eilenberg's varieties theorem to power series over a commutative field. However, as shown in [24], equational theory only works for power series over finite fields.

Finally, let us mention two new promising approaches to recognisability, using respectively categories [1, 2] and monads [7, 8]. For the time being, these two approaches do not seem to apply to cost functions, but we hope our paper will serve as a test bench for future developments of this new point of view.

2 Regular cost functions and stabilisation monoids

In this section, we introduce the notions of cost functions and of stabilisation monoids. For a more complete and detailed presentation, the reader is referred to [10].

Let A be a finite alphabet and let $\mathcal{F}$ be the set of all functions from A^* to $\mathbb{N} \cup \{\infty\}$. Colcombet [9] introduced the following equivalence relation on $\mathcal{F}$: two elements f and g of $\mathcal{F}$ are *equivalent* (denoted by $f \approx g$) if, for each subset S of A^*, f is bounded on S if and only if g is bounded on S. A *cost function* is a $\approx$-class. In practice, cost functions are always represented by one of their representatives in $\mathcal{F}$.

The equivalence relation $\approx$ behaves well with respect to the operations $\min$ and $\max$. Indeed for all $f, g, h \in \mathcal{F}$, if $f \approx g$, then $\min(f, h) \approx \min(g, h)$ and $\max(f, h) \approx \max(g, h)$ [9]. It follows that the minimimum and the maximum of two cost functions are well-defined notions.

Example 1. *Let $A = \{a, b\}$. Given a word u, let $|u|$ denote the length of u and $|u|_a$ the number of occurrences of the letter a in u. Let us define three functions f, g and h from A^* to $\mathbb{N} \cup \{\infty\}$ by setting $f(u) = |u|$, $g(u) = |u|_a$ and $h(u) = 2|u|_a$. Then g is equivalent to h and they represent the same cost function, whereas g is not equivalent to f. Indeed g is bounded on b^* and f is not since for all n, $g(b^n) = 0$ and $f(b^n) = n$.*

The *characteristic function* of a language L on A^* is the function $\chi_L : A^* \to \mathbb{N} \cup \{\infty\}$ defined by $\chi_L(u) = 0$ if $u \in L$ and ∞ otherwise. The crucial observation that $\chi_L \approx \chi_{L'}$ if and only if $L = L'$ allows one to identify a language with the cost function defined by its characteristic function.

Stabilisation monoids were introduced in [9] in order to extend the classical notion of monoids recognising a language to the setting of cost functions. Recall that an *ordered monoid* is a set equipped with an associative binary product, a neutral element and an order compatible with the product. We denote by $E(M)$ the set of idempotents of a monoid M.

A *stabilisation monoid* is an ordered monoid M together with a *stabilisation operator* $\sharp : E(M) \to E(M)$ satisfying the following properties:

(S_1) for all $s, t \in M$ such that $st \in E(M)$ and $ts \in E(M)$, one has $(st)^\sharp s = s(ts)^\sharp$,

(S_2) for all $e \in E(M)$, one has $(e^\sharp)^\sharp = e^\sharp e = ee^\sharp = e^\sharp \leqslant e$,

(S_3) for all $e, f \in E(M)$, $e \leqslant f$ implies $e^\sharp \leqslant f^\sharp$,

(S_4) $1^\sharp = 1$.

Given two stabilisation monoids M and N, a *morphism* φ from M to N is a monoid morphism which is order-preserving and $\sharp$-preserving: if $e \in E(M)$, then $\varphi(e)^\sharp = \varphi(e^\sharp)$.

Just like finite (ordered) monoids recognise regular languages, finite stabilisation monoids can be used to recognise regular cost functions. However, the formal definition of recognition is more involved for cost functions than for languages and relies on the notion of factorisation trees. Let M be a stabilisation monoid and let $h : A \to M$ be a function, called the *labelling map*.

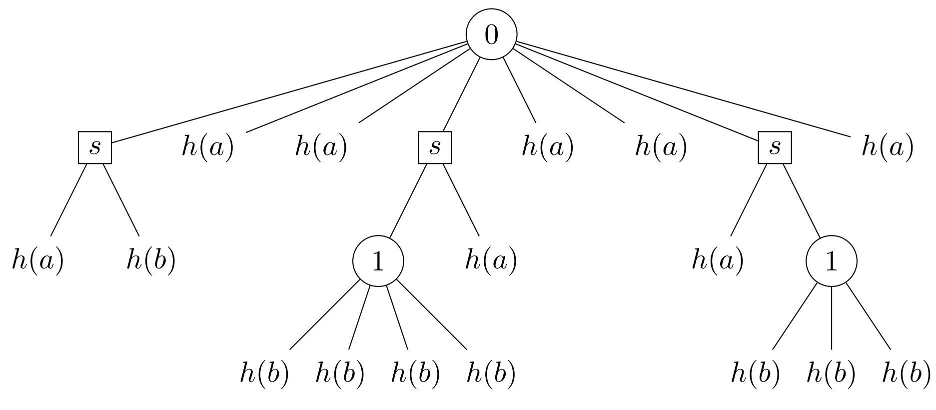
Definition 1. *Let $w = a_1 a_2 \cdots a_k$ be a word of A^* where each a_i is a letter. An h-*factorisation tree *of threshold n for w is a finite tree labelled by the elements of M and such that:*

(T_1) the tree has exactly k leaves, labelled by $h(a_1), \ldots, h(a_k)$, respectively,

(T_2) each binary node is labelled by the product of its left child's label by its right child's label,

(T_3) if a node has arity > 2, then all its children are labelled by the same idempotent e. If the arity of the node is $\leqslant n$, then the node is labelled by e, otherwise it is labelled by $e^\sharp$.

Example 2. *Let S_1 be the stabilisation monoid containing three idempotent elements $\{1, s, 0\}$ with $0 < s < 1$, $0^\sharp = s^\sharp = 0$, 0 is a zero of the monoid and 1 the neutral element. Let $h(a) = s$ and $h(b) = 1$. Here is a factorisation tree of threshold 5 for the word abaabbbbaaaabbba:*



A variant of Simon's factorisation theorem [9] guarantees the existence of trees of bounded height to evaluate input words. More precisely, for each labelling function $h : A \to M$, there is a positive integer K $(= 3|M|)$ such that for all words w and for all integers $n \geqslant 3$, there is an h-factorisation tree of threshold n for w with height at most K. We can now give the formal definition of a regular cost function recognised by a finite stabilisation monoid. Recall that a subset D of a partially ordered set is a *downset* if the conditions $t \in D$ and $s \leqslant t$ imply $s \in D$.

Definition 2. *Let M be a finite stabilisation monoid, $h : A \to M$ a labelling map and I a downset of M. The cost function* recognised *by (M, h, I) is the equivalence class of the function that maps a word u to the maximal threshold n such that there exists an h-factorisation tree for u of threshold n, height at most $3|M|$ and root in I. Such an equivalence class of functions is called a* regular cost function.

We keep the reversal order notation as it was introduced in [9].

Example 3. *Consider the stabilisation monoid S_1 defined in Example 2 and let $I = \{0\}$. Let $h : \{a, b\} \to M$ be the labelling map defined by $h(a) = s$ and $h(b) = 1$. Then S_1 is a stabilisation monoid that can make a distinction between products with no s (that are 1), products containing "few" s (that are s) and products containing "a lot of" s (that are $0 = s^\sharp$). The cost function recognised by (S_1, h, I) is the equivalence class of the function $u \mapsto |u|_a$.*

For instance the tree from example 2 of height 4 and threshold 5 has root labelled by $s^\sharp = 0$ because it is a witness that there are more than 5 occurrences of a in the input word. Conversely, such a factorisation tree of threshold n and height k would have its root labelled by 1 if the input word contains no a, and labelled by s if there are at most n^k occurrences of a. Because k is a fixed constant, these trees can be used to recognise the cost function $u \mapsto |u|_a$, since it is equivalent to $u \mapsto (|u|_a)^k$.

Regular cost functions can also be recognised by generalised forms of nondeterministic finite automata, regular expressions or monadic second-order logical formulas. See [11] for a complete introduction. Moreover, every regular cost function f has a unique *syntactic stabilisation monoid* M, in the sense that:

(1) there is a unique pair (h, I) where $h : A \to M$ is a labelling function and I a downset of M such that f is recognised by (M, h, I),

(2) for any (M', h', I') recognising f, there is a surjective morphism $\varphi : M \to M'$ such that $h = \varphi \circ h'$ and $I' = \varphi^{-1}(I)$.

3 Stabilisation algebras

The goal of the present work is to study algebraic properties of stabilisation monoids and cost functions. In particular, we would like to define regular cost functions as particular subsets of a free stabilisation monoid. However, since in a stabilisation monoid, the $\sharp$-operator is only defined on idempotents, the notion of a free stabilisation monoid cannot be defined directly and requires the introduction of a new algebraic structure, in which idempotents are directly defined in the signature of the algebra: stabilisation algebras.

Given a countable set of variables X, let $T(X)$ be the free term algebra of signature $\{\cdot, \omega, \sharp, 1\}$ over X. An *identity over $T(X)$* is an equation of the form $s \leqslant t$, where t, s are terms of $T(X)$.

A finite stabilisation monoid M *satisfies the identity* $s \leqslant t$ if the equation holds for any instantiation of the variables by elements of M, where 1 is interpreted as the neutral element of M, ω is interpreted as the idempotent power in M, and $\sharp$ is replaced with $\omega\sharp$ (to guarantee that $\sharp$ is only applied to idempotents). A finite stabilisation monoid M *satisfies the identity* $s = t$ if it satisfies the identities $s \leqslant t$ and $t \leqslant s$.

We can now define the structure of a stabilisation algebra in the following way.

Definition 3. *A* stabilisation algebra *is an ordered algebra M with signature $\langle 1, \leqslant, \cdot, \omega, \sharp \rangle$ satisfying the following axioms:*

(A_1) all identities that are satisfied by all finite stabilisation monoids,

(A_2) a description of the behaviour of ω on idempotent elements: $x^2 = x$ implies $x^\omega = x$,

(A$_3$) the three properties expressing that the order $\leqslant$ is compatible with the operations $\cdot, \omega, \sharp$: $x_1 \leqslant x_2$ and $y_1 \leqslant y_2$ imply $x_1y_1 \leqslant x_2y_2$, and $x \leqslant y$ implies $x^\omega \leqslant y^\omega$ and $x^\sharp \leqslant y^\sharp$.

In particular, (A$_1$) implies that a stabilisation algebra is a monoid with neutral element 1. A *morphism* between two stabilisation algebras is a monoid morphism which is order-preserving, ω-preserving and $\sharp$-preserving. Let M and N be two stabilisation algebras. Then N is a *stabilisation subalgebra* of M if $N \subseteq M$, and N is a *quotient* of M if there exists a surjective morphism $M \to N$. The product of two stabilisation algebras M and N is defined on the set product $M \times N$, with operations defined componentwise.

Recall that in a finite monoid, every element x has a unique idempotent power, denoted by x^ω. This fact allows one to identify finite stabilisation monoids and finite stabilisation algebras.

Proposition 1. *Finite stabilisation algebras are in one-to-one correspondence with finite stabilisation monoids, by interpreting ω as the idempotent power.*

We now build a free stabilisation algebra $F(A)$ on each finite alphabet A. Recall that $T(A)$ is the set of terms of the free algebra of signature $\{\cdot, \omega, \sharp, 1\}$ over the alphabet A. Given s and t two elements of $T(A)$, we write $s \equiv t$ if and only if the identity $t = s$ holds in all finite stabilisation monoids. This defines an equivalence relation and we denote by $F(A)$ the set of $\equiv$-classes. Furthermore we denote by $\overline{t}$ the equivalence class of t in $F(A)$.

