Profiles of dynamical systems and their algebra

Caroline Gaze-Maillot[1] and Antonio E. Porreca[2]

[1] Aix Marseille Université, Université de Toulon, CNRS, LIS, Marseille, France
caroline.gaze@gmail.com
[2] Université Publique
antonio.porreca@lis-lab.fr

Abstract. The commutative semiring $\mathbf{D}$ of finite, discrete-time dynamical systems was introduced in order to study their (de)composition from an algebraic point of view. However, many decision problems related to solving polynomial equations over $\mathbf{D}$ are intractable (or conjectured to be so), and sometimes even undecidable. In order to take a more abstract look at those problems, we introduce the notion of "topographic" profile of a dynamical system (A, f) with state transition function $f\colon A \to A$ as the sequence $\operatorname{prof} A = (|A|_i)_{i\in\mathbb{N}}$, where $|A|_i$ is the number of states having distance i, in terms of number of applications of f, from a limit cycle of (A, f). We prove that the set of profiles is also a commutative semiring $(\mathbf{P}, +, \times)$ with respect to operations compatible with those of $\mathbf{D}$ (namely, disjoint union and tensor product), and investigate its algebraic properties, such as its irreducible elements and factorisations, as well as the computability and complexity of solving polynomial equations over $\mathbf{P}$.

1 Introduction

Given a description of a dynamical system, it is often interesting for scientific or engineering purposes to analyse its dynamics, in order to detect its asymptotic behaviour, such as the number and size of limit cycles or fixed points, or other interesting behaviours, such as the reachability of states or its transient paths. However, these problems are often computationally demanding when the system is described in a succinct way, as one normally does, e.g., for Boolean automata networks or cellular automata [5,11]. It is useful, then, to decompose the system into smaller systems before applying such algorithms; if an appropriate decomposition is chosen, the global behaviour of the system may be deduced from the behaviour of its components [8].

Let us now consider finite, discrete-time dynamical systems in the most general sense: as finite sets A of states (including the empty set) together with a state transition function $f\colon A \to A$. The (countably infinite) set of finite dynamical systems up to isomorphism is a semiring $(\mathbf{D}, +, \times)$ with the operations [2]

$$(A, f) + (B, g) = (A \uplus B, f + g) \qquad \text{where } (f+g)(x) = \begin{cases} f(x) & \text{if } x \in A \\ g(x) & \text{if } x \in B \end{cases}$$

$$(A, f) \times (B, g) = (A \times B, f \times g) \qquad \text{where } (f \times g)(a, b) = (f(a), g(b)).$$

These operations can also be defined in terms of the graphs of the dynamics as disjoint union $+$ and graph tensor product $\times$, which equivalently corresponds to the Kronecker product of the adjacency matrices [6].

Given this algebraic structure, one can try to decompose dynamical systems in terms of factoring, or in terms of polynomial equations over $\mathbf{D}[\vec{X}]$ in several variables. The decomposition of a dynamical systems in terms of the operations $+$ and $\times$ does indeed allow us to detect several interesting dynamical behaviours of the system in terms of its components. For instance, the limit cycles in a sum are just the union of the limit cycles of the addends, while in a product one can predict the number and length of limit cycles as a function of the GCD and LCM of the lengths of the cycles of the factors [3].

However, solving equations over $\mathbf{D}[\vec{X}]$ is not easy either, even if the dynamical systems are given in input explicitly, either as a transition table, or equivalently in terms of the graph of its dynamics $G(A, f) = (A, \{(a, f(a)) : a \in A\})$. General polynomial equations are even undecidable [2], systems of linear equations are NP-complete, and single equations (even linear) with a constant side are also suspected to be NP-complete [9].

When (de)composing dynamical systems as products, one frequently works starting from the limit cycles and backwards towards the gardens of Eden (states without preimages). It is then useful to know how many states there are at distance $0, 1, \ldots$ from the limit cycles, as that gives us, for instance, necessary conditions for the compositeness of a system. In this paper we formalise this as the notion of *profile* of a dynamical system, in order to analyse systems from a more abstract point of view. We obtain another semiring $(\mathbf{P}, +, \times)$ with the "natural" operations derived from those of $(\mathbf{D}, +, \times)$, analyse some of its algebraic properties (notably, the majority of profiles are irreducible) and prove that working with equivalence classes of systems (with respect to profile equality) ultimately does *not* reduce the complexity of equation problems: general polynomial equations remain undecidable, and even solving a single linear equation is NP-complete.

2 Profiles of dynamical systems

Any finite dynamical system (A, f) consists of one or more disjoint *limit cycles*, which constitute the asymptotic behaviour of the system. Each cycle of length 1 is called a *fixed point*, and its only state x satisfies $f(x) = x$. The transient (non-asymptotic) behaviour of the system consists of zero or more directed trees of least two nodes having a state of a limit cycle as its root. The existence of limit cycles, which does *not* hold in general for infinite dynamical systems, gives a (pre)ordering to the states, with respect to their distance (in terms of number of applications of f) from the limit cycle in the same connected component of the graph of the dynamics, which we will call its *height*.

