Automatic pre- and postconditions for partial differential equations

Michele Boreale
Università di Firenze, Dipartimento di Statistica, Informatica, Applicazioni (DiSIA) "G. Parenti", Viale Morgagni 65, I-50134 Firenze, Italy
michele.boreale@unifi.it

Abstract

Based on a simple automata-theoretic and algebraic framework, we study equational reasoning for Initial Value Problems (IVPs) of polynomial Partial Differential Equations (PDEs). In order to represent IVPs in their full generality, we introduce *stratified* systems, where function definitions can be decomposed into distinct subsystems, focusing on different subsets of independent variables. Under a certain coherence condition, for such stratified systems we prove existence and uniqueness of formal power series solutions, which conservatively extend the classical analytic ones. We then give a — in a precise sense, complete — algorithm to compute weakest preconditions and strongest postconditions for such systems. To some extent, this result reduces equational reasoning on PDE initial value (and boundary) problems to algebraic reasoning. We illustrate some experiments conducted with a proof-of-concept implementation of the method.

2012 ACM Subject Classification Theory of computation → Invariants; Theory of computation → Operational semantics

Keywords and phrases coalgebra, partial differential equations, polynomials, formal power series

1 Introduction

Techniques for reasoning on ordinary differential equations (ODEs) are at the heart of current formal methods and tools for continuous and hybrid systems, which form an active research area, see e.g. [27, 28, 18, 12, 13, 4] and references therein. Although examples of hybrid systems whose continuous dynamics is described by *partial* differential equations (PDEs) abund, formal techniques for reasoning on PDEs have comparably received much less attention. Existing proposals mostly focus on specific types of equations, such as the Hamilton-Jacobi equations [9, 19]. The present paper, building on [5], is meant as a contribution to developing formal methods for reasoning on PDEs. Our approach is *formal*, in the sense of being entirely based on simple coalgebra (automata theory) and algebra (polynomials), rather than on calculus like most of the previous proposals. Nevertheless, the resulting notion of PDE solution conservatively extends the classical analytic one, in a sense made precise below.

In [5] we have shown that, subject to a certain coherence condition, a system Σ of polynomial PDEs, given an arbitrary initial data specification, admits a unique solution in the set of commutative formal power series (CFPSs; Section 2). Most important, this solution can be expressed operationally, in terms of the transition function of a suitable automaton. This lays the basis for mechanical checking of equations: that is, check that a given (polynomial) expression involving the PDE variables becomes identically 0 when the solution is plugged into it. The corresponding procedure is similar in spirit to an on-the-fly bisimulation checking algorithm. Pragmatically, these CFPS solutions conservatively extend classical ones: if an analytic solution of Σ in the classical sense exists, then its Taylor expansion from 0, seen as a formal power series, coincides with the unique CFPS solution.

In the present paper, we make two substantial steps forward. First, we introduce *stratified systems*, by which one can represent fairly complicated initial value problems — and, through changes of coordinates, also boundary problems. Second and most crucial, we give a complete algorithm to automatically compute *pre-* and *postconditions* of a given system. In particular, this allows one to automatically compute *all* valid polynomial equations that fit a user-specified format (e.g., all conservation laws up to a given degree), rather than just checking the validity of given ones.

More in detail, in a stratified system we have distinct sets of equations (subsystems): in each of them, a distinct subset of the independent variables is fixed to zero. This way, in a system with, say, two independent variables x and y, the solution, $f(x,y)$, can be made dependent on constraints involving not only $f(x,y)$ and its derivatives, but also $f(x,0)$ and its x-derivatives, and $f(0,y)$ and its y-derivatives. This is how initial value problems are formulated in their generality. Under a syntactic acyclicity condition among subsystems, we prove existence and uniqueness of solutions for stratified systems and an automata-theoretic representation of the corresponding Taylor coefficients (Section 3).

This result lays the basis of an algorithm to automatically compute both weakest *preconditions* (= sets of initial data specifications) and strongest *postconditions* (= valid polynomial equations). The method is complete, subject to certain assumptions (Section 4). This way one can, for example, automatically *discover* all polynomial equations up to a given degree, valid under a given set of initial data specifications. Or vice-versa, compute the largest set of initial data specifications for given equations to be valid. The original IVP is therefore reduced to a purely algebraic system, which can be used for equational reasoning and, in some cases, to find explicit solutions. Concepts from algebraic geometry are used to prove the termination and correctness of this algorithm. Using a proof-of-concept implementation (Section 5), we illustrate this algorithm on well-known examples drawn from mathematical physics. Relations with other works, in particular on ODEs [2, 3], is discussed in the concluding section (Section 6). Proofs and additional technical material omitted from the main body of the paper are reported in a separate Appendix (Appendix A).

2 Background

We review some notation and terminology from the theory of formal power series and from the formal theory of PDEs, including the main result of [5].

Commutative formal power series and polynomials Assume a finite set $X = \{x_1, ..., x_n\}$ of *independent variables* is given. The set X, ranged over by $t, x, ...$, will be kept fixed for the rest of the paper. Let $X^\otimes$, ranged over by $\tau, \xi, ...$, be the set of *monomials*[1] that can be formed from the elements of X, in other words, the commutative monoid freely generated by X. A *commutative formal power series* (CFPS) with indeterminates in X and coefficients in $\mathbb{R}$ is a total function $f : X^\otimes \to \mathbb{R}$. The set of these CFPSs will be denoted by $\mathbb{R}[[X]]$. We will sometimes use the suggestive notation $\sum_\tau f(\tau)\cdot\tau$ to denote a CFPS $f = \lambda\tau.f(\tau)$. By slight abuse of notation, for each $\mu \in \mathbb{R}$, we will denote the CFPS that maps ϵ to μ and anything else to 0 simply as μ; while x_i will denote the i-th identity, the CFPS that maps x_i to 1 and anything else to 0. In the sequel, $\delta(f, x) \triangleq \frac{\partial f}{\partial x}$ denotes the CFPS obtained by the usual (formal) partial derivative of f w.r.t. x. For a more workable formulation of this definition, let us introduce the following notation. Let us fix any total order $\mathbf{x} = (x_1, ..., x_n)$ of the variables in X. Given a vector $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_n)$ of nonnegative integers (a *multi-index*), we let $\mathbf{x}^{\boldsymbol{\alpha}}$ denote the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Then $\frac{\partial f}{\partial x_i}$ is defined by the following, for each $\tau = \mathbf{x}^{(\alpha_1,...,\alpha_n)}$

$$\frac{\partial f}{\partial x_i}(\tau) \triangleq (\alpha_i + 1) f(x_i\tau)\,. \tag{1}$$

We recall now the sum and product operations on $\mathbb{R}[[X]]$. For any $\xi = \mathbf{x}^{\boldsymbol{\alpha}}$ and $\tau = \mathbf{x}^{\boldsymbol{\beta}}$, let $\xi \le \tau$ if for each $i = 1, ..., n$, $\alpha_i \le \beta_i$; in this case τ/ξ denotes the monomial $\mathbf{x}^{(\beta_1-\alpha_1,...,\beta_n-\alpha_n)}$. We have the following definitions of sum and product. For each $\tau \in X^\otimes$:

$$(f+g)(\tau) \triangleq f(\tau) + g(\tau) \qquad\qquad (f\cdot g)(\tau) \triangleq \sum_{\xi\le\tau} f(\xi)\cdot g(\tau/\xi)\,. \tag{2}$$

These operations correspond to the usual sum and product of functions, when (convergent) CFPSs are interpreted as analytic functions. These operations enjoy associativity, commutativity and distributivity, which make $\mathbb{R}[[X]]$ a ring. Moreover, if $f(\epsilon) \neq 0$ there exists a unique CFPS $f^{-1} \in \mathbb{R}[[X]]$ that is a multiplicative inverse of f, that is $f \cdot f^{-1} = 1$. Finally, the following familiar rules of differentiation are satisfied:

$$\frac{\partial(f+g)}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \qquad\qquad \frac{\partial(f\cdot g)}{\partial x} = \frac{\partial f}{\partial x}\cdot g + f\cdot\frac{\partial g}{\partial x}\,. \tag{3}$$

If the *support* of f, $\mathrm{supp}(f) \triangleq \{\tau : f(\tau) \neq 0\}$, is finite, we will call f a *polynomial*. The set of polynomials, denoted by $\mathbb{R}[X]$, is closed under the above defined operations of sum, product (which make it a ring) and partial derivative, but in general not inverse. It is important to note that, when confining to polynomials, sum, product and partial derivative are well defined even in case the cardinality of the set X of indeterminates is infinite.

Partial differential equations The definitions in this paragraph are standard, or slight variations of the standard ones; see e.g. [20, 16]. A finite, nonempty set U of *dependent variables*, disjoint from X and ranged over by $u, v, ...$, is given. We let $\mathcal{D} \triangleq \{u_\tau : u \in U, \tau \in X^\otimes\}$ be the set of the *derivatives*. Informally, a symbol $u \in U$ represents a function, and u_τ its partial derivative $\frac{\partial u}{\partial \tau}$; u_ϵ will be identified with u. We let $\mathcal{P} \triangleq \mathbb{R}[X \cup \mathcal{D}]$, ranged over by $E, F, ...$, denote the set of *(differential,*

[1] In general, we shall adopt for monomials the same notation we use for strings, as the context is sufficient to disambiguate. In particular, we overload the symbol ϵ to denote both the empty string and the empty monomial. When $X = \emptyset$, $X^\otimes \triangleq \{\epsilon\}$.

multivariate) polynomials with coefficients in $\mathbb{R}$ and indeterminates in $X \cup \mathcal{D}$. Considered as formal objects, polynomials are just finite-support CFPSs in $\mathbb{R}[[X \cup \mathcal{D}]]$, as per previous paragraph. As such, they inherit from $\mathbb{R}[[X \cup \mathcal{D}]]$ the operations of sum, product and partial derivative, along with the corresponding properties. Syntactically, we shall write polynomials as expressions of the form $\sum_{\gamma \in M} \lambda_\gamma \cdot \gamma$, for $0 \neq \lambda_\gamma \in \mathbb{R}$ and $M \subseteq_{\text{fin}} (X \cup \mathcal{D})^\otimes$. Note that this notation is consistent with the sum and product operations introduced in (2). For example, $E = v_z u_{xy} + v_y^2 + u + 5x$ is a polynomial[2]. For an independent variable $x \in X$, the *total derivative* of $E \in \mathcal{P}$ w.r.t. x is just the derivative of E w.r.t x, taking into account that $\frac{\partial u_\tau}{\partial x} = u_{x\tau}$ and the chain rule. Formally, the operator $D_x : \mathcal{P} \to \mathcal{P}$ is defined by (note $\sum$ below has only finitely many nonzero terms)

$$D_x E \triangleq \frac{\partial E}{\partial x} + \sum_{u,\tau} u_{x\tau} \cdot \frac{\partial E}{\partial u_\tau}$$

where $\frac{\partial E}{\partial a}$ denotes the partial derivative of polynomial E along $a \in X \cup \mathcal{D}$. D_x inherits differentiation rules for sum and product that are the analog of (3). As an example, for the polynomial E above, we have $D_x E = v_{xz} u_{xy} + v_z u_{xxy} + 2 v_y v_{xy} + u_x + 5$. In particular, $D_x u_\tau = u_{x\tau}$ and $D_x x^k = k x^{k-1}$. Just as partial derivatives, total derivatives commute with each other, that is $D_x D_y F = D_y D_x F$. This suggests to extend the notation to monomials: for any monomial $\tau = x_1 \cdots x_m$, we let $D_\tau F$ be $D_{x_1} \cdots D_{x_m} F$, where the order of the variables is irrelevant. We formally introduce systems of PDEs below, along with the key notions of *parametric* and *principal* derivatives. Here, the intuition is that parametric derivatives play a role similar to the lower order derivatives in ODEs initial value problems: once we fix the values of those functions at the origin, the solution will be uniquely determined. On the other hand, the definition of the principal derivatives depends on the parametric ones, just like higher order derivatives in ODEs depend on the lower order ones.

▶ **Definition 2.1** (system of PDEs)**.** *A* system of PDEs *is a nonempty set Σ of equations (pairs) of the form $u_\tau = E$, with $E \in \mathcal{P}$. The set of derivatives u_τ that appear as left-hand sides of equations in Σ is denoted by* $\mathrm{dom}(\Sigma)$. *Based on Σ, the set $\mathcal{D}$ is partitioned into the sets of* principal *and* parametric *derivatives, defined as follows.*

$$\mathcal{P}\mathrm{r}(\Sigma) \triangleq \{u_{\tau\xi} : u_\tau \in \mathrm{dom}(\Sigma) \text{ and } \xi \in X^\otimes\} \qquad \mathcal{P}\mathrm{a}(\Sigma) \triangleq \mathcal{D} \setminus \mathcal{P}\mathrm{r}(\Sigma).$$

We let $\mathcal{P}_0(\Sigma) \triangleq \mathbb{R}[X \cup \mathcal{P}\mathrm{a}(\Sigma)]$ be the set of Σ-normal forms*.*

▶ **Example 2.2** (Heat equation)**.** The Heat equation in one spatial dimension, $u_t(t,x) = u_{xx}(t,x)$, corresponds to $X = \{t,x\}$, $U = \{u\}$ and $\Sigma = \{u_t = u_{xx}\}$. Here we have $\mathcal{P}\mathrm{r}(\Sigma) = \{u_{t\tau} : \tau \in X^\otimes\}$ and $\mathcal{P}\mathrm{a}(\Sigma) = \{u_{x^j} : j \geq 0\}$. See Figure 1, left.