Proposition 2. *$F(A)$ can be equipped with a structure of stabilisation algebra.*

Proof. This result on general ordered algebras is mentioned without proof in [6]. We give in appendix an explicit proof for completeness. □

A recent result [18] states that the equivalence of two $\sharp$-free terms of $T(A)$ is decidable. Actually, the result is more general and also covers the case of $\omega - 1$ powers. However, deciding the equivalence of arbitrary terms in $T(A)$ seems to still be an open problem.

The following theorem shows that $F(A)$ is a free object, by making explicit the corresponding universal property.

Theorem 1 (Universal Property)**.** *For any stabilisation algebra M and any function $h : A \to M$, there exists a unique morphism of stabilisation algebra $\overline{h} : F(A) \to M$ extending h.*

Corollary 1. *Every A-generated stabilisation algebra is a quotient of $F(A)$.*

4 Recognisability

We now define the notion of recognisable downsets in the free stabilisation algebra. We will later see how a regular cost function can be identified with a recognisable downset. This will allow us to generalise the classical notions of syntactic congruence and syntactic monoid. We identify terms $t \in T(A)$ and their class $\overline{t} \in F(A)$ for more readability.

Let I be a downset of $F(A)$, let M be a stabilisation algebra and let $h : F(A) \to M$ be a surjective morphism. We say that I is *recognised by h* if there exists a downset J of M such that $I = h^{-1}(J)$. A downset I of $F(A)$ is said to be *recognisable* if it is recognised by some morphism onto a finite stabilisation algebra.

Syntactic congruence and syntactic stabilisation algebra A *context* on A is an element of $T(A \cup \{x\})$, where $x \notin A$. In other words, a context is a term $T(A)$ with possible occurrences of the free variable x. Given a context C on A and an element t of $T(A)$, we denote by $C[t]$ the element of $T(A)$ obtained by replacing all the occurrences of x by t in C, i.e. $C[t] = C[x \leftarrow t]$. Let $\mathrm{Ctx}(A)$ denote the set of contexts on A.

Given a downset I of $F(A)$ and two elements t and s of $F(A)$, we write that $s \sim_I t$ if for any context C, $C[s] \in I$ is equivalent to $C[t] \in I$. This equivalence relation $\sim_I$ is a congruence on $F(A)$, called the *syntactic congruence* of I and the quotient morphism $h : F(A) \to F(A)/{\sim_I}$ is the *syntactic morphism* of I and the quotient algebra $F(A)/{\sim_I}$ is the *syntactic stabilisation algebra* of I.

Given a recognisable downset I of M, there is a natural preorder among the morphisms recognising I: given $h_1 : M \to N_1$ and $h_2 : M \to N_2$, we set $h_1 \leqslant h_2$ if there exists a surjective morphism $h : N_1 \to N_2$ such that $h_2 = h \circ h_1$. Then we can state:

Proposition 3. *The syntactic morphism is a minimal element for this preorder.*

The analog of this property in the framework of stabilisation monoids is given in [16].

Regular cost functions versus recognisable downsets We have seen that regular cost functions are recognised by finite stabilisation monoids and that recognisable downsets are recognised by finite stabilisation algebras. Now, Proposition 1 shows that finite stabilisation algebras correspond exactly to finite stabilisation monoids. These results indicate that regular cost functions and recognisable downsets are closely related.

One can make this relation a bijection as follows. Let f be a regular cost function and let M be its syntactic stabilisation monoid. Let also (h, I) be the unique pair (where $h : A \to M$ is a labelling function and I is a downset of M) such that f is recognised by (M, h, I). Then one can view M as a stabilisation algebra and extend h to a morphism $h : F(A) \to M$. Then the recognisable downset associated with f is $h^{-1}(I) \subseteq F(A)$.

Conversely, for every recognisable downset $I \subseteq F(A)$, we can define the regular cost function f associated to I by considering the minimal triplet (M, h, J) of I, and see M as a stabilisation monoid.

It is interesting to consider the link with regular languages. Let L be a regular language and let $\sim_L$ be its syntactic congruence. Let I_L be the downset corresponding to the regular cost function χ_L, i.e. the downset representing the language L. Since downsets represents cost functions via their unbounded elements, I_L is actually the set of elements of $T(A)$ representing words **not** in L. Any element of $T(A)$ can be tested for membership in L by evaluating it in the finite stabilisation algebra $M = A^*/{\sim_L}$ where $\sharp = id$ and ω is the idempotent power. Remark now that for all $u, v \in A^*$, we have $u \sim_L v$ if and only if $u \sim_{I_L} v$ (where u and v are interpreted in $T(A)$). In this sense, $\sim_I$ is an extension of the classical syntactic congruence on languages.

Example 4. *Let $A = \{a, b\}$. Consider the downset of $T(A)$ of elements containing an occurrence of a under the scope of $\sharp$. This describes words containing a large number of a. It is of finite index, and it represents the cost function g given in Example 1 that counts the number of a's in a word. It is also recognised by the stabilisation monoid given in Example 2 with elements $1 \geqslant a \geqslant 0$, all idempotent, and with $a^\sharp = 0^\sharp = 0$.*

Proposition 4. *The lattice of regular cost functions under* min *and* max *is isomorphic to the lattice of recognisable downsets under union and intersection.*

5 Varieties

We now generalise the notion of varieties of regular languages and some proofs from [13, 19].

A *lattice of recognisable downsets* is a set of recognisable downsets containing $\emptyset$ and $F(A)$, and closed under finite union and finite intersection. A lattice $\mathcal{L}$ of regular downsets of $F(A)$ is *closed under contexts* if, for every $I \in F(A)$ and each context $C \in \mathrm{Ctx}(A)$, the condition $I \in \mathcal{L}$ implies $C^{-1}[I] \in \mathcal{L}$, where $C^{-1}[I] = \{t \in F(A) \mid C[t] \in I\}$.

A *variety of recognisable downsets* associates with each finite alphabet A a lattice $\mathcal{V}(A)$ of recognisable downsets of $F(A)$ satisfying the following properties:

(V_1) For each alphabet A, $\mathcal{V}(A)$ is closed under contexts.

(V_2) For each morphism $\varphi : F(B) \to F(A)$, the condition $I \in \mathcal{V}(A)$ implies $\varphi^{-1}(I) \in \mathcal{V}(B)$.

Varieties of downsets generalise positive varieties of languages [19], as there is no complementation for downsets.

Example 5. *A recognisable downset I is* aperiodic *if for all $t \in F(A)$, the relation $t^\omega \sim_I t^\omega t$ holds. It is not too difficult to show that aperiodic downsets form a variety of recognisable downsets.*

We now define varieties of stabilisation algebras.

Definition 4. *A* variety of finite stabilisation algebras *is a class of finite stabilisation algebras closed under taking stabilisation subalgebras, quotients and finite products.*

Notice that this notion is often called *pseudovarieties* in the literature, as opposed to Birkhoff varieties which are also closed under arbitrary products.

Example 6. *A finite stabilisation algebra M is aperiodic if for all $x \in M$, we have $x^\omega = x^\omega x$. Aperiodic stabilisation algebras form a variety of finite stabilisation algebras.*

Let M, N be two stabilisation algebras. We say that M *divides* N if M is a quotient of a stabilisation subalgebra of N. It follows from the definition that varieties of stabilisation algebras are closed under division. If S is a (possibly infinite) set of finite stabilisation algebras, the *variety generated by S* is the smallest variety of finite stabilisation algebras containing the elements of S.

Lemma 1. *Let $\mathbf{V}$ be a variety generated by a set S of finite stabilisation algebras, and M be a finite stabilisation algebra. Then $M \in \mathbf{V}$ if and only if M divides a finite product of elements of S.*

Given a variety $\mathbf{V}$ of finite stabilisation algebras, let $\mathcal{V}(A)$ denote the set of recognisable downsets over A whose syntactic stabilisation algebra belongs to $\mathbf{V}$. The correspondence $\mathbf{V} \to \mathcal{V}$ associates with each variety of finite stabilisation algebras a class of recognisable downsets.

Thus, each variety of recognisable downsets $\mathcal{V}$ is associated to the variety of finite stabilisation algebras $\mathbf{V}$ generated by the syntactic stabilisation algebras of downsets in $\mathcal{V}$. This defines a correspondence $\mathcal{V} \to \mathbf{V}$. The analog of the ordered version of Eilenberg's theorem can now be stated as follows:

Theorem 2. *The correspondences $\mathcal{V} \to \mathbf{V}$ and $\mathbf{V} \to \mathcal{V}$ define mutually inverse bijective correspondences between varieties of finite stabilisation algebras and varieties of recognisable downsets.*

6 Profinite stabilisation algebra

The free profinite monoid on A, denoted by $\widehat{A^*}$, can be defined as the completion of A^* for the *profinite metric*. See [4, 20, 21] for more information on this space.

We now prove the existence of free profinite stabilisation algebras. Taking the construction of free profinite monoids as a model, we define it as the completion $\widehat{F(A)}$ of $F(A)$ for an appropriate metric.

Definition 5. *A stabilisation algebra M* separates *two elements s and t of $F(A)$ if there is a morphism $\varphi : F(A) \to M$ such that $\varphi(s) \neq \varphi(t)$. For $s, t \in F(A)$, define*

$$d(s,t) = \begin{cases} 0 & \text{if } s = t \\ 2^{-n(s,t)} & \text{otherwise} \end{cases}$$

where $n(s,t)$ is the minimum size of a finite stabilisation algebra separating s and t.

Note that d is well defined, since if $s \neq t \in F(A)$, then there is by (A_1) a finite stabilisation monoid in which the identity $s = t$ fails. Such a monoid can be viewed as a finite stabilisation algebra separating s and t. The following proposition gathers the properties of d and $\widehat{F(A)}$:

Proposition 5.

(1) *d is an ultrametric distance.*

(2) *The operations on $F(A)$ are uniformly continuous and thus extend by continuity to $\widehat{F(A)}$.*

(3) *The resulting stabilisation algebra $\widehat{F(A)}$ is compact.*

The idempotent power. If M is a finite monoid, then for any $m \in M$ and $n \geqslant |M|$, we have $m^\omega = m^{n!}$ (where ω is the idempotent power). Since finite stabilisation algebras are in particular monoids where ω is the idempotent power, we obtain that for any $u \in F(A)$ and $n > 0$, $d(u^{n!}, u^\omega) \leqslant 2^{-n}$. Therefore, for any element $u \in F(A)$, the sequence $(u^{n!})_{n \in \mathbb{N}}$ converges in $\widehat{F(A)}$, to u^ω.

Proposition 6. *The morphism $\varphi : F(A) \to \widehat{A^*}$ defined by $\varphi(a) = a$, $\varphi(t^\omega) = t^\omega$ and $\varphi(t^\sharp) = t^\omega$ is uniformly continuous, and therefore can be uniquely extended into a continuous morphism of stabilisation algebras $\widehat{\varphi} : \widehat{F(A)} \to \widehat{A^*}$.*

Notice however that this morphism does not coincide with the interpretation of regular cost functions as subsets of $\widehat{A^*}$ as done in [28].