Definition 1. *Let (A, f) be a dynamical system and let $x \in A$. We say that the* height of x*, in symbols $h_A(x)$ or even $h(x)$ if A is implied, is the minimum h such that $f^h(x)$ is a periodic state, that is, the length of a path from x to the nearest periodic state in the graph of the dynamics of A.*

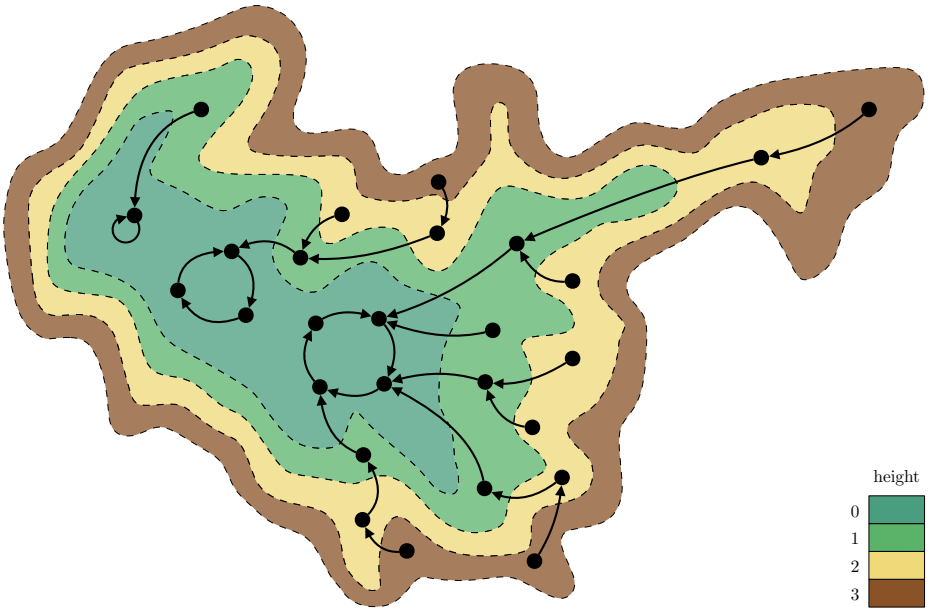

Fig. 1. Visual representation of the contour lines (isolines) of a dynamical system having profile $(8, 7, 8, 4)$: there are 8 states in limit cycles, 7 states of height 1, 8 of height 2, and 4 of height 3.

It is also possible to generalise the notion of height to a dynamical systems.

Definition 2. *Let (A, f) be a dynamical system. We say that the* height of A*, in symbols $h(A)$, is the maximum height of its states: $h(A) = \max\{h_A(x) : x \in A\}$.*

We can now introduce the notion of *profile* of a dynamical system, whose name is inspired by the topographic profile of a terrain (Fig. 1).

Definition 3. *Let (A, f) be a dynamical system of height h, and let $|A|_i$ be the number of states of A having height i. This implies $|A|_i \geq 1$ for $0 \leq i \leq h$ and $|A|_i = 0$ for $i > h$. Then, the* profile of A *is the eventually null sequence of natural numbers $\mathrm{prof}(A, f) = (|A|_i)_{i\in\mathbb{N}}$ counting the states of each height in A, in order of height starting from the limit cycles (height 0). For brevity, we often write a profile as a finite sequence $\mathrm{prof}(A, f) = (|A|_0, |A|_1, \ldots, |A|_h)$ by omitting the null terms, except for the profile (0) of the empty dynamical system.*

Taking the profile of a dynamical system allows us to work at a higher level of abstraction, since it corresponds to taking a whole equivalence class of dynamical systems. In the rest of this paper, we will denote profiles with bold letters and its elements in italics, such as $\mathbf{p} = (p_i)_{i\in\mathbb{N}}$.

3 The semiring of profiles

It is easy to define a sum operation + over the (countably infinite) set of profiles $\mathbf{P}$ that is compatible with the sum over $\mathbf{D}$: since $(A, f) + (B, g)$ is the disjoint union, the profile of this sum is just the elementwise sum of $\mathrm{prof}(A, f)$ and $\mathrm{prof}(B, g)$.

Definition 4. *Given two profiles* $\mathbf{p} = (p_i)_{i\in\mathbb{N}}$ *and* $\mathbf{q} = (q_i)_{i\in\mathbb{N}}$*, define their sum as* $\mathbf{p} + \mathbf{q} = (p_i + q_i)_{i\in\mathbb{N}}$.

Lemma 5. *For each* $A, B \in \mathbf{D}$ *we have* $\mathrm{prof}(A + B) = \mathrm{prof}\, A + \mathrm{prof}\, B$. □

It is less immediate to define a product over $\mathbf{P}$, but it is indeed possible with a little more work. First, we show that, in order to compute the height of a state in a product, it suffices to take the maximum height of the two terms.

Lemma 6. *Let* $(A, f), (B, g) \in \mathbf{D}$ *and let* $(C, t) = (A, f) \times (B, g)$*. Then, for each* $(a, b) \in C$*, we have* $h(a, b) = \max(h(a), h(b))$.

Proof. Notice that the limit cycles of C consist exactly of the states (x, y) such that y is a periodic state of A, and y a periodic state of B.

Suppose that $h(a) \geq h(b)$. Then $f^{h(a)}(a)$ is a periodic state of A, and this is not the case for $f^i(a)$ whenever $i < h(a)$, by definition of height. Furthermore, $g^{h(a)}(b)$ is a periodic state of B, since this is the case for all $g^i(b)$ with $i \geq h(b)$. Then, $t^{h(a)}(a, b) = (f^{h(a)}(a), g^{h(a)}(b))$ is a periodic state of C, and this is not the case for $t^i(a, b)$ whenever $i < h(a)$. This means that $h(a, b) = h(a)$.

Analogously, if $h(a) \leq h(b)$ we obtain $h(a, b) = h(b)$, and the statement of the lemma follows. □

The following lemma allows us to count the states of height k in a product as a function of the number of states of height k and at most k of the two factors.