Note that we do *not* insist that each derivative occurs at most once as left-hand side in Σ. The *infinite prolongation* of a system Σ, denoted Σ^∞, is the system of PDEs of the form $u_{\xi\tau} = D_\xi F$, where $u_\tau = F$ is in Σ and $\xi \in X^\otimes$. Of course, $\Sigma^\infty \supseteq \Sigma$. Moreover, Σ and Σ^∞ induce the *same* sets of principal and parametric derivatives.

We can now introduce the concept of *solution* of PDEs, which is based on a PDE's analog of initial value problems (IVPs). In what follows, we say a function $\psi : \mathcal{P} \to \mathbb{R}[[X]]$ is a *homomorphism* if it preserves sum and product, as expected, and additionally: preserves derivatives, that is $\psi(u_\tau) = \frac{\partial}{\partial \tau}\psi(u)$, and maps each $x_i \in X$ to the i-th identity CFPS. For any function $\psi : U \to \mathbb{R}[[X]]$, its homomorphic extension $\mathcal{P} \to \mathbb{R}[[X]]$, is defined as expected; by slight abuse of notation, we will still denote by "ψ" the homomorphic extension of ψ over $\mathcal{P}$. In the definition below, it is useful to bear in mind that,

[2] Real arithmetic expressions will be used as a meta-notation for polynomials: e.g. $(u + u_x + 1) \cdot (x + u_y)$ denotes the polynomial $xu + uu_y + xu_x + u_x u_y + x + u_y$.

informally, for a parametric derivative u_τ, the initial data value $\rho(u_\tau)$ specifies the value of $\frac{\partial u}{\partial \tau}$ at the origin.

▶ **Definition 2.3** (initial value problem). *Let* Σ *be a system of PDEs. An* initial data specification *is a mapping* $\rho : \mathcal{P}\mathrm{a}(\Sigma) \to \mathbb{R}$. *An* initial value problem (IVP) *is a pair* $\mathbf{iP} = (\Sigma, \rho)$.

A solution *of* $\mathbf{iP}$ *is a homomorphism* $\psi : \mathcal{P} \to \mathbb{R}[[X]]$ *such that: (a) the initial value conditions are satisfied, that is* $\psi(u_\tau)(\epsilon) = \rho(u_\tau)$ *for each* $u_\tau \in \mathcal{P}\mathrm{a}(\Sigma)$*; and (b) all equations are satisfied, that is* $\psi(u_\tau) = \psi(F)$ *for each* $u_\tau = F$ *in* Σ^∞.

For Σ to have a solution, a few syntactic conditions must be imposed, whose purpose is to avoid inconsistencies in the equational theory generated by Σ. A *ranking* is a total order $\prec$ of $\mathcal{D}$ such that: (a) $u_\tau \prec u_{x\tau}$, and (b) $u_\tau \prec v_\xi$ implies $u_{x\tau} \prec v_{x\xi}$, for each $x \in X$, $\tau, \xi \in X^\otimes$ and $u, v \in U$. Dickson's lemma [10] implies that $\mathcal{D}$ with $\prec$ is a well-order, and in particular that there is no infinite descending chain in it. The system Σ is $\prec$*-normal* if, for each equation $u_\tau = E$ in Σ, $u_\tau \succ v_\xi$, for each v_ξ appearing in E. An easy but important consequence of condition (b) above is that if Σ is normal then also its prolongation Σ^∞ is normal.

Now, consider the equational theory over $\mathcal{P}$ induced by the equations in Σ^∞. More precisely, write $E \to_\Sigma F$ if F is the polynomial that is obtained from E by replacing one occurrence of u_τ with G, for some equation $u_\tau = G \in \Sigma^\infty$. Note, in particular, that $E \in \mathcal{P}$ cannot be rewritten if and only if $E \in \mathcal{P}_0(\Sigma)$. We let $=_\Sigma$ denote the reflexive, symmetric and transitive closure of $\to_\Sigma$. The following definition formalizes the key concepts of consistency and coherence of Σ. Basically, as shown in [5], under the natural requirement of normality, consistency is a necessary and sufficient condition for Σ to admit a unique solution under *arbitrary* initial conditions.

▶ **Definition 2.4** (coherence). *Let* Σ *be a system of PDEs.*

- Σ *is* consistent *if for each* $E \in \mathcal{P}$ *there is a unique* $F \in \mathcal{P}_0(\Sigma)$ *such that* $E =_\Sigma F$.
- *Let* $\prec$ *be a ranking. A system* Σ *is* $\prec$-coherent *if it is* $\prec$*-normal and consistent.*

As an example, the Heat equation in Example 2.2 is obviously consistent, as it features just one equation. Moreover, it is $\prec$-coherent w.r.t. the ranking $u_\tau \prec u_\xi$ iff $\tau \prec_{\mathrm{lex}} \xi$, where $\prec_{\mathrm{lex}}$ is the lexicographic monomial order induced by $t > x$. For any consistent system, we can define a *normal form function*

$$S_\Sigma : \mathcal{P} \to \mathcal{P}_0(\Sigma)$$

by letting $S_\Sigma E \triangleq F$, for the unique $F \in \mathcal{P}_0(\Sigma)$ such that $E =_\Sigma F$. The term $S_\Sigma E$ will be often abbreviated as $S\,E$, if Σ is understood from the context. Deciding if a (finite) system Σ is coherent, for a suitable ranking $\prec$, is of course a nontrivial problem. Since $\prec$ is a well-order, there are no infinite sequences of rewrites $E_1 \to_\Sigma E_2 \to_\Sigma E_3 \to_\Sigma \cdots$: therefore it is possible to rewrite any E into some $F \in \mathcal{P}_0(\Sigma)$ in a finite number of steps. Proving coherence reduces then to proving $\to_\Sigma$ confluent. For our purposes, it is enough to know that completing a given system of equations to make it coherent, or deciding that this is impossible, can be achieved by one of many existing computer algebra algorithms, like those in [20, 16]; see the discussion and the references in [5]. In many cases arising from applications, say mathematical physics, transforming the system into a coherent form for an appropriate ranking can be accomplished manually, without much difficulty: see the examples in Section 5.

We can now characterize explicitly the solutions of a coherent Σ. Informally, for any fixed ρ, the CFPS associated with $E \in \mathcal{P}$ takes each monomial $\tau \in X^\otimes$ to the real obtained by evaluating the τ-derivative of E under ρ, once this derivative is written in normal form. Formally, the characterization is based on a transition function, $\delta_\Sigma : \mathcal{P} \times X \to \mathcal{P}_0(\Sigma)$, defined as

$$\delta_\Sigma(E, x) \triangleq S_\Sigma D_x E \,. \tag{4}$$

It can be shown (see [5]) that δ_Σ satisfies the following commutation property: $\delta_\Sigma(\delta_\Sigma(E,x),y) = \delta_\Sigma(\delta_\Sigma(E,y),x)$ for all $x,y \in X$. This justifies the notation $\delta_\Sigma(E,\tau)$ for $\tau \in X^\otimes$, with $\delta_\Sigma(E,\epsilon) \stackrel{\triangle}{=} S_\Sigma E$. Next, an initial data specification $\rho : \mathcal{P}\mathrm{a}(\Sigma) \to \mathbb{R}$ can be extended homomorphically to a function $\mathcal{P}_0(\Sigma) \to \mathbb{R}$, interpreting $+$ and $\cdot$ as the usual sum and product over $\mathbb{R}$, and letting $\rho(x) \stackrel{\triangle}{=} 0$ for each independent variable $x \in X$. The following theorem of existence and uniqueness of solutions is the main result of [5]. For the sake of completeness, the proof is also reproduced in Appendix A.1. Below, for $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$, we define $\alpha! \stackrel{\triangle}{=} \alpha_1! \cdots \alpha_n!$.

▶ **Theorem 2.5** (existence and uniqueness of solution, [5]). *Let Σ be coherent. For any initial data specification ρ, there is a unique solution $\phi_{\mathbf{iP}} : \mathcal{P} \to \mathbb{R}[[X]]$ of the IVP $\mathbf{iP} = (\Sigma, \rho)$. Moreover, $\phi_{\mathbf{iP}}$ satisfies the following formula, for each $E \in \mathcal{P}$ and $\tau = \mathbf{x}^\alpha \in X^\otimes$.*

$$\phi_{\mathbf{iP}}(E)(\tau) = \frac{\rho(\delta_\Sigma(E,\tau))}{\alpha!}\,. \tag{5}$$

We remark that our concept of solution of a PDE IVP conservatively extends the classical solution concept, in the following sense: if a classical solution exists that is analytic around the origin, then its Taylor expansion, seen as a formal power series, coincides with the CFPS solution (Appendix A.3).

3 Stratified systems

Consider the Heat equation of Example 2.2. Suppose we want to specify that the temperature at time $t = 0$ varies along the x-line according to, say, $u(0,x) = \exp(-x)$. With the pure PDE formalism introduced so far, the only way to do so is to describe the function $u(0,x)$ explicitly through the initial data ρ, that is: $\rho(u_{x^j}) = (\frac{d}{dx^j}u(0,x))_{|x=0} \stackrel{\triangle}{=} (-1)^j/j!$, for each $j \geq 0$. Such a ρ is an infinite object which does not obviously lend itself to equational and algorithmic manipulations. It would be more natural, instead, to specify $u(0,x)$ simply via a subsystem $\Sigma_0 = \{u_x = -u\}$ (plus the single initial condition $\rho(u) = 1$), somehow prescribing that this equation applies when fixing $t = 0$, so that the resulting function only depends on x. More generally, a pure PDE system Σ alone cannot express general initial value problems, where one wants to specify constraints on the functions obtained by keeping the value of certain independent variables fixed. This limitation is overcome by stratified systems, introduced below.

We first introduce *subsystems*. Let us fix once and for all a nonempty set of dependent variables U, and a finite set of independent variables X. For $Y \subseteq X$, a Y-subsystem defines, informally, functions where variables outside Y have been zeroed. In particular, derivatives can be taken only along variables in Y. We need now some standard notation on partial orders. For a partial order $\preceq$ defined over some universe set A and for $B \subseteq A$, we will let $\uparrow_\preceq(B) \stackrel{\triangle}{=} \{a \in A : a \succeq b \text{ for some } b \in B\}$ denote the upward closure of B w.r.t $\preceq$; similarly, we will let $\downarrow_\preceq(B)$ denote the downward closure of B. Moreover, we will let $\min_\preceq(B) \stackrel{\triangle}{=} \{b \in B : \text{ whenever } b' \in B \text{ and } b' \preceq b \text{ then } b' = b\}$ denote the set of $\preceq$-minimal element of B. Additionally, we define the following partial order $\leq_Y$ on the set of derivatives $\mathcal{D}$, depending on $Y \subseteq X$: $u_\tau \leq_Y u_{\tau'}$ if and only if $\tau' = \tau\xi$ for some $\xi \in Y^\otimes$. In the definition of subsystem given below, the intuition is that the $\leq_Y$-minimal derivatives, the set U_Γ, act as the dependent variables of a new system of PDEs with independent variables in Y and derivatives in $\mathcal{D}_\Gamma$.

▶ **Definition 3.1** (subsystem). *Let Σ a set of equations and $Y \subseteq X$. For $\Gamma = (\Sigma, Y)$, we define the following subsets of $\mathcal{D}$.*

$$\begin{aligned} U_\Gamma &\stackrel{\triangle}{=} \min_{\leq_Y}(\downarrow_{\leq_Y}\{u_\tau : u_\tau \textit{ occurs in } \Sigma\}) & \qquad \mathcal{D}_\Gamma &\stackrel{\triangle}{=} \uparrow_{\leq_Y}(U_\Gamma) \\ \mathcal{P}\mathrm{r}(\Gamma) &\stackrel{\triangle}{=} \uparrow_{\leq_Y}(\mathrm{dom}(\Sigma)) & \qquad \mathcal{P}\mathrm{a}(\Gamma) &\stackrel{\triangle}{=} \mathcal{D}_\Gamma \setminus \mathcal{P}\mathrm{r}(\Gamma)\,. \end{aligned}$$

Figure 1 Derivatives partially ordered under $\leq$. *Left*: system Σ of Example 2.2, where dark-shaded region = $\mathcal{P}\mathrm{r}(\Sigma)$, white region = $\mathcal{P}\mathrm{a}(\Sigma)$. *Right*: stratified system $H = \{\Gamma_1, \Gamma_2\}$ of Example 3.3, where dark-shaded region = $\mathcal{P}\mathrm{r}(\Gamma_1)$, light-shaded region = $\mathcal{P}\mathrm{r}(\Gamma_2)$, white region = $\mathcal{P}\mathrm{a}(H)$.