The profinite metric can be relativised to any variety of stabilisation algebras to obtain the so-called pro-$\mathbf{V}$ metric. For $s, t \in F(A)$ and $\mathbf{V}$ a variety of stabilisation algebras, define $d_{\mathbf{V}}(s,t) = 2^{-|M|}$ where M is one of the smallest stabilisation algebras from $\mathbf{V}$ separating s and t and $d_{\mathbf{V}}(s,t) = 0$ if there is no such M. Remark that the metric d of Definition 7 corresponds to $d_{\mathbf{V}}$ where $\mathbf{V}$ is the variety of all finite stabilisation algebras. We also define an equivalence $s \sim_{\mathbf{V}} t$ by $d_{\mathbf{V}}(s,t) = 0$.

Proposition 7. *For any variety $\mathbf{V}$, $\sim_{\mathbf{V}}$ is a congruence on $F(A)$ and d_V is an ultrametric distance on $F(A)/{\sim_{\mathbf{V}}}$.*

We now define the pro-$\mathbf{V}$ stabilisation algebra $\widehat{F_{\mathbf{V}}(A)}$ as the completion of $F(A)/{\sim_{\mathbf{V}}}$ with respect to $d_{\mathbf{V}}$. As before, we can show that $\widehat{F_{\mathbf{V}}(A)}$ is compact and can be equipped with a structure of stabilisation algebra. The following result now follows from general results on profinite algebras.

Theorem 3. *A finite A-generated stabilisation algebra belongs to $\mathbf{V}$ if and only if it is a continuous quotient of $\widehat{F_{\mathbf{V}}(A)}$.*

7 Duality, equations and identities

Stone duality tells us that every bounded distributive lattice $\mathcal{L}$ has an associated compact Hausdorff space, called its *dual space*. The dual space of the Boolean algebra of all *regular* languages of A^* [3] is the free profinite monoid on A.

A similar result holds for the lattice of regular cost functions, which, by Proposition 4, is isomorphic to the lattice of recognisable downsets under union and intersection.

Theorem 4. *The dual space of the lattice of recognisable downsets of $F(A)$ is the space $\widehat{F(A)}$.*

7.1 Equations of lattices

It is shown in [14] that any lattice of regular languages can be defined by a set of equations of the form $u \to v$, where u and v are profinite words. This result can also be extended to recognisable downsets.

Let $u, v \in \widehat{F(A)}$. We say that a recognisable downset I of $F(A)$ *satisfies the equation* $u \to v$ if $u \in \overline{I}$ implies $v \in \overline{I}$, where $\overline{I}$ denotes the topological closure of I.

A set $\mathcal{L}$ of recognisable downsets is *defined by a set E of equations* if the following property holds: a recognisable downset belongs to $\mathcal{L}$ if and only if it satisfies all the equations of E. We can now state our second main result.

Theorem 5. *A set of recognisable downsets of $F(A)$ is a lattice of recognisable downsets if and only if it is defined by a set of equations of the form $u \to v$.*

The case of lattices of languages closed under quotients was also considered in [14]. The corresponding notion for lattices of downsets is to be closed under contexts.

A *profinite context* C on the finite alphabet A is an element of $F(\widehat{A \cup \{x\}})$ where $x \notin A$. If u is an element of $\widehat{F(A)}$, then $C[u]$ is also an element of $\widehat{F(A)}$, defined by replacing x by u and evaluating the operations ω and $\sharp$ in the stabilisation algebra $\widehat{F(A)}$.

Definition 6. *A recognisable downset of $F(A)$* satisfies the equation $u \leqslant v$ *if, for all profinite contexts C, it satisfies the equation $C[u] \to C[v]$.*

Equivalently, a stabilisation algebra *satisfies the equation* $u \leqslant v$ if, for all downsets J of M and for all contexts C, $C[v] \in J$ implies $C[u] \in J$. By density of $F(A)$ in $\widehat{F(A)}$, it is enough to consider contexts in Ctx_A for this definition. The notation $u = v$ is used as a shortcut for $u \leqslant v$ and $v \leqslant u$. We can now state:

Theorem 6. *A set of recognisable downsets of $F(A)$ is a lattice of recognisable downsets closed under contexts if and only if it is defined by a set of equations of the form $u \leqslant v$.*

7.2 Identities of varieties

Condition (V_2) of the definition of a variety allows one to use identities instead of equations.

Let B be an alphabet and let u and v be two elements of $\widehat{F(B)}$. We say that a recognisable downset I of $F(A)$ *satisfies the profinite identity* $u \leqslant v$ if, for each morphism $\gamma : F(B) \to F(A)$, I satisfies the equation $\widehat{\gamma}(u) \leqslant \widehat{\gamma}(v)$.

We use the term *identity* because, in this case, each letter of B can be replaced (through the morphism γ) by *any* element of $F(A)$.

In practice, it is more convenient to use the following characterisation. Let I be a recognisable downset and let M be its syntactic stabilisation algebra. Then I satisfies the identity $u \leqslant v$ if and only if for every continuous morphism $h : F(B) \to M$, one has $\widehat{h}(u) \leqslant_M \widehat{h}(v)$, where $\leqslant_M$ is the order of M.

Theorem 7. *A class of finite stabilisation algebras (resp. recognisable downsets) is a variety if and only if it is defined by a set of identities of the form $u \leqslant v$.*

8 Examples of equational descriptions of varieties and lattices

In this section, we gather examples of varieties of regular cost functions and of sets of equations. First, it is interesting to see how the identification of regular languages with cost functions extends to varieties.

Proposition 8. *If a positive variety of regular languages is defined by a set of identities E, then it is a variety of regular cost functions, defined by the set of identities $E \cup \{x^\sharp = x^\omega\}$. Conversely, if a variety of regular cost functions is defined by a set of identities E, then the variety of regular cost functions defined by $E \cup \{x^\sharp = x^\omega\}$ can be identified with a positive variety of languages.*

For instance, the variety of regular cost functions defined by $x^\omega = x^{\omega+1}$ and $x^\sharp = x^\omega$ contains only the characteristic functions of star-free languages.

Aperiodic cost functions The variety of *aperiodic cost functions* is defined by the identity $x^\omega = x^{\omega+1}$. It contains recognisable downsets that are not languages, like $u \mapsto |u|_a$. This variety has a nice connection with the logics CFO and CLTL, first introduced in [16, 17] as a generalisation to cost functions of the logics FO and LTL on words. Indeed, the results of [16, 17] can be reformulated as follows:

Theorem 8. *The variety of aperiodic cost functions coincides with the variety of CFO-definable cost functions and with the variety of CLTL-definable cost functions.*

Note that given a finite stabilisation algebra M, one can effectively test whether it verifies equations like $x^\omega = x^{\omega+1}$ or $x^\sharp = x^\omega$: it suffices to check that it stands for each x in M. It follows that one can effectively decide whether a regular cost function is CFO-definable (respectively CLTL-definable).

Temporal cost functions Another interesting example is the class of temporal cost functions, first introduced in [11]. These functions allow one to count the number of occurrences of consecutive events. Many equivalent characterizations of these functions are known. In [11], the algebraic characterization is expressed in terms of the interplay between Green relations and stabilisation in the syntactic monoid, but it can be formulated in terms of equations as follows:

Theorem 9. *Temporal cost functions over A form a lattice of regular cost functions, defined by the equations $(xy^\sharp z)^\sharp = (xy^\sharp z)^\omega$, for all $x, z \in F(A)$ and all $y \in F(A) - \{1\}$.*

Proof. Let M be the syntactic stabilisation monoid of a regular cost function f. An idempotent e is called *stable* if $e^\sharp = e$. The algebraic characterization from [11] states that f is temporal if and only if an idempotent $\mathcal{J}$-below a stable idempotent different from 1 is itself stable. Recall that the $\mathcal{J}$-order is defined by $e \leqslant_\mathcal{J} s$ if there exist $x, y \in M$ such that $e = xsy$. To show that our set of equations is equivalent to this characterization, it suffices to observe that an element is a stable idempotent if and only if it is of the form $s^\sharp$ for some s. This means that the characterization from [11] specifies that the idempotents of the form $(xs^\sharp z)^\omega$, with $s \neq 1$ are stable. Using Corollary 1, one can now lift these properties to $F(A)$, yielding the equations of the statement. □

Commutative cost functions The description of the variety of languages corresponding to commutative monoids is one of the first known examples of Eilenberg's correspondence between varieties of languages and varieties of monoids [13]. We prove below a similar result for cost functions.

Let us say that a finite stabilisation algebra M is *commutative* if for all $x, y \in M$, we have $xy = yx$. We will say that M is $\sharp$-*commutative* if it is commutative and for all $x, y \in M$, $x^\sharp y^\sharp =$

$(xy)^\sharp$. A cost function is called *commutative* (resp. $\sharp$-*commutative*) if its syntactic stabilisation algebra is commutative (resp. $\sharp$-commutative).

Example 7. *The cost function* $\mathrm{maxblock}_a$ *on the alphabet* $\{a,b\}$ *defined by:*

$$\mathrm{maxblock}_a(a^{n_1}ba^{n_2}b\cdots ba^{n_k}) = \max(n_1,n_2,\dots,n_k)$$

is commutative but not $\sharp$-commutative. The downset representing this regular cost function consists in $\omega\sharp$-expressions on $\{a,b\}$ containing a subexpression of the form $u^\sharp$, where u is an $\omega\sharp$-expression on the alphabet $\{a\}$ with at least one occurrence of a. Its syntactic stabilisation algebra M has four elements: $0 \leqslant a \leqslant 1 \leqslant b$, all elements are idempotent and commute, and we have $a^\sharp = 0$, $x^\sharp = x$ for $x \neq a$, and $ab = b = ba$. It is not $\sharp$-commutative because $a^\sharp b^\sharp = 0b = 0$ and $(ab)^\sharp = b^\sharp = b$.

We will use freely the following useful lemma.

Lemma 2. *Any finite $\sharp$-commutative stabilisation algebra divides the product of its monogenic stabilisation subalgebras.*

The stabilisation monoid S_1 defined in Example 2 has three idempotent elements $0 < a < 1$ such that $0^\sharp = a^\sharp = 0$. It is also the syntactic stabilisation algebra of the function $f = u \mapsto |u|_a$.

Let U_1^+ denote the stabilisation monoid with two idempotent elements $0 \leqslant 1$ such that $0^\sharp = 0$.

Proposition 9. *Let $\mathbf{J_1^+}$ be the variety of finite stabilisation algebras defined by the equations $x \leqslant 1$, $x^2 = x$, $xy = yx$ and $x^\sharp y^\sharp = (xy)^\sharp$. Then the corresponding variety of cost functions is generated by the functions $u \mapsto |u|_a$ for all letters a.*

Proof. Since S_1 is the syntactic stabilisation algebra of the function $u \mapsto |u|_a$, it is equivalent to prove that the variety of finite stabilisation algebras $\mathbf{V}$ generated by S_1 is equal to $\mathbf{J_1^+}$. Since S_1 satisfies all the equations of $\mathbf{J_1^+}$, the relation $\mathbf{V} \subseteq \mathbf{J_1^+}$ holds. To prove the opposite inclusion, consider a finite stabilisation algebra M of $\mathbf{J_1^+}$. By Lemma 2, M divides the product of its monogenic stabilisation subalgebras. But if $m \in M$, the stabilisation algebra generated by m is $\{1, m, m^\sharp\}$: indeed, the equations $m^2 = m$, $(m^\sharp)^2 = m^\sharp$ and the properties of a stabilisation monoid imply that $mm^\sharp = m^\sharp m = m^\sharp$. Thus, this stabilisation algebra is either $\{1\}, U_1^+$ or S_1. Since U_1^+ is a quotient of S_1, M actually divides a product of copies of S_1, and therefore $M \in \mathbf{V}$. Thus $\mathbf{V} = \mathbf{J_1^+}$. □

As stated earlier, if we add the equation $x^\sharp = x^\omega$, we obtain the positive variety of regular languages corresponding to the variety of ordered monoids generated by the ordered monoid U_1^+ (see [19]).