Lemma 7. *Let* $A, B \in \mathbf{D}$ *be dynamical systems and let* $C = A \times B$*. Then*

$$|C|_k = |A|_k \times |B|_{\leq k} + |B|_k \times |A|_{\leq k} - |A|_k \times |B|_k \tag{1}$$

for each height k*, where* $|D|_{\leq k} = |D|_0 + |D|_1 + \cdots + |D|_k$ *for each* $D \in \mathbf{D}$.

Proof. Let $a \in A$ with $h(a) = k$ and $b \in B$ with $h(b) \leq k$. Then, by Lemma 6, we have $h(a, b) = k$. This corresponds to the term $|A|_k \times |B|_{\leq k}$ of the sum. Analogously, we have $h(a, b) = k$ if $h(a) \leq k$ and $h(b) = k$, which corresponds to the term $|B|_k \times |A|_{\leq k}$. This way, we have counted twice the states (a, b) with $h(a) = k$ and $h(b) = k$, thus it is necessary to subtract the term $|A|_k \times |B|_k$, and that gives us the correct result. Notice that (1) works even if $|A|_k = 0$ or $|B|_k = 0$ (i.e., if $k > h(A)$ or $k > h(B)$), giving the expected $|C|_k = |B|_k \times |A|_{\leq k}$ and $|C|_k = |A|_k \times |B|_{\leq k}$, or even $|C|_k = 0$ if $k > \max(h(A), h(B))$. □

Notice how Lemma 7 does *not* depend on the exact shapes of A and B, but only on their profile. This allows us to define the product of profiles as follows (Fig. 2):

×	(0)	(1)	(2)	(1,1)	(3)	(1,2)	(2,1)	(1,1,1)
(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)	(0)
(1)	(0)	(1)	(2)	(1,1)	(3)	(1,2)	(2,1)	(1,1,1)
(2)	(0)	(2)	(4)	(2,2)	(6)	(2,4)	(4,2)	(2,2,2)
(1,1)	(0)	(1,1)	(2,2)	(1,3)	(3,3)	(1,5)	(2,4)	(1,3,2)
(3)	(0)	(3)	(6)	(3,3)	(9)	(3,6)	(6,3)	(3,3,3)
(1,2)	(0)	(1,2)	(2,4)	(1,5)	(3,6)	(1,8)	(2,7)	(1,5,3)
(2,1)	(0)	(2,1)	(4,2)	(2,4)	(6,3)	(2,7)	(4,5)	(2,4,3)
(1,1,1)	(0)	(1,1,1)	(2,2,2)	(1,3,2)	(3,3,3)	(1,5,3)	(2,4,3)	(1,3,5)

Fig. 2. Multiplication table for profiles of size 0, 1, 2, and 3.

Definition 8. *For two profiles* $\mathbf{p} = (p_i)_{i\in\mathbb{N}}$ *and* $\mathbf{q} = (q_i)_{i\in\mathbb{N}}$*, let their product be*

$$\mathbf{p} \times \mathbf{q} = \left(p_i \times \sum_{j=0}^{i} q_i + q_i \times \sum_{j=0}^{i} p_i - p_i \times q_i\right)_{i\in\mathbb{N}}.$$

From Definition 8 and Lemma 7 we obtain the expected result:

Lemma 9. *For each* $A, B \in \mathbf{D}$ *we have* $\operatorname{prof}(A \times B) = \operatorname{prof} A \times \operatorname{prof} B$. □

This finally gives us the algebraic structure of $\mathbf{P}$.

Theorem 10. $(\mathbf{P}, +, \times)$ *is a commutative semiring.*

Proof. The operation $+$ inherits the associative and commutative properties from $\mathbb{N}$ and has, as the neutral element, the null profile $0_{\mathbf{P}} = (0)$, i.e., the profile of the empty dynamical system $0_{\mathbf{D}} = \varnothing$. Thus $(\mathbf{P}, +)$ is a commutative monoid.

Let $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbf{P}$, and let $A, B, C \in \mathbf{D}$ such that $\operatorname{prof} A = \mathbf{p}, \operatorname{prof} B = \mathbf{q}$, and $\operatorname{prof} C = \mathbf{r}$. Then we have $\mathbf{p} \times \mathbf{q} = \operatorname{prof} A \times \operatorname{prof} B = \operatorname{prof}(A \times B)$ by Lemma 9, then $\operatorname{prof}(A \times B) = \operatorname{prof}(B \times A)$ by commutativity of $\times$ in $\mathbf{D}$, and $\operatorname{prof}(B \times A) = \operatorname{prof} B \times \operatorname{prof} A = \mathbf{q} \times \mathbf{p}$, and thus $\times$ is commutative. Similarly, we have the associative property $\mathbf{p} \times (\mathbf{q} \times \mathbf{r}) = (\mathbf{p} \times \mathbf{q}) \times \mathbf{r}$, the neutral element $1_{\mathbf{P}} = \operatorname{prof} 1_{\mathbf{D}} = (1)$, i.e., the profile of a the dynamical system consisting of a single fixed point, and the distributive property $\mathbf{p} \times (\mathbf{q} + \mathbf{r}) = \mathbf{p} \times \mathbf{q} + \mathbf{p} \times \mathbf{r}$. □

From Lemmata 5 and 9 we also obtain that "taking the profile" does indeed respect the semiring operations and, being surjective, it gives us, in a sense, a good abstraction of dynamical systems.

Corollary 11. *The function* $\mathrm{prof}\colon \mathbf{D} \to \mathbf{P}$ *is a surjective semiring homomorphism.* □

A very important and useful result is that $\mathbf{P}$ contains an isomorphic copy of the naturals, the initial semiring of its category:

Lemma 12. $(\mathbf{P}, +, \times)$ *has a subsemiring isomorphic to* $(\mathbb{N}, +, \times)$.