We let $\mathcal{P}_\Gamma \triangleq \mathbb{R}[Y \cup \mathcal{D}_\Gamma]$. *We say* $\Gamma = (\Sigma, Y)$ *is a* Y-subsystem *if* U_Γ *is finite, and for each polynomial* E *appearing in* Σ, $E \in \mathcal{P}_\Gamma$. *We call* Γ *a* main *subsystem if* $Y = X$ *and* $U_\Gamma = U$. *Finally,* $\Gamma^\infty \triangleq \{u_{\tau\xi} = D_\xi G : u_\tau = G \in \Sigma$ *and* $\xi \in Y^\otimes\}$.

Stratified systems can encode initial value problems in their general form. A precedence relation among components, $\Gamma_i \prec \Gamma_j$, formalizes that Γ_j depends on parametric variables that are defined (are principal) in Γ_i.

▶ **Definition 3.2** (stratified system)**.** *A* stratified system *is a finite set of subsystems* $H = \{\Gamma_1, ..., \Gamma_m\}$ *(*$m \geq 1$, $\Gamma_i = (\Sigma_i, X_i)$, $\Sigma_i \neq \emptyset$, $X_i \subseteq X$*) such that:*

(a) for some $1 \leq j \leq m$, Γ_j *is a main subsystem; we will conventionally take* $j = 1$*;*

(b) for any $i \neq j$, $\mathcal{P}\mathrm{r}(\Gamma_i) \cap \mathcal{P}\mathrm{r}(\Gamma_j) = \emptyset$*;*

(c) the binary relation over $\{1, ..., m\}$ *defined as* $i \prec j$ *iff* $\mathcal{P}\mathrm{r}(\Gamma_i) \cap \mathcal{P}\mathrm{a}(\Gamma_j) \neq \emptyset$*, is acyclic.*

The parametric derivatives and normal forms of H *are* $\mathcal{P}\mathrm{a}(H) \triangleq \mathcal{D} \setminus (\cup_{i=1}^m \mathcal{P}\mathrm{r}(\Gamma_i))$ *and* $\mathcal{P}_0(H) \triangleq \mathbb{R}[X \cup \mathcal{P}\mathrm{a}(H)]$*, respectively.* H *is* coherent *if all of its subsystems are coherent w.r.t. one and the same ranking on* $\mathcal{D}$.

Note that each H features a unique main subsystem.

▶ **Example 3.3** (Heat equation with initial temperature)**.** Consider the Heat equation of Example 2.2, with an initial temperature exponentially decaying from the origin, $u_x(0, x) = -u(0, x)$. The corresponding stratified system is $H = \{\Gamma_1, \Gamma_2\} = \{(\Sigma_1, X_1), (\Sigma_2, X_2)\}$ with $\Sigma_1 = \{u_t = u_{xx}\}$, $X_1 = X = \{t, x\}$ and $\Sigma_2 = \{u_x = -u\}$, $X_2 = \{x\}$. We have (see Fig. 1, right):

$$\begin{array}{llllllll} U_{\Gamma_1} = \{u\} & \mathcal{D}_{\Gamma_1} = \{u_\tau : \tau \in X^\otimes\} & \mathcal{P}\mathrm{r}(\Gamma_1) = \{u_{t\tau} : \tau \in X^\otimes\} & \mathcal{P}\mathrm{a}(\Gamma_1) = \{u_{x^j} : j \geq 0\} \\ U_{\Gamma_2} = \{u\} & \mathcal{D}_{\Gamma_2} = \{u_{x^j} : j \geq 0\} & \mathcal{P}\mathrm{r}(\Gamma_2) = \{u_{x^j} : j \geq 1\} & \mathcal{P}\mathrm{a}(\Gamma_2) = \{u\}. \end{array}$$

Note that $\mathcal{D}_{\Gamma_1} = \mathcal{D}$, so Γ_1 is the main subsystem, and that $\mathcal{P}\mathrm{a}(H) = \{u\}$. Clearly, $2 \prec 1$, as $\mathcal{P}\mathrm{r}(\Gamma_2) \cap \mathcal{P}\mathrm{a}(\Gamma_1) \neq \emptyset$; on the other hand, $1 \nprec 2$, as $\mathcal{P}\mathrm{r}(\Gamma_1) \cap \mathcal{P}\mathrm{a}(\Gamma_2) = \emptyset$; so the relation $\prec$ is acyclic. Finally, fixing the lexicographic order induced by $t > x$, H is trivially seen to be coherent.

In order to define solutions of stratified systems, let us introduce some additional notation about CFPSs. For a CFPS $f \in \mathbb{R}[[X]]$ and $Y \subseteq X$, we can consider the CFPS $f_{|Y^\otimes} \in \mathbb{R}[[Y]]$. For an intuitive explanation of this concept, assume e.g. f represents $f(x_1, x_2)$ and $Y = \{x_2\}$: recalling that we take the origin as the expansion point, $f_{|Y^\otimes}$ represents $f(0, x_2)$, that is, f where the variables not in Y have been replaced by 0. Formally, for $\psi : \mathcal{P} \to \mathbb{R}[[X]]$ and a subsystem $\Gamma = (\Sigma, Y)$, we let $\psi_\Gamma : \mathcal{P}_\Gamma \to \mathbb{R}[[Y]]$ be defined as: $\psi_\Gamma(E) \triangleq \psi(E)_{|Y^\otimes}$ for each $E \in \mathcal{P}_\Gamma$.

▶ **Definition 3.4** (solutions of H)**.** *Let H be a stratified system.*

1. *A* solution *of* H *is a homomorphism* $\psi : \mathcal{P} \to \mathbb{R}[[X]]$ *such that for each* $\Gamma_i \in H$, $\psi_{\Gamma_i} : \mathcal{P}_{\Gamma_i} \to \mathbb{R}[[X_i]]$ *respects all the equations in* Γ_i^∞.

2. *Let $\rho : \mathcal{P}\mathrm{a}(H) \to \mathbb{R}$ be an initial data specification. Let $\Sigma_0 = \{u_\tau = \rho(u_\tau) : u_\tau \in \mathcal{P}\mathrm{a}(H)\}$ and $\Gamma_0 = (\Sigma_0, \emptyset)$. A* solution of the initial value problem $\mathbf{iP} = (H,\rho)$ *is solution of the stratified system $H \cup \{\Gamma_0\}$.*

We can linearly order the subsystems of H according to a total order compatible with $\prec$ and then lift inductively existence and uniqueness (Theorem 2.5) to H.

▶ **Theorem 3.5** (existence and uniqueness for H)**.** *Let H be a coherent stratified system. For any initial data specification ρ for H, there is a unique solution of* $\mathbf{iP} = (H,\rho)$*.*

We illustrate the idea behind the proof of Theorem 3.5 on the Heat equation of Example 3.3.

▶ **Example 3.6** (Example 3.3, cont.)**.** Let us fix any initial data specification $\rho(u) = u_0 \in \mathbb{R}$ for H. As prescribed by Def. 3.4(2), we consider the extended system $\overline{H} \stackrel{\triangle}{=} H \cup \{\Gamma_0\}$, where $\Gamma_0 = (\{u = u_0\}, \emptyset)$. Note that $U_{\Gamma_0} = \mathcal{D}_{\Gamma_0} = \mathcal{P}\mathrm{r}(\Gamma_0) = \{u\}$ and $\mathcal{P}\mathrm{a}(\Gamma_0) = \emptyset$. Now we build a sequence of IVPs $\mathbf{iP}_i$, and corresponding solutions $\psi_i : \mathcal{P}_{\Gamma_i} \to \mathbb{R}[[X_i]]$, for the subsystems Γ_i's in $\overline{H}$. The construction proceeds inductively on a linear order compatible with $\prec$, that is: $0 \prec 2 \prec 1$. The definition of each initial data specification $\rho_i : \mathcal{P}\mathrm{a}(\Gamma_i) \to \mathbb{R}$ relies on the solutions ψ_j for $j \prec i$. The existence of such solutions is guaranteed by Theorem 2.5. In particular:

- $\mathbf{iP}_0 = (\{u = u_0\}, \rho_0)$, with $\rho_0(u) \stackrel{\triangle}{=} \emptyset$ (empty function), has solution[3] $\psi_0 : \mathcal{P}_{\Gamma_0}(= \mathbb{R}[u]) \to \mathbb{R}[[\emptyset]]$;
- $\mathbf{iP}_2 = (\{u_x = -u\}, \rho_2)$, with $\rho_2(u) \stackrel{\triangle}{=} \psi_0(u)(\epsilon)$, has solution $\psi_2 : \mathcal{P}_{\Gamma_2}(= \mathbb{R}[x,u]) \to \mathbb{R}[[x]]$;
- $\mathbf{iP}_1 = (\{u_t = u_{xx}\}, \rho_1)$, with $\rho_1(u_{x^k}) \stackrel{\triangle}{=} \psi_2(u_{x^k})(\epsilon)$ $(k \geq 0)$, has solution $\psi_1 : \mathcal{P}_{\Gamma_1}(= \mathcal{P}) \to \mathbb{R}[[t,x]]$.

It can be shown — and this is the nontrivial part of Theorem 3.5 — that the solution of the main subsystem, ψ_1, is a solution of $\overline{H}$ (Def. 3.4(1)), and in particular: $(\psi_1)_{\Gamma_i} = \psi_i$ for each i. Hence ψ_1 is the (unique) solution of (H,ρ).

In view of the subsequent algorithmic developments, the next step is to obtain a formula for the Taylor coefficients of the solutions of H, in analogy with the formula (5) for pure systems. This formula will be based on the transition function of the main subsystem, δ_{Σ_1}. However, a pivotal role will now be also played by a reduction function S_H, introduced below: it will allow one to rewrite any $E \in \mathcal{P}$ to a normal form in $\mathcal{P}_0(H)$, where it can be evaluated for any given initial data specification ρ for H. Below, $\to_{\Sigma_i}$ denotes the rewrite relation over $\mathcal{P}$ induced by the equations in Γ_i^∞.

▶ **Definition 3.7** (reduction S_H)**.** *Let $H = \{\Gamma_1, ..., \Gamma_m\}$ be a coherent stratified system. Let $=_H$ denote the reflexive, symmetric and transitive closure over $\mathcal{P}$ of $\to_{\Sigma_1} \cup \cdots \cup \to_{\Sigma_m}$. For each $E \in \mathcal{P}$, we let $S_H E$ denote an arbitrarily fixed $F \in \mathcal{P}_0(H)$ such that $E =_H F$.*

In the definition above, note that, due to normality, each $E \in \mathcal{P}$ must have an $=_H$-equivalent term in $\mathcal{P}_0(H)$, so $S_H E$ is well defined. Now, let ϕ be a solution of an IVP (H,ρ). If $E =_H F$, it is *not* true in general that $\phi(E) =_H \phi(F)$. It *is* true, however, that $\phi(E)(\epsilon) =_H \phi(F)(\epsilon)$; moreover if $F \in \mathcal{P}_0(H)$ then $\phi(F)(\epsilon) = \rho(F)$. This fact is quite intuitive, recalling the informal interpretation of $f(\epsilon)$ as $f(0)$ for a CFPS f. For instance, in the Heat equation system of Example 3.3, one would have $u_t(0,0) = u_{xx}(0,0) = u(0,0)(= \rho(u))$, where the first and second equality follow from applying Σ_1 and Σ_2 (twice), respectively. Formally, we have the following formula, giving the Taylor coefficients of $\phi(E)$. This is also key to the algorithm in the next section.

▶ **Corollary 3.8** (Taylor coefficients)**.** *Let H be a coherent stratified system. Denote by δ_{Σ_1} the transition function of the main subsystem of H. For any initial data specification ρ for H, the unique solution ϕ of (H,ρ) enjoys the following, for every $E \in \mathcal{P}$ and $\tau = \mathbf{x}^\alpha \in X^\otimes$.*

$$\phi(E)(\tau) = \frac{\rho(S_H(\delta_{\Sigma_1}(E,\tau)))}{\alpha!}. \tag{6}$$

[3] Specifically, $\psi_0(E)(\epsilon) = E(u_0)$ for each $E \in \mathbb{R}[u]$.