Proposition 1. *Let $\mathbf{Acom}$ be the variety of finite stabilisation algebras defined by the equations $x^\omega = x^{\omega+1}$, $xy = yx$ and $x^\sharp y^\sharp = (xy)^\sharp$. Then the corresponding variety of cost functions is generated by the functions $u \mapsto |u|_a$ and $\chi_{L_{a,k}}$ where $L_{a,k} = \{u \mid |u|_a = k\}$ for each $k \geqslant 0$ and each letter a.*

Proof. The proof is similar to the proof of Proposition 9 and mimics the corresponding proof for varieties of finite ordered monoids. Let $\mathbf{V}$ be the variety generated by the stabilisation algebras of the functions given in the proposition. It suffices to show that $\mathbf{V} = \mathbf{Acom}$. Let $M_k = \{1, a, a^2, a^k, \dots, a^{k+1} = 0\}$ be the syntactic (stabilisation) monoid of $\chi_{L_{a,k}}$. First, notice that S_1 and every M_k satisfy the equations of $\mathbf{Acom}$, so that $\mathbf{V} \subseteq \mathbf{Acom}$. We now define the stabilisation algebra $M'_k = \{1, a, a^2, a^k, a^{k+1} = a^{k+2}, 0\}$ with $(a^{k+1})^\sharp = 0$. Notice that M'_k divides $S_1 \times M_k$, as witnessed by the morphism $\varphi(1, a^i) = \varphi(a, a^i) = a^i$, and $\varphi(0, a^i) = 0$.

Therefore, $M'_k \in \mathbf{V}$. Let M be any monoid in $\mathbf{Acom}$, then M is $\sharp$-commutative, and by Lemma 2, it divides the product of its monogenic stabilisation subalgebras. But according to the equations of $\mathbf{Acom}$, if $m \in M$, $m^\omega = m^{\omega+1}$ is idempotent and $(m^\omega)^\sharp$ is a zero according to the definition of a stabilisation monoid ($m^i(m^\omega)^\sharp = (m^\omega)^\sharp$) and the equations ($((m^\omega)^\sharp)^2 = (m^\omega m^\omega)^\sharp$). Thus, the stabilisation subalgebra generated by m is either of the form M_k or M'_k, which allows us to conclude that $M \in \mathcal{V}$. It follows that $\mathbf{Acom} \subseteq \mathbf{V}$, and finally $\mathbf{Acom} = \mathbf{V}$. □

Proposition 2. *Let* $\mathbf{Com}$ *be the variety of finite stabilisation algebras defined by the equations* $xy = yx$ *and* $x^\sharp y^\sharp = (xy)^\sharp$*. Then the corresponding variety of cost functions is generated by the functions* $u \mapsto |u|_a$, $\chi_{L_{a,k}}$ *and* $\chi_{L_{k,n}}$*, where* $L_{a,k,n} = \{u \mid |u|_a \equiv k \bmod n\}$.

Proof. Let $\mathbf{V}$ be the variety generated by the stabilisation algebras of the functions given in the proposition. It suffices to show that $\mathbf{V} = \mathbf{Com}$. Let $M_k = \{1, a, a^2, a^k, \dots, a^{k+1} = 0\}$ be the syntactic (stabilisation) monoid of $\chi_{L_{a,k}}$. The syntactic (stabilisation) monoid of $\chi_{L_{a,k,n}}$ is $\mathbb{Z}/n\mathbb{Z}$ (with $x^\sharp = x^\omega$). First notice that for all k, the stabilisation monoids S_1, M_k and $\mathbb{Z}/k\mathbb{Z}$ satisfy the equations of $\mathbf{Com}$, so we obtain $\mathbf{V} \subseteq \mathbf{Com}$. Let now M be a stabilisation monoid in $\mathbf{Com}$. Then M is $\sharp$-commutative, and by Lemma 2, it divides the product of its monogenic stabilisation subalgebras. Let N be a monogenic stabilisation submonoid of M generated by a. By using only products, we get a generic monogenic monoid $P = \{1, a, a^2, \dots, a^k, \dots, a^{k+n} = a^k\}$, containing an idempotent a^j with $0 < n \leqslant j$ and $a^{j+n} = a^j$, and for any $i < n$, $a^{j+i} \neq a^j$. Then, using stabilisation, we obtain a new element $(a^j)^\sharp$, that can again be multiplied with a, forming $(a^j)^\sharp a, (a^j)^\sharp a^2, \dots, (a^j)^\sharp a^{n-1}$. We cannot form new elements because $(a^j)^\sharp a^n = (a^j)^\sharp a^j a^n = (a^j)^\sharp a^j = (a^j)^\sharp$, and the only idempotents are $1, a^j$, and $(a^j)^\sharp$. Therefore, we get that N is isomorphic to $P \cup \mathbb{Z}/n\mathbb{Z}$, where the element i of the $\mathbb{Z}/n\mathbb{Z}$ component represents $(a^j)^\sharp a^i$. Let M'_k be the stabilisation algebra from the proof of Proposition 1, and notice that $M'_k \in \mathcal{V}$. We can now notice that N is isomorphic to a stabilisation subalgebra of $M'_k \times \mathbb{Z}/n\mathbb{Z} \in \mathcal{V}$, via the injection $\varphi : N \to M'_k \times \mathbb{Z}/n\mathbb{Z}$ defined by

$$\begin{cases} \varphi(a^i) = (a^i, 0) & \text{for } i \leqslant j, \\ \varphi(a^{j+i}) = (a^j, i) & \text{for } i < n, \\ \varphi((a^j)^\sharp a^i) = (0, i) & \text{for } i < n, \end{cases}$$

which implies that $N \in \mathbf{V}$. Thus M divides a product of elements of $\mathbf{V}$, so finally $M \in \mathbf{V}$. We obtained $\mathbf{Com} \subseteq \mathbf{V}$, and finally $\mathbf{Com} = \mathbf{V}$. □

9 Conclusion

We provide a new representation of regular cost functions as downsets of a free stabilisation algebra, an ordered algebraic structure. This new representation allows us to extend Eilenberg's variety theory, in its ordered version: varieties of regular cost functions correspond to varieties of finite stabilisation algebras and are characterised by profinite identities. Furthermore, we also extend the duality approach of [14] to this new setting, leading to profinite equational descriptions of lattices of regular cost functions. Finally, we give several examples of equational characterisations of classes of cost functions related to logic. We also investigate the extensions of commutative languages to regular cost functions. We uncover the role of a new identity, $x^\sharp y^\sharp = (xy)^\sharp$, in the study of these extensions.

These results confirm the pertinence and the usefulness of the theory of regular cost functions as a well-behaved quantitative generalisation of regular languages. They also open new perspective for the study of cost functions.

For instance, it would be interesting to extend other known characterisations of varieties of languages to the setting of cost functions. An emblematic example would be Simon's characterisation of piecewise testable languages.

Acknowledgments

The authors would like to thank Thomas Colcombet, Nathanaël Fijalkow and Sam Van Gool for many helpful discussions.

References

[1] J. ADÁMEK, S. MILIUS, R. S. R. MYERS AND H. URBAT, Generalized Eilenberg Theorem I: Local Varieties of Languages, in *FOSSACS 2014*, A. Muscholl (ed.), pp. 366–380, *Lecture Notes in Comput. Sci.* vol. 8412, Springer, 2014.

[2] J. ADÁMEK, R. S. R. MYERS, H. URBAT AND S. MILIUS, Varieties of Languages in a Category, in *30th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2015, Kyoto, Japan, July 6-10, 2015*, pp. 414–425, IEEE, 2015.

[3] J. ALMEIDA, Residually finite congruences and quasi-regular subsets in uniform algebras, *Portugaliæ Mathematica* **46** (1989), 313–328.

[4] J. ALMEIDA, *Finite semigroups and universal algebra*, World Scientific Publishing Co. Inc., River Edge, NJ, 1994. Translated from the 1992 Portuguese original and revised by the author.

[5] J. BERSTEL AND C. REUTENAUER, *Noncommutative rational series with applications*, *Encyclopedia of Mathematics and its Applications* vol. 137, Cambridge University Press, Cambridge, 2011.

[6] S. L. BLOOM, Varieties of ordered algebras, *J. Comput. System Sci.* **13**,2 (1976), 200–212.

[7] M. BOJANCZYK, Recognisable Languages over Monads, in *Developments in Language Theory - 19th International Conference, DLT 2015*, I. Potapov (ed.), pp. 1–13, *Lecture Notes in Comput. Sci.* vol. 9168, Springer, 2015.

[8] M. BOJANCZYK, Recognisable languages over monads, *CoRR* **abs/1502.04898** (2015).

[9] T. COLCOMBET, The theory of stabilisation monoids and regular cost functions, in *Automata, languages and programming. Part II*, Berlin, 2009, pp. 139–150, *Lecture Notes in Comput. Sci.* vol. 5556, Springer.

[10] T. COLCOMBET, Regular cost functions, Part I: Logic and algebra over words, *Log. Methods Comput. Sci.* **9**,3 (2013), 3:3, 47.

[11] T. COLCOMBET, D. KUPERBERG AND S. LOMBARDY, Regular temporal cost functions, in *Automata, languages and programming, ICALP 2010, Part II*, S. Abramsky, C. Gavoille, C. Kirchner, F. M. auf der Heide and P. G. Spirakis (eds.), pp. 563–574, *Lecture Notes in Comput. Sci.* vol. 6199, Springer, Berlin, 2010.

[12] M. Droste, W. Kuich and H. Vogler (eds.), *Handbook of weighted automata*, *Monographs in Theoretical Computer Science. An EATCS Series*, Springer-Verlag, Berlin, 2009.

[13] S. EILENBERG, *Automata, languages, and machines. Vol. B*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1976.

[14] M. GEHRKE, S. GRIGORIEFF AND J.-E. PIN, Duality and equational theory of regular languages, in *ICALP 2008, Part II*, L. Aceto and al. (eds.), Berlin, 2008, pp. 246–257, *Lecture Notes in Comput. Sci.* vol. 5126, Springer.

[15] D. KROB, The equality problem for rational series with multiplicities in the tropical semiring is undecidable, *Internat. J. Algebra Comput.* **4**,3 (1994), 405–425.

[16] D. KUPERBERG, Linear temporal logic for regular cost functions, in *28th International Symposium on Theoretical Aspects of Computer Science*, pp. 627–636, *LIPIcs. Leibniz Int. Proc. Inform.* vol. 9, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2011.

[17] D. KUPERBERG AND M. VANDEN BOOM, On the expressive power of cost logics over infinite words, in *Automata, languages, and programming. Part II*, pp. 287–298, *Lecture Notes in Comput. Sci.* vol. 7392, Springer, Heidelberg, 2012.

[18] I. MIKHAILOVA, A proof of Zhil'tsov's theorem on decidability of equational theory of epigroups. http://arxiv.org/abs/1310.5023, 2013.

[19] J.-E. PIN, A variety theorem without complementation, *Russian Mathematics (Izvestija vuzov.Matematika)* **39** (1995), 80–90.

