Logical Methods in Computer Science
Volume 16, Issue 1, 2020, pp. 24:1–24:29
https://lmcs.episciences.org/

Submitted Mar. 01, 2018
Published Feb. 21, 2020

CONTINUITY OF FUNCTIONAL TRANSDUCERS: A PROFINITE STUDY OF RATIONAL FUNCTIONS

MICHAËL CADILHAC [a], OLIVIER CARTON [b], AND CHARLES PAPERMAN [c]

[a] DePaul University
e-mail address: michael@cadilhac.name

[b] IRIF, Université Paris Diderot
e-mail address: olivier.carton@irif.fr

[c] Université de Lille
e-mail address: charles.paperman@univ-lille.fr

Abstract. A word-to-word function is continuous for a class of languages $\mathcal{V}$ if its inverse maps $\mathcal{V}$-languages to $\mathcal{V}$. This notion provides a basis for an algebraic study of transducers, and was integral to the characterization of the sequential transducers computable in some circuit complexity classes.

Here, we report on the decidability of continuity for functional transducers and some standard classes of regular languages. To this end, we develop a robust theory rooted in the standard profinite analysis of regular languages.

Since previous algebraic studies of transducers have focused on the sole structure of the underlying input automaton, we also compare the two algebraic approaches. We focus on two questions: When are the automaton structure and the continuity properties related, and when does continuity propagate to superclasses?

1. Introduction

The algebraic theory of regular languages is tightly interwoven with fundamental questions about the computing power of Boolean circuits and logics. The most famous of these braids revolves around $\mathcal{A}$, the class of *aperiodic* or *counter-free* languages. Not only is it expressed using the logic FO[<], but it can be seen as the basic building block of AC^0, the class of languages recognized by circuit families of polynomial size and constant depth. This class is in turn expressed by the logic FO[arb] (see [25] for a lovely account). This pervasive interaction naturally suggests lifting this study to the functional level, hence to *rational functions*. This was started in [5], where it was shown that a subsequential (i.e., input-deterministic) transducer computes an AC^0 function iff it preserves the regular languages of AC^0 by inverse image. Buoyed by this clean, semantic characterization, we wish to further investigate this latter property for different classes: say that a function $f\colon A^* \to B^*$ is $\mathcal{V}$-continuous, for a class of languages $\mathcal{V}$, if for every language $L \subseteq B^*$ of $\mathcal{V}$, the language $f^{-1}(L)$ is also a

Key words and phrases: Transducers, rational functions, language varieties, continuity.
Extended version of the paper "Continuity and Rational Functions" appearing in *Proceedings ICALP'17*.

LOGICAL METHODS IN COMPUTER SCIENCE
DOI:10.23638/LMCS-16(1:24)2020

language of $\mathcal{V}$. Our main focus will be on deciding $\mathcal{V}$-continuity for rational functions; before listing our main results, we emphasize two additional motivations.

First, there has been some historical progression towards this goal. Noting, in [14], that inverse rational functions provide a uniform and compelling view of a wealth of natural operations on regular languages, Pin and Sakarovitch initiated in [15] a study of regular-continuous functions. It was already known at the time, by a result of Choffrut (see [3, Theorem 2.7]), that regular-continuity together with some uniform continuity property *characterize* functions computed by subsequential transducers. This characterization was instrumental in the study of Reutenauer and Schützenberger [20], who already noticed the peculiar link between uniform continuity for some distances on words and continuity for certain classes of languages. This link was tightened by Pin and Silva [16] who formalized a topological approach and generalized it to rational relations. More recently [17], the same authors made precise the link unveiled by Reutenauer and Schützenberger, and developed a fascinating and robust framework in which language continuity has a topological interpretation (see the beginning of Section 3, as we build upon this theory). Pin and Silva [18] notably proposed thereafter a study of functions for which continuity for a class is propagated to subclasses. In addition, Daviaud et al. [7, 6] recently explored continuity notions in the spirit of Choffrut's characterization to study weighted automata and cost-register automata.

Second, the interweaving between languages, circuits, and logic that was alluded to previously can in fact be formally stated (see again [25, 26]). A central property towards this formalization is the correspondence between "cascade products" of automata, stacking of circuits, and nesting of formulas, respectively. Strikingly, these operations can all be seen as inverse rational functions [26]. These operations are intrinsic in the construction of complex objects: languages, circuits, and formulas are often given as a sequence of simple objects to be composed (see, e.g., [24, Section 5.5]). We remark that a sufficient condition for the result of the composition to be in some given class (of languages, circuits, or logic formulas), is that each rational function be continuous for that class. Hence deciding continuity allows to give a sufficient condition for this membership question *without* computing the result of the composition, which is subject to combinatorial blowup.

Here, we report on three questions, the first two relating continuity to the other main algebraic approach to transducers, while allowing a more gentle introduction to the evaluation of *profinite words* by transducers:

- When does the transducer *structure* (i.e., its so-called *transition monoid*) impact its continuity? The results of Reutenauer and Schützenberger [20] can indeed be seen as the starting point of two distinct algebraic theories for rational functions; on the one hand the study of continuity, and on the other the study of the transition monoid of the transducer (disregarding the output). This latter avenue was explored by [8]. We show in Section 4.1:

  Theorem 1.1. *Let $\mathcal{V}$ be a variety of languages among $\mathcal{J}$, $\mathcal{R}$, $\mathcal{L}$, $\mathcal{DA}$, $\mathcal{A}$, $\mathcal{C}om$, $\mathcal{A}b$, $\mathcal{G}_{\text{nil}}$, $\mathcal{G}_{\text{sol}}$, or $\mathcal{G}$.*
  - *The statement "Any rational function* structurally *in $\mathcal{V}$ is continuous for $\mathcal{V}$" holds for $\mathcal{V} \in \{\mathcal{A}, \mathcal{G}_{\text{sol}}, \mathcal{G}\}$ and does not otherwise;*
  - *The statement "Any rational function continuous for $\mathcal{V}$ is* structurally *in $\mathcal{V}$" holds for $\mathcal{V} \in \{\mathcal{G}_{\text{nil}}, \mathcal{G}_{\text{sol}}, \mathcal{G}\}$ and does not otherwise.*

- What is the impact of *variety inclusion* on the inclusion of the related classes of continuous rational functions? When focusing on transducer structure alone, there is a natural propagation to superclasses; when is it the case for continuity? We show in Section 4.2:

Theorem 1.2. *Let $\mathcal{V}$ and $\mathcal{W}$ be two different varieties of languages among $\mathcal{J}$, $\mathcal{R}$, $\mathcal{L}$, $\mathcal{DA}$, $\mathcal{A}$, $\mathcal{C}om$, $\mathcal{A}b$, $\mathcal{G}_{\rm nil}$, $\mathcal{G}_{\rm sol}$, or $\mathcal{G}$. The statement "all rational functions continuous for $\mathcal{V}$ are continuous for $\mathcal{W}$" holds only when one of these properties is satisfied:*
– $\mathcal{V}, \mathcal{W} \in \{\mathcal{G}_{\rm nil}, \mathcal{G}_{\rm sol}, \mathcal{G}\}$ and $\mathcal{V} \subseteq \mathcal{W}$;
– $\mathcal{V} = \mathcal{A}b$ and $\mathcal{W} = \mathcal{C}om$;
– $\mathcal{V} = \mathcal{DA}$ and $\mathcal{W} = \mathcal{A}$.

- When is $\mathcal{V}$-continuity decidable for rational functions? We show in Section 5:

Theorem 1.3. *Let $\mathcal{V}$ be a variety of languages among $\mathcal{J}$, $\mathcal{R}$, $\mathcal{L}$, $\mathcal{DA}$, $\mathcal{A}$, $\mathcal{C}om$, $\mathcal{A}b$, $\mathcal{G}_{\rm sol}$, or $\mathcal{G}$. It is decidable, given an unambiguous rational transducer, whether it realizes a function continuous for $\mathcal{V}$.*

This constitutes our main contribution; note that the case $\mathcal{G}_{\rm nil}$ is left open.

Contents

2. Preliminaries

We assume some familiarity with the theory of automata and transducers, and concepts related to metric spaces (see, e.g., [3, 13] for presentations pertaining to our topic). We first settle the notation for these prerequisites.

We will use A and B for alphabets, and A^* for words over A, with 1 the empty word. For each word u, there is a smallest v, called the *primitive root* of u, such that $u = v^c$ for some c; if $c = 1$, then u is itself *primitive*. We write $|u|$ for the length of a word $u \in A^*$ and $\mathsf{alph}(u)$ for the set of letters that appear in u.

Let $L \subseteq A^*$ be a language. We write $L^{\rm c}$ for the complement of L. For a word $u \in A^*$, we write $u^{-1}L$ for $\{v \mid u \cdot v \in L\}$, and symmetrically for Lu^{-1}, these two operations being called the left and right quotients of L by u, respectively. We naturally extend concatenation and quotients to binary relations, in a component-wise fashion, e.g., for $R \subseteq A^* \times A^*$ and a pair $\rho \in A^* \times A^*$, we may use $\rho^{-1}R$ and $R\rho^{-1}$. More generally, with Id the identity relation, we will write, e.g., $Id \cdot ((x^*, x^*)\rho^{-1})$ for the pairs (ww_1, ww_2) of words such that $w_i \in x^*\rho_i^{-1}$, $i = 1, 2$.

A *variety* is a mapping $\mathcal{V}$ which associates with each alphabet A a set $\mathcal{V}(A^*)$ of regular languages closed under the Boolean operations and quotient, and such that for any morphism $h\colon A^* \to B^*$ and any $L \in \mathcal{V}(B^*)$, we have that $h^{-1}(L) \in \mathcal{V}(A^*)$. Reg is the variety that maps every alphabet A to the set $\mathrm{Reg}(A^*)$ of regular languages over A.

Given two languages $K, L \subseteq A^*$, we say that they are $\mathcal{V}$*-separable* if there is a $S \in \mathcal{V}(A^*)$ such that $K \subseteq S$ and $L \cap S = \emptyset$. Since $\mathcal{V}$ is closed under complement, K, L are $\mathcal{V}$-separable iff L, K are. Naturally, $L \in \mathcal{V}(A^*)$ iff L is $\mathcal{V}$-separable from L^c.

Transducers. A transducer τ is a 9-tuple $(Q, A, B, \delta, I, F, \lambda, \mu, \rho)$ where (Q, A, δ, I, F) forms a nondeterministic automaton (i.e., Q is a state set, A an input alphabet, $\delta \subseteq Q \times A \times Q$ a transition set, $I \subseteq Q$ a set of initial states, and $F \subseteq Q$ a set of final states), and additionally, B is an output alphabet and $\lambda\colon I \to B^*, \mu\colon \delta \to B^*, \rho\colon F \to B^*$ are the output functions. We write $\tau_{q,q'}$ for τ with $I := \{q\}$ and $F := \{q'\}$, adjusting λ and ρ to output 1 if they were undefined on these states. Similarly, $\tau_{q,\bullet}$ is τ with $I := \{q\}$ and F unchanged, and symmetrically for $\tau_{\bullet,q}$. For $q \in Q$ and $u \in A^*$, we write $q.u$ for the set of states reached from q by reading u. We assume that all the transducers and automata under study have no useless state, that is, all states appear in some accepting path.

With $w \in A^*$, let $t_1 t_2 \cdots t_{|w|} \in \delta^*$ be an accepting path for w, starting in a state $q \in I$ and ending in some $q' \in F$. The output of this path is $\lambda(q)\mu(t_1)\mu(t_2)\cdots\mu(t_n)\rho(q')$, and we write $\tau(w)$ for the set of outputs of such paths. We use τ for both the transducer and its associated partial function from A^* to subsets of B^*. Relations of the form $\{(u, v) \mid v \in \tau(u)\}$ are called *rational relations*.

The transducer τ is *unambiguous* if there is at most one accepting path for each word. In that case $\tau_{q,q'}$ is also an unambiguous transducer for any states q, q'. When τ is unambiguous, it realizes a (partial) word-to-word function: the set of functions computed by unambiguous transducers is the set of *rational functions*. Further restricting, if the underlying automaton is deterministic, we say that τ is a *subsequential* transducer. If τ is a finite union of subsequential transducers of disjoint domains, we say that τ is *plurisubsequential*.

Word distances, profinite words. For a variety $\mathcal{V}$ of regular languages, we define a distance between words for which, intuitively, two words are close if it is hard to separate them with $\mathcal{V}$ languages. Define $d_{\mathcal{V}}(u, v)$, for words $u, v \in A^*$, to be 2^{-r} where r is the size of the smallest automaton that recognizes a language of $\mathcal{V}(A^*)$ that separates $\{u\}$ from $\{v\}$; if no such language exists, then $d_{\mathcal{V}}(u, v) = 0$. It can be shown that this distance is a *pseudo-ultrametric* [13, Section VII.2]; we make only implicit and innocuous use of this fact.

The complete metric space that is the completion of (A^*, d_{Reg}) is denoted $\widehat{A^*}$ and is called the *free profinite monoid*, its elements being the *profinite words*, and the concatenation being naturally extended. By definition, if $(u_n)_{n>0}$ is a Cauchy sequence, it should hold that for any regular language L, there is a N such that either all u_n with $n > N$ belong to L, or none does. For any $x \in A^*$, define the profinite word $x^\omega = \lim x^{n!}$, and more generally, for any $c > 0$, $x^{\omega - c} = \lim x^{n! - c}$. That $(x^{n!})_{n>0}$ is a Cauchy sequence is a starting point of the profinite theory [13, Proposition VI.2.10]; it is also easily checked that $x^{c \times \omega} = \lim x^{c \times n!}$

is equal to x^ω for any integer $c \geq 1$. Given a language $L \subseteq A^*$, we write $\overline{L} \subseteq \widehat{A^*}$ for its closure, and we note that if L is regular, $\overline{L}^{\mathrm{c}} = \overline{L^{\mathrm{c}}}$—the complement being taken in $\widehat{A^*}$ in the left-hand side and in A^* in the right-hand side. Furthermore, for L' regular, $\overline{L \cup L'} = \overline{L} \cup \overline{L'}$, and similarly for intersection (see [13, Theorem VI.3.15]).

Equations. For $u, v \in \widehat{A^*}$, a language $L \subseteq A^*$ *satisfies the (profinite) equation* $u = v$ if for any words $s, t \in A^*$, $[s \cdot u \cdot t \in \overline{L} \Leftrightarrow s \cdot v \cdot t \in \overline{L}]$. Similarly, a class of languages satisfies an equation if all the languages of the class satisfy it. For a variety $\mathcal{V}$, we write $u =_{\mathcal{V}} v$, and say that u is equal to v in $\mathcal{V}$, if $\mathcal{V}(A^*)$ satisfies $u = v$. For a partial function f, $f(u) =_{\mathcal{V}} f(v)$ means that either both $f(u)$ and $f(v)$ are undefined, or they are both defined and equal in $\mathcal{V}$.

Given a set E of equations over $\widehat{A^*}$, the class of languages *defined* by E is the class of languages over A^* that satisfy all the equations of E. Reiterman's theorem shows in particular that for any variety $\mathcal{V}$ and any alphabet A, $\mathcal{V}(A^*)$ is defined by a set of equations (the precise form of which being studied in [9]).