▶ **Example 3.9** (Example 3.3, cont.). Consider any initial data specification $\rho(u) = u_0 \in \mathbb{R}$ for H, let ψ be the solution of (H,ρ) and $f = \psi(u)$. We compute the first few coefficients of f by applying (6) with $E = u$. Let us first compute a few $S_H(\delta_{\Sigma_1}(u,\tau))$ s. Recall that the definition of $=_{\Sigma_i}$ is based on Γ_i^∞ $(i = 1,2)$.

$$S_H(\delta_{\Sigma_1}(u,\epsilon)) = S_H u = u, \quad S_H(\delta_{\Sigma_1}(u,t)) = S_H u_{xx} = S_H(-u_x) = u, \quad S_H(\delta_{\Sigma_1}(u,x)) = S_H u_x = -u$$
$$S_H(\delta_{\Sigma_1}(u,tt)) = S_H u_{x^4} = u, \quad S_H(\delta_{\Sigma_1}(u,tx)) = S_H u_{x^3} = -u, \quad S_H(\delta_{\Sigma_1}(u,xx)) = S_H u_{xx} = u\,.$$

In general, one can check that for $\tau = (t,x)^\alpha$, $\alpha = (\alpha_1,\alpha_2) \in \mathbb{N}^2$, $S_H(\delta_{\Sigma_1}(u,\tau)) = (-1)^{\alpha_2}u$. Hence, by (6), we have the CFPS: $f = u_0 + u_0 t - u_0 x + (u_0/2)t^2 - u_0 tx + (u_0/2)x^2 \cdots = \sum_{\tau=\mathbf{x}^\alpha}(-1)^{\alpha_2}(u_0/\alpha!)\tau$.

4 Algorithms for pre- and postconditions

We will first recall some terminology and some basic facts from algebraic geometry, then introduce pre- and postconditions and finally the Post algorithm to compute them.

Preliminaries From now on, we will restrict our attention to the subset of *autonomous* polynomials, $\mathcal{P}^{\mathrm{a}} \triangleq \mathbb{R}[\mathcal{D}]$.

▶ **Definition 4.1** (autonomous and finite parameters). *A stratified system H is* autonomous *if the polynomials occurring in H are autonomous. H is* finite-parameter *if $\mathcal{P}\mathrm{a}(H)$ is finite. It is* afp *if it is both finite-parameter and autonomous.*

For instance, the system in Example 3.3 is afp, while the system consisting of Γ_1 only is autonomous, but not finite-parameter. In concrete applications, one expects that most systems are afp[4]. Let us now introduce some additional notation and terminology about polynomials. According to (6), the calculation of the Taylor coefficients of a solution of an autonomous IVP $\mathbf{iP} = (H,\rho)$ involves evaluating expressions in $\mathcal{P}^{\mathrm{a}}_0(H) = \mathbb{R}[\mathcal{P}\mathrm{a}(H)]$. As $k \triangleq |\mathcal{P}\mathrm{a}(H)| < +\infty$, elements of $\mathcal{P}^{\mathrm{a}}_0(H)$ can be treated as usual multivariate polynomials in a *finite* number of indeterminates. In particular, we can identify initial data specifications ρ for H with points in $\mathbb{R}^k$. Accordingly, for polynomials $E \in \mathcal{P}^{\mathrm{a}}_0(H)$ and initial data specification $\rho \in \mathbb{R}^k$, it is notationally convenient to write $\rho(E)$ as $E(\rho)$, that is the value in $\mathbb{R}$ obtained by evaluating the polynomial E at point ρ.

In what follows, we shall use a few elementary notions from algebraic geometry. In particular, an *ideal* $J \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ is a nonempty set of polynomials closed under addition, and under multiplication by polynomials in $\mathcal{P}^{\mathrm{a}}_0(H)$. For $P \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$, $\langle\, P \,\rangle \triangleq \{\sum_{i=1}^m F_i \cdot E_i \,:\, m \geq 0, F_i \in \mathcal{P}^{\mathrm{a}}_0(H), E_i \in P\}$ denotes the smallest ideal which includes P, and $\mathbf{V}(P) \subseteq \mathbb{R}^k$ the *(affine) variety* induced by P: $\mathbf{V}(P) \triangleq \{\rho \in \mathbb{R}^k \,:\, E(\rho) = 0 \text{ for each } E \in P\} \subseteq \mathbb{R}^k$. For $W \subseteq \mathbb{R}^k$, $\mathbf{I}(W) \triangleq \{E \in \mathcal{P}^{\mathrm{a}}_0(H) : E(\rho) = 0 \text{ for each } \rho \in V\}$. We will use a few basic facts about ideals and varieties: (a) both $\mathbf{I}(\cdot)$ and $\mathbf{V}(\cdot)$ are inclusion reversing: $P_1 \subseteq P_2$ implies $\mathbf{V}(P_1) \supseteq \mathbf{V}(P_2)$ and $W_1 \subseteq W_2$ implies $\mathbf{I}(W_1) \supseteq \mathbf{I}(W_2)$; (b) any ascending chain of ideals $I_0 \subseteq I_1 \subseteq \cdots \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ stabilizes in a finite number of steps (Hilbert's basis theorem); (c) for finite $P \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$, the problem of deciding if $E \in \langle\, P \,\rangle$ is decidable, by computing a Gröbner basis (a set of generators with special properties) of $\langle\, P \,\rangle$. See [10] for a comprehensive treatment.

Preconditions and postconditions. Informally, computing the *preconditions* of a given set $Q \subseteq \mathcal{P}^{\mathrm{a}}$ means finding all the initial data specifications $\rho \in \mathbb{R}^k$ under which all the polynomials in Q represent valid equations for the system H — that is, they become identically zero when one plugs the solution of (H,ρ) into them. Dually, computing the *postconditions* of a given set of initial data

[4] In particular, the restriction to autonomous systems implies no loss of generality, as independent variables can always be encoded as dependent ones. See Example 5.1.

specifications $W \subseteq \mathbb{R}^k$ means finding the set $Q \subseteq \mathcal{P}^{\mathrm{a}}$ of all polynomial equations that are valid under all initial data $\rho \in W$. Here, we shall confine ourselves to *algebraic* sets W, that is $W = \mathbf{V}(P)$ for some $P \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$. Formally, we have the following definition. Recall that, for a coherent H and an initial data specification $\rho \in \mathbb{R}^k_0$, we let $\phi_{(H,\rho)} : \mathcal{P} \to \mathbb{R}[[X]]$ denote the unique solution of the IVP (H,ρ).

▶ **Definition 4.2** (pre- and postconditions). *Let H be coherent and* AFP*. Let P and Q be sets of polynomials such that $P \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ and $Q \subseteq \mathcal{P}^{\mathrm{a}}$. We define the sets of* weakest preconditions $\mathrm{wp}_H(Q) \subseteq \mathbb{R}^k$ *and of the* strongest postconditions $\mathrm{sp}_H(P) \subseteq \mathcal{P}^{\mathrm{a}}$ *as follows.*

$$\mathrm{wp}_H(Q) \triangleq \{\rho \in \mathbb{R}^k \,:\, \phi_{(H,\rho)}(E) = 0 \text{ for each } E \in Q\}$$

$$\mathrm{sp}_H(P) \triangleq \{E \in \mathcal{P}^{\mathrm{a}} \,:\, \phi_{(H,\rho)}(E) = 0 \text{ for each } \rho \in \mathbf{V}(P)\}.$$

Any $W \subseteq \mathrm{wp}_H(Q)$ will be called an (algebraic) precondition for Q, any $R \subseteq \mathrm{sp}_H(P)$ a postcondition for $\mathbf{V}(P)$. We focus here on computing strongest postconditions, which, as we shall see, can be used to compute preconditions as well. Actually, it is computationally convenient to introduce a *relativized* version of this problem.

Given user-specified sets P and R ($P \subseteq_{\mathrm{fin}} \mathcal{P}^{\mathrm{a}}_0(H)$ and $R \subseteq \mathcal{P}^{\mathrm{a}}$), find a finite characterization of $\mathrm{sp}_H(P) \cap R$. (7)

By 'finding a finite characterization', we mean effectively computing a finite set of generators, of an appropriate algebraic type, for the set in question (see next paragraph). Following a well-established tradition in the field of continuous and hybrid system, the set R will be represented by means of a polynomial template, to be introduced shortly.

A double chain algorithm. We first introduce *polynomial templates* [27], that is, polynomials in $\mathrm{Lin}(\mathbf{a})[\mathcal{D}]$, where $\mathrm{Lin}(\mathbf{a})$ are (formal) linear combinations of the parameters in $\mathbf{a} = (a_1, ..., a_s)$ (for fixed $s \geq 1$) with real coefficients. For instance, $\ell = 5a_1 + 42a_2 - 3a_3$ is one such expression[5]. In other words, a polynomial template has the form $\pi = \sum_i \ell_i \gamma_i$ for distinct monomials $\gamma_i \in \mathcal{D}^{\otimes}$, and ℓ_i linear expressions in the parameters a_i s. For example, the following is a template: $\pi = (5a_1 + (3/4)a_3)u_x v^2 + (7a_1 + (1/5)a_2)uv_{xy} + (a_2 + 42a_3)$. A *parameter evaluation* is a vector $v = (v_1, ..., v_s) \in \mathbb{R}^s$; we denote by $\pi[v] \in \mathcal{P}^{\mathrm{a}}$ the polynomial obtained from π by replacing each occurrence of a_i with v_i in the linear expressions of π and evaluating them. For $V \subseteq \mathbb{R}^s$, $\pi[V] \triangleq \{\pi[v] \,:\, v \in V\} \subseteq \mathcal{P}^{\mathrm{a}}$. In particular, for a user specified π, we will set $R \triangleq \pi[\mathbb{R}^s]$ in the relativized strongest postcondition problem (7). We extend δ_{Σ_1} and S_H to templates as expected: for $\pi = \sum_i \ell_i \gamma_i$, $\delta_{\Sigma_1}(\pi, x) \triangleq \sum_i \ell_i \delta_{\Sigma_1}(\gamma_i, x)$ and $S_H \pi \triangleq \sum_i \ell_i S_H \gamma_i$, seen as a polynomials in $\mathrm{Lin}(\mathbf{a})[\mathcal{D}]$ and $\mathrm{Lin}(\mathbf{a})[\mathcal{P}\mathrm{a}(H)]$, respectively. We shall make use of the following substitution properties of templates, which hold true in coherent systems (Lemma A.17 in Appendix A.4). For each $x \in X$ and $v \in \mathbb{R}^s$:

$$\delta_{\Sigma_1}(\pi[v], x) = \delta_{\Sigma_1}(\pi, x)[v] \qquad\qquad S_H(\pi[v]) = (S_H \pi)[v]. \tag{8}$$

We are now set to introduce the Post algorithm. Given $P \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ and a template π, fix P_0 s.t. $I_0 \triangleq \langle P_0 \rangle \subseteq \mathbf{I}(\mathbf{V}(P))$ ($P_0 = P$ is a possible choice). The algorithm consists in generating two sequences of sets, $V_i \subseteq \mathbb{R}^s$ and $J_i \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$, for $i \geq 0$, defined as follows. The idea is that, at step i, V_i collects those $v \in \mathbb{R}^s$ such that $S_H(\pi[v])$, and its derivatives up to order i, vanish on $\mathbf{V}(P)$, that is belong to $\mathbf{I}(\mathbf{V}(P))$. The J_i's are used to detect stabilization. We use π_τ as an abbreviation of $\delta_{\Sigma_1}(\pi, \tau)$.

$$V_i \triangleq \bigcap_{\tau : |\tau| \leq i} \{v \in \mathbb{R}^s \,:\, (S_H \pi_\tau)[v] \in I_0\} \tag{9}$$

$$J_i \triangleq \Big\langle \bigcup_{\tau : |\tau| \leq i} (S_H \pi_\tau)[V_i] \Big\rangle. \tag{10}$$

[5] Linear expressions with a constant term, such as $2 + 5a_1 + 42a_2 - 3a_3$ are not allowed.

Consider the least m such that *both* $V_m = V_{m+1}$ and $J_m = J_{m+1}$: we let $\mathrm{Post}_H(P_0,\pi) \triangleq (V_m, J_m)$. Note that m is well defined. Indeed, $V_0 \supseteq V_1 \supseteq \cdots$ forms a descending chain of finite-dimensional vector spaces in $\mathbb{R}^s$, which must stabilize at some m'; then $J_{m'} \subseteq J_{m'+1} \subseteq \cdots$ forms an ascending chain of ideals in $\mathcal{P}_0^{\mathrm{a}}(H)$, which must stabilize at some $m \geq m'$. We remark that the condition $V_{m+1} = V_m$ alone does *not* imply stabilization in general. The next theorem states correctness and relative completeness of Post.