[20] J.-É. PIN, Profinite methods in automata theory, in *26th International Symposium on Theoretical Aspects of Computer Science (STACS 2009)*, S. Albers and J.-Y. Marion (eds.), pp. 31–50, Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2009.

[21] J.-É. PIN, Equational descriptions of languages, *Int. J. Found. Comput. S.* **23** (2012), 1227–1240.

[22] C. REUTENAUER, Variétés d'algèbres et de séries rationnelles, in *1er Congrès Math. Appl. AFCET-SMF*, vol. 2, pp. 93–102, AFCET, 1978.

[23] C. REUTENAUER, Séries formelles et algèbres syntactiques, *J. Algebra* **66**,2 (1980), 448–483.

[24] C. REUTENAUER, *Séries rationnelles et algèbres syntactiques*, PhD thesis, Université Paris 6, 1980. Thèse de Doctorat d'État.

[25] M. P. SCHÜTZENBERGER, On the definition of a family of automata, *Information and Control* **4** (1961), 245–270.

[26] M.-P. SCHÜTZENBERGER, Finite counting automata, *Information and Control* **5** (1962), 91–107.

[27] M.-P. SCHÜTZENBERGER, On a theorem of R. Jungen, *Proc. Amer. Math. Soc.* **13** (1962), 885–890.

[28] S. TORUŃCZYK, *Languages of profinite words and the limitedness problem*, PhD thesis, Computer Science Department, University of Warsaw, 2011.

Appendix

This appendix contains the proofs of all statements which are not proved in the paper. The statements labelled by **A.xx** only appear in the Appendix. The other ones already appeared in the article with the same label.

3 Stabilisation algebras

Any finite stabilisation monoid can be viewed as a stabilisation algebra, by interpreting ω as the idempotent power, and defining the new $\sharp$ as $\omega\sharp$. Conversely, any finite stabilisation algebra can be seen as a finite stabilisation monoid, by forgetting the operator ω and restricting $\sharp$ to the idempotent elements. This is formalised in the following lemma:

Proposition 1. *Finite stabilisation algebras are in one-to-one correspondence with finite stabilisation monoids, by interpreting ω as the idempotent power.*

Proof. Let $(M, \leqslant, \cdot, \sharp)$ be a finite stabilisation monoid. We show that if we define $\omega : M \to M$ as the idempotent power, then $(M, \leqslant, \cdot, \omega, \omega\sharp)$ is an stabilisation algebra. All axioms from (A_1) are satisfied by definition. Since ω is the idempotent power, it satisfies (A_2). Finally, the order $\leqslant$ satisfies all implications from (A_3) by axioms of stabilisation monoids.

Conversely, let $(M, \leqslant, \cdot, \omega, \sharp)$ be a finite stabilisation algebra. We want to show that this structure forms a stabilisation monoid, where ω is interpreted as the idempotent power. First, let us prove that for all $a \in M$, a^ω is the idempotent power of a. Let $a \in M$ and $n = |M|!$, then a^n is the idempotent power of a. We want to show that $a^\omega = a^n$. The identity $(x^n)^\omega = x^\omega$ is in (A_1), so $(a^n)^\omega = a^\omega$. Moreover, $a^n a^n = a^n$, so by (A_2), we have $(a^n)^\omega = a^n$. We finally obtain $a^\omega = a^n$. We turn to the axioms of stabilisation monoids (S_1) to (S_4):

(S_1): We prove that the identity $(xy)^{\omega\sharp}x = x(yx)^{\omega\sharp}$ is in (A_1). Let M' be any finite stabilisation monoid and $a, b \in M'$. Let $n = |M'|!$, $s = a$ and $t = b(ab)^{n-1}$. We have $st = (ab)^n = (ab)^\omega \in E(M)$, and $ts = b(ab)^{n-1}a = (ba)^n = (ba)^\omega \in E(M)$, where ω is the idempotent power in M. Therefore, we get by (S_1) that $(st)^\sharp s = s(ts)^\sharp$, i.e. $(ab)^{\omega\sharp}a = a(ba)^{\omega\sharp}$. This is true for any a, b in any M', so the identity $(xy)^{\omega\sharp}x = x(yx)^{\omega\sharp}$ is in (A_1). We now take $s, t \in M$ such that $st \in E(M)$ and $ts \in E(M)$. By (A_1), we have $(st)^\sharp s = s(ts)^\sharp$, therefore M satisfies (S_1).

(S_2): It follows from the fact that the identities $(x^{\omega\sharp})^\sharp = x^{\omega\sharp}x^\omega = x^\omega x^{\omega\sharp} = x^{\omega\sharp} \leqslant x^\omega$ are in (A_1).

(S_3): Axiom (A_3) guarantees that the order is compatible with all operations.

(S_4): $1^\sharp = 1$ is an identity in (A_1). □

This means that the finite stabilisation algebras are exactly finite stabilisation monoids, up to these identifications.

Proposition 2. *$F(A)$ can be equipped with a structure of stabilisation algebra.*

Proof. As usual, we denote the product of two elements s and t by st instead of $s \cdot t$.

We first define the operations and the order on $F(A)$ and prove that it is an stabilisation algebra. We set $\overline{t}\overline{s} = \overline{ts}$, $\overline{t}^\omega = \overline{t^\omega}$ and $\overline{t}^\sharp = \overline{t^\sharp}$. These operators are well-defined, since if $s \equiv s'$ and $t \equiv t'$, the identities $s = s'$ and $t = t'$ hold in any finite stabilisation monoid M, and thus $ss' = tt'$, $t^\omega = t'^\omega$ and $t^\sharp = t'^\sharp$ also hold in M.

We set $\overline{s} \leqslant \overline{t}$ if the identity $s \leqslant t$ holds in all finite stabilisation monoids. First of all, this definition is equivalent to the following one: $\overline{s} \leqslant \overline{t}$ if for all $s' \in \overline{s}$ and $t' \in \overline{t}$, the identity $s' \leqslant t'$ holds in all finite stabilisation monoids. Indeed, if $s' \in \overline{s}$ and $t' \in \overline{t}$, then the sequence of identities

$s' = s \leqslant t = t'$ holds in all finite stabilisation monoids and thus $s' \leqslant t'$. It is now easy to see that $\leqslant$ is an order relation.

Finally, we prove that $F(A)$ is a stabilisation algebra. By definition, $F(A)$ satisfies (A_1). As for (A_2), let $t \in F(A)$ such that $\overline{tt} = \overline{t}$. Then $tt = t$ holds in all finite stabilisation monoids. Thus, t is idempotent and $t^\omega = t$ holds in all finite stabilisation monoids, whence $\overline{t}^\omega = \overline{t^\omega} = \overline{t}$. As for ($A_3$), let us prove the first item (the other ones are similar). Let $\overline{s} \leqslant \overline{t}$ and $\overline{s'} \leqslant \overline{t'}$. The identities $s \leqslant t$ and $s' \leqslant t'$ hold in all finite stabilisation monoids. Thus, $ss' \leqslant tt'$ also holds in all finite stabilisation monoids and $\overline{s}\overline{s'} = \overline{ss'} \leqslant \overline{tt'} = \overline{t}\overline{t'}$. □

The following theorem shows that $F(A)$ is a free object, by expliciting the corresponding universal property.

Theorem 1 (Universal Property)**.** *Let M be an stabilisation algebra. Any function $h : A \to M$ can be extended into an stabilisation algebra morphism $\overline{h} : F(A) \to M$ in a unique way.*

Proof. Since $T(A)$ is a free algebra of signature $\{\cdot, \omega, \sharp, 1\}$, there is a morphism $\widetilde{h}$ from $T(A)$ to M. Let us prove that if $t \equiv s$ in $T(A)$ then $\widetilde{h}(t) = \widetilde{h}(s)$. By definition, if $t \equiv s$ then by morphism the identity $t = s$ holds in the finite stabilisation monoid M and then $\widetilde{h}(t) = \widetilde{h}(s)$ also holds in M.

Thus there is a morphism $\overline{h}$ from $F(A)$ to M such that for all $t \in T(A)$, $\overline{h}(\overline{t}) = \widetilde{h}(t)$. Futhermore, by definition $\overline{h}$ is a monoid morphism, ω-preserving and $\sharp$-preserving.

It remains to prove that it is order-preserving. If $\overline{t} \leqslant \overline{s}$ in $F(A)$, then in all finite stabilisation monoids, the identity $t \leqslant s$ holds. Since operations in stabilisation monoids are order-preserving, then $\widetilde{h}(t) \leqslant \widetilde{h}(s)$ also holds in all finite stabilisation monoids. Finally, $\overline{h}(\overline{t}) \leqslant \overline{h}(\overline{s})$.

For uniqueness, it suffices to remark that any morphism $F(A) \to M$ is completely defined by its values on A. □

4 Recognisability

Theorem A.1. *If I is a recognisable downset, then its minimal triplet is $(F(A)/{\sim_I}, h, I/ \sim_I)$, where h is the natural projection $A \to F(A)/{\sim_I}$.*

This result in the context of stabilisation algebras is given in [16], where the only differences with our statement are:

- the set that is being quotiented is $T(A)$ (called Oexpr in [16]) instead of $F(A)$,
- The syntactic object is an stabilisation algebra instead of a stabilisation algebra.

Moreover, ω is already interpretd as an idempotent power. Therefore, by Lemma **??**, all we have left to show is that the function $\pi : T(A) \to M$ as defined in [16] induces a well-defined function $\overline{\pi} : F(A) \to M$. This is straightforwad from the definition of $F(A)$: two terms of $T(A)$ correspond to the same element of $F(A)$ if and only if they cannot be distinguished by any finite stabilisation algebra. Therefore, if t_1, t_2 are equivalent in $T(A)$, we have $\pi(t_1) = \pi(t_2)$. This shows that the image of the equivalence class of t_1 in $F(A)$ is well-defined, and $\overline{\pi}$ is indeed a stabilisation algebra morphism: $F(A) \to M$.

5 Varieties

Example 8. *A recognisable downset I is* aperiodic *if for all $t \in F(A)$, the relation $t^\omega \sim_I t^\omega t$ holds.*

Proposition A.2. *Aperiodic downsets form a variety of recognisable downsets.*

Proof. First of all $\emptyset$ and $F(A)$ are aperiodic. Let $I, J \subseteq F(A)$ be aperiodic downsets. Let $t \in F(A)$. Then $t^\omega \sim_I t^\omega t$ and $t^\omega \sim_J t^\omega t$. Thus for each context C, one has $C[t^\omega] \in I$ (resp. J) if and only if $C[t^\omega t] \in I$ (resp. J). We can conclude that for any context C, we have $C[t^\omega] \in I \cup J$ (resp. $I \cap J$) if and only if $C[t^\omega t] \in I \cup J$ (resp. $I \cap J$). This means that both $I \cup J$ and $I \cap J$ are aperiodic.

Moreover, for all contexts D, $C[D[t^\omega]] \in I$ if and only if $C[D[t^\omega t]] \in I$, therefore $t^\omega \equiv_{C^{-1}[I]} t^\omega t$. So $C^{-1}[I]$ is aperiodic.

Finally, let $I \subseteq F(A)$ be an aperiodic downset and $\varphi : F(B) \to F(A)$ be a morphism. Let $t \in F(B)$, we have $\varphi(t)^\omega \sim_I \varphi(t)^\omega \varphi(t)$, so $\varphi(t^\omega) \sim_I \varphi(t^\omega t)$, and then $t^\omega \sim_{\varphi^{-1}(I)} t^\omega t$. So $\varphi^{-1}[I]$ is also aperiodic. □

Lemma 1. *Let* $\mathbf{V}$ *be a variety generated by a set* S *of finite stabilisation algebras, and* M *be a finite stabilisation algebra. Then* $M \in \mathbf{V}$ *if and only if* M *divides a finite product of elements of* S.