More on varieties. Borrowing from Almeida and Costa [2], we say that a variety $\mathcal{V}$ is *supercancellative* when for any alphabet A, any $u, v \in \widehat{A^*}$ and $x, y \in A$, if $u \cdot x =_{\mathcal{V}} v \cdot y$ or $x \cdot u =_{\mathcal{V}} y \cdot v$, then $u =_{\mathcal{V}} v$ and $x = y$. This implies in particular that for any word $w \in A^*$, both $w \cdot A^*$ and $A^* \cdot w$ are in $\mathcal{V}(A^*)$. We further say that a variety $\mathcal{V}$ *separates words* if for any $s, t \in A^*$, $\{s\}$ and $\{t\}$ are $\mathcal{V}$-separable.

Our main applications revolve around some classical varieties, that we define over any possible alphabet A as follows, where x, y range over all of A^*, and a, b over A:

- $\mathcal{J}$, def. by $(xy)^\omega \cdot x = y \cdot (xy)^\omega = (xy)^\omega$
- $\mathcal{R}$, def. by $(xy)^\omega \cdot x = (xy)^\omega$
- $\mathcal{L}$, def. by $y \cdot (xy)^\omega = (xy)^\omega$
- $\mathcal{DA}$, def. by $x^\omega \cdot z \cdot x^\omega = x^\omega$ for all $z \in \mathsf{alph}(x)^*$
- $\mathcal{A}$, def. by $x^{\omega+1} = x^\omega$
- $\mathcal{C}\textsc{om}$, def. by $ab = ba$
- $\mathcal{A}\textsc{b}$, def. by $ab = ba$ and $a^\omega = 1$
- $\mathcal{G}_{\text{nil}}$, the languages rec. by nilpotent groups
- $\mathcal{G}_{\text{sol}}$, the languages rec. by solvable groups
- $\mathcal{G}$, the languages rec. by groups

The varieties included in $\mathcal{A}$ are called *aperiodic varieties* and those in $\mathcal{G}$ are called *group varieties*. Precise definitions, in particular for the group varieties, can be found in [25, 19]; we simply note that in group varieties, x^ω equals 1 for all $x \in A^*$. All these varieties except for $\mathcal{A}\textsc{b}$ and $\mathcal{C}\textsc{om}$ separate words, and only $\mathcal{DA}$ and $\mathcal{A}$ are supercancellative. They satisfy:

$$\begin{array}{ccccc} & \subsetneq & \mathcal{R} & \subsetneq & \\ \mathcal{J} = \mathcal{R} \cap \mathcal{L} & & & & \mathcal{DA} \subsetneq \mathcal{A} \\ & \subsetneq & \mathcal{L} & \subsetneq & \end{array} \qquad \begin{array}{l} \quad\;\; \subsetneq \mathcal{C}\textsc{om} \\ \mathcal{A}\textsc{b} = \mathcal{G} \cap \mathcal{C}\textsc{om} \subsetneq \mathcal{G}_{\text{nil}} \subsetneq \mathcal{G}_{\text{sol}} \subsetneq \mathcal{G} \end{array}$$

On transducers and profinite words. For a profinite word u and a state q of an unambiguous transducer τ, the set $q.u$ is well defined; indeed, with $u = \lim u_n$, the set $q.u_n$ is eventually constant, as otherwise for some state q', the domain of $\tau_{q,q'}$ would be a regular language that separates infinitely many u_n's.

A transducer $\tau\colon A^* \to B^*$ is a $\mathcal{V}$-*transducer*,[1] for a variety $\mathcal{V}$, if for some set of equations E defining $\mathcal{V}(A^*)$, for all $(u = v) \in E$ and all states q of τ, the equality $q.u = q.v$ holds. A rational function is $\mathcal{V}$-*realizable* if it is realizable by a $\mathcal{V}$-transducer.

Continuity. For a variety $\mathcal{V}$, a function $f\colon A^* \to B^*$ is $\mathcal{V}$-continuous[2] iff for any $L \in \mathcal{V}(B^*)$, $f^{-1}(L) \in \mathcal{V}(A^*)$. We mostly restrict our attention to rational functions. Since they are computed by transducers, there are countably many such functions. We note that many more Reg-continuous functions exist, in particular uncomputable ones:

Proposition 2.1. *There are uncountably many Reg-continuous functions.*

Proof. Consider a strictly increasing function $g\colon \mathbb{N} \to \mathbb{N}$. Define $f\colon \{a\}^* \to \{a\}^*$ by $f(a^n) = a^{g(n)!}$. Recall that any regular language over a unary alphabet is a finite union of languages of the form $a^i(a^j)^*$. Moreover, we have that $f^{-1}(a^i(a^j)^*)$ is finite when $i \not\equiv 0 \bmod j$, and cofinite otherwise, thus f is Reg-continuous. There are however uncountably many increasing functions g, hence uncountably many Reg-continuous functions f. □

Continuity is a formal notion of "functions being compatible with a class of languages." An equally valid notion could be to consider classes of functions that contain the characteristic functions of the languages, and closed under composition; it turns out that the largest such class coincides with the class of continuous functions. Indeed, writing $\chi_L\colon A^* \to \{0,1\}$ for the characteristic function of a language $L \subseteq A^*$:

Proposition 2.2. *Let $\mathcal{V}$ be a variety such that $\{1\} \in \mathcal{V}(\{0,1\}^*)$. Let $\mathcal{F}$ be the* largest *class of functions such that:*

(1) *For any alphabet A, $\mathcal{F} \cap \{0,1\}^{A^*} = \{\chi_L \mid L \in \mathcal{V}(A^*)\}$;*
(2) *$\mathcal{F}$ is closed under composition.*

The class $\mathcal{F}$ is well defined and it coincides with the class of $\mathcal{V}$-continuous functions.

Proof. We say that a class of functions is *good* if it satisfies properties (1) and (2). We show that the $\mathcal{V}$-continuous functions form a good class, and that any good class is included in the $\mathcal{V}$-continuous functions. This implies that there *is* a largest good class, and that it coincides with the class of $\mathcal{V}$-continuous functions, as claimed.

(Continuous functions form a good class.) Clearly, the class of $\mathcal{V}$-continuous functions is closed under composition. Now consider a $\mathcal{V}$-continuous function $f\colon A^* \to \{0,1\}$. By continuity, $L = f^{-1}(\{1\})$ is in $\mathcal{V}(A^*)$, since by hypothesis $\{1\} \in \mathcal{V}(\{0,1\}^*)$. Hence $f = \chi_L$ for some $L \in \mathcal{V}(A^*)$, concluding this step.

[1] The usual definition of $\mathcal{V}$-transducer is based on the so-called transition monoid of τ, see, e.g., [20]; the definition here is easily seen to be equivalent by [1, Lemma 3.2] and [5, Lemma 1].

[2] A note on terminology: There has been some fluctuation on the use of the term "continuous" in the literature, mostly when a possible incompatibility arises with topology. In [18], the authors use the term "preserving" in the more general context of functions from monoids to monoids. In our study, we focus on word to word functions, in which the natural topological context provides a solid basis for the use of "continuous," as used in [16, 5].

(Functions in good classes are continuous.) Let $f\colon A^* \to B^*$ be in a good class, and let $L \in \mathcal{V}(B^*)$; we ought to show that $f^{-1}(L)$ is in $\mathcal{V}(A^*)$. We have:

$$\begin{aligned} f^{-1}(L) &= f^{-1}(\chi_L^{-1}(1)) \\ &= (\chi_L \circ f)^{-1}(1) \\ &= g^{-1}(1) \ . && \text{(with } g = \chi_L \circ f\text{)} \end{aligned}$$

Note that χ_L is by hypothesis in the good class, and it being closed under composition, g also belongs to the good class. Since $g \in \{0,1\}^{A^*}$, there is a $L' \in \mathcal{V}(A^*)$ such that $g = \chi_{L'}$. This implies that $f^{-1}(L) = L'$, and it thus belongs to $\mathcal{V}(A^*)$. □

3. Continuity: The profinite approach

We build upon the work of Pin and Silva [16] and develop tools specialized to rational functions. In Section 3.1, we present a lemma asserting the equivalence between $\mathcal{V}$-continuity and the "preservation" of the defining equations for $\mathcal{V}$. In the sections thereafter, we specialize this approach to rational functions. As noted in [16], it often occurs that results about rational functions can be readily applied to the larger class of Reg-continuous functions; here, this is in particular the case for the Preservation Lemma of Section 3.1.

The connection to the classical notion of continuity is given by the next Theorem.

Theorem 3.1 [17, Theorem 4.1]**.** *Let $f\colon A^* \to B^*$. It holds that f is $\mathcal{V}$-continuous iff f is uniformly continuous for the distance $d_{\mathcal{V}}$.*

Consequently, if f is Reg-continuous then it has a unique continuous extension to the free profinite monoid with domain $\overline{f^{-1}(B^*)}$, written $\widehat{f}\colon \widehat{A^*} \to \widehat{B^*}$. The salient property of this mapping is that it is continuous in the *topological sense* (see, e.g., [13]). For our specific needs, we simply mention that it implies that for any regular language L, we have that $\widehat{f}^{-1}(\overline{L})$ is closed (that is, it is the closure of some set).

3.1. The Preservation Lemma: Continuity is equivalent to preserving equations.

The Preservation Lemma gives us a key characterization in our study: it ties together continuity and some notion of preservation of equations. This can be seen as a generalization for functions of the notion of equation satisfaction for languages. We will need the following technical lemma that extends [13, Proposition VI.3.17] from morphisms to arbitrary Reg-continuous functions; interestingly, this relies on a quite different proof.

Lemma 3.2. *Let $f\colon A^* \to B^*$ be a Reg-continuous function and L a regular language. The equality $\widehat{f}^{-1}(\overline{L}) = \overline{f^{-1}(L)}$ holds.*

Proof. First note that $f^{-1}(L) \subseteq \widehat{f}^{-1}(\overline{L})$, and that the right-hand side of this inclusion is closed. Hence $\overline{f^{-1}(L)} \subseteq \widehat{f}^{-1}(\overline{L})$.

For the converse inclusion, first write $D = f^{-1}(B^*)$, a regular language by hypothesis. We have that $\widehat{f}^{-1}(\overline{L}) = (\widehat{f}^{-1}(\overline{L}^{\mathrm{c}}))^{\mathrm{c}} \cap \overline{D}$, and similarly, $f^{-1}(L) = (f^{-1}(L^{\mathrm{c}}))^{\mathrm{c}} \cap D$. This latter equality implies that $\overline{f^{-1}(L)} = \overline{f^{-1}(L^{\mathrm{c}})}^{\mathrm{c}} \cap \overline{D}$, since $f^{-1}(L^{\mathrm{c}})$ and D are regular.

Hence the inclusion to be shown, that is, $\widehat{f}^{-1}(\overline{L}) \subseteq \overline{f^{-1}(L)}$, is equivalent to:

$$(\widehat{f}^{-1}(\overline{L}^{\mathrm{c}}))^{\mathrm{c}} \cap \overline{D} \subseteq \overline{f^{-1}(L^{\mathrm{c}})}^{\mathrm{c}} \cap \overline{D} \ ,$$

or equivalently,

$$\overline{f^{-1}(L^c)} \cup \overline{D}^c \subseteq \widehat{f}^{-1}(\overline{L}^c) \cup \overline{D}^c \ .$$

The inclusion to be shown is thus implied by $\overline{f^{-1}(L^c)} \subseteq \widehat{f}^{-1}(\overline{L}^c)$, that is, since L is regular, by $\overline{f^{-1}(L^c)} \subseteq \widehat{f}^{-1}(\overline{L^c})$. As in the proof of the converse inclusion, the right-hand side being closed, this inclusion holds. ☐

Lemma 3.3 (Preservation Lemma)**.** *Let $f\colon A^* \to B^*$ be a Reg-continuous function and E a set of equations that defines $\mathcal{V}(A^*)$. The function f is $\mathcal{V}$-continuous iff for all $(u = v) \in E$ and words $s, t \in A^*$, $\widehat{f}(s \cdot u \cdot t) =_{\mathcal{V}} \widehat{f}(s \cdot v \cdot t)$.*

Proof. (Only if.) Suppose f is $\mathcal{V}$-continuous. Let $u, v \in \widehat{A^*}$ such that $u =_{\mathcal{V}} v$, and $s, t \in A^*$. Since by $\mathcal{V}$-continuity $f^{-1}(B^*) \in \mathcal{V}(A^*)$, either both $s \cdot u \cdot t$ and $s \cdot v \cdot t$ belong to the closure of this language, or they both do not. The latter case readily yields the result, hence suppose we are in the former case.

By definition, $u = \lim u_n$ and $v = \lim v_n$ for some Cauchy sequences of words $(u_n)_{n>0}$ and $(v_n)_{n>0}$. Since $s \cdot u \cdot t =_{\mathcal{V}} s \cdot v \cdot t$, the hypothesis yields that $d_{\mathcal{V}}(s \cdot u_n \cdot t, s \cdot v_n \cdot t)$ tends to 0. By Theorem 3.1, f is uniformly continuous for $d_{\mathcal{V}}$, hence $d_{\mathcal{V}}(f(s \cdot u_n \cdot t), f(s \cdot v_n \cdot t))$ also tends to 0 (note that both $f(s \cdot u_n \cdot t)$ and $f(s \cdot v_n \cdot t)$ are defined for all n big enough). This shows that $\widehat{f}(s \cdot u \cdot t) =_{\mathcal{V}} \widehat{f}(s \cdot v \cdot t)$.

(If.) Suppose that f preserves the equations of E as in the statement. Let $L \in \mathcal{V}(B^*)$, we wish to verify that $L' = f^{-1}(L) \in \mathcal{V}(A^*)$, or equivalently by definition, that L' satisfies all the equations of E. Let $(u = v) \in E$ be one such equation, and $s, t \in A^*$; we must show that $s \cdot u \cdot t \in \overline{L'} \Leftrightarrow s \cdot v \cdot t \in \overline{L'}$.

By Lemma 3.2, since f is Reg-continuous, $\widehat{f}(\overline{L'}) = \widehat{f}(\widehat{f}^{-1}(\overline{L})) \subseteq \overline{L}$. Now let $s \cdot u \cdot t \in \overline{L'}$, we thus have that $\widehat{f}(s \cdot u \cdot t) \in \overline{L}$ (observe that $\widehat{f}(s \cdot u \cdot t)$ is indeed defined). By hypothesis, $\widehat{f}(s \cdot u \cdot t) =_{\mathcal{V}} \widehat{f}(s \cdot v \cdot t)$; now since $L \in \mathcal{V}(B^*)$, it must hold that $\widehat{f}(s \cdot v \cdot t) \in \overline{L}$. Taking the inverse image of $\widehat{f}$ on both sides, it thus holds that $s \cdot v \cdot t \in \widehat{f}^{-1}(\overline{L})$, and Lemma 3.2 then shows that $s \cdot v \cdot t \in \overline{L'}$. As the argument works both ways, this shows that $s \cdot u \cdot t \in \overline{L'} \Leftrightarrow s \cdot v \cdot t \in \overline{L'}$, concluding the proof. ☐

Continuity can be seen as preserving *membership* in $\mathcal{V}$ (by inverse image); this is where the nomenclature "$\mathcal{V}$-preserving function" of [18] stems from. Strikingly, this could also be worded as preserving *nonmembership* in $\mathcal{V}$:

Proposition 3.4. *A Reg-continuous total*[3] *function $f\colon A^* \to B^*$ is $\mathcal{V}$-continuous iff for all $L \subseteq A^*$ that do* not *belong to $\mathcal{V}(A^*)$, $f(L)$ and $f(L^c)$ are not $\mathcal{V}$-separable.*

Proof. We rely on a characterization due to Almeida [1, Lemma 3.2]: two languages K and L are $\mathcal{V}$-separable iff for all $u \in \overline{K}, v \in \overline{L}$, we have that $u \neq_{\mathcal{V}} v$.