▶ **Theorem 4.3** (relative completeness of Post)**.** *Let H be coherent and* AFP*. Let $P \subseteq \mathcal{P}_0^{\mathrm{a}}(H)$ and π be a template. Fix P_0 s.t. $I_0 \triangleq \langle P_0 \rangle \subseteq \mathbf{I}(\mathbf{V}(P))$. Let $\mathrm{Post}_H(P_0,\pi) = (V_m, J_m)$.*

(a) $\pi[V_m] \subseteq \pi[\mathbb{R}^s] \cap \mathrm{sp}_H(P)$*, with equality if* $I_0 = \mathbf{I}(\mathbf{V}(P))$*;*

(b) $\mathbf{V}(X \cup J_m) = \mathrm{wp}_H(\pi[V_m])$.

Proof. In the proof we shall make use of the following stabilization property of the sequence of the (V_i, J_i)s (Lemma A.18 in the Appendix A.4).

$$\mathrm{Post}_H(P_0,\pi) = (V_m, J_m) \text{ implies that for each } j \geq 1,\ V_m = V_{m+j} \text{ and } J_m = J_{m+j}. \tag{11}$$

Let us consider part (a) of the theorem. Fix any $v \in V_m$, we must prove that $\pi[v] \in \mathrm{sp}_H(P)$, that is $\phi_{(H,\rho)}(\pi[v]) = 0$ for each $\rho \in \mathbf{V}(P)$. By Corollary 3.8, our task reduces to showing that, for each τ, $(S_H(\pi[v]_\tau))(\rho) = (S_H\pi_\tau)[v](\rho) = 0$ (here we have used (8)), for each $\rho \in \mathbf{V}(P)$. That is, for each τ, $(S_H\pi_\tau)[v] \in \mathbf{I}(\mathbf{V}(P))$. The latter is implied by $(S_H\pi_\tau)[v] \in I_0 \subseteq \mathbf{I}(\mathbf{V}(P))$. By definition (9), this holds for each τ such that $v \in V_{|\tau|}$. Hence for each τ, as $v \in V_0 \supseteq \cdots \supseteq V_m = V_{m+1} = \cdots$ (by (11)). Assume now that $I_0 = \mathbf{I}(\mathbf{V}(P))$ and consider $v \in \mathbb{R}^s$ such that $\pi[v] \in \mathrm{sp}_H(P)$: we show that $v \in V_m$. Our task is showing that for each τ with $|\tau| \leq m$, $(S_H\pi_\tau)[v] \in \mathbf{I}(\mathbf{V}(P))$. The latter means precisely that $(S_H\pi_\tau)[v](\rho) = 0$ for each $\rho \in \mathbf{V}(P)$. But this holds by definition of $\pi[v] \in \mathrm{sp}_H(P)$ and Corollary 3.8: indeed, for each τ, $(S_H(\pi[v]_\tau))(\rho) = (S_H\pi_\tau)[v](\rho) = 0$ (here we have used (8)), for each $\rho \in \mathbf{V}(P)$.

Let us consider part (b). First, consider any $\rho \in \mathrm{wp}_H(\pi[V_m])$. By definition and Corollary 3.8 (and using (8)), this is equivalent to $(S_H\pi_\tau)[v](\rho) = 0$ for each $v \in V_m$ and τ. By definition of ideal J_m, this implies $F(\rho) = 0$ for each $F \in J_m$, that is $\rho \in \mathbf{V}(J_m)$. On the other hand, consider any $\rho \in \mathbf{V}(J_m)$ and any $v \in V_m$. Clearly $\rho \in \mathbb{R}^k$. Then proving that $\rho \in \mathrm{wp}_H(\pi[V_m])$, that is $\phi_{(H,\rho)}(\pi[v]) = 0$, is equivalent, via Corollary 3.8 (and again (8)), to showing that $(S_H\pi_\tau)[v](\rho) = 0$, for each τ. Consider any such τ: for $k \geq m$ large enough, by definition of J_k and the fact that $V_m = V_k$, we have $J_k \supseteq (S_H\pi_\tau)[V_m]$, hence $J_m = J_k \supseteq (S_H\pi_\tau)[V_m]$ (by (11)), therefore $(S_H\pi_\tau)[v](\rho) = 0$, as required. ◀

The vector spaces V_i s in (9) can be effectively represented by the successive linear constraints imposed on the parameters in $\mathbf{a} = (a_1, ..., a_s)$ by (9). In turn, this permits computing finite sets of generators for the ideals J_i s in (10). This is illustrated with an example below. For a set of linear expressions $L \subseteq \mathrm{Lin}(\mathbf{a})$, we let $\mathrm{span}(L) \triangleq \{v \in \mathbb{R}^s : \ell[v] = 0 \text{ for each } \ell \in L\} \subseteq \mathbb{R}^s$ be the vector space of parameter evaluations that annihilate all expressions in L.

▶ **Example 4.4** (Example 3.3, cont.)**.** Fix $P = P_0 = \emptyset$, hence $\mathbf{V}(P) = \mathbb{R}$ (here $k = |\{u\}| = 1$): we impose no constraints on the initial data. We seek for linear relations between u and u_x, considering the template $\pi \triangleq a_1 u + a_2 u_x$. We compute $\mathrm{Post}_H(P_0,\pi) = (V_m, J_m)$ as follows. Below we reuse the equalities for $S_H(\delta_{\Sigma_1}(u,\tau))$ already computed in Example 3.9.

- $(i = 0)$. $S_H\pi = (a_1 - a_2)u$. Therefore $V_0 = \mathrm{span}(\{a_1 - a_2\}) = \{(\lambda,\lambda) : \lambda \in \mathbb{R}\}$ and $J_0 = \{0\}$.
- $(i = 1)$. $S_H\pi_x = S_H(a_1 u_x + a_2 u_{xx}) = (a_2 - a_1)u$ and $S_H\pi_t = S_H(a_1 u_{xx} + a_2 u_{x^3}) = (a_1 - a_2)u$. Therefore $V_1 = \mathrm{span}(\{a_2 - a_1, a_1 - a_2\}) = V_0$ and similarly $J_1 = J_0$.

Hence the algorithm stabilizes already at $m = 0$, returning $V_0 = \{(\lambda,\lambda) : \lambda \in \mathbb{R}\}$ and $J_0 = \{0\}$. This means that the valid instances of π are of the form $\lambda(u + u_x)$, for all $\lambda \in \mathbb{R}$. Or, equivalently, that $u_x = -u$ is a valid equation, under any initial data specification.

Suppose $\mathrm{Post}_H(P_0,\pi) = (V_m, J_m)$. Given a parameter evaluation $v \in \mathbb{R}^s$, checking if $\pi[v] \in \pi[V_m]$ is equivalent to checking if $v \in V_m$: this can be effectively done knowing a basis B_m of the vector space V_m. In practice, it is more convenient to represent the whole set $\pi[V_m]$ returned by $\textsc{Post}_H$ compactly in terms of a new *result template* π' with $s' \le s$ parameters, such that $\pi'[\mathbb{R}^{s'}] = \pi[V_m]$. In the example above, $\pi' = a_1(u + u_x)$. The result template π' can in fact be computed directly from π, by propagating, via substitutions, the linear constraints on $\mathbf{a}$ arising from (9) as they are generated (further details in Appendix A.5).

5 Examples

We have put a proof-of-concept implementation of the $\textsc{Post}$ algorithm of Section 4 at work on some IVPs drawn from mathematical physics. We illustrate two cases below[6].

▶ **Example 5.1** (Burgers' equation)**.** We consider the inviscid case of the Burgers' equation [1, 7], with a linear initial condition at $t = 0$ (for b, c arbitrary real constants).

$$u_t = -u \cdot u_x \qquad\qquad u(0,x) = bx + c\,.$$

We fix $X = \{t, x\}$ and $U = \{u, b, c, \hat{t}, \hat{x}\}$ ($\hat{t}, \hat{x}$ will encode t, x as dependent variables). The above IVP is encoded by the stratified system $H = \{\Gamma_1, \Gamma_2\}$, where

$$\Gamma_1 = (\{u_t = -uu_x\} \cup \Sigma_{aux1},\ \{t, x\}) \qquad\qquad \Gamma_2 = (\{u_x = b\} \cup \Sigma_{aux2},\ \{x\})\,.$$

$\Sigma_{aux1} = \{b_t = 0, c_t = 0, c_x = 0, \hat{t}_t = 1, \hat{t}_x = 0, \hat{x}_t = 0, \hat{x}_t = 1\}$ and $\Sigma_{aux2} = \{b_x = 0\}$ just encode that b, c are constants and the obvious equations for $\hat{t}, \hat{x}$. As $\mathcal{P}\mathrm{a}(H) = \{u, b, c, \hat{t}, \hat{x}\}$, the system is $\textsc{afp}$. Moreover, H, with the lexicographic order induced by $u > b > c > \hat{t} > \hat{x}$ and $t > x$, is coherent. We fix the set of possible initial data specifications to $\mathbf{V}(P)$ where $P = \{u - c, \hat{t}, \hat{x}\}$: this just ensures that $u(0,0) = c$ and that $\hat{t} = t, \hat{x} = x$. In order to discover interesting postconditions of P, we consider a complete polynomial template of total degree 3 over the indeterminates $\mathcal{P}\mathrm{a}(H)$, $\pi = \sum_{\gamma_i \in \mathcal{P}\mathrm{a}(H)^{\otimes}, |\gamma_i| \le 3} a_i \gamma_i$, which consists of $s = 56$ terms. Letting $P_0 = P$, we run $\textsc{Post}_H(P, \pi)$, which halts at the iteration $m = 5$, returning (V_5, J_5). This took about 10s in our experiment. The algorithm returns V_5 in the form of a 1-parameter result template π', such that $\pi'[\mathbb{R}] = \pi[V_5]$: the set of all instances of π' forms a valid postcondition of P. In this case Theorem 4.3(a) implies that $\pi'[\mathbb{R}] = \mathrm{sp}_H(P) \cap \pi[\mathbb{R}^s]$. Specifically, we find, for a_1 a template parameter:

$$\pi' = a_1 \cdot (c\hat{t}u + u - b - c\hat{x})\,.$$

In other words, up to the multiplicative constant a_1, $c\hat{t}u + u = b + c\hat{x}$ is the only equation of degree ≤ 3 satisfied by the solutions of H, for initial data specifications $\rho \in \mathbf{V}(P)$. This equation can be easily solved algebraically for u — note that we are actually manipulating CFPSs— and yields the unique solution of the IVP:

$$u = \frac{c\hat{x} + b}{c\hat{t} + 1} = \frac{cx + b}{ct + 1}\,.$$

▶ **Example 5.2** (Heat equation)**.** We consider an IVP for the heat equation in one spatial dimension, with a (generic) sinusoidal initial condition at $t = 0$ (with b, c representing arbitrary real constants).

$$u_t = b \cdot u_{xx} \qquad\qquad u(0,x) = \sin(cx)\,. \tag{12}$$

We seek for all solutions u of the form: (sinusoid of x) $\times$ (exponential of t). Let us code this problem into a stratified system. We fix $U = \{u, f, g, h, a, b, c, d, i, j\}$ and $X = \{t, x\}$. Here, f, g, h will code $\cos(cx)$, $\sin(cx)$ and $\exp(-dt)$, respectively, while a, b, c, d, i, j will act as a supply of generic constants.

---

[6] Additional examples, concerning boundary problems and conservation laws, are reported in Appendices A.6-A.7. Code and examples are available at `https://github.com/micheleatunifi/PDEPY/blob/master/PDEv2.py`. Execution times reported here are for a Python Anaconda distribution running under Windows 10 on a Surface Pro laptop.

We let $H = \{\Gamma_1, \Gamma_2, \Gamma_3\}$, where

$$\Gamma_1 = (\{u_t = bu_{xx}\} \cup \Sigma_{aux1},\ \{t, x\}) \qquad \Gamma_2 = (\{u = g,\ f_x = -cg,\ g_x = cf\} \cup \Sigma_{aux2},\ \{x\})$$

$$\Gamma_3 = (\{h_t = -dh\} \cup \Sigma_{aux3},\ \{t\}).$$

The auxiliary equations in Σ_{auxi} encode that a, b, c, d, i, j are constants, like in the previous example, and moreover that $f_t = g_t = h_x = 0$:

$$\Sigma_{aux1} = \{a_t = 0, a_x = 0, b_t = 0, b_x = 0, c_t = 0, d_x = 0, f_t = 0, g_t = 0, h_x = 0, i_t = 0, i_x = 0, j_t = 0, j_x = 0\}$$

$$\Sigma_{aux2} = \{c_x = 0\} \qquad \Sigma_{aux3} = \{d_t = 0\}.$$

By inspection, $2 \prec 1$, $3 \prec 1$ and $1 \not\prec 2, 3$, which ensures that H is stratified; also $\mathcal{P}\mathrm{a}(H) = U \setminus \{u\}$ is finite. Moreover, the system is consistent: apart from the trivial case of constants, each subsystem features at most one equation per dependent variable. As for normality, hence coherence, we order the independent variables as $t > x$ and consider a ranking $\prec$ such that: (a) $v_\xi \prec u_\tau$ if either $v \neq u$ or ($v = u$ and $\xi \prec_{\mathrm{lex}} \tau$); (b) the remaining pairs, not involving u, are ordered according to an arbitrary graded[7] ranking.