Proof. Let $\mathbf{W}$ denote the set of stabilisation algebras dividing a finite product of elements of S. It is clear that $S \subseteq \mathbf{W} \subseteq \mathbf{V}$. Since $\mathbf{V}$ is minimal, it is sufficient to prove that $\mathbf{W}$ is a variety to obtain $\mathbf{W} = \mathbf{V}$. The set $\mathbf{W}$ is stable under division, hence under stabilisation subalgebras and quotients. It remains to show that it is stable under product.

Let M, N be in $\mathbf{W}$, there is a surjective morphism $\varphi : M' \to M$ (resp. $\psi : N' \to N$) where M' (resp. N') is a stabilisation subalgebra of a product $M_1 \times \cdots \times M_k$ (resp. $N_1 \times \cdots \times N_p$) of elements of S. Therefore, $M' \times N'$ is a stabilisation subalgebra of $M_1 \times \cdots \times M_k \times N_1 \times \cdots \times N_p$, and $\gamma : M' \times N' \to M \times N$ defined by $\gamma(x, y) = (\varphi(x), \psi(y))$ is a surjective morphism. We obtain that $M \times N$ divides $M_1 \times \cdots \times M_k \times N_1 \times \cdots \times N_p$, and hence $M \times N \in \mathbf{W}$. □

Example 9. *A finite stabilisation algebra* M *is aperiodic if for all* $x \in M$, *we have* $x^\omega = x^\omega x$. *Aperiodic stabilisation algebras form a variety of finite stabilisation algebras.*

Proof. Let M be a finite aperiodic stabilisation algebra. If M' is a stabilisation subalgebra of M', it is clear that M' is also aperiodic. If $\varphi : M \to N$ is a surjective morphism of stabilisation algebras and $x \in N$, then there is $m \in M$ such that $x = \varphi(m)$. We then have $x^\omega = \varphi(m)^\omega = \varphi(m^\omega) = \varphi(m^\omega m) = \varphi(m^\omega)\varphi(m) = x^\omega x$. This shows that aperiodic stabilisation algebras are closed under quotient. Finally, let M_1, M_2 be finite aperiodic stabilisation algebras, and $x = (m_1, m_2) \in M_1 \times M_2$. Then $x^\omega = (m_1^\omega, m_2^\omega) = (m_1^\omega m_1, m_2^\omega m_2) = (m_1^\omega, m_2^\omega)(m_1, m_2) = x^\omega x$. This shows that finite stabilisation algebras are closed under finite products. □

Lemma A.3. *Let* $\mathbf{V}$ *be a variety of finite stabilisation algebras and let* $M \in \mathbf{V}$. *Then there exist a finite alphabet* A *and recognisable downsets* $I_1, \ldots, I_\ell \in \mathcal{V}(A)$ *such that* M *is isomorphic with a stabilisation subalgebra of* $M_1 \times \cdots \times M_\ell$, *where* M_i *is the syntactic stabilisation algebra of* I_i.

Proof. Let A be a finite alphabet and $\varphi : F(A) \to M$ a surjective morphism.

For each downset $J \subseteq M$, the set $I_J = \varphi^{-1}(J)$ is a downset of $F(A)$ of finite index, so I_J is a recognisable downset. Moreover M recognises I_J, so I_J is in $\mathcal{V}(A)$. Let M_J be the syntactic stabilisation algebra of I_J. Then there exists a surjective morphism $\pi_J : M \to M_J$. Let π be the morphism $M \to \prod_{\text{downset } J \subseteq M} M_J$ the morphism of stabilisation algebras defined by $\pi(x) = (\pi_J(x))_J$.

Let us show that π is injective. If $\pi(x) = \pi(y)$, then in particular $\pi_{\downarrow y}(x) = \pi_{\downarrow y}(y)$. This means that for all context C, $C[x] \in \downarrow y$ if and only if $C[y] \in \downarrow y$. Applying this result with the identity context $C[z] = z$, we get $x \in \downarrow y$, that means $x \leqslant y$. Similarly, $y \in \downarrow x$ and $y \leqslant x$. Thus $x = y$. This proves that π is injective and M is isomorphic with a stabilisation subalgebra of $\prod_{\text{downset } J \subseteq M} M_J$. □

Proposition A.4. *Let* $\mathbf{V}$ *and* $\mathbf{W}$ *be two varieties of finite stabilisation algebras. Suppose that* $\mathbf{V} \to \mathcal{V}$ *and* $\mathbf{W} \to \mathcal{W}$*. Then* $\mathbf{V} \subseteq \mathbf{W}$ *if and only if, for every finite alphabet* A*,* $\mathcal{V}(A) \subseteq \mathcal{W}(A)$*. In particular* $\mathbf{V} = \mathbf{W}$ *if and only if* $\mathcal{V} = \mathcal{W}$*.*

Proof. If $\mathbf{V} \subseteq \mathbf{W}$ then, by definition, for every finite alphabet A, $\mathcal{V}(A) \subseteq \mathcal{W}(A)$. For the reverse inclusion, suppose that for every finite alphabet A, $\mathcal{V}(A) \subseteq \mathcal{W}(A)$, and let $M \in \mathbf{V}$. Then by Lemma A.3, there exist a finite alphabet A and recognisable downsets $I_1, \ldots, I_\ell \in \mathcal{V}(A)$ (and hence in $\mathcal{W}(A)$) such that M is isomorphic with a stabilisation subalgebra of $M_1 \times \cdots \times M_\ell$, where M_i is the syntactic stabilisation algebra of I_i and hence M_i is in $\mathbf{W}$. Then M is also in $\mathbf{W}$. □

Lemma A.5. *Let* $\mathbf{V}$ *be a variety of finite stabilisation algebras. If* $\mathbf{V} \to \mathcal{V}$ *then* $\mathcal{V}$ *is a variety of recognisable downsets.*

Proof. Let us prove the closure properties of varieties of recognisable downsets: Let A be a finite alphabet, $I, J \in \mathcal{V}(A)$, $M_I, M_J \in \mathbf{V}$ be their syntactic stabilisation algebras and h_I, h_J be their syntactic morphisms.

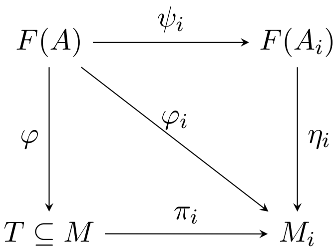
The downsets $I \cap J$ and $I \cup J$ are both recognisable by $M_I \times M_J \in \mathbf{V}$, so they are in $\mathcal{V}(A)$.

Let $C \in \mathrm{Ctx}_A$ be a context. Denote $h = h_I$. We show that $C^{-1}[I]$ is also recognisable by M, and therefore $C^{-1}[I] \in \mathcal{V}(A)$. Let $J = \{m \in M \mid C[m] \in h(I)\}$. Let $t \in F(A)$, we have $t \in C^{-1}[I]$ iff $C[t] \in I$ iff $h(C[t]) \in h(I)$ iff $C[h(t)] \in h(I)$ iff $h(t) \in J$. Therefore, (M, h, J) recognises $C^{-1}[I]$.

Let B be a finite alphabet and $\varphi : F(B) \to F(A)$ be a morphism. Then $\varphi^{-1}(I)$ is recognised by $(M, h \circ \varphi, h(I))$, so $\varphi^{-1}(I) \in \mathcal{V}(B)$. □

Theorem 2. *The correspondences* $\mathcal{V} \to \mathbf{V}$ *and* $\mathbf{V} \to \mathcal{V}$ *define mutually inverse bijective correspondences between varieties of finite stabilisation algebras and varieties of recognisable downsets.*

Proof. Suppose that $\mathcal{V} \to \mathbf{V}$ and $\mathbf{V} \to \mathcal{W}$ and let us prove that $\mathcal{V} = \mathcal{W}$. First, let $I \in \mathcal{V}(A)$, then the syntactic stabilisation algebra of I is in $\mathbf{V}$, and then $I \in \mathcal{W}(A)$. Thus $\mathcal{V} \subseteq \mathcal{W}$. Conversely, let $I \in \mathcal{W}(A)$. Then the syntactic stabilisation algebra of I (denoted by M_I) belongs to $\mathbf{V}$. Besides $\mathbf{V}$ is the variety generated by the syntactic stabilisation algebras of downsets belonging to $\mathcal{V}$, thus by Lemma 1 there is a positive integer n, and for all $1 \leqslant i \leqslant n$, a finite alphabet A_i, and a downset $I_i \in \mathcal{V}(A_i^*)$ (let us denote its syntactic monoid by M_i and its syntactic morphism by η_i) such that M_I divides $M = M_1 \times \cdots \times M_n$. Thus, there exists T a stabilisation subalgebra of M such that M_I is a quotient of T, and hence such that T recognises I. Thus, there is a morphism $\varphi : F(A) \to T$ and a downset J of T, such that $\varphi^{-1}(J) = I$. Let $\pi_i : M \to M_i$ be the i-th projection and set $\varphi_i = \pi_i \circ \varphi$. For all $i \in \{1, \ldots, n\}$ and $a \in A$, we have $\varphi_i(a) \in M_i$, and since η_i is onto, there is (at least one) $a_i \in A_i$ such that $\eta_i(a_i) = \varphi_i(a)$. We define the morphism $\psi_i : F(A) \to F(A_i)$ by $\psi_i(a) = a_i$. Remark that $\varphi_i = \eta_i \circ \psi_i$.



Let us prove that $I \in \mathcal{V}(A)$. First, we show that:

$$I = \varphi^{-1}(J) = \bigcup_{a \in J} \varphi^{-1}(\downarrow a)$$

Indeed, if $t \in \varphi^{-1}(J)$, then $\varphi(t) = a_t \in J$, so $t \in \varphi^{-1}(\{a_t\}) \subseteq \varphi^{-1}(\downarrow a_t) \subseteq \bigcup_{a \in J} \varphi^{-1}(\downarrow a)$. Conversely, if there is $a \in J$ such that $t \in \varphi^{-1}(\downarrow a)$, then $\varphi(t) \in \downarrow a$. Moreover, since J is a downset, then $\varphi(t) \in J$, and finally $t \in \varphi^{-1}(J)$.

Therefore, we just need to prove that for all $a \in M$, $\varphi^{-1}(\downarrow a) \in \mathcal{V}(A)$. Denote $a = (a_1, \ldots, a_n)$. One has $\downarrow a = \bigcap_{1 \leqslant i \leqslant n} \pi_i^{-1}(\downarrow a_i)$. So we get:

$$\varphi^{-1}(\downarrow a) = \varphi^{-1}(\bigcap_{1 \leqslant i \leqslant n} \pi_i^{-1}(\downarrow a_i)) = \bigcap_{1 \leqslant i \leqslant n} \varphi^{-1}(\pi_i^{-1}(\downarrow a_i)) = \bigcap_{1 \leqslant i \leqslant n} \varphi_i^{-1}(\downarrow a_i)$$

It is now sufficient to prove that for all i, $\varphi_i^{-1}(\downarrow a_i) \in \mathcal{V}(A)$. Since $\varphi_i^{-1}(\downarrow a_i) = \psi_i^{-1}(\eta_i^{-1}(\downarrow a_i))$ and varieties of recognisable downsets are closed under inverses of morphisms, we just need to prove that $\eta_i^{-1}(\downarrow a_i) \in \mathcal{V}(A_i)$. This result is given in Lemma A.7. Lemma A.6 is used in the proof of Lemma A.7.