(Only if.) Suppose f is $\mathcal{V}$-continuous, and let $L \subseteq A^*$ be a language outside $\mathcal{V}(A^*)$. There must be two profinite words $u, v \in \widehat{A^*}$ such that $u =_{\mathcal{V}} v$, $u \in \overline{L}$ and $v \in \overline{L}^c$. By $\mathcal{V}$-continuity and the Preservation Lemma, $\widehat{f}(u) =_{\mathcal{V}} \widehat{f}(v)$, and moreover, $\widehat{f}(u) \in \overline{f(L)}$ and $\widehat{f}(v) \in \overline{f(L^c)}$. The characterization above thus implies that $f(L)$ and $f(L^c)$ are not $\mathcal{V}$-separable.

[3]In all the varieties we are interested in, one can easily modify any partial function into a total function while preserving its continuity properties.

(If.) Assume that for all $L \subseteq A^*$, if $f(L)$ and $f(L^c)$ are $\mathcal{V}$-separable, then $L \in \mathcal{V}(A^*)$. For all $K \in \mathcal{V}(B^*)$, we show that $L = f^{-1}(K) \in \mathcal{V}(A^*)$. Now $f(L) \subseteq K$, and $f(L^c) = f(f^{-1}(K)^c) = f(f^{-1}(K^c)) \subseteq K^c$. Since $K \in \mathcal{V}(B^*)$, it is $\mathcal{V}$-separable from its complement, hence $f(L)$ and $f(L^c)$ are $\mathcal{V}$-separable, and our assumption implies that $L \in \mathcal{V}(A^*)$. □

3.2. **The profinite extension of rational functions.** The Preservation Lemma already hints at our intention to see transducers as computing functions from and to the free profinite monoids. Naturally, if τ is a rational function, its being Reg-continuous allows us to do so (by Theorem 3.1). For $u = \lim u_n$ a profinite word, we will write $\tau(u)$ for $\widehat{\tau}(u)$, i.e., the limit $\lim \tau(u_n)$, which exists by continuity. In this section, we develop a slightly more combinatorial approach to the evaluation of $\widehat{\tau}$, and address two classes of profinite words: those expressed as $s \cdot u \cdot t$ for s, t words and u a profinite word, and those expressed as x^ω for x a word.

Let τ be an unambiguous transducer. Recall that for any state q of τ and any profinite word u, $q.u$ is well defined. As a consequence, if s and t are words, then there is at most one initial state q_0, one $q \in q_0.s$ and one $q' \in q.u$ such that $q'.t$ is final, and these states exist iff $\tau(s \cdot u \cdot f)$ is defined. Thus:

Lemma 3.5. *Let τ be an unambiguous transducer from A^* to B^*, $s, t \in A^*$ and $u \in \widehat{A^*}$. Suppose $\tau(s \cdot u \cdot t)$ is defined, and let q_0, q, q' be the unique states such that q_0 is initial, $q \in q_0.s$, $q' \in q.u$, and $q'.t$ is final. The following holds:*

$$\tau(s \cdot u \cdot t) = \tau_{\bullet,q}(s) \cdot \tau_{q,q'}(u) \cdot \tau_{q',\bullet}(t) \ \ .$$

Let us now turn to the evaluation of ω-terms:

Lemma 3.6. *Let τ be an unambiguous transducer from A^* to B^* and $x \in A^*$. If $\tau(x^\omega)$ is defined, then there are words $s, y, t \in B^*$ such that:*

$$\tau(x^\omega) = s \cdot y^{\omega-1} \cdot t \ \ .$$

Proof. Consider a large value n; we study the behavior of $x^{n!}$ on τ. There is an initial state q_0, a state q, and a final state q_1 such that $x^{n!}$ is accepted by a path going from q_0 to q reading x^i, from q to q reading x^k with $k < n$, and from q to q_1 reading x^j. Thus the accepting path for any word of the form $x^{m!}, m > n$ is similar to the one for $x^{n!}$: from q_0 to q, looping $(m! - n!)/k + 1$ times on q, and then from q to q_1. Let thus $s = \tau_{q_0,q}(x^i)$, $z = \tau_{q,q}(x^k)$, and $t = \tau_{q,q_1}(x^j)$. It then holds that $\tau((x^k)^{m!}) = s \cdot z^{m!-(n!/k)+1} \cdot t$. Letting $c = n!/k - 1$, this shows that $\tau((x^k)^\omega) = s \cdot z^{\omega-c} \cdot t$. Now on the one hand, $(x^k)^\omega = x^\omega$, and on the other hand, we similarly have that $z^{\omega-c} = y^{\omega-1}$ by letting $y = z^c$. We thus obtain that $\tau(x^\omega) = s \cdot y^{\omega-1} \cdot t$. □

These constitute our main ways to effectively evaluate the image of profinite words through transducers. Since they are ubiquitous in our study, we will frequently apply these lemmas without explicitly citing them.

3.3. **The Syncing Lemma: Preservation Lemma applied to transducers.** We apply the Preservation Lemma on transducers and deduce a slightly more combinatorial characterization of transducers describing continuous functions. This does not provide an immediate decidable criterion, but our decidability results will often rely on it. The goal of the forthcoming lemma is to decouple, when evaluating $s \cdot u \cdot t$ (with the notations of the Preservation Lemma), the behavior of the u part and that of the s, t part. This latter part will be tested against an *equalizer* set:

Definition 3.7 (Equalizer set)**.** Let $u, v \in \widehat{A^*}$. The *equalizer set* of u and v in $\mathcal{V}$ is:

$$\mathsf{Equ}_{\mathcal{V}}(u, v) = \{(s, s', t, t') \in (A^*)^4 \mid s \cdot u \cdot t =_{\mathcal{V}} s' \cdot v \cdot t'\} \enspace .$$

Remark 3.8. The complexity of equalizer sets can be surprisingly high. For instance, letting $\mathcal{V}$ be the class of languages defined by $\{x^2 = x^3 \mid x \in A^*\}$, there is a profinite word u for which $\mathsf{Equ}_{\mathcal{V}}(u, u)$ is undecidable (this relies on the existence of arbitrarily long square-free words). On the other hand, equalizer sets quickly become less complex for common varieties; for instance, Lemma 3.12 will provide a simple form for the equalizer sets of aperiodic supercancellative varieties. ■

Definition 3.9 (Input synchronization)**.** Let $R, S \subseteq A^* \times B^*$. The *input synchronization* of R and S is defined as the relation over $B^* \times B^*$ obtained by synchronizing the first component of R and S:

$$R \bowtie S = \{(u, v) \mid (\exists s)[(s, u) \in R \land (s, v) \in S]\} \ \big(= S \circ R^{-1} \big) \enspace .$$

Naturally, the input synchronization of two rational functions is a rational relation.

Lemma 3.10 (Syncing Lemma)**.** *Let τ be an unambiguous transducer from A^* to B^* and E a set of equations that defines $\mathcal{V}(A^*)$. The function τ is $\mathcal{V}$-continuous iff:*

(1) *$\tau^{-1}(B^*) \in \mathcal{V}(A^*)$, and*
(2) *For any $(u = v) \in E$, any states p, q, any $p' \in p.u$, and any $q' \in q.v$, and letting $u' = \tau_{p,p'}(u)$ and $v' = \tau_{q,q'}(v)$:*

$$(\tau_{\bullet,p} \bowtie \tau_{\bullet,q}) \times (\tau_{p',\bullet} \bowtie \tau_{q',\bullet}) \subseteq \mathsf{Equ}_{\mathcal{V}}(u', v') \enspace .$$

Proof. We rely on the Preservation Lemma, since τ is Reg-continuous.

(Only if.) Suppose that τ is $\mathcal{V}$-continuous, the first point is immediate. For the second, we use the notation of the statement. Let $(s, s', t, t') \in (\tau_{\bullet,p} \bowtie \tau_{\bullet,q}) \times (\tau_{p',\bullet} \bowtie \tau_{q',\bullet})$. This implies that there are words $x, y \in A^*$ such that:

- $s = \tau_{\bullet,p}(x), s' = \tau_{\bullet,q}(x)$;
- $t = \tau_{p',\bullet}(y), t' = \tau_{q',\bullet}(y)$.

By Lemma 3.5, we have that $\tau(x \cdot u \cdot y) = s \cdot u' \cdot t$ and $\tau(x \cdot v \cdot y) = s' \cdot v' \cdot t'$. The Preservation Lemma then asserts that $s \cdot u' \cdot t =_{\mathcal{V}} s' \cdot v' \cdot t'$, showing that $(s, s', t, t') \in \mathsf{Equ}_{\mathcal{V}}(u', v')$.

(If.) Let $(u = v) \in E$ and $x, y \in A^*$. We must show that $\tau(x \cdot u \cdot y) =_{\mathcal{V}} \tau(x \cdot v \cdot y)$. Since $\tau^{-1}(B^*) \in \mathcal{V}(A^*)$, either τ is defined on both $x \cdot u \cdot y$ and $x \cdot v \cdot y$, or on neither; in this latter case, the equality is satisfied by definition. We thus suppose that both values are defined. This implies that there are states p, q, p', q' as in the statement, and using the same notation, letting s, s', t, t' just as above, the hypothesis yields that $s \cdot u' \cdot t =_{\mathcal{V}} s' \cdot v' \cdot t'$, showing the claim. □

3.4. **A profinite toolbox for the aperiodic setting.** In this section, we provide a few lemmas pertaining to our study of aperiodic continuity. We show that the equalizer sets of aperiodic supercancellative varieties are well behaved. Intuitively, the larger the varieties are, the more their nonempty equalizer sets will be similar to the identity. For instance, if $s \cdot x^\omega =_{\mathcal{A}} x^\omega$, for words s and x, it should hold that s and x have the same primitive root. We first note the following easy fact that will only be used in this section; it is reminiscent of the notion of *equidivisibility*, studied in the profinite context by Almeida and Costa [2].

Lemma 3.11. *Let u, v be profinite words over an alphabet A and $\mathcal{V}$ be a supercancellative variety. Suppose that there are $s, t \in A^*$ such that $u \cdot t =_{\mathcal{V}} s \cdot v$, then there is a $w \in \widehat{A^*}$ such that $u =_{\mathcal{V}} s \cdot w$ and $v =_{\mathcal{V}} w \cdot t$. If moreover $u = v$ and $\mathcal{V}$ is aperiodic, then $u =_{\mathcal{V}} s \cdot u \cdot t$.*

Proof. Let $u = \lim u_n$; if $(u_n)_{n>0}$ is ultimately constant, then this is immediate, so we assume that $|u_n|$ is unbounded. From $u \cdot t =_{\mathcal{V}} s \cdot v$, and the fact that $s \cdot A^* \in \mathcal{V}(A^*)$ by supercancellativity, we obtain that for n large enough, $u_n \cdot t \in s \cdot A^*$. Since u is nonfinite, $|u_n| > |s|$ for n large enough, in which case $u_n = s \cdot w_n$ for some sequence $(w_n)_{n>0}$. Let $w \in \widehat{A^*}$ be a limit point of this sequence, that exists by compactness (this is an essential property of the free profinite monoid, see, e.g., [13, Theorem VI.2.5]). It holds that $u = s \cdot w$. Replacing u by this value in the equation of the hypothesis, we thus have that $s \cdot w \cdot t =_{\mathcal{V}} s \cdot v$, and since $\mathcal{V}$ is supercancellative, that $v =_{\mathcal{V}} w \cdot t$.

For the last point, with $u = v$, we iterate the previous construction on w, since in that case, $u =_{\mathcal{V}} w \cdot t =_{\mathcal{V}} s \cdot w$. This provides a sequence $w = w_1, w_2, w_3, \ldots$ such that $u =_{\mathcal{V}} s^n \cdot w_n =_{\mathcal{V}} w_n \cdot t^n$. Taking a limit point x of $(w_n)_{n>0}$, it thus holds that $u =_{\mathcal{V}} s^\omega \cdot x =_{\mathcal{V}} x \cdot t^\omega$, showing, by aperiodicity, that $u =_{\mathcal{V}} s \cdot u =_{\mathcal{V}} u \cdot t$. □

Lemma 3.12. *Let u, v be profinite words over an alphabet A and $\mathcal{V}$ be an aperiodic supercancellative variety. Suppose $\mathsf{Equ}_{\mathcal{V}}(u, v)$ is nonempty. There are words $x, y \in A^*$ and two pairs $\rho_1, \rho_2 \in (A^*)^2$ such that:*

$$\mathsf{Equ}_{\mathcal{V}}(u, v) = \Big(Id \cdot \big((x^*, x^*)\rho_1^{-1}\big)\Big) \times \Big(\big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id\Big) \ .$$

Proof. Let us first establish the property for $u = v$. Assume that there are *nonempty primitive* words x, y such that $x \cdot u \cdot y =_{\mathcal{V}} u$; we show the statement of the lemma with these x and y, and $\rho_1 = \rho_2 = (1, 1)$. Note that $x^\omega \cdot u \cdot y^\omega =_{\mathcal{V}} u$, hence, since $x^{\omega+1} = x^\omega$ and similarly for y, we have that $x \cdot u =_{\mathcal{V}} u \cdot y =_{\mathcal{V}} u$. This and the fact that $\mathcal{V}$ is supercancellative show the right-to-left inclusion.

For the left-to-right inclusion, let s, s', t, t' be such that $s \cdot u \cdot t =_{\mathcal{V}} s' \cdot u \cdot t'$. Since $\mathcal{V}$ is supercancellative, this implies that the equation also holds if common prefixes of s and s' and common suffixes of t and t' are removed. We may thus assume that we are in two possible situations, by symmetry:

(1) Suppose $s' = t' = 1$, that is, $s \cdot u \cdot t =_{\mathcal{V}} u$, and that s, t are nonempty. By the same token as above, this shows that $s \cdot u =_{\mathcal{V}} u \cdot t =_{\mathcal{V}} u$. In particular, this implies that:

$$s^{|x|} \cdot u \cdot t^{|y|} =_{\mathcal{V}} x^{|s|} \cdot u \cdot y^{|t|} \ ,$$

which implies, since $\mathcal{V}$ is supercancellative, that $s^{|x|} = x^{|s|}$ and $t^{|y|} = y^{|t|}$. As x and y are primitive, this shows that $s \in x^*$ and $t \in y^*$. (Note that this holds even if one of s or t is empty.)