To search for solutions of the wanted form[8], we consider an ansatz represented by $E \triangleq a \cdot (u + igh + jfh)$ and look for the weakest precondition $\mathrm{wp}_H(\{E\})$, that is, the largest algebraic set of initial data specification under which the solutions of H satisfy $E = 0$. We will then solve algebraically for a, d, i, j (considering b, c as given), replace the corresponding values in E and find u. To compute $\mathrm{wp}_H(\{E\})$, we use the Post algorithm. We consider $P = \{a\}$, that is, we pose $a = 0$ in the precondition, and $\pi = a_1 \cdot E$, for a dummy template parameter a_1. Then $\mathrm{sp}_H(P) \cap \pi[\mathbb{R}]$ is nonempty, as $a = 0$ trivially implies $E = 0$ is valid, and consists in fact of all scalar multiples of E. We then run $\mathrm{Post}_H(P, \pi)$, which halts at iteration $m = 3$, returning (V_3, J_3). This took about 3s in our experiment. Theorem 4.3(b) ensures that $\mathbf{V}(J_3) = \mathrm{wp}_H(\pi[V_3]) = \mathrm{wp}_H(\{E\})$. A Gröbner basis of $J_3 \subseteq \mathcal{P}_0(H)$ consists of 22 polynomials. To pick up a specific solution, we impose further conditions on some variables in $\mathcal{P}\mathrm{a}(H)$: $a = 1$ (as E is defined up to a multiplicative constant), $f = 1$, $g = 0$, $h = 1$ (initial values of cos, sin and exp) and $c \neq 0$ (rules out trivial solutions), we solve the resulting algebraic equations for d, i, j and find: $d = bc^2$, $i = -1$ and $j = 0$. We replace these values in E and, recalling that f, g, h encode $\cos(cx)$, $\sin(cx)$ and $\exp(-dt)$, we find

$$u = \sin(cx) \cdot \exp(-bc^2 t)$$

which is the classical solution obtained when applying the separation of variables method.

6 Conclusion and related work

Conceptually, the present development parallels and extends previous work on polynomial ODEs, especially the works [2, 3]. Technically, the case of PDEs is remarkably more challenging, for the following reasons. (a) Existence of solutions, and of the transition structure itself, depends now on coherence, which is trivial in ODEs. (b) In stratified systems and the related IVPs, a prominent role is played by their (acyclic) hierarchical structure, which is again trivial in ODEs. (c) In PDEs, differential polynomials live in the infinite-indeterminates space $\mathcal{P}$, which requires reduction to $\mathcal{P}_0(H)$ via S_H, and a finiteness assumption on parametric derivatives; in ODEs, $\mathcal{P} = \mathcal{P}_0(\Sigma)$ has always finitely many indeterminates.

Our work is related to the field of Differential Algebra (DA), see [6, 16, 22, 20, 25] and references therein. In particular, Boulier et al.'s RosenfeldGröbner algorithm [6], computes the ideal of the

[7] That is, $\deg(\xi) < \deg(\tau)$ implies $v_\xi \prec w_\tau$.

[8] (sinusoid of x) $\times$ (exponential of t).

differential and polynomial consequences of a system Σ. While this ideal is clearly related to our $\mathrm{sp}_{\{(\Sigma,X)\}}(\emptyset)$, how to encode general pre- or postconditions in their format is far from trivial, if possible at all. More generally, while DA techniques can be used to reduce systems to a coherent form, which is required by our approach, they do not seem concerned with IVPs or boundary problems as such. The only exceptions we are aware of are [23, 24], which focus on linear ODEs.

References

1 H. Bateman. Some recent researches on the motion of fluids. *Monthly Weather Review*, 43(4), 163-170, 1915.

2 M. Boreale. Algebra, coalgebra, and minimization in polynomial differential equations. In *Proc. of FoSSACS 2017*, LNCS 10203:71-87, Springer, 2017. Full version in *Logical Methods in Computer Science* 15(1), 2019. arXiv.org:1710.08350

3 M. Boreale. Complete algorithms for algebraic strongest postconditions and weakest preconditions in polynomial ODE's. *SOFSEM 2018: Theory and Practice of Computer Science - 44th International Conference on Current Trends in Theory and Practice of Computer Science*, LNCS 10706:442-455, Springer, 2018.

4 M. Boreale. Algorithms for exact and approximate linear abstractions of polynomial continuous systems. *HSCC 2018*: 207-216, ACM, 2018.

5 M. Boreale. On the Coalgebra of Partial Differential Equations. *MFCS 2019*, LIPIcs 138: 24:1-24:13, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. `http://www.dagstuhl.de/dagpub/978-3-95977-117-7`

6 F. Boulier, D. Lazard, F. Ollivier, M. Petitot. Computing representations for radicals of finitely generated differential ideals. *Appl. Algebra Engrg. Comm. Comput.*20(1), 73-121, 2009.

7 J.M. Burgers. A mathematical model illustrating the theory of turbulence. In *Advances in applied mechanics*, Vol. 1, pp. 171-199, Elsevier, 1948.

8 S. Chandrasekhar. On the decay of plane shock waves. *Ballistic Research Laboratories* 423, 1943.

9 Ch.G. Claudel, A.M. Bayen. Solutions to Switched Hamilton-Jacobi Equations and Conservation Laws Using Hybrid Components. *HSCC 2008*:101-115, ACM, 2008.

10 D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Undergraduate Texts in Mathematics, Springer, 2007.

11 C. Evans. *Partial Differential Equations*. American Mathematical Society. 2nd edition, 2010.

12 K. Ghorbal, A. Platzer. Characterizing Algebraic Invariants by Differential Radical Invariants. *TACAS 2014*, LNCS 8413: 279-294, 2014. Extended version available from `http://reports-archive.adm.cs.cmu.edu/anon/2013/CMU-CS-13-129.pdf`.

13 H. Kong, S. Bogomolov, Ch. Schilling, Yu Jiang, Th.A. Henzinger. Safety Verification of Nonlinear Hybrid Systems Based on Invariant Clusters. *HSCC 2017*:163-172, ACM, 2017.

14 F. Lemaire. An orderly linear pde system with analytic initial conditions with a non-analytic solution. *J. Symb. Comput.*, 35(5):487–498, May 2003. URL: `http://dx.doi.org/10.1016/S0747-7171(03)00017-8`.

15 J. Levandosky. *Lecture Notes for Partial Differential Equations of Applied Mathematics (Math 220A)*. Stanford University, 2002. `https://web.stanford.edu/class/math220a/lecturenotes.html`.

16 M. Marvan. Sufficient Set of Integrability Conditions of an Orthonomic System. *Foundations of Computational Mathematics*, 9(6):651-674, 2009

17 P.J. Olver. *Applications of Lie Groups to Differential Equations*, 2/E. Graduate Texts in Mathemathics. Springer, 1993.

18 A. Platzer. Logics of dynamical systems. *LICS 2012*: 13-24, IEEE, 2012.

19 A. Platzer. Differential hybrid games. *ACM Trans. Comput. Log.*, 18(3):19–44, 2017.

20 G. Reid, A. Wittkopf, A. Boulton. Reduction of systems of nonlinear partial differential equations to simplified involutive forms. *European Journal of Applied Mathematics*, 7(6), 635-666, 1996.

21 C. Riquier. *Les systèmes d'équations aux dérivèes partielles*. Gauthiers-Villars, Paris, 1910.

22 D. Robertz. *Formal Algorithmic Elimination for PDEs*. Lectures Notes in Mathematics, Springer, 2014.

23 M. Rosenkranz, G. Regensburger. Solving and factoring boundary problems for linear ordinary differential equations in differential algebras. *J. Symb. Comput.* 43(8): 515-544, 2008.

24 M. Rosenkranz, G. Regensburger, L. Tec, B. Buchberger. Symbolic Analysis for Boundary Problems: From Rewriting to Parametrized Gröbner Bases. *CoRR abs/1210.2950*, 2012.

25 C.J. Rust, G.J. Reid, A.D. Wittkopf. Existence and Uniqueness Theorems for Formal Power Series Solutions of Analytic Differential Systems. *ISSAC* 1999: 105-112, 1999.

26 J.J.M.M. Rutten. Behavioural differential equations: a coinductive calculus of streams, automata, and power series. *Theoretical Computer Science*, 308(1–3): 1–53, 2003.

27 S. Sankaranarayanan, H. Sipma, and Z. Manna. Non-linear loop invariant generation using Gröbner bases. *POPL 2004*, ACM, 2004.

28 S. Sankaranarayanan. Automatic invariant generation for hybrid systems using ideal fixed points. *HSCC 2010*: 221-230, ACM, 2010.

A Proofs and additional technical material

A.1 Proofs of Section 2

We give here a self-contained proof of Theorem 2.5. This result is reproduced from [5]. The proof is based on simple coalgebraic concepts, which are recalled below.

A.1.1 Commutative coalgebras

Let X be a finite set of *actions (or variables)*, ranged over by $x, y, ...$ and O a nonempty set. We recall that a *(Moore) coalgebra*[9] with actions in X and outputs in O is a triple $C = (\mathcal{S}, \delta, o)$ where: $\mathcal{S}$ is a set of *states*, $\delta : \mathcal{S} \times X \to \mathcal{S}$ is a *transition* function, and $o : \mathcal{S} \to O$ is an *output* function (see e.g. [26]). A *bisimulation* in C is a binary relation $R \subseteq \mathcal{S} \times \mathcal{S}$ such that whenever $s\,R\,t$ then: (a) $o(s) = o(t)$, and (b) for each x, $\delta(s,x)\,R\,\delta(t,x)$. It is an (easy) consequence of the general theory of bisimulation that a largest bisimulation over C, called bisimilarity and denoted by $\sim_C$, exists, is the union of all bisimulation relations, and is an equivalence relation over $\mathcal{S}$. Given two coalgebras with actions in X and outputs in O, C_1 and C_2, a *morphism* from C_1 to C_2 is a function $\mu : \mathcal{S}_1 \to \mathcal{S}_2$ that: (1) preserves outputs ($o_1(s) = o_2(\mu(s))$), and (2) preserves transitions ($\mu(\delta_1(s,x)) = \delta_2(\mu(s),x)$), for each state s and action x). It is an easy consequence of this definition that a morphism preserves bisimulation in both directions, that is: $s \sim_{C_1} t$ if and only if $\mu(s) \sim_{C_2} \mu(t)$.

We introduce now the subclass of Moore coalgebras we will focus on. We say a coalgebra C has *commutative actions* (or just that is *commutative*) if for each state s and actions x, y, it holds that $\delta(\delta(s,x),y) \sim_C \delta(\delta(s,y),x)$. We will introduce below an example of commutative coalgebra. In what follows, we let σ range over X^*, and, for any state s, let $s(\sigma)$ be defined inductively as: $s(\epsilon) \triangleq s$ and $s(x\sigma) \triangleq \delta(s,x)(\sigma)$.

▶ **Lemma A.1.** *Let C be a commutative coalgebra. If $\sigma, \sigma' \in X^*$ are permutation of one another then for any state $s \in S$, $s(\sigma) \sim_C s(\sigma')$.*

We define the coalgebra of CFPSs, C_{F}

$$C_{\mathrm{F}} \quad \triangleq \quad (\mathbb{R}[[X]], \delta_{\mathrm{F}}, o_{\mathrm{F}})$$

where $\delta_{\mathrm{F}}(f,x) = \frac{\partial f}{\partial x}$ and $o_{\mathrm{F}}(f) = f(\epsilon)$ (the constant term of f). Bisimilarity in C_{F}, denoted by $\sim_{\mathrm{F}}$, coincides with equality. It is easily seen that for each x, y, $\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}$, so that C_{F} is a commutative coalgebra. Now fix any commutative coalgebra $C = (S, \delta, o)$. We define the function $\mu : S \to \mathbb{R}[[X]]$ as follows. For each $\tau = \mathbf{x}^{\alpha}$

$$\mu(s)(\tau) \quad \triangleq \quad \frac{o(s(\tau))}{\alpha!} \tag{13}$$

where $\alpha! \triangleq \alpha_1! \cdots \alpha_n!$. Here, abusing slightly notation, we let $o(s(\tau))$ denote $o(s(\sigma))$, for some string σ obtained by arbitrarily ordering the elements in τ: the specific order does not matter, in view of Lemma A.1 and of condition (a) in the definition of bisimulation.

▶ **Lemma A.2.** *Let C be a commutative coalgebra and $f = \mu(s)$. For each x, $\frac{\partial f}{\partial x} = \mu(\delta(s,x))$.*

[9] In the paper, we only consider Moore coalgebras. For brevity, we shall omit the qualification "Moore".