Lemma A.6. *For any stabilisation algebra M and morphism $h : A \to M$ extended into a surjective morphism $h : F(A) \to M$, there is a finite set of contexts $C_M \subseteq \mathrm{Ctx}_A$ such that for all $C \in \mathrm{Ctx}_A$, there is $C' \in C_M$ that satisfies for all $t \in F(A)$: $h(C[t]) = h(C'[t])$.*

Proof. Let us define an equivalence relation over the contexts of Ctx_A: $C \equiv_c C'$ if for all $t \in F(A)$, $h(C[t]) = h(C'[t])$. For all $m \in M$, let t_m be an element of $F(A)$ such that $h(t_m) = m$. Let $C, C' \in \mathrm{Ctx}_A$ such that for all $m \in M$, $h(C[t_m]) = h(C'[t_m])$. Let us prove that $C \equiv_c C'$: let $t \in F(A)$ and $m = h(t)$. One has $h(C[t]) = h(C[t_m]) = h(C'[t_m]) = h(C'[t])$. It means that it is sufficient to consider the elements t_m to characterise equivalence classes of $\equiv_c$. In other words, the $\equiv_c$-equivalence class of C is uniquely defined by the function $f_C : M \to M$ defined by $f_C(m) = h(C[t_m])$. So $F(A)/\equiv_c$ is finite, and C_M can be defined by any finite set of representatives. ☐

Lemma A.7. *Let $\mathcal{V}$ be a variety of recognisable downsets and let $\eta : F(A) \to M$ be the syntactic morphism of a downset $I \in \mathcal{V}(A)$. Then for all $a \in M$, $\eta^{-1}(\downarrow a)$ belongs to $\mathcal{V}(A)$.*

Proof. We consider here the finite set of contexts C_M given in Lemma A.6. For more readability, we suppose that the contexts are always evaluated in the monoid, and that they can be applied directly on elements of the monoid (to avoid the use of expressions t_m). Therefore, we omit the function h in this proof.

Let $K = \eta(I)$ a downset of M and $E = \{C \in C_M \mid C[a] \in K\}$. Let us prove that:

$$\downarrow a = \bigcap_{C \in E} C^{-1}[K].$$

Indeed, if $b \in \downarrow a$, then for all $C \in E$, we have $C[b] \leqslant C[a] \in K$, so $C[b] \in K$, and $b \in C^{-1}(K)$. Conversely, suppose that $b \in \bigcap_{C \in E} C^{-1}[K]$. Let $C \in C_M$, if $C[a] \in K$ then $C \in E$ and then $b \in C^{-1}(K)$ and $C[b] \in K$. Thus $b \leqslant a$.

Finally:

$$\eta^{-1}(\downarrow a) = \bigcap_{C \in E} \eta^{-1}(C^{-1}[K]) = \bigcap_{C \in E} C^{-1}[\eta^{-1}(K)] = \bigcap_{C \in E} C^{-1}[I].$$

For any C in E, we have $C^{-1}[I] \in \mathcal{V}(A)$, and since $\mathcal{V}(A)$ is closed by finite intersection, we finally obtain that $\eta^{-1}(\downarrow a) \in \mathcal{V}(A)$. ☐

☐

6 Profinite stabilisation algebra and duality

In this section, we define a profinite monoid over $F(A)$, in the same way as it is done for A^* in [14].

6.1 Profinite stabilisation algebra

We define now a distance over $F(A)$ and we define the profinite monoid over $F(A)$ as the completion of $F(A)$ with respect to this distance.

Definition 7. *A stabilisation algebra M* separates *two elements s and t of $F(A)$ if there is a morphism $\varphi : F(A) \to M$ such that $\varphi(s) \neq \varphi(t)$.*

For $s, t \in F(A)$, define $d(s,t) = \begin{cases} 0 & \text{if } s = t \\ 2^{-n_{s,t}} & \text{otherwise} \end{cases}$
where $n_{s,t}$ is the minimum of the sizes of finite stabilisation algebras that separate s and t.

First of all, d is well defined, since if $s \neq t \in F(A)$, then by (A_1) of stabilisation algebras, that there is a finite stabilisation monoid where the identity $s = t$ is false. Such a monoid can be viewed as a finite stabilisation algebra M that separates s and t.

Moreover, d is an ultrametric distance. Indeed, $d(s,t) = 0$ if and only if $s = t$ by the previous argument. We also have $d(s,t) = d(t,s)$ by definition. And finally, it remains to prove that $d(s,r) \leqslant \max\{d(s,t), d(t,r)\}$, for $s, t, r \in F(A)$. If $s = r$ then we immediately obtain the expected result. Otherwise, let M be a stabilisation algebra of minimal size separating s and r, then either M separates s and t, or it separates t and r. If M separates s and t, then $d(s,r) \leqslant d(s,t)$, otherwise $d(s,r) \leqslant d(t,r)$.

Thus, $(F(A), d)$ is a metric space, and we denote by $\widehat{F(A)}$ its completion. Recall that $\widehat{F(A)}$ is the set of Cauchy sequences of $F(A)$, up to the equivalence relation: $(s_n) \sim (t_n)$ if $\lim_n d(s_n, t_n) = 0$.

Proposition 10. *The operations on $F(A)$ are uniformly continuous.*

Proof. Let us prove that the product, ω and $\sharp$ are uniformly continuous. Let $s_1, s_2, t_1, t_2 \in F(A)$. Since d is an ultrametric distance then we have $d(s_1s_2, t_1t_2) \leqslant \max(d(s_1s_2, t_1s_2), d(t_1s_2, t_1t_2))$. Moreover, if a finite stabilisation algebra separates s_1s_2 from t_1s_2, it also separates s_1 from t_1, so $d(s_1s_2, t_1s_2) \leqslant d(s_1, t_1)$. With a similar observation on the second pair, we get $d(s_1s_2, t_1t_2) \leqslant \max(d(s_1,t_1), d(s_2,t_2))$, which implies that the product is uniformly continuous.

Similarly, $d(s_1^\omega, s_2^\omega) \leqslant d(s_1, s_2)$ and $d(s_1^\sharp, s_2^\sharp) \leqslant d(s_1, s_2)$ since every finite stabilisation algebra that separates s_1^ω and s_2^ω (resp. $s_1^\sharp$ and $s_2^\sharp$) also separates s_1 and s_2. Therefore, ω and $\sharp$ are also uniformly continuous. $\square$

Proposition 11. *The space $\widehat{F(A)}$ is compact and can be provided with a structure of stabilisation algebra.*

Proof. First, let us prove that $\widehat{F(A)}$ is totally bounded, i.e. for any n, it can be covered by finitely many balls of size 2^{-n}. Let n be a positive integer. Define $s \sim_n t$ if s and t are not separated by any stabilisation algebra of size n. Since there is a finite number of morphisms from A^* to an stabilisation algebra of size n, this relation is of finite index. Moreover, classes for this relation are exactly the open balls of radius 2^{-n}. So $\widehat{F(A)}$ is totally bounded. A standard result in topology states that a complete totally bounded metric space is compact, that concludes the proof.

Moreover, by Proposition 10, all the operations and axioms of $F(A)$ as an stabilisation algebra can be extended to $\widehat{F(A)}$ by uniform continuity. $\square$

The stabilisation algebra $\widehat{F(A)}$ generalises the classical profinite monoid $\widehat{A^*}$ in the following way:

Proposition 12. *The morphism $\varphi : F(A) \to \widehat{A^*}$ defined by $\varphi(a) = a$, $\varphi(t^\omega) = t^\omega$ and $\varphi(t^\sharp) = t^\omega$ is uniformly continuous, and therefore can be uniquely extended into a continuous morphism of stabilisation algebras $\widehat{\varphi} : \widehat{F(A)} \to \widehat{A^*}$.*

Proof. We just have to prove that φ is a uniformly continuous morphism, with $\widehat{A^*}$ considered as an stabilisation algebra equipped with $\sharp = \omega = \pi$. It is a morphism and it is uniformly continuous because if M separates $\varphi(t)$ from $\varphi(s)$, then the same M also separates t from s, and therefore $d(\varphi(t), \varphi(s)) \leqslant d(t, s)$. □

Notice however that this morphism does not coincide with the interpretation of regular cost functions as subsets of $\widehat{A^*}$ as done in [28].

Proposition 7. *For any variety $\mathbf{V}$, $\sim_\mathbf{V}$ is a congruence on $F(A)$ and d_V is an ultrametric distance on $F(A)/\sim_\mathbf{V}$.*

Proof. First, let us prove that $\sim_\mathbf{V}$ is an equivalence relation: reflexivity and symmetry are trivial. Finally, the relation $d_\mathbf{V}(s,t) \leqslant \max(d_\mathbf{V}(s,r), d_\mathbf{V}(r,t))$ implies that $\sim_\mathbf{V}$ is transitive.

Let us prove now that $\sim_\mathbf{V}$ is a congruence. Let $s, t, r \in F(A)$ such that $d_\mathbf{V}(s,t) = 0$. This means that the image of s and t in any finite stabilisation algebra of $\mathbf{V}$ is the same. One gets $d_\mathbf{V}(sr, tr) = d_\mathbf{V}(rs, rt) = d_\mathbf{V}(s^\omega, t^\omega) = d_\mathbf{V}(s^\sharp, t^\sharp) = 0$.

Finally, by using arguments given above and the fact that $d_\mathbf{V}(s,t) = 0$ if and only if $s \sim_\mathbf{V} t$, $d_\mathbf{V}$ defines an ultrametric distance on $F(A)/\sim_\mathbf{V}$. □

7 Duality, equations and identities

7.1 Equations of lattices

An equation is said to be *explicit* if both u and v belong to $F(A)$. Given a recognisable downset I, I satisfies an explicit equation $u \to v$ if and only if $u \in I$ implies $v \in I$.

The main result, stating that a set of recognisable downsets is a lattice if and only if it is defined by a set of equations, is given in Theorem 10. The three following lemmas are used in the proof of this theorem.