(2) Suppose $s = t' = 1$, that is, $u \cdot t =_{\mathcal{V}} s' \cdot u$, and that s', t are nonempty. By Lemma 3.11, we have that $u =_{\mathcal{V}} s' \cdot u \cdot t$, and we can appeal to the previous situation, showing that $s' \in x^*$ and $t \in y^*$.

(The cases where one of s, t is empty, in the first point, or one of s', t is empty, in the second, are treated similarly. Note that it is not possible for both s and s' to be nonempty, since that would imply that they start with different letters, falsifying the assumed equation by supercancellativity.)

We assumed that the x, y existed, we ought to show the other cases satisfy the claim. The two situations above show that if $\mathsf{Equ}_{\mathcal{V}}(u, u)$ is nonempty, then such x, y exist, although without the guarantee that they be nonempty. Now if $x \cdot u =_{\mathcal{V}} u$ and there are no nonempty y such that $x \cdot u \cdot y =_{\mathcal{V}} u$, this implies that there are no nonempty y such that $u \cdot y =_{\mathcal{V}} u$. Consequently, in the above case, $t = t' = 1$, and the analysis stands. This concludes the proof for the case $u = v$.

We will reduce the case $u \neq v$ to this one. Indeed, suppose that $s \cdot u \cdot v = s' \cdot v \cdot t'$. Again, by stripping away common prefixes and suffixes, we are faced with two cases:

(1) Suppose $s' = t' = 1$, that is, $s \cdot u \cdot t =_{\mathcal{V}} v$. We have that $\mathsf{Equ}_{\mathcal{V}}(u, v) = \mathsf{Equ}_{\mathcal{V}}(u, s \cdot u \cdot t)$, hence $(m, m', n, n') \in \mathsf{Equ}_{\mathcal{V}}(u, v)$ iff $(m, m' \cdot s, n, t \cdot n') \in \mathsf{Equ}_{\mathcal{V}}(u, u)$, and the result follows.
(2) Suppose $s = t' = 1$, that is, $u \cdot t =_{\mathcal{V}} s' \cdot v$. By Lemma 3.11, there is a profinite word w such that $u =_{\mathcal{V}} s' \cdot w$ and $v =_{\mathcal{V}} w \cdot t$, hence $(m, m', n, n') \in \mathsf{Equ}_{\mathcal{V}}(u, v)$ iff $(m \cdot s', m', n, t \cdot n') \in \mathsf{Equ}_{\mathcal{V}}(w, w)$, concluding the proof. □

Lemma 3.13. *Let x, y be words. For every aperiodic supercancellative variety $\mathcal{V}$, the equality $\mathsf{Equ}_{\mathcal{V}}(x^\omega, y^\omega) = \mathsf{Equ}_{\mathcal{A}}(x^\omega, y^\omega)$ holds.*

Proof. The inclusion from right to left is clear, since all equations true in $\mathcal{A}$ hold in $\mathcal{V}$.

In the other direction, let us write $u = x^{|y|}$ and $v = y^{|x|}$; we have that $u^\omega = x^\omega$ and $v^\omega = y^\omega$. Suppose $s \cdot u^\omega t =_{\mathcal{V}} s' \cdot v^\omega \cdot t'$. In particular, since $\mathcal{V}$ is supercancellative, this means that $s \cdot u^n$ is a prefix of $s' \cdot v^n$, or vice-versa, depending on whether $|s| > |s'|$ or the opposite. This implies that $u \cdot u_{\mathrm{p}} = v_{\mathrm{s}} \cdot v$ for some prefix u_{p} of u and suffix v_{s} of v. Hence (by, e.g., [11, Proposition 1.3.4]) u and v are conjugate. Their respective primitive roots are thus conjugate (by [11, Proposition 1.3.3]); writing $z \cdot z'$ and $z' \cdot z$ for them, we have that $u^\omega = (z \cdot z')^\omega$ and $v^\omega = (z' \cdot z)^\omega$.

Thus the equation above reads: $s \cdot (z \cdot z')^\omega \cdot t =_{\mathcal{V}} s' \cdot (z' \cdot z)^\omega \cdot t'$. As in the proof of Lemma 3.12, removing the common prefixes and suffixes (which we can do both in $\mathcal{V}$ and $\mathcal{A}$), we are left with two possibilities:

- Suppose $s' = t' = 1$, that is, $s \cdot (z \cdot z')^\omega \cdot t =_{\mathcal{V}} (z' \cdot z)^\omega$. The same argument as in Lemma 3.12 shows that $s \in z' \cdot (z \cdot z')^*$ and $t \in (z \cdot z')^* \cdot z$, and hence the equation holds in $\mathcal{A}$ too;
- Suppose $s = t' = 1$, that is, $(z \cdot z')^\omega \cdot t =_{\mathcal{V}} s' \cdot (z' \cdot z)^\omega$. Similarly, as $\mathcal{V}$ is supercancellative and aperiodic, this shows that $s' \in z \cdot (z' \cdot z)^*$ and $t \in (z \cdot z')^* \cdot z$, and the equation holds in $\mathcal{A}$ too, concluding the proof. □

Remark 3.14. For two aperiodic supercancellative varieties $\mathcal{V}$ and $\mathcal{W}$, we could further show that if both $\mathsf{Equ}_{\mathcal{V}}(u, v)$ and $\mathsf{Equ}_{\mathcal{W}}(u, v)$ are nonempty, then they are equal, for any profinite words u, v. It may however happen that one equalizer set is empty while the other is not; for instance, with $u = (ab)^\omega$ and $v = (ab)^\omega \cdot a \cdot (ab)^\omega$, the equalizer set of u and v in $\mathcal{DA}$ is nonempty, while it is empty in $\mathcal{A}$. ■

4. Intermezzos

We present a few facts of independent interest on continuous rational functions. Through this, we develop a few examples, showing in particular how the Preservation and Syncing Lemmas can be used to show (non)continuity. In a first part, we study when the structure of the transducer is relevant to continuity, and in a second, when the (non)inclusion of variety relates to (non)inclusion of the class of continuous rational functions.

4.1. Transducer structure and continuity.

As noted by Reutenauer and Schützenberger [20, p. 231], there exist numerous natural varieties $\mathcal{V}$ for which any $\mathcal{V}$-realizable rational function is $\mathcal{V}$-continuous. Indeed:

Proposition 4.1. *Let $\mathcal{V}$ be a variety of languages closed under inverse $\mathcal{V}$-realizable rational function. Any $\mathcal{V}$-realizable rational function is $\mathcal{V}$-continuous. This holds in particular for the varieties $\mathcal{A}$, $\mathcal{G}_{\text{sol}}$, and $\mathcal{G}$.*

Proof. This is due to a classical result of Sakarovitch [22] (see also [15]), stating, in modern parlance, that a variety $\mathcal{V}$ is closed under *block product* iff it is closed under inverse $\mathcal{V}$-realizable rational functions (note that there has been some fluctuation on vocabulary, since wreath product was used at some point to mean block product). That $\mathcal{A}$, $\mathcal{G}_{\text{sol}}$, and $\mathcal{G}$ are closed under block product is folklore. □

This naturally fails for all our other varieties, since they are not closed under inverse $\mathcal{V}$-realizable rational functions. For completeness, we give explicit constructions in the proof of the following Proposition.

Proposition 4.2. *For $\mathcal{V} \in \{\mathcal{J}, \mathcal{L}, \mathcal{R}, \mathcal{DA}, \mathcal{A}b, \mathcal{G}_{\text{nil}}, \mathcal{C}om\}$, there are $\mathcal{V}$-realizable rational functions that are not $\mathcal{V}$-continuous.*

Proof. We devise simple counter examples with $A = \{a, b\}$.

(The $\mathcal{J}, \mathcal{R}$ and $\mathcal{C}om$ cases.) Recall that $A^*a \notin \mathcal{R}(A^*) \cup \mathcal{C}om(A^*)$. The minimal unambiguous two-state transducer τ that erases all of its input except for the last letter is a $(\mathcal{J} \cap \mathcal{C}om)$-transducer; indeed, a acts in the same way as b and they are idempotent on the transducer. However, $\tau^{-1}(a) = A^*a$.

(The $\mathcal{R}$ and $\mathcal{DA}$ cases.) Consider the *Dyck language* D over A; this is the (nonregular) language of well-parenthesized expressions where a is the opening and b the closing parenthesis. Write $D^{(k)}$ for the Dyck language where parentheses are nested at most k times, for instance $D^{(0)} = 1, D^{(1)} = (ab)^*$ and $D^{(2)} = (a(ab)^*b)^*$. These languages have great importance in algebraic language theory, as they separate each level of the *dot-depth hierarchy* [4]. It holds in particular that $D^{(1)} \notin \mathcal{DA}(A^*)$.

Let τ be the rational function that removes the first letter of each block of a's and each block of b's; naturally, τ is $\mathcal{L}$-realizable. However, $\tau^{-1}(D^{k-1}) = D^k$, showing not only that τ is not continuous for $\mathcal{DA}$, but also not continuous for *any* level of the dot-depth hierarchy.

(The $\mathcal{G}_{\text{nil}}$ and $\mathcal{A}b$ cases.) Consider the two-state transducer τ where a loops on both states, and a b on one state goes to the other. When a is read on the first state, it produces a x, while all the other productions are the identity. This is an $\mathcal{A}b$-transducer. However, $\tau(aba) = xba \neq baa = \tau(baa)$, hence it is not $\mathcal{A}b$-continuous by the Preservation Lemma, since $aba =_{\mathcal{A}b} baa$. For $\mathcal{G}_{\text{nil}}$, let L be the language over $\{a, b, x\}$ with a number of x congruent

to 0 modulo 3. It can be shown that $\tau^{-1}(L) \notin \mathcal{G}_{\text{nil}}(A^*)$, intuitively since this language needs to differentiate between those a's that are an even number of b's away from the beginning of the word, and those which are not. ◻

The converse concern, that is, whether all $\mathcal{V}$-continuous rational functions are $\mathcal{V}$-realizable, was mentioned by Reutenauer and Schützenberger [20] for $\mathcal{V} = \mathcal{A}$.

Proposition 4.3. *For $\mathcal{V} \in \{\mathcal{J}, \mathcal{L}, \mathcal{R}, \mathcal{DA}, \mathcal{A}, \mathcal{A}b, \mathcal{C}om\}$, there are $\mathcal{V}$-continuous rational functions that are not $\mathcal{V}$-realizable.*

Proof. (The aperiodic cases.) Let $A = \{a\}$, a unary alphabet. Consider the transducer τ that removes every second a: its minimal transducer not being a $\mathcal{A}$-transducer, it is not $\mathcal{A}$-realizable (this is a property of subsequential transducers [20]). However, all the unary languages of $\mathcal{V}$ are either finite or co-finite, and hence for any $L \in \mathcal{V}(A^*)$, $\tau^{-1}(L)$ is either finite or co-finite, hence belongs to $\mathcal{V}(A^*)$.

(The $\mathcal{A}b$ and $\mathcal{C}om$ cases.) Over $A = \{a, b\}$, define τ to map words w in aA^* to $(ab)^{|w|}$, and words w in bA^* to $(ba)^{|w|}$. Clearly, a and b cannot act commutatively on the transducer. Now $\tau(ab) =_{\mathcal{C}om} \tau(ba)$, and moreover $\tau(x^\omega) =_{\mathcal{A}b} (ab)^\omega =_{\mathcal{A}b} 1 = \tau(1)$, hence τ is continuous for both $\mathcal{A}b$ and $\mathcal{C}om$ by the Preservation Lemma. ◻

We delay the positive answers to that question, namely for $\mathcal{G}_{\text{nil}}, \mathcal{G}_{\text{sol}}, \mathcal{G}$, to Corollary 5.8 as they constitute our main lever towards the decidability of continuity for these classes.

4.2. **Variety inclusion and inclusion of classes of continuous functions.** In this section, we study the consequence of variety (non)inclusion on the inclusion of the related classes of continuous rational functions. This is reminiscent of the notion of *heredity* studied by [17], where a function is $\mathcal{V}$-hereditarily continuous if it is $\mathcal{W}$-continuous for each subvariety $\mathcal{W}$ of $\mathcal{V}$. Variety noninclusion provides the simplest study case here:

Proposition 4.4. *Let $\mathcal{V}$ and $\mathcal{W}$ be two varieties. If $\mathcal{V} \not\subseteq \mathcal{W}$ then there are $\mathcal{V}$-continuous rational functions that are not $\mathcal{W}$-continuous.*

Proof. Let $L \in \mathcal{V}(A^*)$ be such that $L \notin \mathcal{W}(A^*)$. Define $f \colon A^* \to A^*$ as the identity function with domain L. Clearly, as $f^{-1}(K) = K \cap L$, the function f is $\mathcal{V}$-continuous. However, $f^{-1}(A^*) = L \notin \mathcal{W}(A^*)$ and $A^* \in \mathcal{W}(A^*)$, thus f is not $\mathcal{W}$-continuous. ◻

The remainder of this section focuses on a dual statement:

If $\mathcal{V} \subsetneq \mathcal{W}$, are all $\mathcal{V}$-continuous rational functions $\mathcal{W}$-continuous?

4.2.1. *The group cases.* We first focus on group varieties. Naturally, if 1. $\mathcal{V}$-continuous rational functions are $\mathcal{V}$-realizable and 2. $\mathcal{W}$-realizable rational functions are $\mathcal{W}$-continuous, this holds. Appealing to the forthcoming Corollary 5.8 for point 1 and Proposition 4.1 for point 2, we then get:

Proposition 4.5. *For $\mathcal{V}, \mathcal{W} \in \{\mathcal{G}_{\text{nil}}, \mathcal{G}_{\text{sol}}, \mathcal{G}\}$ with $\mathcal{V} \subsetneq \mathcal{W}$, all $\mathcal{V}$-continuous rational functions are $\mathcal{W}$-continuous. This however fails for $\mathcal{V} = \mathcal{A}b$ and for any $\mathcal{W} \in \{\mathcal{G}_{\text{nil}}, \mathcal{G}_{\text{sol}}, \mathcal{G}\}$.*

Proof. It remains to show the case $\mathcal{V} = \mathcal{A}_{\mathcal{B}}$. This is in fact the same example as in the proof of Proposition 4.3, to wit, over $A = \{a, b\}$, the rational function τ that maps $w \in aA^*$ to $(ab)^{|w|}$, and words $w \in bA^*$ to $(ba)^{|w|}$. Indeed, we saw that this function is continuous for $\mathcal{A}_{\mathcal{B}}$, but we have that $\tau(a) = ab$ on the one hand, and $\tau(b^\omega a) = (ba)^\omega ba =_{\mathcal{W}} ba$, but $ab \neq_{\mathcal{W}} ba$. The Preservation Lemma then shows that τ is not continuous for $\mathcal{W}$. □

Proposition 4.6. *All $\mathcal{A}_{\mathcal{B}}$-continuous rational functions are $\mathcal{C}_{\mathcal{OM}}$-continuous.*

Proof. Indeed, if $u =_{\mathcal{A}_{\mathcal{B}}} v$ with u, v words, then $u =_{\mathcal{C}_{\mathcal{OM}}} v$, since these varieties separate the same words. As $\mathcal{C}_{\mathcal{OM}}$ is defined using equations on words, this directly shows the claim by the Preservation Lemma. □

4.2.2. *The aperiodic cases.* We now turn to aperiodic varieties. For less expressive varieties, the property fails:

Proposition 4.7. *For $\mathcal{V} \in \{\mathcal{J}, \mathcal{L}, \mathcal{R}\}$ and $\mathcal{W} \in \{\mathcal{L}, \mathcal{R}, \mathcal{DA}, \mathcal{A}\}$ with $\mathcal{V} \subsetneq \mathcal{W}$, there are $\mathcal{V}$-continuous rational functions that are not $\mathcal{W}$-continuous.*

Proof. Define $\tau\colon \{a\}^* \to \{a, b\}^*$ to be the rational function that changes every other a to b; that is, $\tau(a^{2n}) = (ab)^n$, and $\tau(a^{2n+1}) = (ab)^n \cdot a$. Note that naturally, over a single letter, $a \cdot (aa)^\omega = (aa)^\omega \cdot a = a^{\omega+1}$. Now $\tau(a^\omega) = (ab)^\omega$ and $\tau(a^{\omega+1}) = (ab)^\omega \cdot a$, and since these two profinite words are equal in $\mathcal{J}$ and $\mathcal{R}$, the Preservation Lemma shows that τ is continuous for both $\mathcal{J}$ and $\mathcal{R}$. However, these two profinite words are not equal in $\mathcal{L}$, $\mathcal{DA}$, and $\mathcal{A}$, showing that τ is continuous for none of those varieties.