Proof. Let $\phi\colon \mathbb{N} \to \mathbf{D}$ be defined by $\phi(n) = (n)$, that is, the profile having n as its first component and zero everywhere else. Then $\phi(0) = (0) = 0_{\mathbf{P}}$, $\phi(1) = (1) = 1_{\mathbf{P}}$, $\phi(m+n) = (m+n) = (m) + (n) = \phi(m) + \phi(n)$, and $\phi(m \times n) = (m \times n) = (m) \times (n) = \phi(m) \times \phi(n)$. Thus ϕ is a semiring homomorphism. Furthermore, $\phi(m) = \phi(n)$ implies $m = n$, i.e., ϕ is injective. As a consequence, its image $\phi(\mathbb{N})$ is a subsemiring of $\mathbf{P}$ isomorphic to $\mathbb{N}$. □

The size of a profile is the number of states of any dynamical system with that profile, and it also enjoys some nice properties.

Definition 13. *The* size *of a profile* $\mathbf{p} = (p_i)_{i \in \mathbb{N}}$ *is given by* $|\mathbf{p}| = \sum_{i \in \mathbb{N}} p_i$.

Lemma 14. *The function* $|\cdot|\colon \mathbf{P} \to \mathbb{N}$ *is a semiring homomorphism.*

Proof. The size $|\mathbf{p}|$ of a profile $\mathbf{p}$ is just the number of states $|A|$ of any dynamical system A such that $\mathrm{prof}\, A = \mathbf{p}$. Since the sum and product of dynamical systems have the disjoint union and the Cartesian product as their set of states, respectively, we have $|A + B| = |A| + |B|$ and $|A \times B| = |A| \times |B|$ for $A, B \in \mathbf{D}$. Furthermore, we have $|0_{\mathbf{P}}| = 0$ and $|1_{\mathbf{P}}| = 1$. The result then follows from Lemmata 5 and 9. □

Since the profiles $\mathbf{P}$ contain the naturals (Lemma 12), they are not only a semiring, but also an $\mathbb{N}$-semimodule, a "vector space" with the naturals as its "scalars" [7], with the semimodule axioms satisfied as a direct consequence of the semiring axioms. This will be useful later when analysing the complexity of solving linear equations over $\mathbf{P}$ (Section 5).

Theorem 15. $(\mathbf{P}, +)$ *is an* $\mathbb{N}$*-semimodule with its ordinary multiplication restricted to* $\mathbb{N} \times \mathbf{P} \to \mathbf{P}$*, that is,* $n \times (p_i)_{i \in \mathbb{N}} = (np_i)_{i \in \mathbb{N}}$. □

Unfortunately, profiles are not particularly nice as a semimodule, since there is only one minimal generating set, and it is not linearly independent.

Theorem 16. $\mathbf{P}$ *as an* $\mathbb{N}$*-semimodule has a unique, countably infinite minimal generating set* $G = \{\mathbf{p} \in \mathbf{P} : p_0 = 1\}$*, the set of profiles starting with* 1*, which is linearly dependent.*

Proof. The set G is a generating set, since any profile $\mathbf{q} = (q_0, q_1, q_2, \ldots)$ can be written as $(q_0 - 1) \times (1) + 1 \times (1, q_1, q_2, \ldots)$, that is, any element of $\mathbf{P}$ is a linear combination of at most two elements of G. This set is countably infinite.

To prove that any generating set of $\mathbf{P}$ must contain G, consider any element $\mathbf{p} \in G$. If it is a linear combination $\mathbf{p} = \sum_{i=1}^{m} a_i \mathbf{q}_i$ of profiles $\mathbf{q}_i \in \mathbf{P}$ with

coefficients $a_i \in \mathbb{N}$, then one of the $\mathbf{q}_i$ must start with 1 (i.e., $q_{i,0} = 1$) and its coefficient a_i must be 1 as well, otherwise we would have either $p_0 = 0$ or $p_0 > 1$; furthermore, we must have $(a_j \mathbf{q}_j)_0 = 0$ for all $j \neq i$, which implies $a_j \mathbf{q}_j = 0$ (since all elements of a profile are null after the first 0) and thus $a_j = 0$. But then we have $\mathbf{p} = 1\mathbf{q}_i$, that is, $\mathbf{q}_i = \mathbf{p}$: any linear combination giving $\mathbf{p}$ as its result must contain $\mathbf{p}$ itself, and thus it must belong to any generating set.

However, G is not linearly independent, that is, there exist m profiles $\mathbf{p}_i \in \mathbf{P}$ and corresponding natural numbers $a_i, b_i \in \mathbb{N}$ such that $\sum_{i=1}^{m} a_i \mathbf{p}_i = \sum_{i=1}^{m} b_i \mathbf{p}_i$ but $a_i \neq b_i$ for some i. For instance, we have $(1,1) + (1,2) = (1) + (1,3)$. □

4 Reducibility and factorisation of profiles

When studying the reducibility of profiles with respect to semiring product, a simple sufficient condition for irreducibility is given by the primality of its size.