Lemma A.8. *A set of equations defines a lattice of recognisable downsets.*

Proof. Let γ be a set of equations. The recognisable dowsets $\emptyset$ and $F(A)$ satisfies γ because they satisfy all the equations $u \to v$ since $u \notin \emptyset$, and $v \in \widehat{F(A)}$. If I and J satisfy $u \to v \in \gamma$, then $I \cap J$ satisfy $u \to v$: if $u \in \overline{I \cap J} = \overline{I} \cap \overline{J}$ (according to Lemma **??**) then $u \in \overline{I}$ so $v \in \overline{I}$ and $u \in \overline{J}$ so $v \in \overline{J}$, and finally $v \in \overline{I \cap J}$. A similar reasoning holds for the union. □

Lemma A.9. *Let $I, I_1, \ldots, I_n$ be recognisable downsets. If I satisfies the explicit equations that are satisfied by $I_1, I_2 \ldots I_n$ then I belongs to the lattice generated by $I_1, I_2 \ldots I_n$.*

Proof. First, if $I = F(A)$ then I belongs to the lattice generated by $I_1, I_2 \ldots I_n$. Let us suppose now that $I \neq F(A)$. For $t \in I$, let us define $K_t = \{i \in \{1, \ldots, n\} \mid t \in I_i\}$. Then:

$$I = \bigcup_{t \in I} \bigcap_{i \in K_t} I_i$$

Indeed, first let us prove that $I \subseteq \bigcup_{t \in I} \bigcap_{i \in K_t} I_i$. Let $t \in I$. If for all i, $t \notin I_i$ then for all s, the equations $t \to s$ are satisfied by $I_1, \ldots I_n$. Thus these equations are satisfied by I and since

$t \in I$ then $I = F(A)$ that contradicts the hypothesis. So, there is i such that $t \in I_i$ and thus $t \in \bigcap_{i \in K_t} I_i$. As for the reverse inclusion, if t belongs to $\bigcup_{t \in I} \bigcap_{i \in K_t} I_i$ then there is $s \in I$ such that t belongs to $\bigcap_{i \in K_s} I_i$. So the equation $s \to t$ is satisfied by $I_1, I_2 \ldots I_n$ (either $s \notin I_i$ and the equation is satisfied, either $s \in I_i$ and so $i \in K_s$ and since t belongs to $\bigcap_{i \in K_s} I_i$, then $t \in I_i$). Thus, I satisfies $s \to t$, and $t \in I$. □

Given I a recognisable downset, let us define $E_I = \{(u, v) \mid I \text{ satisfies } u \to v\}$.

Lemma A.10. *If I is a recognisable downset, then E_I is a clopen subset of $\widehat{F(A)} \times \widehat{F(A)}$.*

Proof.

$$\begin{aligned} E_I &= \{(u, v) \mid I \text{ satisfies } u \to v\} \\ &= \{(u, v) \mid v \in \bar{I} \text{ or } u \notin \bar{I}\} \\ &= (\bar{I}^c \times \widehat{F(A)}) \cup (\widehat{F(A)} \times \bar{I}) \end{aligned}$$

It is thus sufficient to prove that $\bar{I}$ is a clopen of $\widehat{F(A)}$. Indeed, in this case, $\bar{I}^c$ would also be a clopen of $\widehat{F(A)}$, and so does E_I. Let φ be the syntactic morphism of I, by Lemma **??**, we have $\bar{I} = \widehat{\varphi}^{-1}(\varphi(I))$. Moreover, $\varphi(I)$ is a clopen subset of M since M is finite, and thus $\bar{I}$ is also a clopen subset of $\widehat{F(A)}$. □

Theorem 10. *A set of recognisable downsets is a lattice if and only if it is defined by a set of equations.*

Proof. Let Γ be a lattice of recognisable downsets and let γ be the set of equations satisfied simultaneously by all the downsets of Γ. Let us prove that γ defines Γ. Let I be a recognisable downset that satisfies γ.

First by Lemma A.10, E_I and for all K, E_K^c are open sets of $\widehat{F(A)} \times \widehat{F(A)}$. Moreover, they form a covering of $\widehat{F(A)} \times \widehat{F(A)}$ Indeed, if (u, v) does not belong to $\cup_{K \in \Gamma} E_K$, then for all $K \in \Gamma$, $u \to v$ is satisfied by K. Thus $u \to v$ is satisfied by I and (u, v) belongs to E_I. Thus, since $\widehat{F(A)} \times \widehat{F(A)}$ is compact, there are $K_1, \ldots K_n \in \Gamma$ such that $E_I \cup E_{K_1}^c \cup \ldots \cup E_{K_n}^c = \widehat{F(A)} \times \widehat{F(A)}$. Let $u \to v$ be an explicit equation satisfied by $K_1, \ldots, K_n$. Then $(u, v) \notin E_{K_1}^c \cup \ldots \cup E_{K_n}^c$ and so $(u, v) \in E_I$. So I satisfies the explicit equations satisfied by $K_1, \ldots, K_n$. By applying Lemma A.9, I is in the lattice generated by $K_1, \ldots, K_n$. Since by definition, the lattice generated by $K_1, \ldots, K_n$ is contained in Γ, then $I \in \Gamma$, and γ defines Γ. □

7.2 Identities of varieties

Therefore we can define the set of pro-$\mathbf{V}$ expressions $\widehat{F_{\mathbf{V}}(A)}$ as the completion of $F(A)/\sim_{\mathbf{V}}$ with respect to $d_{\mathbf{V}}$. As before, we can show that $\widehat{F_{\mathbf{V}}(A)}$ is compact and can be provided with a structure of stabilisation algebra.

Theorem 7. *A set of finite stabilisation algebras (resp. recognisable downsets) is a variety if and only if it is defined by a set of identities of the form $u \leqslant v$.*

Proof. First, let us prove that a set of equations defines a variety. Since the class of varieties is closed under arbitrary intersections, it suffices to prove the result for a single equation $u \leqslant v$. Let M be an stabilisation algebra satisfying $u \leqslant v$. It is clear that a stabilisation subalgebra of M also satisfies $u \leqslant v$. Let $\varphi : M \to N$ be a surjective morphism, we show that N also satisfies $u \leqslant v$. Let $h_N : A \to N$, then we can choose $h_M : A \to M$ such that $h_N = \varphi \circ h_M$, by choosing $h_M(a)$

to be a φ-antecedent of $h_N(a)$ for each $a \in A$. Since M satisfies $u \leqslant v$, we get $\widehat{h_M}(u) \leqslant \widehat{h_M}(v)$, and so $\widehat{h_N}(u) = \varphi(\widehat{h_M}(u)) \leqslant \varphi(\widehat{h_M}(v)) = \widehat{h_N}(v)$, which shows that N satisfies $u \leqslant v$. Finally, let $M_1, \dots, M_k$ satisfying $u \leqslant v$ and M be their product. Let $\varphi_i : M \to M_i$ be the natural projection of M onto M_i, for each i. Let $h : A \to M$, then $\varphi_i \circ h$ is a morphism $F(A) \to M_i$, and we have for all i, $\varphi_i \circ \widehat{h}(u) \leqslant \varphi_i \circ \widehat{h}(v)$. Since the order on M is defined component-wise, we can conclude that $\widehat{h}(u) \leqslant \widehat{h}(v)$. So any set of equations of the form $u \leqslant v$ defines a variety of stabilisation algebras.

Conversely, let $\mathbf{V}$ be a variety of stabilisation algebras. Let γ be the set of equations of the form $u \leqslant v$ satisfied by all the elements of $\mathbf{V}$, and let $\mathbf{W}$ be the variety generated by γ. It is clear that $\mathbf{V} \subseteq \mathbf{W}$. Let $M \in \mathbf{W}$, we want to show that it is in $\mathbf{V}$.

Let A be a finite alphabet such that there is a surjective morphism $h : F(A) \to M$, naturally extended to a continuous $h : \widehat{F(A)} \to M$. Let π be the natural projection $\widehat{F(A)} \to \widehat{F_{\mathbf{V}}(A)}$.

Notice that by definition of $\sim_{\mathbf{V}}$, for any $u, v \in \widehat{F(A)}$, if $\pi(u) = \pi(v)$ then $h(u) = h(v)$, since u and v cannot be distinguished by any stabilisation algebra from $\mathbf{V}$. This allows us to define the continuous surjective function $g : \widehat{F_{\mathbf{V}}(A)} \to M$ as $g(\pi(u)) = h(u)$. We also define the continuous surjective function $g' : \widehat{F_{\mathbf{V}}(A)}^2 \to M^2$ by $g'(u,v) = (g(u), g(v))$. Let $D = \{(x,x) | x \in M\}$ and $E = \{(u,v) \in \widehat{F_{\mathbf{V}}(A)}^2 | g(u) = g(v)\}$, we have that $E = g'^{-1}(D)$. Since M is discrete and g' is continuous, E is a clopen in $\widehat{F_{\mathbf{V}}(A)}^2$. Let $\mathcal{M}$ be the class of morphisms from $\widehat{F_{\mathbf{V}}(A)}$ to a monoid of $\mathbf{V}$. For each $\varphi \in \mathcal{M}$, let $C_\varphi = \{(u,v) \in \widehat{F_{\mathbf{V}}(A)}^2 | \varphi(u) \neq \varphi(v)\}$. Since φ is continuous, C_φ is an open set. Observe that $E \cup (C_\varphi)_{\varphi \in \mathcal{M}}$ is an open cover of $\widehat{F_{\mathbf{V}}(A)}^2$: for any $(u,v) \in \widehat{F_{\mathbf{V}}(A)}^2$, either $g(u) = g(v)$ or there is a morphism $\varphi \in \mathcal{M}$ that can separate u and v. Since $\widehat{F_{\mathbf{V}}(A)}^2$ is compact, we can extract a finite family $F \subseteq \mathcal{M}$ such that $E \cup (C_\varphi)_{\varphi \in F}$ covers $\widehat{F_{\mathbf{V}}(A)}^2$. Therefore, if for all $\varphi \in F$ we have $\varphi(u) = \varphi(v)$, it follows that $g(u) = g(v)$. Let $S \subseteq \prod_{\varphi \in F} \varphi(\widehat{F_{\mathbf{V}}(A)})$ be the set $\{(\varphi_1(u), \varphi_2(u), \dots, \varphi_{|F|}(u)) | u \in \widehat{F_{\mathbf{V}}(A)}\}$. The set S is a stabilisation subalgebra of the finite stabilisation algebra $\prod_{\varphi \in F} \varphi(\widehat{F_{\mathbf{V}}(A)})$, and therefore $S \in \mathbf{V}$. We can now define the surjective morphism $\psi : S \to M$ by $\psi(\varphi_1(u), \varphi_2(u), \dots, \varphi_{|F|}(u)) = g(u)$. We finally get that M is a quotient of S, and therefore $M \in \mathbf{V}$. We have proved that $\mathbf{V} = \mathbf{W}$. □

8 Examples of equational descriptions of varieties and lattices

Lemma 2. *Any finite $\sharp$-commutative stabilisation algebra divides the product of its monogenic stabilisation subalgebras.*

Proof. Let M be a finite $\sharp$-commutative monoid, and N be the product of its monogenic stabilisation subalgebras. Let $\varphi : N \to M$ defined by $\varphi(x_1, \dots, x_k) = x_1 \dots x_k$. We show that φ is a surjective morphism of stabilisation algebras, which implies the wanted result. Let $x = (x_1, \dots, x_k) \in N$ and $y = (y_1, \dots, y_k) \in N$. We have $\varphi(xy) = \varphi(x_1y_1, \dots, x_ky_k) = x_1y_1 \dots x_ky_k = x_1x_2 \dots x_ky_1 \dots y_k = \varphi(x)\varphi(y)$, by commutativity of M.

If additionally $x \leqslant y$, it means that for all i we have $x_i \leqslant y_i$, and therefore by compatibility of $\leqslant$ with the product in M, we get $\varphi(x) \leqslant \varphi(y)$.

Let $x \in N$, we recall that since for any n big enough, $x^\omega = x^{n!}$ and $\varphi(x)^\omega = \varphi(x)^{n!}$, we have $\varphi(x^\omega) = \varphi(x)^\omega$.

Moreover, if x is idempotent, we have $\varphi(x^\sharp) = \varphi(x_1^\sharp, \dots, x_k^\sharp) = x_1^\sharp \dots x_k^\sharp = (x_1 \dots x_k)^\sharp = \varphi(x)^\sharp$, by $\sharp$-commutativity of M.

It is not hard to verify that φ is surjective: every element m of M is in a monogenic stabilisation submonoid of M, so we have $m = \varphi(1, \ldots, 1, m, 1, \ldots, 1)$. $\square$