The remaining case, that is, showing the existence of a $\mathcal{J}$-continuous rational function that is not $\mathcal{R}$-continuous is done symmetrically, with the function mapping a^{2n} to $(ab)^n$ and a^{2n+1} to $b \cdot (ab)^n$. □

Proposition 4.8. *Any $\mathcal{DA}$-continuous rational function is $\mathcal{A}$-continuous.*

Proof. First note that both $\mathcal{DA}$ and $\mathcal{A}$ satisfy the hypotheses of Lemma 3.12. Consider a $\mathcal{DA}$-continuous rational function $\tau\colon A^* \to B^*$. By the Syncing Lemma, to show that it is $\mathcal{A}$-continuous, it is enough to show that 1. $\tau^{-1}(B^*) \in \mathcal{A}(A^*)$, and 2. That some input synchronizations of τ, based on equations of the form $x^\omega =_{\mathcal{A}} x^{\omega+1}$, belong to an equalizer set of the form (by Lemma 3.6):

$$\mathsf{Equ}_{\mathcal{A}}(\alpha \cdot y^\omega \cdot \beta,\ \alpha' \cdot z^\omega \cdot \beta') = \{(s, s', t, t') \mid (s \cdot \alpha,\ s' \cdot \alpha',\ \beta \cdot t,\ \beta' \cdot t') \in \mathsf{Equ}_{\mathcal{A}}(y^\omega,\ z^\omega)\}\ .$$

Applying the Syncing Lemma on τ for the variety $\mathcal{DA}$, we get that point 1 is true, since $\tau^{-1}(B^*) \in \mathcal{DA}(A^*)$. Similarly, point 2 is true since $x^\omega = x^{\omega+1}$ is an equation of $\mathcal{DA}$, and Lemma 3.13 implies that the equalizer set of the equation above is the same in $\mathcal{DA}$ and $\mathcal{A}$. □

However, this property does not hold beyond *rational* functions:

Proposition 4.9. *There are nonrational functions that are continuous for both $\mathcal{DA}$ and Reg but are not $\mathcal{A}$-continuous.*

Proof. Define $f\colon \{a\}^* \to \{a, b\}^*$ by $f(a^{2n}) = (ab)^n$ and $f(a^{2n+1}) = (ab)^n \cdot a \cdot (ab)^n$. We first have to check that f is indeed Reg-continuous. Given a regular language L, we define a pushdown automaton over $\{a\}^*$ that recognizes $f^{-1}(L)$; since all unary context-free languages are regular, by Parikh's theorem, this shows the claim. If the input is of the form a^{2n}, then the pushdown automaton may check that $(ab)^n \in L$, by simulating the automaton for L. If the input is of the form a^{2n+1}, then the pushdown automaton can guess the middle position of the input, and accordingly check that $(ab)^n \cdot a \cdot (ab)^n \in L$, again using the automaton for L. This concludes the construction.

The function f being Reg-continuous, consider its extension $\widehat{f}$. As in Proposition 4.7, checking that f is $\mathcal{DA}$-continuous amounts to checking that $\widehat{f}(a^\omega) =_{\mathcal{DA}} \widehat{f}(a^{\omega+1})$. The left-hand side being $(ab)^\omega$ while the right-hand side is $(ab)^\omega \cdot a \cdot (ab)^\omega$, this holds. However, these two profinite words are equal in $\mathcal{DA}$ but not in $\mathcal{A}$, hence this function is not $\mathcal{A}$-continuous, again appealing to the Preservation Lemma. ☐

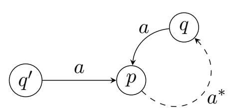
5. Deciding continuity for transducers

5.1. **Deciding continuity for group varieties.** Reutenauer and Schützenberger showed in [20] that a rational function is $\mathcal{G}$-continuous iff it is $\mathcal{G}$-realizable. Since this is proven effectively, it leads to the decidability of $\mathcal{G}$-continuity. In Proposition 4.1, we saw that the right-to-left statement also holds for $\mathcal{G}_{\text{sol}}$; we now show that the left-to-right statement holds for all group varieties $\mathcal{V}$ that contain $\mathcal{G}_{\text{nil}}$—this odd-looking condition on $\mathcal{V}$ is here to ensure that the *free group* is embedded, in a precise sense, in $\mathcal{V}$ [21, § 6.1.9]. As in [20], but with sensibly different techniques, we show that $\mathcal{V}$-continuous transducers are plurisubsequential. The Syncing Lemma will then imply that such transducers are $\mathcal{V}$-transducers. Both properties rely on an omnibus normal form which uses the three following notions:

Definition 5.1. • A transducer τ is *utilitarian* if for any two states p, q:

$$\Big[(\exists x, y)\big[\emptyset \neq (\tau_{p,\bullet} \bowtie \tau_{q,\bullet}) \subseteq (x, y) \cdot Id\big]\Big] \Rightarrow p = q \ .$$

- Within a transducer, a *pan* is a triplet (q, q', p) of the form:



  This is regardless of the outputs. The pan is *proper* if $q \neq q'$.
- Within a transducer, and for a letter a, a state q is *a-recurrent* if $p.a^\omega = \{p\}$. The transducer itself is said to be a-recurrent if all of its states are, and recurrent if it is a-recurrent for all letters.

Lemma 5.2. *Let τ be a transducer. There is a utilitarian transducer τ' computing the same function. Additionally:*

- *If τ is $\mathcal{V}$-continuous for some group variety that contains $\mathcal{G}_{\text{nil}}$, then τ' has no proper pan;*
- *For any letter a, if τ is a-recurrent, so is τ'.*

Proof. We start with the utilitarianism, then focus on the first of the "additionally" parts. The second point therein is easily satisfied by construction. Utilitarianism is shown in three steps.

(Step 1.) Define τ' to be the Cartesian product of τ with the powerset automaton of its reversal. By construction, we get the following fact, writing R_p for the language of words accepted from p, that is $\{w \mid p.w \cap F \neq \emptyset\}$:

Fact 5.3. For any states p, p' of τ', $R_p \cap R_{p'} \neq \emptyset \Rightarrow R_p = R_{p'}$.

(Step 2.) Write simply τ for the result of Step 1. In this step, we make sure that outputs are produced "as soon as possible", a process known as *normalization* (e.g., [12, Section 1.5.2]) that we sketch for completeness. For every state q, write π_q for the longest string such that $\tau_{q,\bullet}(A^*) \subseteq \pi_q \cdot B^*$. Now define the new output function (λ', μ', ρ') by letting:

$$\mu'(1,q) = \mu(1,q) \cdot \pi_q, \quad \mu'(q,a,q') = \pi_q^{-1}\mu(q,a,q') \cdot \pi_{q'}, \quad \mu'(q,1) = \pi_q^{-1}\mu(q,1) \ .$$

Write τ' for τ equipped with the output function μ'. For no state q there is a letter $b \in B^*$ such that $\tau_{q,\bullet}(A^*) \subseteq b \cdot B^*$.

(Step 3.) Write again τ for the result of the previous step. Naturally, τ still satisfies Fact 5.3. Consider two states p, q such that there are $x, y \in B^*$ satisfying $\emptyset \neq (\tau_{p,\bullet} \bowtie \tau_{q,\bullet}) \subseteq (x, y) \cdot Id$. The first part of this assumption implies that $R_p \cap R_q \neq \emptyset$, and thus, by Fact 5.3, $R_p = R_q$. In other words, $\tau_{p,\bullet}$ and $\tau_{q,\bullet}$ have the same domain. The second part of the assumption thus indicates that every production of p (resp. of q) starts with x (resp. with y), and Step 2 asserts that $x = y = 1$. Hence $\tau_{p,\bullet}$ and $\tau_{q,\bullet}$ actually compute the same function. We can thus merge them into a single state without changing the function realized. Repeating this operation results in a transducer τ' that is utilitarian.

(No proper pans.) Consider a pan (q, q', p) on a in τ'. As p can be reached from both q and q' reading a, the product $P = \tau'_{q,\bullet} \bowtie \tau'_{q',\bullet}$ is nonempty. Write $x = \tau'_{q',p}(a), y = \tau'_{p,q}(a^n), z = \tau'_{q,p}(a)$, for some n such that y is defined. We let h be the longest common suffix of x and z, and $x = x' \cdot h$ and $z = z' \cdot h$.

As τ is $\mathcal{G}$-continuous, let us apply point 2 of the Syncing Lemma on τ', the equation $(a^\omega = 1)$, from the pair of states (q', q') to (q, q'). With $Z = \tau_{\bullet,q'} \bowtie \tau_{\bullet,q'}$, a nonempty subset of the identity, we have that $Z \times P \subseteq \mathsf{Equ}_{\mathcal{V}}(x(yz)^{\omega-1}y, 1)$. We write, in the following, ν^{-1} for $\nu^{\omega-1}$, to convey the fact that ν^{-1} is the inverse of ν in $\mathcal{V}$; that is: $\nu \cdot \nu^{-1} =_{\mathcal{V}} 1$ (this analogy naturally carries further, since for instance, $(\nu\eta)^{-1} =_{\mathcal{V}} \eta^{-1} \cdot \nu^{-1}$). Let $(s, s, u, u') \in Z \times P$, then:

$$\begin{aligned} s \cdot u' &=_{\mathcal{V}} s \cdot x \cdot (yz)^{\omega-1} \cdot y \cdot u && \text{(By the Syncing Lemma)} \\ &=_{\mathcal{V}} s \cdot x \cdot z^{-1} \cdot y^{-1} \cdot y \cdot u \\ &=_{\mathcal{V}} s \cdot x \cdot z^{-1} \cdot u \\ &=_{\mathcal{V}} s \cdot (x'h) \cdot (z'h)^{-1} \cdot u =_{\mathcal{V}} s \cdot x' \cdot z'^{-1} \cdot u \ . \end{aligned}$$

By cancellation, this shows that $u' =_{\mathcal{V}} x' \cdot z'^{-1} \cdot u$. Since u' is a word and x' and z' do not share a common suffix, there is a word w such that $u = z' \cdot w$ (this is true in the *free group*, which is embedded, in a precise sense, in $\mathcal{V}$ [21, § 6.1.9]). This implies that $u' =_{\mathcal{V}} x' \cdot w$ and shows that $P \subseteq (z', x') \cdot Id$, hence that $q = q'$ by the main property of this lemma, and the pan is not proper. □ *of Lemma 5.2.*

Lemma 5.4. *Let $\mathcal{V}$ be a variety of group languages that contains $\mathcal{G}_{\text{nil}}$. For any $\mathcal{V}$-continuous unambiguous transducer, there is an equivalent plurisubsequential $\mathcal{V}$-transducer.*

Proof. Let τ be a $\mathcal{V}$-continuous unambiguous transducer. The proof is split in three facts:

Fact 5.5. There is a utilitarian recurrent transducer τ' that defines the same function as τ.

Fact 5.6. Any recurrent transducer is plurisubsequential.

Fact 5.7. Any $\mathcal{V}$-continuous utilitarian plurisubsequential transducer is a $\mathcal{V}$-transducer.

These facts together naturally imply the lemma.

Proof of Fact 5.5. We first apply Lemma 5.2 on τ. We turn the resulting transducer—which we call τ again—into a recurrent transducer one letter at a time. After each letter, we apply again Lemma 5.2, thus obtaining at the end of the process a utilitarian and recurrent transducer equivalent to τ.

In the following, for any state q, we write L_q for $\{w \mid q \in I.w\}$ and R_q for $\{w \mid q.w \cap F \neq \emptyset\}$. The preimage of τ is denoted by $L = \tau^{-1}(B^*)$. Since τ is $\mathcal{V}$-continuous, then L is in $\mathcal{V}$. Finally, we say that a state p is weakly a-recurrent if $p \in p.a^w$.

Let a be a letter. We first perform the direct product of the automaton with an automaton remembering the last weakly a-recurrent state seen. Let p be a non weakly a-recurrent state. Now τ is $\mathcal{V}$-continuous and has no proper pan, hence for any u, v such that $u \in L_p$ and $v \in R_p$, there exist q, q' such that:

$$I \xrightarrow{u} q \xrightarrow{a^\omega} q \xrightarrow{a^n} q' \xrightarrow{v} F \ .$$

By the same argument used in showing the absence of proper pan in Lemma 5.2, it holds that $p = q'$. Since p is connected to some weakly a-recurrent state, there is some $k > 0$ such that the above path can be decomposed as follows:

$$I \xrightarrow{u} q \xrightarrow{a^\omega} q \xrightarrow{a^{\omega-k}} q'' \xrightarrow{a^k} p \xrightarrow{v} F \ .$$

Since p contains the information of the last weakly a-recurrent state seen, the choice of q'' is independent from u and v. Hence the choice of q too is independent from both u and v. In particular, this shows that $L_p \subseteq L_q$. Furthermore, for any word in $u \in L_q$ and any word $v \in R_q$, we have $ua^\omega v \in \overline{L}$, and thus $uv \in L$ as well. Hence there exists a state r such that:

$$I \xrightarrow{u} r \xrightarrow{v} F \ .$$

By the same argument as above, we necessary have $r = p$, proving that $L_p = L_q$.

We now merge together all states p, q satisfying the property that $L_p = L_q$. Since, by Lemma 5.2, Fact 5.3 we have either $R_p \cap R_q = \emptyset$ or $R_p = R_q$, merging p and q will not change the function computed. After these merges, all states are a-recurrent.

We now ensure that merging these states preserves that states were recurrent for other letters. Assume that the transducer was b-recurrent. Consider p and q to be merged; since $L_p = L_q$, they are both in a cycle of b's of the exact same length, say n. Let $p' \in p.b^k$ and $q' \in q.b^k$. We finish this proof by showing that $L_{p'} = L_{q'}$, which implies that p' and q' are also merged during the process.