Proof. Let $x = x_i$. For each $\tau = \mathbf{x}^\alpha$ in $X^\otimes$ we have

$$\begin{array}{rcl}\frac{\partial f}{\partial x_i}(\tau) & = & (\alpha_i+1)f(x_i\tau) \\ & = & (\alpha_i+1)\frac{o(s(x_i\tau))}{\boldsymbol{\alpha}!(\alpha_i+1)} \\ & = & \frac{o(\delta(s,x_i)(\tau))}{\boldsymbol{\alpha}!} \\ & = & \mu(\delta(s,x_i))(\tau)\end{array}$$

where the first and second equality follow from (1) and (13), respectively, and the third one from the definition of $s(x_i\tau)$. This proves the wanted statement. ◂

Based on the above lemma and the fact that $\sim_F$ is equality, we can prove the following corollary, saying that C_F is *final* in the class of *commutative* coalgebras.

▸ **Corollary A.3** (coinduction and finality of C_F)**.** *Let C be a commutative coalgebra. The function μ in* (13) *is the unique coalgebra morphism from C to C_F. Moreover, the following* coinduction *principle is valid: $s \sim_C t$ if and only if $\mu(s) = \mu(t)$ in $\mathbb{R}[[X]]$.*

Proof. We have: (1) $o(s) = \mu(s)(\epsilon)$ by definition of μ, and (2) $\mu(\delta(s,x)) = \delta_F(\mu(s),x)$, by Lemma A.2. This proves that μ is a coalgebra morphism. Next, we prove that $\sim_F$ coincides with equality in $\mathbb{R}[[X]]$. More precisely, we prove that for each τ and for each f,g: $f \sim_F g$ implies $f(\tau) = g(\tau)$. Proceeding by induction on the length of τ, we see that the base case is trivial, while for the induction step $\tau = x_i\tau'$ we have: $f \sim_F g$ implies $\frac{\partial f}{\partial x_i} \sim_F \frac{\partial g}{\partial x_i}$ (bisimilarity), which in turn implies $\frac{\partial f}{\partial x_i}(\tau') = \frac{\partial g}{\partial x_i}(\tau')$ (induction hypothesis); but by (1), $f(x_i\tau') = (\frac{\partial f}{\partial x_i}(\tau'))/(\alpha_i+1)$ and $g(x_i\tau') = (\frac{\partial g}{\partial x_i}(\tau'))/(\alpha_i+1)$, and this completes the induction step. From the coincidence of $\sim_F$ with equality in $\mathbb{R}[[X]]$, and the fact that any morphism preserves bisimilarity in both directions, the last part of the statement (coinduction) follows immediately. Finally, let ν be any morphism from C to C_F. From the definitions of bisimulation and morphism it is easy to see that for each s, $\mu(s) \sim_F \nu(s)$: this implies $\mu(s) = \nu(s)$ by coinduction, and proves uniqueness of μ. ◂

A.1.2 Proof of Theorem 2.5

We need a few technical lemmas. First, a result about normal forms in coherent systems.

▸ **Lemma A.4.** *Let Σ be coherent. For each $x \in X$ and $F \in \mathcal{P}$, $SD_xSF = SD_xF$.*

Proof. The *leading derivative* of an expression $E \in \mathcal{P} \setminus \mathcal{P}_0(\Sigma)$ is the principal derivative u_τ of highest ranking occurring in E. Let us define the *rank* of F, $\mathrm{rk}(F)$, as 0 if $F \in \mathcal{P}_0(\Sigma)$, and as the leading derivative of F otherwise. The set of ranks is well ordered according to $\prec$, augmented with the rule $0 \prec u_\tau$. The proof goes by induction on the rank of F.

The base case $F \in \mathcal{P}_0(\Sigma)$ and is trivial, as $SF = F$ by consistency. Assume now that $\mathrm{rk}(F) = u_\tau$, where u_τ is the leading derivative of F: then F has the form $\sum_j c_j \cdot u_\tau^{k_j}\gamma_j + F'$, where $0 \neq c_j \in \mathbb{R}$, $k_j \geq 1$ and u_τ does not occur in the monomials γ_j and in the polynomial F'. Let $u_\tau = G \in \Sigma^\infty$, so that $u_{x\tau} = D_xG \in \Sigma^\infty$ as well. We have the following.

- Applying (repeatedly) $u_\tau = G$ from left to right, we have by equational reasoning $F =_\Sigma E \triangleq \sum_j c_j \cdot G^{k_j}\gamma_j + F'$. Hence $SF = SE$, where, by normality, $\mathrm{rk}(E) \prec \mathrm{rk}(F)$. Then, using the induction hypothesis in the second equality below, and then the rules for total differentiation, which imply

$D_xE = \sum_j c_jk_jG^{k_j-1}D_xG\gamma_j + c_jG_j^kD_x\gamma_j + D_xF'$, we have

$$\begin{aligned} SD_xSF &= SD_xSE \\ &= SD_xE \\ &= S(\sum_j c_jk_jG^{k_j-1}D_xG\gamma_j + c_jG_j^kD_x\gamma_j + D_xF'). \end{aligned} \tag{14}$$

- On the other hand, by total differentiation and then by applying (repeatedly) both $u_\tau = G$ and $u_{x\tau} = D_xG$, we have

$$\begin{aligned} SD_xF &= S(\sum_j c_jk_ju_\tau^{k_j-1}u_{x\tau}\gamma_j + c_ju_\tau^{k_j}D_x\gamma_j + D_xF') \\ &= S(\sum_j c_jk_jG^{k_j-1}D_xG\gamma_j + c_jG_j^kD_x\gamma_j + D_xF') \end{aligned}$$

where the last term above is the same as (14).

◀

Next, a result about solutions.

▶ **Lemma A.5.** *Let* $\mathbf{iP} = (\Sigma,\rho)$ *and* ψ *a solution of* $\mathbf{iP}$*. For each* $E,F \in \mathcal{P}$*,* $E =_\Sigma F$ *implies* $\psi(E) = \psi(F)$*.*

Proof. If $E \to_\Sigma F$, the thesis is a consequence of property (b) of the definition of solution, and the fact that ψ is a homomorphism over $\mathcal{P}$. The proof for the general case follows from this fact and from the definition of $=_\Sigma$. ◀

With any coherent (w.r.t. some ranking) Σ and initial data specification ρ, $\mathbf{iP} = (\Sigma,\rho)$, we can now associate a coalgebra as follows.

$$C_{\mathbf{iP}} \triangleq (\mathcal{P},\delta_\Sigma,o_\rho)$$

where δ_Σ is defined in (4) and $o_\rho(E) \triangleq \rho(SE)$. We will denote by $\sim_{\mathbf{iP}}$ bisimilarity in $C_{\mathbf{iP}}$. As a consequence of Lemma A.4, $\delta_\Sigma(\delta_\Sigma(E,x),y) = \delta_\Sigma(\delta_\Sigma(E,y),x)$, so that for any monomial τ, the notation $\delta_\Sigma(E,\tau)$ is well defined. As an example of transition, for the heat equation $\Sigma = \{u_{xx} = au_t\}$, one has $\delta_\Sigma(u_{xx},t) = au_{tt}$.

As expected, $C_{\mathbf{iP}}$ is a commutative coalgebra. Moreover, each expression is bisimilar to its normal form. This is the content of the following lemma.

▶ **Lemma A.6.** *Let* $\mathbf{iP} = (\Sigma,\rho)$*, with* Σ *coherent. Then: (1)* $C_{\mathbf{iP}}$ *is commutative; and (2) For each* $E \in \mathcal{P}$*,* $E \sim_{\mathbf{iP}} SE$*.*

Proof. For what concerns part 1, for each x,y and F, we have

$$\begin{aligned} \delta_\Sigma(\delta_\Sigma(F,x),y) &= SD_xSD_yF \\ &= SD_xD_yF \\ &= SD_yD_xF \\ &= SD_ySD_xF \\ &= \delta_\Sigma(\delta_\Sigma(F,y),x) \end{aligned} \tag{15}$$

where the second equality and fourth follow from Lemma A.4, and the third one is a property of total derivatives.

For what concerns part 2, it is sufficient to show that the relation $R = \{(E,SE) : E \in \mathcal{P}\} \cup Id$, where Id is the identity relation, is a bisimulation. Condition (a) of the definition holds trivially; concerning condition (b), for any x we have that $\delta_\Sigma(E,x) = SD_xE = SD_xSE = \delta_\Sigma(SE,x)$, where the second equality follows again from Lemma A.4. ◀

As a consequence of the previous lemma, part 1, and of Corollary A.3, there exists a unique morphism from $C_{\mathbf{iP}}$ to C_{F}. This morphism is the unique solution of **iP** we are after. We need a lemma, saying that the unique morphism ϕ from $C_{\mathbf{iP}}$ to C_{F} is compositional.

▶ **Lemma A.7.** *Let* $\mathbf{iP} = (\Sigma,\rho)$*, with* Σ *coherent, and let* $\phi_{\mathbf{iP}}$ *be the unique morphism from* $C_{\mathbf{iP}}$ *to* C_{F}*. Then* $\phi_{\mathbf{iP}}$ *is a homomorphism over* $\mathcal{P}$*.*

Proof. Let us denote by ψ the homomorphic extension of $(\phi_{\mathbf{iP}})_{|U}$ to $\mathcal{P}$. One checks that $\psi(E) \sim_{\mathrm{F}} \phi_{\mathbf{iP}}(E)$, by induction on E. The proof also exploits the fact that, by Lemma A.4, $\delta_\Sigma(u,\tau) = S\,u_\tau$, hence $u_\tau \sim_{\mathbf{iP}} \delta_\Sigma(u,\tau)$ by virtue of Lemma A.6(2), therefore $\phi_{\mathbf{iP}}(u_\tau) = \phi_{\mathbf{iP}}(\delta_\Sigma(u,\tau))$ by coinduction. ◀

Proof of Theorem 2.5. Let $\phi_{\mathbf{iP}}$ denote the unique morphism from $C_{\mathbf{iP}}$ to C_{F}. We prove that $\phi_{\mathbf{iP}}$ is the unique solution of **iP**. By virtue of Lemma A.7, $\phi_{\mathbf{iP}}$ coincides with the homomorphic extension of $(\phi_{\mathbf{iP}})_{|U}$. We first prove that that $\phi_{\mathbf{iP}}$ respects the initial data specification. Let u_τ be parametric. By the definition of coalgebra morphism and of output functions in C_{F} and $C_{\mathbf{iP}}$, we have

$$\begin{array}{rcccl} \phi_{\mathbf{iP}}(u_\tau)(\epsilon) & = & o_{\mathrm{F}}(\phi_{\mathbf{iP}}(u_\tau)) & = & o_\rho(u_\tau) \\ \rho(S\,u_\tau) & = & \rho(u_\tau) & & \end{array}$$

which proves the wanted condition. Next, we have to prove that $\phi_{\mathbf{iP}}$ satisfies the equations in Σ^∞. But for each such equation, say $u_\tau = F$, we have $S\,u_\tau =_\Sigma S\,F$ by the definition of $=_\Sigma$, hence $u_\tau \sim_{\mathbf{iP}} F$ by Lemma A.6(2), hence the thesis by coinduction (Corollary A.3). We finally prove uniqueness of the solution. Assume ψ is a solution of **iP**. We prove that ψ is a coalgebra morphism from $C_{\mathbf{iP}}$ to C_{F}, hence $\psi = \phi_{\mathbf{iP}}$ will follow by coinduction (Corollary A.3). Let $E \in \mathcal{P}$. There are two steps in the proof.

- $\psi(E)(\epsilon) = \rho(S\,E) = o_\rho(E)$. This follows directly from Lemma A.5, since $\psi(E) = \psi(S\,E)$.
- For each x, $\frac{\partial\psi(E)}{\partial x} = \psi(\delta_\Sigma(E,x))$. First, we note that $\frac{\partial\psi(E)}{\partial x} = \psi(D_x E)$. This is proven by induction on the size[10] of E: in the base case when $E = u_\tau$, just use the fact that, by the definition of solution, $\frac{\partial\psi(u_\tau)}{\partial x} = \frac{\partial}{\partial x}\frac{\partial\psi(u)}{\partial\tau} = \frac{\partial\psi(u)}{\partial\tau x} = \psi(u_{\tau x}) = \psi(D_x u_\tau)$; in the induction step, use the fact that ψ is an homomorphism over $\mathcal{P}$, and the differentiation rules of D_x and $\frac{\partial}{\partial x}$ for sum and product. Now applying Lemma A.5, we get $\psi(D_x E) = \psi(S\,D_x E) = \psi(\delta_\Sigma(E,x))$, which is the wanted equality.

Finally, formula (5) is an immediate consequence of the definition of coalgebra $C_{\mathbf{iP}}$ and of the final morphism $\phi_{\mathbf{iP}} = \mu$ in (13). ◀

A.2 Proofs of Section 3

We first state a simple property of solutions of simple IVPs (Σ,ρ).