Lemma 17. *Let $\mathbf{p} \in \mathbf{P}$ be a profile such that $|\mathbf{p}|$ is prime. Then $\mathbf{p}$ is irreducible.*

Proof. Suppose $\mathbf{p} = \mathbf{q} \times \mathbf{r}$. Since $|\cdot|$ is a semiring homomorphism (Lemma 14), we have $|\mathbf{p}| = |\mathbf{q}| \times |\mathbf{r}|$. But $|\mathbf{p}|$ is prime, thus either $|\mathbf{q}| = 1$, or $|\mathbf{r}| = 1$. Since the only profile of size 1 is $1_{\mathbf{P}} = (1)$, this factorisation is trivial. □

It is easy to check by inspection of the product table (Fig. 2) that some profiles admit multiple factorisations into irreducibles, a property that they share with the semiring of dynamical systems [2].

Theorem 18. $\mathbf{P}$ *is* not *a unique factorisation semiring.*

Proof. The smallest counterexample is the profile $(2,4)$, which has two distinct factorisations: $(2,4) = (2) \times (1,2) = (1,1) \times (2,1)$. All the factors (2), $(1,2)$, $(1,1)$, and $(2,1)$ are irreducible because of their prime size (Lemma 17). □

Another property in common with dynamical systems [3] is that most profiles, a fraction asymptotically equal to 1, are irreducible.

Theorem 19. *The majority of profiles is irreducible; specifically,*

$$\lim_{n \to \infty} \frac{\textit{number of reducible profiles of size at most } n}{\textit{number of profiles of size at most } n} = 0.$$

Proof. There are as many profiles of size i as there are ordered tuples of strictly positive naturals having sum i, which correspond to the ways of writing i as an ordered sum of strictly positive integers. The latter are the *compositions* of i, and there are 2^{i-1} of them for $i \geq 1$ [1, Chapter 4], and 1 for $i = 0$. Hence, the number of profiles of size at most n is given by $1 + \sum_{i=1}^{n} 2^{i-1} = 1 + 2^n - 1 = 2^n$.

Suppose that $\mathbf{p} \in \mathbf{P}$ has size $i = |\mathbf{p}|$ and that $i = \ell m$. Then, there are *at most* $2^{\ell-1} \times 2^{m-1} = 2^{\ell+m-2}$ ways of choosing profiles $\mathbf{q}$ and $\mathbf{r}$, of sizes ℓ and m respectively, such that $\mathbf{p} = \mathbf{q} \times \mathbf{r}$.

Let k be the number of distinct factorisations of i into products of two integers $\ell_j \geq m_j > 1$. The number of ways of decomposing $\mathbf{p}$ into a product of two non-trivial divisors is at most $\sum_{j=1}^{k} 2^{\ell_j + m_j - 2}$. Observe[3] that $\ell + m \leq \ell m/2 + 2$ for

[3] This can be proved by induction on any of the two variables.

all $\ell, m > 1$; this implies $\sum_{j=1}^{k} 2^{\ell_j + m_j - 2} \leq \sum_{j=1}^{k} 2^{\ell_j m_j / 2} = \sum_{j=1}^{k} 2^{i/2} = k\sqrt{2^i}$. We have $k < i$, since the number of non-trivial divisors of i is strictly less than i (at least 1 and i have to be thrown out), hence the number of ways of obtaining a profile of size i as a product of two non-trivial profiles, which is the same as the number of reducible profiles of size i, is bounded by $i\sqrt{2^i}$ for $k \geq 1$, and it is 1 for $i = 0$. The number of reducible profiles of size at most n is then bounded by $1 + \sum_{i=1}^{n} i\sqrt{2^i} \leq 1 + n\sum_{i=0}^{n} \sqrt{2^i} = 1 + n\frac{\sqrt{2^{n+1}}-1}{\sqrt{2}-1}$.

By dividing that by the number of profiles of size n we obtain

$$\lim_{n\to\infty} \frac{1 + n\frac{\sqrt{2^{n+1}}-1}{\sqrt{2}-1}}{2^n} = 0$$

as required. □

Thus, the semiring of profiles is quite complex from the point of view of reducibility: most profiles are not reducible at all, but those that are sometimes admit multiple factorisations. Furthermore, since height-1 profiles behave as the natural numbers, we also obtain a complexity lower bound to profile factorisation.

Theorem 20. *The problem of profile factorisation, that is, given a profile* $\mathbf{p} \in \mathbf{P}$*, finding a divisor* $\mathbf{d}$ *of* $\mathbf{p}$ *with* $\mathbf{d} \neq 1_{\mathbf{P}}$ *and* $\mathbf{d} \neq \mathbf{p}$ *(or answering that* $\mathbf{p}$ *is irreducible, if this is the case) is at least as hard as integer factorisation.*

Proof. Given a natural number n, let us consider the profile $\mathbf{n} = (n)$, that is, n followed by zeros. This profile can only be divided by profiles $\mathbf{d} = (d)$ and $\mathbf{q} = (q)$ of height 0 by Lemma 6. Then, $\mathbf{n} = \mathbf{d} \times \mathbf{q} = (d \times q + d \times q - d \times q) = (d \times q)$ for some profile $\mathbf{d}$ with $\mathbf{d} \neq 1_{\mathbf{P}}$ and $\mathbf{d} \neq \mathbf{n}$ if and only if $n = d \times q$ for some d with $d \neq 1$ and $d \neq n$, which is the integer factorisation problem. □

5 Solving polynomial equations over profiles

One of the reasons for introducing profiles is to abstract away from the exact shape of dynamical systems, with the hope of making polynomial equations easier to solve. As we will show in this section, this is not at all the case. First of all, let us prove that polynomial equations with *natural* coefficients do sometimes have *non-natural* solutions in $\mathbf{P}$ (e.g., $3X = Z$ has solution $X = (1, 2), Z = (3, 6)$), but only if there also exist natural ones (e.g., $X = 3, Z = 9$), as in the semiring $\mathbf{D}$ [2].