For any $u \in L_{p'}$, we have $ub^{n-k} \in L_p = L_q$. Hence, ub^n is in $L_{q'}$. Since the transducer is b-recurrent, then necessarily the unique state that can reach q by reading n letters b backwards is itself, proving that $u \in L_{q'}$ and so $L_{p'} \subseteq L_{q'}$. Symmetrically, $L_{q'} \subseteq L_{p'}$, entailing $L_{p'} = L_{q'}$ and concluding the proof. □ *of Fact 5.5.*

Proof of Fact 5.6. Let a be some letter and q be a state of the transducer. Consider two states p, p' in $q.a$. Since $q.a^\omega = \{q\}$, it holds that $p.a^{\omega-1} = \{q\}$, and thus that $p' \in p.a^\omega$. Hence $p = p'$, since p is a-recurrent. □ *of Fact 5.6.*

Proof of Fact 5.7. Let τ be a $\mathcal{V}$-continuous utilitarian plurisubsequential transducer. Consider an equation $u =_{\mathcal{V}} v$, a state q of τ, and let $p = q.u$ and $p' = q.v$. We show that $p = p'$, concluding this point. We rely on the Syncing Lemma, since τ is $\mathcal{V}$-continuous; it ensures in particular that:

$$(\tau_{\bullet,q} \bowtie \tau_{\bullet,q}) \times (\tau_{p,\bullet} \bowtie \tau_{p',\bullet}) \subseteq \mathsf{Equ}_{\mathcal{V}}(u', v') \quad \text{with } u' = \tau_{q,p}(u), v' = \tau_{q,p'}(v) \ . \tag{5.1}$$

Let (s, s, t_1, t_2) be in the left-hand side. We have that $s \cdot u' \cdot t_1 =_{\mathcal{V}} s \cdot v' \cdot t_2$, thus $u' \cdot t_1 =_{\mathcal{V}} v' \cdot t_2$ (here and in the following, we derive equivalent equations by appealing to the fact that the *free group* is embedded, in a precise sense, in $\mathcal{V}$ [21, § 6.1.9]). Now consider another tuple (s', s', t'_1, t'_2) again in the left-hand side of Equation (5.1). It also holds that $u' \cdot t'_1 =_{\mathcal{V}} v' \cdot t'_2$, hence we obtain that $t_1 \cdot t_2^{-1} =_{\mathcal{V}} t'_1 \cdot t_2'^{-1}$. This is in turn equal in $\mathcal{V}$ to some $\alpha \cdot \beta^{-1}$ such that α and β are words that do not share the same last letter. This shows that $t_1 = \alpha \cdot t$ and $t_2 = \beta \cdot t$ for some word t, and similarly for t'_1 and t'_2. More generally: $(\tau_{p,\bullet} \bowtie \tau_{p',\bullet}) \subseteq (\alpha, \beta) \cdot Id$, and since τ is utilitarian, $p = p'$. □ *of Fact 5.7 and Lemma 5.4.*

As an immediate corollary:

Corollary 5.8. *For $\mathcal{V} \in \{\mathcal{G}_{\text{nil}}, \mathcal{G}_{\text{sol}}, \mathcal{G}\}$, any $\mathcal{V}$-continuous rational function is $\mathcal{V}$-realizable.*

Theorem 5.9. *Let $\mathcal{V}$ be a decidable variety of group languages that includes $\mathcal{G}_{\text{nil}}$ and that is closed under inverse $\mathcal{V}$-realizable rational functions. It is decidable, given an unambiguous transducer, whether it realizes a $\mathcal{V}$-continuous function. This holds in particular for $\mathcal{G}_{\text{sol}}$ and $\mathcal{G}$.*

Proof. Lemma 5.4 together with Proposition 4.1 shows that a transducer is $\mathcal{V}$-continuous iff its equivalent transducer effectively computed by Lemma 5.4 is a $\mathcal{V}$-transducer. This latter property being testable, the result follows. □

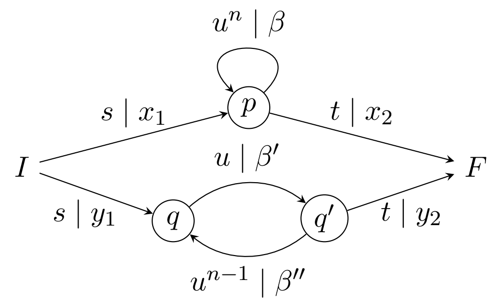
5.2. **Deciding continuity for aperiodic varieties.** We saw in Section 4.1 that the approach of the previous section cannot work: there is no correspondence between continuity and realizability for aperiodic varieties. Herein, we use the Syncing Lemma to decide continuity in two main steps. First, note that all of our aperiodic varieties are defined by an infinite number of equations for each alphabet. The Syncing Lemma would thus have us check an infinite number of conditions; our first step is to reduce this to a finite number, which we stress through the forthcoming notion of "pertaining triplet" of states. Second, we have to show that the inclusion of the second point of the Syncing Lemma can effectively be checked. This will be done by simplifying this condition, and showing a decidability property on rational relations.

We will need the following technical result in combinatorics on words in the proof of the forthcoming Lemma 5.12:

Lemma 5.10. *Let $u, v, x, y, s, t \in A^*$ be words satisfying:*
(1) $u \cdot v,\ x \cdot y \in s^*$;
(2) $v \cdot u,\ y \cdot x \in t^*$;
(3) *s and t are primitive.*
There exist $z, z' \in A^*$ such that:
(1) $s = z \cdot z'$ *and* $t = z' \cdot z$;
(2) $u,\ x \in s^* \cdot z$;
(3) $v,\ y \in t^* \cdot z'$.

Proof. Write $u = s^c \cdot z$ and $v = z' \cdot s^{c'}$ such that $s = z \cdot z'$. It follows that $v \cdot u \in (z' \cdot z)^*$, and since $z' \cdot z$ is primitive, we have that $t = z' \cdot z$. We can do the same with x and y, letting $x = s^{\dot{c}} \cdot \dot{z}$ and $y = \dot{z}' \cdot s^{\dot{c}'}$, with $s = \dot{z} \cdot \dot{z}'$. The same reasoning then shows that $t = \dot{z}' \cdot \dot{z}$. Since s has precisely $|s|$ conjugates (by [11, Proposition 1.3.2]), we have that $\dot{z} = z$ and $\dot{z}' = z'$, and the properties of the Lemma follow. □

Definition 5.11. A triplet of states (p, q, q') is *pertaining* if there are words s, u, t and an integer n such that:



where $\cdot$ means "any word." Further, a pertaining triplet is *empty* if, in the above picture, $\beta = \beta'\beta'' = 1$ and *full* if both words are nonempty; it is *degenerate* if only one of β or $\beta'\beta''$ is empty.

It is called "pertaining" as the second point of the Syncing Lemma elaborates on properties of such a triplet, in particular, since $u^\omega = u^{\omega+1}$ is an equation of $\mathcal{A}$. The following characterization of $\mathcal{A}$-continuity is then made *without appeal* to equations or profinite words:

Lemma 5.12. *A transducer* $\tau\colon A^* \to B^*$ *is* $\mathcal{A}$*-continuous iff all of the following hold:*
(1) $\tau^{-1}(B^*) \in \mathcal{A}(A^*)$*;*
(2) *For all full pertaining triplets* (p, q, q')*, there exist* $x, y \in B^*$ *and* $\rho_1, \rho_2 \in (B^*)^2$ *such that* $\tau_{\bullet,p} \bowtie \tau_{\bullet,q} \subseteq Id \cdot \big((x^*, x^*)\rho_1^{-1}\big)$ *and* $\tau_{p,\bullet} \bowtie \tau_{q',\bullet} \subseteq \big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id$*;*
(3) *For all empty pertaining triplets* (p, q, q')*, we have that* $(\tau_{\bullet,p} \bowtie \tau_{\bullet,q}) \cdot (\tau_{p,\bullet} \bowtie \tau_{q',\bullet}) \subseteq Id$*;*
(4) *No pertaining triplet is degenerate.*

Proof. (Only if.) Suppose τ is $\mathcal{A}$ -continuous, and let us appeal to the Syncing Lemma. Point 1 is then immediate. Point 2 is a direct consequence of the second point of the Syncing Lemma and of Lemma 5.12.

We shall now check point 3, by contradiction. Let (p, q, q') be an empty pertaining triplet; we use the notations of Definition 5.11. Then by functionality of τ, we have that $\tau(s \cdot u^\omega \cdot t) = x_1 \cdot x_2$ and $\tau(s \cdot u^{\omega+1} \cdot t) = y_1 \cdot y_2$. By the Preservation Lemma, and since $s \cdot u^\omega \cdot t =_{\mathcal{A}} s \cdot u^{\omega+1} \cdot t$, it should hold that $x_1 \cdot x_2 = y_1 \cdot y_2$, proving point 3.

Point 4 is proven using similar ideas as point 3: with (p, q, q') a degenerate pertaining tuple, and using the same notations as above, either the production of $s \cdot u^\omega \cdot t$ going through p is a not a finite word while the production of $s \cdot u^{\omega+1} \cdot t$ through q, q' is, or vice-versa. In both cases, it is not possible for these productions to be equal in $\mathcal{A}$, hence if such a case happens, τ cannot be $\mathcal{A}$-continuous.

(If.) We again rely on the Syncing Lemma, the first point of which being satisfied by hypothesis. Let $u^\omega = u^{\omega+1}$ be an equation of $\mathcal{A}$ with u a word; the set of such equations defines $\mathcal{A}(A^*)$. Let p, q, p', q' be states such that $p' \in p.u^\omega$ and $q' \in q.u^{\omega+1}$, and let s, t be

words with $p, q \in q_0.s$ and $p'.t, q'.t \in F$. To conclude and apply the Syncing Lemma, we need to show that:

$$\tau_{\bullet,p}(s) \cdot \tau_{p,p'}(u^\omega) \cdot \tau_{p',\bullet}(t) =_{\mathcal{A}} \tau_{\bullet,q}(s) \cdot \tau_{q,q'}(u^{\omega+1}) \cdot \tau_{q',\bullet}(t) \ . \tag{5.2}$$

(This is a direct consequence of the way profinite words are evaluated in a transducer, as per Lemma 3.5.)

Consider a large number $N = n!$, so that $p' \in p.u^N$ and $q' \in q.u^{N+1}$. With a large enough N, there must be two states P and Q, and integers i, j with $i + j = N$, such that:

- $P \in p.u^i$, $p' \in P.u^j$, and $P \in P.u^N$ (i.e., P is "between" p and p', and belongs to a loop);
- $Q \in q.u^i$, $q' \in Q.u^{j+1}$, and $Q \in Q.u^N$.

(That such a pair exists can easily be seen on the product automaton of τ by itself: The path from (p, q) to (p', q'') with $q' \in q''.u$ reading u^N must go twice through the same pair of states (P, Q), and this pair respects the above requirements.)

Now define the following words:

- $\alpha = \tau_{\bullet,P}(s \cdot u^i)$, $\beta = \tau_{P,P}(u^N)$, $\gamma = \tau_{P,\bullet}(u^j \cdot t)$,
- $\alpha' = \tau_{\bullet,Q}(s \cdot u^i)$, $\beta' = \tau_{Q,Q}(u^N)$, $\gamma' = \tau_{Q,\bullet}(u^{j+1} \cdot t)$.

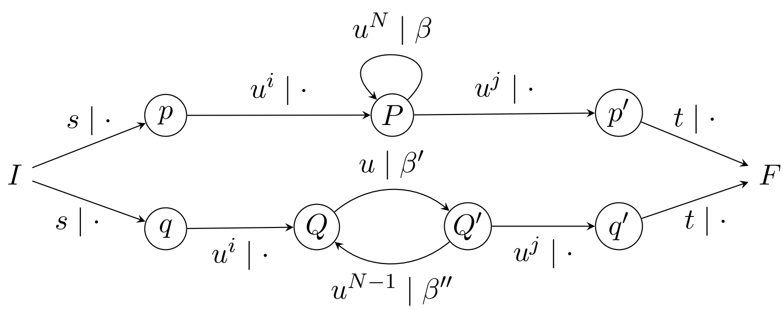
Using the same reasoning as Lemma 3.6, and the unambiguity of τ, Equation (5.2) is equivalent to:

$$\alpha \cdot \beta^{\omega-N} \cdot \gamma =_{\mathcal{A}} \alpha' \cdot \beta'^{\omega-N} \cdot \gamma' \ .$$

Naturally, since $\alpha \cdot \beta^{\omega-N} \cdot \gamma =_{\mathcal{A}} \alpha \cdot \beta^\omega \cdot \gamma$, and similarly for the right-hand side, Equation (5.2) is equivalent to:

$$\alpha \cdot \beta^\omega \cdot \gamma =_{\mathcal{A}} \alpha' \cdot \beta'^\omega \cdot \gamma' \ .$$

To make use of the hypotheses of the present Lemma, define Q' to be in $Q.u$ and such that $Q \in Q'.u^{N-1}$: (P, Q, Q') is thus pertaining. The situation is then:



Since by hypothesis this triplet cannot be degenerate, either both of β and β' are empty, or none are. Suppose they are both empty, then the hypothesis on empty triplets shows that:

$$\tau_{\bullet,P}(s \cdot u^i) \cdot \tau_{P,\bullet}(u^{N+j} \cdot t) = \tau_{\bullet,Q}(s \cdot u^i) \cdot \tau_{Q',\bullet}(u^{N+j} \cdot t) \ .$$

The left-hand side evaluates to $\alpha \cdot \gamma$. Since $\tau_{Q',\bullet}(u^{N+j} \cdot t) = \tau_{Q',Q}(u^{N-1}) \cdot \tau_{Q,\bullet}(u^{j+1} \cdot t) = \gamma'$, the right-hand side evaluates to $\alpha' \cdot \gamma'$, and Equation (5.2) is thus satisfied.

Let us thus suppose that both β and β' are nonempty. We divide β' into $b_1 b_2$ such that $b_1 = \tau_{Q,Q'}(u)$ and $b_2 = \tau_{Q',Q}(u^{N-1})$. Now let $x, y \in B^*$ and $\rho_1, \rho_2 \in (B^*)^2$ be the (pairs of) words provided by point 2 for the triplet (P, Q, Q'). Define $L = Id \cdot \big((x^*, x^*)\rho_1^{-1}\big)$ and $R = \big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id$. For any $k \geq 1$, and letting $\eta = s \cdot u^{i+k\times N}$ and $\eta' = u^{k\times N+j} \cdot t$, it holds by hypothesis that:

- $(\tau_{\bullet,P}(\eta), \tau_{\bullet,Q}(\eta)) = (\alpha \cdot \beta^k,\ \alpha' \cdot \beta'^k) \in L$; (a)
- $(\tau_{P,\bullet}(\eta), \tau_{Q',\bullet}(\eta)) = (\beta^k \cdot \gamma,\ b_2 \cdot \beta'^{k-1} \cdot \gamma') \in R$. (b)

Let us first emphasize an easy property of L and R:

Fact 5.13. If $(w \cdot w', w \cdot w'') \in L$ with $|w'|, |w''| > |x|$, then $(w', w'') \in L$. Moreover, if $(w, w') \in L$, then w is a prefix of w' or vice-versa.

Similarly, if $(w' \cdot w, w'' \cdot w) \in R$ with $|w'|, |w''| > |y|$, then $(w', w'') \in R$. Moreover, if $(w, w') \in R$, then w is a suffix of w' or vice-versa.

Proof. We only show this for L, the case for R being similar.

For the first part of the statement, the hypothesis ensures the existence of a word z, integers n', n'', and two prefixes x', x'' of x such that $w \cdot w' = z \cdot x^{n'} \cdot x'$ and $w \cdot w'' = z \cdot x^{n''} \cdot x''$. If w is a prefix of z, the property is easy to verify. In the other cases, $w = z \cdot x^n \cdot \chi$ for some integer $n < n', n''$ (strictness coming from the hypothesis) and $x = \chi\chi'$. Hence $w' = \chi' \cdot x^{n'-n-1} \cdot x'$ and $w'' = \chi' \cdot x^{n''-n-1} \cdot x'$, and thus both belong to L. The case of R is similar.

For the second part of the statement, w and w' start with a common word z, then some repetitions of x, and a prefix of x. Clearly, one has to be a prefix of the other. □ *of Fact 5.13.*

We first focus on the consequences of (a). First, since either $\alpha \cdot \beta^k$ is a prefix of $\alpha' \cdot \beta'^k$ or vice-versa, we have that either α is a prefix of $\alpha' \cdot \beta'^k$, or α' a prefix of $\alpha \cdot \beta^k$, for some k. Suppose for instance that $\alpha' = \alpha \cdot \beta^c \cdot \beta_{\mathrm{p}}$, with $\beta = \beta_{\mathrm{p}} \cdot \beta_{\mathrm{s}}$; the other case will be treated later. Appealing to Fact 5.13, for k big enough, factoring out α' yields that $((\beta_{\mathrm{s}}\beta_{\mathrm{p}})^{k-c-1}\beta_{\mathrm{s}}, \beta'^k) \in L$. Hence $(\beta_{\mathrm{s}}\beta_{\mathrm{p}})^*$ and β'^* share common prefixes of unbounded length, implying that $\beta_{\mathrm{s}}\beta_{\mathrm{p}}$ and β' are powers of a same primitive word z_1 (by, e.g., [11, Proposition 1.3.5]).

Now similarly focusing on (b), we obtain that γ is a suffix of $\beta'^k \cdot \gamma'$ or γ' is a suffix of $\beta^k \cdot \gamma$, for some k. Suppose for instance that $\gamma = \beta'_{\mathrm{s}} \cdot \beta'^{c'} \cdot \gamma'$, with $\beta' = \beta'_{\mathrm{p}} \cdot \beta'_{\mathrm{s}}$, again delaying the other case. It follows, just as above, that β and $\beta'_{\mathrm{s}}\beta'_{\mathrm{p}}$ are powers of a same primitive word z_2. Noting that $(\eta^c)^\omega = \eta^\omega$, for any η, Equation (5.2) is thus equivalent to:

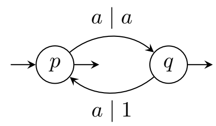
$$\alpha \cdot z_2^\omega \cdot \beta'_{\mathrm{s}} \cdot \beta'^{c'} \cdot \gamma' =_{\mathcal{A}} \alpha \cdot \beta^c \cdot \beta_{\mathrm{p}} \cdot z_1^\omega \cdot \gamma' \ .$$

Lemma 5.10 indicates that there exist words z, z' such that $z_1 = z \cdot z'$, $z_2 = z' \cdot z$, and $\beta'_{\mathrm{s}} \in z_2^* \cdot z', \beta_{\mathrm{p}} \in z' \cdot z_1^*$. By eliminating α and γ' we thus obtain that there are some integers n_1, n_2 such that Equation (5.2) is equivalent to $z_2^\omega \cdot z' \cdot z_1^{n_1 \times c'} =_{\mathcal{A}} z_2^{n_2 \times c} \cdot z' \cdot z_1^\omega$, which clearly holds as both sides evaluate to $(z' \cdot z)^\omega \cdot z'$.

(Remaining cases.) We made two suppositions: α' is a prefix of α, and γ is a suffix of γ'. The case where α is a prefix of α' and γ' a suffix of γ is entirely symmetric. Let us keep our supposition on α' and assume that γ' is a suffix of γ; the last remaining case is similar to this one.

Let us thus write $\gamma' = \dot\beta_{\mathrm{s}} \cdot \beta^{c'} \cdot \gamma$, with $\beta = \dot\beta_{\mathrm{p}} \cdot \dot\beta_{\mathrm{s}}$. We then obtain, factoring out γ' this time, that $(\beta^{k-c-1}\dot\beta_{\mathrm{p}}, \beta'^k) \in R$. This implies that $(\dot\beta_{\mathrm{s}}\dot\beta_{\mathrm{p}})$ and β' are powers of the same primitive word, which can only be z_1. Writing z_2 for the primitive root of β, Lemma 5.10 shows the existence of words z, z' such that $z_1 = z \cdot z'$, $z_2 = z' \cdot z$, and $\beta_{\mathrm{p}} \in z_2^* \cdot z', \dot\beta_{\mathrm{s}} \in z \cdot z_2^*$. By eliminating α and γ, we similarly obtain that Equation (5.2) is equivalent, for some n_1, n_2, to $z_2^\omega =_{\mathcal{A}} z_2^{n_1 \times c} \cdot \beta_{\mathrm{p}} \cdot z_1^\omega \cdot \dot\beta_{\mathrm{s}} \cdot z_2^{n_2 \times c'}$. Then both sides evaluate to z_2^ω, hence Equation (5.2) holds. □ *of Lemma 5.12.*

Example 5.14. We show that the transducer of Proposition 4.3 is $\mathcal{A}$-continuous. Let τ be:



First, the function is total, hence the first point of Lemma 5.12 is satisfied. Second, there are no empty nor degenerate pertaining triplets, hence the third and fourth points are satisfied. Now the full pertaining triplets are (p,p,p), (p,p,q), (q,q,q), and (q,q,p). We check that the pertaining triplet (p,p,q) satisfies the second condition of Lemma 5.12, the other cases being similar or clear. The first half of the condition is immediate. Now $\tau_{p,\bullet} \bowtie \tau_{q,\bullet} = \{(a^{\lfloor n+1/2 \rfloor}, a^{\lfloor n/2 \rfloor}) \mid n \geq 0\}$ which satisfies the condition. ∎

We now show that the property of Lemma 5.12 is indeed decidable:

Proposition 5.15. *It is decidable, given a rational relation $R \subseteq A^* \times A^*$, whether there is a word $x \in A^*$ and a pair $\rho \in (A^*)^2$, such that $R \subseteq Id \cdot \big((x^*, x^*)\rho^{-1}\big)$.*

Proof. We rely on the classical result that it is decidable whether a rational relation is included in the identity [23, p. 650].

We first tackle a related, simpler decision problem: Given a rational relation $R \subseteq (A^* \times A^*)$ and a word $x \in A^*$, check whether $R \subseteq Id \cdot (x^*, x^*)$. Write $f \colon A^* \to A^*$ for the function that removes the longest suffix in x^* of its argument, and note that f is a rational function. Closure under inverse and composition of rational relations implies that $R' = \{(f(u), f(v)) \mid (u,v) \in R\}$ is a rational relation computable from R. We have that $R' \subseteq Id$ if and only if $R \subseteq Id \cdot (x^*, x^*)$, hence the decision problem at hand is equivalent to checking whether $R' \subseteq Id$, which is decidable.

We now reduce the main decision problem to the previous one.

First, we note that if a solution (x, ρ) exists, then there is another solution $(\dot{x}, \dot{\rho})$ with one component of $\dot{\rho}$ empty. Indeed, write $x' = x\rho_1^{-1}, x'' = x\rho_2^{-1}$. Assume x' is also a prefix of x'' (the symmetric case being similar). We may thus write $x = x'y$ and $x'' = x'z$, and have that $R \subseteq Id \cdot \big((yx')^*, (yx')^* z\big)$, showing that $\dot{x} := yx'$ and $\dot{\rho} := (1, z^{-1}yx')$ fit the requirements.

We thus task ourselves with finding a solution (x, ρ) with ρ_2 empty (the symmetric case being similar). We first check that $R \subseteq Id$. If this is not the case, we can compute a pair $(u,v) \in R \setminus Id$ (again by [23, p. 650]). All the suffixes of u are candidates for $x\rho_1^{-1}$; we go through all these candidates x' (including the empty word). We say that x' is a *valid choice* if it stems from a valid solution (x, ρ).

Next, we verify that all pairs $(u,v) \in R$ are such that u ends with x' (this is decidable, e.g., since $R \cap \big((A^*x')^{\mathrm{c}} \times A^*\big) = \emptyset$ is decidable [3, Proposition 2.6, Proposition 8.2]). If this is not the case, then x' is not a valid choice. Otherwise, let us write $R' = R \cdot (x', 1)^{-1}$, a rational relation.

We now check again that $R' \subseteq Id$; if it is the case, we are done and values of (x, ρ) can be deduced. Otherwise, we are given a pair $(u,v) \in R' \setminus Id$. Now if x' is a valid choice, then either u is a prefix of v, or vice-versa. In the former case, write $v = u \cdot z$; if x' is a valid choice, then $z \in x^*$, and this provides us with candidates for x: all the possible roots z' of z. We may now test that one such z' starts with x', and check whether $R \subseteq Id \cdot (z'^*, z'^*)$ using the above decision problem. If this holds, then there do exist an x and a ρ satisfying $R \subseteq Id \cdot \big((x^*, x^*)\rho^{-1}\big)$. Moreover, if such words exist, this procedure will find them. □

Remark 5.16. In general, the problem of deciding, given a rational relation R and a *recognizable* relation K, whether $R \subseteq Id \cdot K$, is undecidable. Indeed, testing $R \cap Id = \emptyset$ is undecidable [3], and equivalent to testing:

$$R \subseteq Id \cdot \Big((A^+ \times \{1\}) \cup (\{1\} \times A^+) \cup \bigcup_{a \neq b \in A} (a \cdot A^* \times b \cdot A^*) \Big) \enspace ,$$

the right-hand side being of the form $Id \cdot K$. ■

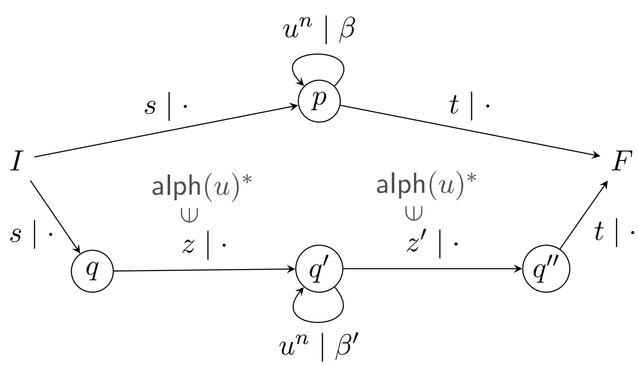
Theorem 5.17. *It is decidable, given an unambiguous transducer, whether it realizes an $\mathcal{A}$-continuous function.*

Proof. This is a consequence of Lemma 5.12: Given a transducer, one can list all its pertaining triplets, and whether they are empty, full, or degenerate. For full pertaining triplets, the property of Lemma 5.12 is checked with Proposition 5.15 and the same Proposition applied on the reverse of the transducer. The property for empty triplets can be checked since the inclusion of a rational relation in Id is decidable. □

The rest of this section focuses on conditions *à la* Lemma 5.12 for $\mathcal{J}$, $\mathcal{R}$, $\mathcal{L}$, and $\mathcal{DA}$. In each of these cases, we define the proper notion of "pertaining" and rewrite the conditions of Lemma 5.12 to match the defining equations. Since the proofs are simple variants of that of Lemma 5.12, we omit them; we note that in each case, the conditions are effectively verifiable.

5.2.1. *The case of $\mathcal{J}$.* We use a different set of equations to define $\mathcal{J}$, that can easily be proved to be equivalent to the one given in the Preliminaries. Specifically, $\mathcal{J}$ is defined over any alphabet A by the set of equations $x^\omega = y \cdot x^\omega \cdot z$, with $y, z \in \mathsf{alph}(x)$. The definition of "pertaining" then reads as follows:

Definition 5.18. For two alphabets C, D, a quadruplet of states (p, q, q', q'') is (C, D)-*pertaining* if there are words s, u, t with $\mathsf{alph}(u) = C$, words $z, z' \in C^*$, and an integer n such that:



and moreover $\mathsf{alph}(\beta) \cup \mathsf{alph}(\beta') = D$. The pertaining quadruplet is *empty* if $D = \emptyset$; it is *full* if $\mathsf{alph}(\beta) = \mathsf{alph}(\beta') \neq \emptyset$, and *degenerate* otherwise.

Lemma 5.19. *A transducer $\tau \colon A^* \to B^*$ is $\mathcal{J}$-continuous iff all of the following hold:*
(1) $\tau^{-1}(B^*) \in \mathcal{J}(A^*)$;

(2) *For all full (C,D)-pertaining quadruplets (p,q,q',q''):*

$$\tau_{\bullet,p} \bowtie \Big(\tau_{\bullet,q} \cdot \big(\varepsilon, \tau_{q,q'}(C^*)\big) \cdot \tau_{q',q'}\Big) \subseteq Id \cdot (D^*, D^*) \ \text{ and}$$
$$\tau_{p,\bullet} \bowtie \Big(\tau_{q',q'} \cdot \big(\varepsilon, \tau_{q',q''}(C^*)\big) \cdot \tau_{q'',\bullet}\Big) \subseteq (D^*, D^*) \cdot Id \ ;$$

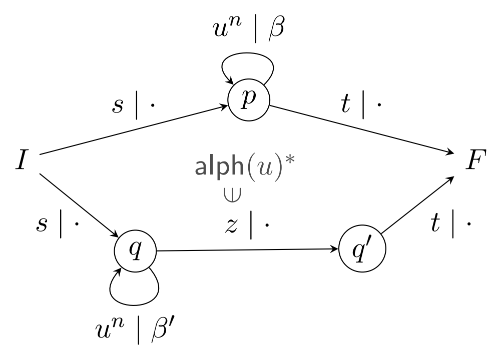
(3) *For all empty (C,D)-pertaining quadruplets (p,q,q',q''):*

$$(\tau_{\bullet,p} \cdot \tau_{p,\bullet}) \bowtie \Big(\tau_{\bullet,q} \cdot \big(\varepsilon, \tau_{q,q'}(C^*)\big) \cdot \tau_{q',q'} \cdot \big(\varepsilon, \tau_{q',q''}(C^*)\big) \cdot \tau_{q'',\bullet}\Big) \subseteq Id \ ;$$

(4) *No pertaining quadruplet is degenerate.*

5.2.2. *The case of $\mathcal{R}$.* Again, we slightly diverge from the usual equations for $\mathcal{R}$, as presented in the Preliminaries. Indeed, $\mathcal{R}$ is also defined, over any alphabet A, by $x^\omega = x^\omega \cdot y$ with $y \in \mathsf{alph}(x)$. We turn to the definition of "pertaining:"

Definition 5.20. For an alphabet C, a triplet of states (p,q,q') is *C-pertaining* if there are words s,u,t with $\mathsf{alph}(u) = C$, words $z \in C^*$, and an integer n such that:



The pertaining triplet is *empty* if, in the above picture, $\beta = \beta' = 1$; it is *full* if none of β, β' is empty, and *degenerate* otherwise.