Lemma 21. *Let* $p, q \in \mathbb{N}[\vec{X}]$ *be polynomials with natural coefficients over the variables* $\vec{X} = (X_1, \ldots, X_m)$*. Then the equation* $p(\vec{X}) = q(\vec{X})$ *has a solution in* $\mathbf{P}$ *if and only if it has a solution in* $\mathbb{N}$*.*

Proof. If the equation has a solution in $\mathbb{N}$, then this is already a solution in $\mathbf{P}$. Conversely, let $\vec{\mathbf{r}} = (\mathbf{r}_1, \ldots, \mathbf{r}_m)$ be a solution in $\mathbf{P}$. We claim that $|\vec{\mathbf{r}}| = (|\mathbf{r}_1|, \ldots, |\mathbf{r}_m|)$ is also a solution, in $\mathbb{N}$; that is, by replacing each profile by (the dynamical system corresponding to) its size, the equation remains valid.

If the equation $p(\vec{X}) = q(\vec{X})$ is of degree at most n, then it can be written as $\sum_{\vec{i}\in[0,n]^m} \left(a_{\vec{i}} \prod_{j=1}^m X_j^{i_j}\right) = \sum_{\vec{i}\in[0,n]^m} \left(b_{\vec{i}} \prod_{j=1}^m X_j^{i_j}\right)$, that is, we compute all products of the m variables, each variable with an exponent ranging from 0 to n (these exponents are collected in a vector $\vec{i} \in [0,n]^m$), and multiply it by a corresponding coefficient $a_{\vec{i}} \in \mathbb{N}$ or $b_{\vec{i}} \in \mathbb{N}$, and then all these monomials are added together. Any of the coefficients $a_{\vec{i}}$ and $b_{\vec{i}}$ can be 0 (if there is more than one variable, some of them will surely be, in order to keep the degree at most n).

If $\vec{\mathbf{r}}$ is a solution the equation, i.e., if $p(\vec{\mathbf{r}}) = q(\vec{\mathbf{r}})$, then by expanding we obtain $\sum_{\vec{i}\in[0,n]^m} \left(a_{\vec{i}} \prod_{j=1}^m \mathbf{r}_j^{i_j}\right) = \sum_{\vec{i}\in[0,n]^m} \left(b_{\vec{i}} \prod_{j=1}^m \mathbf{r}_j^{i_j}\right)$. By applying the size function $|\cdot|$ to both sides of the equation, and exploiting the fact that it is a semiring homomorphism (Lemma 14) and that $|a_{\vec{i}}| = a_{\vec{i}}$ and $|b_{\vec{i}}| = b_{\vec{i}}$ since they already are natural numbers, we obtain the equation over the naturals $\sum_{\vec{i}\in[0,n]^m} \left(a_{\vec{i}} \prod_{j=1}^m |\mathbf{r}_j|^{i_j}\right) = \sum_{\vec{i}\in[0,n]^m} \left(b_{\vec{i}} \prod_{j=1}^m |\mathbf{r}_j|^{i_j}\right)$, which is nothing else than $p(|\vec{\mathbf{r}}|) = q(|\vec{\mathbf{r}}|)$. Thus $|\vec{\mathbf{r}}|$ is indeed a natural solution to the original equation. □

As a consequence, "Hilbert's 10th problem over $\mathbf{P}$" has a negative answer: there is no algorithm for deciding if a polynomial equation in $\mathbf{P}[\vec{X}]$ is solvable, otherwise you could use the same algorithm for natural equations.

Theorem 22. *Deciding whether an equation $p(\vec{X}) = q(\vec{X})$ with $p, q \in \mathbf{P}[\vec{X}]$ has a solution in $\mathbf{P}$ is undecidable.* □

We can get a subclass of algorithmically solvable equations by having one constant side, that, by considering equations of the form $p(\vec{X}) = \mathbf{q}$ with $\mathbf{q} \in \mathbf{P}$. The constant side makes the search space of the solutions finite, which means that at least a brute-force search algorithm is available.

Lemma 23. *Let $\mathbf{p}, \mathbf{q}, \mathbf{r} \in \mathbf{P}$ be profiles. Then $\mathbf{p} + \mathbf{q} = \mathbf{r}$ implies $p_i \leq r_i$, and $\mathbf{p} \times \mathbf{q} = \mathbf{r}$ implies $p_i \leq r_i$ whenever $\mathbf{q} \neq 0_{\mathbf{P}}$, for all $i \in \mathbb{N}$.*

Proof. If $\mathbf{p} + \mathbf{q} = \mathbf{r}$, then $p_i \leq p_i + q_i = r_i$ for all $i \in \mathbb{N}$, as required. Now suppose $\mathbf{p} \times \mathbf{q} = \mathbf{r}$ and $\mathbf{q} \neq 0_{\mathbf{P}}$. By Definition 8, this means

$$r_i = (\mathbf{p} \times \mathbf{q})_i = p_i \times \sum_{j=0}^{i} q_i + q_i \times \sum_{j=0}^{i} p_i - p_i \times q_i.$$

Since $\mathbf{q} \neq 0_{\mathbf{P}}$, we have $q_j \geq 1$ for at least one $j \leq i$. This means that $p_i \times \sum_{j=0}^{i} q_i \geq p_i$. If $q_i = 0$ then $r_i \geq p_i$ as required. So suppose $q_i \geq 1$; this implies $q_i \times \sum_{j=0}^{i} p_i \geq q_i \times p_i$ and $r_i \geq p_i + q_i \times p_i - p_i \times q_i = p_i$, which completes the proof. □

By applying Lemma 23 repeatedly, we obtain the following result.