Lemma 5.21. *A transducer $\tau\colon A^* \to B^*$ is $\mathcal{R}$-continuous iff all of the following hold:*

(1) $\tau^{-1}(B^*) \in \mathcal{R}(A^*)$;

(2) *For all full C-pertaining triplets (p,q,q'), there exist $x \in B^*$ and $\rho \in (B^*)^2$ such that both inclusions hold:*

$$\tau_{\bullet,p} \bowtie \tau_{\bullet,q} \subseteq Id \cdot \big((x^*, x^*)\rho^{-1}\big) \ ,$$
$$\tau_{p,\bullet} \bowtie \Big(\tau_{q,q} \cdot \big(\varepsilon, \tau_{q,q'}(C^*)\big) \cdot \tau_{q',\bullet}\Big) \subseteq \big(\mathsf{alph}(x)^*, \mathsf{alph}(x)^*\big) \cdot Id \ ;$$

(3) *For all empty C-pertaining triplets (p,q,q'):*

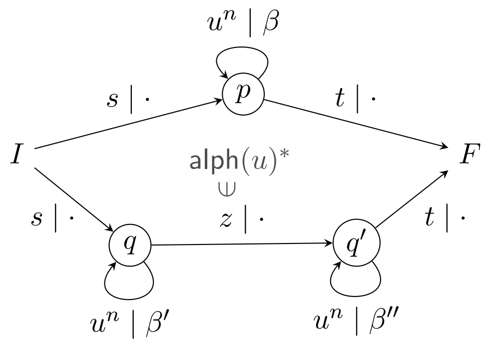
$$(\tau_{\bullet,p} \cdot \tau_{p,\bullet}) \bowtie \Big(\tau_{\bullet,q} \cdot \big(\varepsilon, \tau_{q,q'}(C^*)\big) \cdot \tau_{q',\bullet}\Big) \subseteq Id \ ;$$

(4) *No pertaining triplet is degenerate.*

Note that the x of Proposition 5.15 can be effectively found. The case of $\mathcal{L}$ can be simply seen as the reversal of the previous case.

5.2.3. *The case of $\mathcal{DA}$.* Similarly, we use a slightly less standard equational definition of $\mathcal{DA}$. Indeed, $\mathcal{DA}$ is also defined, over any alphabet A, by $x^\omega = x^\omega \cdot y \cdot x^\omega$ with $y \in \mathsf{alph}(x)$. The definition of "pertaining" reflects these equations:

Definition 5.22. For an alphabet C, a triplet of states (p, q, q') is *C-pertaining* if there are words s, u, t with $\mathsf{alph}(u) = C$, a word $z \in C^*$, and an integer n such that:



Further, a pertaining triplet is *empty* if, in the above picture, $\beta = \beta' = \beta'' = 1$; it is *left-empty* if only β' is empty, *right-empty* if only β'' is empty, *full* if none of β, β', β'' is empty, and *degenerate* in the other cases.

Lemma 5.23. *A transducer $\tau\colon A^* \to B^*$ is $\mathcal{DA}$-continuous iff all of the following hold:*

(1) $\tau^{-1}(B^*) \in \mathcal{DA}(A^*)$*;*

(2) *For all full C-pertaining triplets (p, q, q'), there exist $x, y \in B^*$ and $\rho_1, \rho_2 \in (B^*)^2$ such that these three inclusions hold:*

$$\tau_{\bullet,p} \bowtie \tau_{\bullet,q} \subseteq Id \cdot \big((x^*, x^*)\rho_1^{-1}\big) \ ,$$
$$\tau_{p,\bullet} \bowtie \tau_{q',\bullet} \subseteq \big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id \ ,$$
$$\tau_{q,q'}(C^*) \subseteq \mathsf{alph}(x \cdot y)^* \ ;$$

(3) *For all empty C-pertaining triplets (p, q, q'):*

$$(\tau_{\bullet,p} \bowtie \tau_{\bullet,q}) \cdot (\varepsilon, \tau_{q,q'}(C^*)) \cdot (\tau_{p,\bullet} \bowtie \tau_{q',\bullet}) \subseteq Id \ ;$$

(4) *For all right-empty C-pertaining triplets (p, q, q'), there exist $x, y \in B^*$ and $\rho_1, \rho_2 \in (B^*)^2$ such that:*

$$\tau_{\bullet,p} \bowtie \tau_{\bullet,q} \subseteq Id \cdot \big((x^*, x^*)\rho_1^{-1}\big) \quad \text{and} \quad \tau_{p,\bullet} \bowtie \big((\varepsilon, \tau_{q,q'}(C^*)) \cdot \tau_{q',\bullet}\big) \subseteq \big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id \ ;$$

(5) *For all left-empty C-pertaining triplets (p, q, q'), there exist $x, y \in B^*$ and $\rho_1, \rho_2 \in (B^*)^2$ such that:*

$$\tau_{\bullet,p} \bowtie \big(\tau_{\bullet,q} \cdot (\varepsilon, \tau_{q,q'}(C^*))\big) \subseteq Id \cdot \big((x^*, x^*)\rho_1^{-1}\big) \quad \text{and} \quad \tau_{p,\bullet} \bowtie \tau_{q',\bullet} \subseteq \big(\rho_2^{-1}(y^*, y^*)\big) \cdot Id \ ;$$

(6) *No pertaining triplet is degenerate.*

5.3. **Deciding Com- and Ab-continuity.** The case of $\mathcal{C}om$ and $\mathcal{A}b$ is comparatively much simpler, in particular because these varieties are defined using a finite number of equations for each alphabet. However, the argument relies on different ideas:

Theorem 5.24. *For $\mathcal{V} = \mathcal{C}om, \mathcal{A}b$, it is decidable, given an unambiguous transducer, whether it realizes a $\mathcal{V}$-continuous function.*

Proof. We apply the Syncing Lemma. Its first point is clearly decidable. We reduce its second point to decidable properties about semilinear sets (see, e.g., [10]). We also rely on the notion of Parikh image, that is, the mapping $\mathsf{Pkh}\colon A^* \to \mathbb{N}^A$ such that $\mathsf{Pkh}(w)$ maps $a \in A$ to the number of a's in the word w.

Since every $\mathcal{A}b$-continuous function is $\mathcal{C}om$-continuous (Proposition 4.6), the conditions to test for $\mathcal{A}b$-continuity are included in those for $\mathcal{C}om$-continuity—this can also be seen as a consequence of the fact that if u, v are words, $\mathsf{Equ}_{\mathcal{A}b}(u, v) = \mathsf{Equ}_{\mathcal{C}om}(u, v)$.

Let $\tau\colon A^* \to B^*$ be a given transducer. Consider an equation $ab = ba$ and four states p, p', q, q' of τ. Write $u = \tau_{p,p'}(ab)$ and $v = \tau_{q,q'}(ba)$. We ought to check, by the Syncing Lemma, the inclusion in $\mathsf{Equ}_{\mathcal{C}om}(u, v) = \{(s, s', t, t') \mid s \cdot u \cdot t =_{\mathcal{C}om} s' \cdot v \cdot t'\}$ of some input synchronization. Now this set is the set of (s, s', t, t') such that $\mathsf{Pkh}(s \cdot u \cdot t) = \mathsf{Pkh}(s' \cdot v \cdot t')$, and is thus defined by a simple semilinear property. The input synchronizations themselves, e.g., $\tau_{\bullet,p} \bowtie \tau_{\bullet,q}$, are rational relations, and their component-wise Parikh image is thus a semilinear set. Since the inclusion of semilinear sets is decidable, the inclusion of the second point of the Syncing Lemma is also decidable.

For $\mathcal{A}b$, we should additionally check the equations $a^\omega = 1$. The reasoning is similar. Consider three states (p, p', q), and write $x \cdot u^{\omega-1} \cdot y$ for $\tau_{p,p'}(a^\omega)$. By commutativity and the fact that $u^{\omega-1}$ acts as an inverse of u in the equations holding in $\mathcal{A}b$, we have that $(s, s', t, t') \in \mathsf{Equ}_{\mathcal{A}b}(x \cdot u^{\omega-1} \cdot y, 1)$ iff $s \cdot t =_{\mathcal{A}b} s' \cdot u \cdot t'$. This again reduces the inclusion of the second point of the Syncing Lemma to a decidable semilinear property. □

6. Discussion

We presented a study of continuity in functional transducers, on the one hand focused on general statements (Section 3), on the other hand on continuity for classical varieties. The heart of this contribution resides in decidability properties (Section 5), although we also addressed natural and related questions in a systematic way (Section 4). We single out two main research directions.

First, there is a sharp contrast between the genericity of the Preservation and Syncing Lemma and the technicality of the actual proofs of decidability of continuity. To which extent can these be unified and generalized? We know of two immediate extensions: 1. the generic results of Section 3 readily apply to Boolean algebras of languages closed under quotient, a relaxation of the conditions imposed on varieties, and 2. Proposition 4.1 and Lemma 5.4 can be shown to also hold for the varieties $\mathcal{G}_p$ of languages recognized by p-groups, hence $\mathcal{G}_p$-continuity is decidable for transducers. Beyond these two points, we do not know how to show decidability for $\mathcal{G}_{\text{nil}}$ (which is the *join* of the $\mathcal{G}_p$), and the surprising complexity of the equalizer sets for some Burnside varieties (e.g., the one defined by $x^2 = x^3$, see the Remark on page 10) leads us to conjecture that continuity may be undecidable in that case, hence that no unified way to show the decidability of continuity exists.

Second, the notion of continuity may be extended to more general settings. For instance, departing from regular languages, it can be noted that every recursive function is continuous for the class of recursive languages. Another natural generalization consists in studying $(\mathcal{V}, \mathcal{W})$-continuity, that is, the property for a function to map $\mathcal{W}$-languages to $\mathcal{V}$-languages by inverse image. This would provide more flexibility for a sufficient condition for cascades of languages (or stackings of circuits, or nestings of formulas) to be in a given variety.

Acknowledgment. We are deeply indebted to Shaull Almagor, Jorge Almeida (in particular for the Remark on page 10), Luc Dartois, Bruno Guillon (in particular for the Remark on page 23), Ismaël Jecker, and Jean-Éric Pin for their insightful comments and kind help.

The first and third authors are partly funded by the DFG Emmy Noether program (KR 4042/2); the second author is funded by the DeLTA project (ANR-16-CE40-0007).

References

[1] Jorge Almeida. Some algorithmic problems for pseudovarieties. *Publ. Math. Debrecen*, 54(1):531–552, 1999.

[2] Jorge Almeida and Alfredo Costa. Equidivisible pseudovarieties of semigroups. *Publ. Math. Debrecen*, 90(3-4), 2017.

[3] Jean Berstel. *Transductions and Context-Free Languages*, volume 38 of *Leitfäden der angewandten Mathematik und Mechanik LAMM.* Teubner, 1979.

[4] Janusz A. Brzozowski and Robert Knast. The dot-depth hierarchy of star-free languages is infinite. *Journal of Computer and System Sciences*, 16(1):37 – 55, 1978. `doi:10.1016/0022-0000(78)90049-1`.

[5] Michaël Cadilhac, Andreas Krebs, Michael Ludwig, and Charles Paperman. A circuit complexity approach to transductions. In *MFCS '15*, pages 141–153, 2015. `doi:10.1007/978-3-662-48057-1_11`.

[6] Laure Daviaud, Ismaël Jecker, Pierre-Alain Reynier, and Didier Villevalois. Degree of sequentiality of weighted automata. In *FOSSACS '17*, pages 215–230, 2017. `doi:10.1007/978-3-662-54458-7_13`.

[7] Laure Daviaud, Pierre-Alain Reynier, and Jean-Marc Talbot. A generalised twinning property for minimisation of cost register automata. In *LICS '16*, pages 857–866, 2016. `doi:10.1145/2933575.2934549`.

[8] Emmanuel Filiot, Olivier Gauwin, and Nathan Lhote. First-order definability of rational transductions: An algebraic approach. In *LICS '16*, pages 387–396. ACM, 2016. `doi:10.1145/2933575.2934520`.

[9] Mai Gehrke, Serge Grigorieff, and Jean-Éric Pin. *Duality and Equational Theory of Regular Languages*, pages 246–257. Springer, 2008. `doi:10.1007/978-3-540-70583-3_21`.

[10] Seymour Ginsburg. *The Mathematical Theory of Context-Free Languages.* McGraw-Hill, Inc., New York, NY, USA, 1966.

[11] M. Lothaire. *Combinatorics on Words.* Cambridge University Press, second edition, 1997.

[12] M. Lothaire. *Algebraic combinatorics on words.* Number 90. Cambridge University Press, 2002.

[13] Jean-Éric Pin. *Mathematical Foundations of Automata Theory.* 2016.

[14] Jean-Éric Pin and Jacques Sakarovitch. Some operations and transductions that preserve rationality. In *Theoretical Computer Science*, pages 277–288. Springer, 1982. `doi:10.1007/BFb0036488`.

[15] Jean-Éric Pin and Jacques Sakarovitch. Une application de la representation matricielle des transductions. *Theoretical Computer Science*, 35:271 – 293, 1985. `doi:10.1016/0304-3975(85)90019-2`.

[16] Jean-Éric Pin and Pedro V. Silva. A topological approach to transductions. *Theoretical Computer Science*, 340(2):443 – 456, 2005. `doi:10.1016/j.tcs.2005.03.029`.

[17] Jean-Éric Pin and Pedro V. Silva. On profinite uniform structures defined by varieties of finite monoids. *Int. J. of Algebra and Computation*, 21(01n02):295–314, 2011. `doi:10.1142/S0218196711006170`.

[18] Jean-Éric Pin and Pedro V. Silva. On uniformly continuous functions for some profinite topologies. *Theoretical Computer Science*, 658, Part A:246 – 262, 2017. `doi:10.1016/j.tcs.2016.06.013`.

[19] Jean-Éric Pin and Pascal Weil. Profinite semigroups, Mal'cev products and identities. *Journal of Algebra*, 182(3):604–626, 1996.

[20] Christophe Reutenaeur and Marcel-Paul Schützenberger. Variétés et fonctions rationnelles. *Theoretical Computer Science*, 145(1–2):229–240, July 1995. `doi:10.1016/0304-3975(94)00180-Q`.

[21] Derek J.S. Robinson. *A Course in the Theory of Groups.* Springer, 2 edition, 1995.

[22] Jacques Sakarovitch. Sur la définition du produit en couronne. *Codages et Transductions, G. Pirillo (éd.)*, pages 285–299, 1979.
[23] Jacques Sakarovitch. *Elements of automata theory*. Cambridge University Press, 2009.
[24] Klaus Schneider. *Verification of Reactive Systems: Formal Methods and Algorithms*. Springer, 2004.
[25] Howard Straubing. *Finite Automata, Formal Logic, and Circuit Complexity*. Birkhäuser, Boston, 1994.
[26] Pascal Tesson and Denis Thérien. Logic meets algebra: the case of regular languages. *Logical Methods in Computer Science*, 3(1), 2007.