Lemma 24. *Let $p(\vec{X}) \in \mathbf{P}[\vec{X}]$ over the variables $\vec{X} = (X_1, \ldots, X_m)$, and let $\mathbf{q} \in \mathbf{P}$ be a constant. Then, if $p(\vec{\mathbf{r}}) = \mathbf{q}$ for some $\vec{\mathbf{r}} = (\mathbf{r}_1, \ldots, \mathbf{r}_m) \in \mathbf{P}^m$, there exists a (possibly different) solution $\vec{\mathbf{s}} = (\mathbf{s}_1, \ldots, \mathbf{s}_m) \in \mathbf{P}^m$ such that $p(\vec{\mathbf{s}}) = \mathbf{q}$ and $s_{i,j} \leq q_j$ for all $1 \leq i \leq m$ and $j \in \mathbb{N}$.*

Proof. If all coefficients of p are nonzero, and all profiles $\mathbf{r}_i$ are also nonzero, then let $\vec{\mathbf{s}} = \vec{\mathbf{r}}$, and the result follows from Lemma 23 by induction on the structure of the expression $p(\vec{\mathbf{r}})$.

Otherwise, the expression $p(\vec{\mathbf{r}})$ is a sum with at least one null term, say $\mathbf{a} \times \mathbf{p}_{i_1}^{e_1} \times \cdots \times \mathbf{p}_{i_k}^{e_k}$. If any of the terms $\mathbf{p}_i$ occurring in this product have $p_{i,j} > q_j$ for some $j \in \mathbb{N}$, then it means that $\mathbf{p}_i$ never occurs in a non-null term of the expression $p(\vec{\mathbf{r}})$, since all sums are computed elementwise and this would invalidate the equality. Hence $\mathbf{p}_i$ only occurs multiplied by $0_{\mathbf{P}}$, and it can thus be replaced by *any* profile $\mathbf{p}'_i$ satisfying $p'_{i,j} \leq q_j$ (for instance, $\mathbf{p}'_i = 0_{\mathbf{P}}$ always works) without changing the validity of the equation. By repeating this operation with all profiles $\mathbf{p}_i$ of this kind, we obtain another solution $\vec{\mathbf{s}}$ which satisfies the required inequalities. □

In the rest of the paper we encode profiles, as is natural, as finite sequences of natural numbers $\mathbf{p} = (p_0, \ldots, p_h)$ in binary notation. We can prove that polynomial equation with a constant side can be solved in nondeterministic polynomial time.

Lemma 25. *Deciding whether an equation $p(\vec{X}) = \mathbf{q}$, with $p \in \mathbf{P}[\vec{X}]$ and constant right-hand side $\mathbf{q} \in \mathbf{P}$, has a solution in $\mathbf{P}$ is an* NP *problem. The same holds for a* system *of equations.*

Proof. By Lemma 24, the equation has a solution if and only if it has a solution $\vec{\mathbf{r}} = (\mathbf{r}_1, \ldots, \mathbf{r}_m)$ where each element of $\mathbf{r}_1, \ldots, \mathbf{r}_m$ is bounded by an element of $\mathbf{q}$. Thus, guessing a solution to the equation amounts to guessing, for each $\mathbf{r}_i = (r_{i,0}, \ldots, r_{i,h})$, a natural number $r_{i,j} \in [0, q_j]$ for each height $0, \ldots, h(\mathbf{q})$. This can be performed in nondeterministic polynomial time. Then, the candidate solution can be verified in deterministic polynomial time by evaluating the $p(\vec{X})$ in $\vec{\mathbf{r}}$ and checking equality with $\mathbf{q}$. This proves that the problem belongs to NP.

In the case of multiple equations, after guessing the solution we need to verify that *all* equations are satisfied, which still takes polynomial time. □

Unfortunately, the NP upper bound is strict. We prove that first for systems of linear equations.

Theorem 26. *Deciding whether a* system of linear equations

$$\underbrace{p_1(\vec{X}) = \mathbf{q}_1 \qquad \cdots \qquad p_n(\vec{X}) = \mathbf{q}_n}$$

with $p_i \in \mathbf{P}[\vec{X}]$ and constant right-hand sides $\mathbf{q}_i \in \mathbf{P}$ has a solution in $\mathbf{P}$ is NP*-complete.*

Proof. We prove that the problem is NP-hard by reduction from the NP-complete problem One-in-three 3SAT, the problem of deciding whether a Boolean formula φ in ternary conjunctive normal form has a satisfying assignment which makes only one literal per clause true [10].

For each logical variable x of φ we have a pair of variables X and X' and an equation $X + X' = 1$. This equation forces exactly one between X and X' to 1, and the other to 0. We use X to represent x and X' to represent $\neg x$.

For each clause $(\ell_1 \vee \ell_2 \vee \ell_3)$ of three literals we have an equation $L_1 + L_2 + L_3 = 1$, where $L_i = X_i$ if $\ell_i = x_i$, and $L_i = X'_i$ if $\ell_i = \neg x_i$. This forces exactly one variable corresponding to a literal of the clause to 1, and the other two to 0.

Then, the system of equations obtained from φ is linear and has constant right-hand sides; furthermore, it has a solution if and only if the formula has a satisfying assignment. The satisfying assignments and the solutions to the system of equations are actually the same, if we interpret 0 as false and 1 as true. □

By exploiting the $\mathbb{N}$-semimodule structure of $\mathbf{P}$, we can combine several linear equations together, proving that even a single one is already NP-hard.

Theorem 27. *Deciding whether a* single linear equation $p(\vec{X}) = \mathbf{q}$*, with* $p \in \mathbf{P}[\vec{X}]$ *and constant right-hand side* $\mathbf{q} \in \mathbf{P}$*, has a solution in* $\mathbf{P}$ *is* NP*-complete.*

Proof. We prove this problem NP-complete by adapting the proof of Theorem 26, reducing the system of linear equations $p_1(\vec{X}) = 1, \ldots, p_n(\vec{X}) = 1$ to a single linear equation. Remark that all coefficients of the polynomials p_i, as well as the right-hand sides of the equations, are actually natural numbers in that proof.

As mentioned above (Theorem 15), $\mathbf{P}$ is an $\mathbb{N}$-semimodule. Consider the elements $\mathbf{e}_i = \underbrace{(1, \ldots, 1)}_{i \text{ times}} \in \mathbf{P}$ for $1 \leq i \leq n$; these elements are linearly independent over $\mathbb{N}$, since the $n \times n$ matrix over $\mathbb{R}$ having the length-n prefixes of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ as columns has determinant 1, and if there exists no linear combination with nonzero real coefficients giving the null vector, certainly there does not exist one with natural coefficients. Now consider the following equation:

$$\sum_{i=1}^{n} \mathbf{e}_i p_i(\vec{X}) = \sum_{i=1}^{n} \mathbf{e}_i.$$

It is a linear equation with constant right-hand side in $\mathbf{P}$ and left-hand side in $\mathbf{P}[\vec{X}]$. By linear independence over $\mathbb{N}$ of the elements $\mathbf{e}_i$, this equality holds if and only if $p_i(\vec{X}) = 1$ for all $1 \leq i \leq n$, that is, the equation has exactly the same solutions as the original system of equations, and this completes the proof. □

6 Conclusions

The quest for a suitable algebraic abstraction of dynamical systems where polynomial equations are tractable, such as a semiring R with a surjective homomorphism $\mathbf{D} \to R$ that does not erase too much information, is not over. However, we feel like the semiring $\mathbf{P}$ itself still deserves further investigation. Is the borderline between decidable and undecidable equation problems, for instance in terms of polynomial degree or number of variables, the same as for natural numbers? Are there interesting subclasses of equations that are solvable in polynomial time,

and others decidable but strictly harder than NP? Is there a polynomial-time reducibility test? And, from a more algebraic perspective, what are the prime elements of $\mathbf{P}$? Do they exist at all?

Acknowledgements Caroline Gaze-Maillot was funded by a research internship and Antonio E. Porreca by his salary of French public servant (both affiliated to Aix Marseille Université, Université de Toulon, CNRS, LIS, Marseille, France). This work is an extended version of Caroline Gaze-Maillot's research internship work [4]. We would like to thank Luca Manzoni for several fruitful discussions about the subject of this paper, in particular on the generating sets of $\mathbf{P}$ as an $\mathbb{N}$-semimodule (Theorem 16), Florian Bridoux for having the good idea on how to reduce several linear equations to a single one (Theorem 27), and Ananda Ayu Permatasari for finding an error in the original proof of Theorem 19.

References

1. Andrews, G.E.: The Theory of Partitions. Cambridge University Press (1984), https://doi.org/10.1017/CBO9780511608650
2. Dennunzio, A., Dorigatti, V., Formenti, E., Manzoni, L., Porreca, A.E.: Polynomial equations over finite, discrete-time dynamical systems. In: Mauri, G., El Yacoubi, S., Dennunzio, A., Nishinari, K., Manzoni, L. (eds.) Cellular Automata, 13th International Conference on Cellular Automata for Research and Industry, ACRI 2018. Lecture Notes in Computer Science, vol. 11115, pp. 298–306. Springer (2018), https://doi.org/10.1007/978-3-319-99813-8_27
3. Dorigatti, V.: Algorithms and Complexity of the Algebraic Analysis of Finite Discrete Dynamical Systems. Master's thesis, Università degli Studi di Milano-Bicocca (2018)
4. Gaze-Maillot, C.: Analyse algébrique des systèmes dynamiques finis (2020), mémoire de stage de Master 2 Recherche, Aix-Marseille Université
5. Goles, E., Martínez, S.: Neural and Automata Networks: Dynamical Behavior and Applications. Springer (1990), https://doi.org/10.1007/978-94-009-0529-0
6. Hammack, R., Imrich, W., Klavžar, S.: Handbook of Product Graphs. Discrete Mathematics and Its Applications, CRC Press, second edn. (2011), https://doi.org/10.1201/b10959
7. Hebisch, U., Weinert, H.J.: Semirings: Algebraic Theory and Applications in Computer Science. World Scientific (1998), https://doi.org/10.1142/3903
8. Perrot, K., Perrotin, P., Sené, S.: On Boolean automata networks (de)composition. Fundamenta Informaticae (2020), https://pageperso.lis-lab.fr/sylvain.sene/files/publi_pres/pps20.pdf, accepted
9. Porreca, A.E.: Composing behaviours in the semiring of dynamical systems. In: International Workshop on Boolean Networks (IWBN 2020) (2020), https://doi.org/10.5281/zenodo.3934396, invited talk
10. Schaefer, T.J.: The complexity of satisfiability problems. In: STOC '78: Proceedings of the Tenth Annual ACM Symposium on Theory of Computing. pp. 216–226 (1978), https://doi.org/10.1145/800133.804350
11. Sutner, K.: On the computational complexity of finite cellular automata. Journal of Computer and System Sciences **50**(1), 87–97 (1995), https://doi.org/10.1006/jcss.1995.1009