▶ **Lemma A.8.** *Let* ψ *be the solution of a coherent IVP* $\mathbf{iP} = (\Sigma,\rho)$*. For each* $E \in \mathcal{P}$ *and* $\xi = \mathbf{x}^\alpha$*,* $\psi(E)(\xi) = \frac{\psi(D_\xi E)(\epsilon)}{\alpha!}$*.*

Proof. An immediate application of formula (5) and of the definition of δ_Σ in (4). Note in particular that $\psi(D_\xi E)(\epsilon) = \rho(S_\Sigma D_\xi E) = \rho(\delta_\Sigma(E,\xi))$. ◀

The next lemma basically says that each subsystem $\Gamma_i = (\Sigma_i, X_i)$ in a coherent stratified system can be interpreted as a coherent system in the dependent variables U_{Γ_i} and the independent variables X_i.

▶ **Lemma A.9.** *Let* $H = \{\Gamma_1,\ldots,\Gamma_k\}$ *be a coherent stratified system. Then, for each i,* (Σ_i, X_i)*, seen as a pure system of PDEs with dependent variables in* U_{Γ_i}*, independent variables in* X_i *and derivatives in* $\mathcal{D}_i \triangleq \{v_\xi : v \in U_{\Gamma_i}, \xi \in X_i^\otimes\}$*, is coherent in the sense of Definition 2.4.*

[10] That is, $\sum_{\tau\in\mathrm{supp}(E)}|\tau|$.

Proof. By assumption each Σ_i is $\prec$-normal, for one and the same ranking $\prec$ defined on $\mathcal{D}$. The ranking $\prec$ induces a total order $\prec'$ over $\mathcal{D}_i$ defined as: $(u_\tau)_\xi \prec' (v_{\tau'})_{\xi'}$ iff $u_{\tau\xi} \prec v_{\tau'\xi'}$. The total order $\prec'$ is a ranking over $\mathcal{D}_i$: this immediately stems from $\prec$ being a ranking over $\mathcal{D}$. By the same reasoning, Σ_i is $\prec'$-normal when elements of $\mathcal{D}_{\Gamma_i}$ are interpreted as elements of $\mathcal{D}_i$. ◀

We next prove Theorem 3.5. In fact, it is technically convenient for the subsequent development to prove a slightly more detailed statement, which also provides us with information about the form of the solution.

▶ **Theorem A.10** (Theorem 3.5). *Let H be a coherent stratified system. For any initial data specification ρ for H, there is a unique solution $\phi_{\mathbf{iP}}$ of $\mathbf{iP} = (H,\rho)$. Moreover, for each i, $(\phi_{\mathbf{iP}})_{\Gamma_i}$ is also the unique solution of (Σ_i,ρ_i), for some ρ_i whose restriction to $\mathcal{P}a(H)$ coincides with ρ.*

Proof. Consider the stratified system $\overline{H} \triangleq H \cup \{\Gamma_0\}$. We will define below a set of initial value problems $\mathbf{iP}_i = (\Gamma_i,\rho_i)$ (Definition 2.3), $i = 0,...,k$, where each Γ_i is seen as a pure system of PDEs with independent variables X_i and dependent variables U_{Γ_i}. By Lemma A.9, each Γ_i is coherent, hence $\mathbf{iP}_i$ will have a unique solution ψ_i in the sense of Definition 2.3 (Theorem 2.5). Note that, under the identification $\mathcal{D}_i = \mathcal{D}_{\Gamma_i}$, ψ_i induces a function $\mathcal{P}_{\Gamma_i} \to \mathbb{R}[[X_i]]$: this function, still denoted by ψ_i, respects the equations in Σ_i. Similarly, ρ_i induces a function $\mathcal{P}a(\Gamma_i) \to \mathbb{R}$.

We proceed now to the actual definition of the $\mathbf{iP}_i$s by induction on the relation over subsystem indices ($i \prec j$), which is by definition acyclic. Note that $\mathcal{P}a(\overline{H}) = \emptyset$, so that each $u_\tau \in \mathcal{D}$ is principal for exactly one subsystem.

- The base case is when $\mathcal{P}a(\Gamma_i) = \emptyset$. Then we let $\mathbf{iP}_i \triangleq ((\Sigma_i, X_i),\emptyset)$, where $\emptyset$ denotes here the empty function, and let ψ_i be the corresponding unique solution (Theorem 2.5).
- Assume $\mathcal{P}a(\Gamma_i) \neq \emptyset$. Then we let $\mathbf{iP}_i \triangleq ((\Sigma_i, X_i),\rho_i)$, where $\rho_i : \mathcal{P}a(\Gamma_i) \to \mathbb{R}$ is the initial data specification defined by $\rho_i(u_\tau) \triangleq \psi_j(u_\tau)(\epsilon)$, for each $u_\tau \in \mathcal{P}a(\Gamma_i)$; here j is the unique index such that $j \prec i$ and $u_\tau \in \mathcal{P}r(\Gamma_j)$, and ψ_j is the unique solution of $\mathbf{iP}_j$.

Now we show that $\psi \triangleq \psi_1$ is a solution of $\overline{H}$ (recall that $X_1 = X$ by convention). In fact, we show that for each i, $\psi_{\Gamma_i} = \psi_i$ from which the wanted claim follows. We first show that for each subsystem Γ_i and $u_\tau \in \mathcal{D}_{\Gamma_i}$

$$\psi_{\Gamma_i}(u_\tau)(\epsilon) \quad = \quad \psi_i(u_\tau)(\epsilon)\,. \tag{16}$$

This is obvious if $i = 1$, hence assume $i \neq 1$. We distinguish the case $u_\tau \in \mathcal{P}a(\Gamma_i)$ from the case $u_\tau \in \mathcal{P}r(\Gamma_i)$. In the first case, let j be the unique index such that $u_\tau \in \mathcal{P}r(\Gamma_j)$, so that $j \prec i$. Note that $j \neq 1$: otherwise, one would have $1 \prec i$, which is impossible, due to acyclicity and $i \prec 1$ (as to the latter, note that there must exist $u_{\tau'} \in \mathcal{P}r(\Gamma_i) \cap \mathcal{P}a(\Gamma_1)$; in fact $\mathcal{P}r(\Gamma_i) \neq \emptyset$, as $\Sigma_i \neq \emptyset$). Then the following equalities follow from the definitions of $\psi_{\Gamma_k}, \psi_k, \rho_k$ $(0 \leq k \leq m)$.

$$\begin{aligned}
\psi_{\Gamma_i}(u_\tau)(\epsilon) &= \psi_1(u_\tau)(\epsilon)\\
&= \rho_1(u_\tau)\\
&= \psi_j(u_\tau)(\epsilon)\\
&= \rho_i(u_\tau)\\
&= \psi_i(u_\tau)(\epsilon)\,.
\end{aligned}$$

In the second case, $u_\tau \in \mathcal{P}r(\Gamma_i)$, we have the following.

$$\begin{aligned}
\psi_{\Gamma_i}(u_\tau)(\epsilon) &= \psi_1(u_\tau)(\epsilon)\\
&= \rho_1(u_\tau)\\
&= \psi_i(u_\tau)(\epsilon)\,.
\end{aligned}$$

This proves (16). Now in order to show that $\psi_{\Gamma_i} = \psi_i$, consider the following, for arbitrary $u_\tau \in \mathcal{D}_{\Gamma_i}$ and $\xi \in X_i^{\otimes}$, $\xi = \mathbf{x}^{\boldsymbol{\alpha}}$.

$$\begin{aligned}
\psi_{\Gamma_i}(u_\tau)(\xi) &= \psi_{\Gamma_i}(u_{\tau\xi})(\epsilon)/\boldsymbol{\alpha}! && (17)\\
&= \psi_i(u_{\tau\xi})(\epsilon)/\boldsymbol{\alpha}! && (18)\\
&= \psi_i(u_\tau)(\xi) && (19)
\end{aligned}$$

where (17) and (19) follow from Lemma A.8 applied to ψ_1 and ψ_i respectively, and (18) from (16).

Next, we prove that ψ is the unique solution. Suppose ϕ is a solution of $\overline{H}$. Then it easily follows by induction on $\prec$ that for each i, ϕ_{Γ_i} is a solution of $\mathbf{iP}_i$ as defined above (under the identification $\mathcal{D}_{\Gamma_1} = \mathcal{D}_i$). By uniqueness (Theorem 2.5), ϕ_{Γ_i} is the unique solution of $\mathbf{iP}_i$, hence $\phi_{\Gamma_i} = \psi_i$ as defined above. Moreover, clearly $\phi = \phi_{\Gamma_1}$. Hence $\phi = \phi_{\Gamma_1} = \psi_1 = \psi$.

The last part of the statement follows by construction of $\phi_{\mathbf{iP}}$. ◀

▶ **Lemma A.11.** *Let H be coherent and let ρ be an initial data specification for H. Let ϕ be the unique solution of (H,ρ). For each $E,F \in \mathcal{P}$, $E =_H F$ implies $\phi(E)(\epsilon) = \phi(F)(\epsilon)$.*

Proof. Let ϕ be the unique solution of (H,ρ). Therefore, for each i, $\phi(\cdot)_{|X_i^{\otimes}}$ is the unique solution of $\mathbf{iP}_i = (\Gamma_i,\rho_i)$ with $\Gamma_i = (\Sigma_i, X_i)$ the i-th subsystem (as per proof of Theorem 3.5). By definition of solution and Lemma A.5, for each $u_\tau = G \in \Sigma_i^{\infty}(X_i)$, $\phi(u_\tau)_{|X_i^{\otimes}} = \phi(G)_{|X_i^{\otimes}}$. Now, since ϕ acts as a homomorphism on $\mathcal{P}$ (Lemma A.7), the same does $\phi(\cdot)_{|X_i^{\otimes}}$. As a consequence, for any polynomial $E \in \mathcal{P}$, $\phi(E)_{|X_i^{\otimes}} = \phi(E[G/u_\tau])_{|X_i^{\otimes}}$. This in turn implies that whenever $E \to_H F$ (with $\to_H \overset{\triangle}{=} \cup_i \to_{\Sigma_i}$), where $F = E[G/u_\tau]$, one has $\phi(E)_{|X_i^{\otimes}} = \phi(F)_{|X_i^{\otimes}}$ for some X_i; in particular, $\phi(E)(\epsilon) = \phi(F)(\epsilon)$, as of course $\epsilon \in X_i^{\otimes}$. This finally implies that whenever $E =_H F$ one has $\phi(E)(\epsilon) = \phi(F)(\epsilon)$, as required. ◀

Proof of Corollary 3.8. We use the characterizations of ϕ as the unique solution of the IVP $\mathbf{iP}_1 = ((\Sigma_1, X),\rho_1)$ (Theorem A.10) and as a coalgebra morphism (Theorem 2.5). First, we observe that by Lemma A.8, $\phi(E)(\tau) = \phi(D_\tau E)(\epsilon)/\boldsymbol{\alpha}! = \phi(\delta_{\Sigma_1}(E,\tau))(\epsilon)/\boldsymbol{\alpha}!$, where the last equality stems from the definition of δ_{Σ_1} and Lemma A.4. Second, by Lemma A.11 , we have that $\phi(\delta_{\Sigma_1}(E,\tau))(\epsilon) = \phi(S_H(\delta_{\Sigma_1}(E,\tau)))(\epsilon)$. For brevity, let $F = S_H(\delta_{\Sigma_1}(E,\tau))$. As $F \in \mathcal{P}_0(H) \subseteq \mathcal{P}_0(\Sigma_1)$, we have $\phi(F)(\epsilon) = \rho_1(F)$ by definition of coalgebra morphism (13). But, by Theorem A.10, ρ_1 coincides with ρ on elements of $\mathcal{P}_0(H)$, hence $\phi(F)(\epsilon) = \rho_1(F) = \rho(F)$, which completes the proof of (6). ◀

A.3 Proof of conservative extension

We show that CFPS solutions are a conservative extension of analytic solutions in the classical sense. We first prove this for pure systems, then extend the result to stratified ones.

Pure systems Let $\mathcal{A}$ denote the set of real functions f that are analytic — admit a Taylor expansion — in a neighborhood of $0 \in \mathbb{R}^n$; for definiteness, we take each such function defined over the largest possible open set containing the origin. If $n = 0$, stipulate that $\mathcal{A} \overset{\triangle}{=} \{f : \{0\} \to \mathbb{R}\}$. $\mathcal{A}$ induces a commutative coalgebra $C_{\mathcal{A}} = (\mathcal{A}, \delta_{\mathcal{A}}, o_{\mathcal{A}})$, where $\delta_{\mathcal{A}}(f,x) = \frac{\partial f}{\partial x}$ (conventional partial derivative along x) and $o_{\mathcal{A}}(f) = f(0)$. The unique morphism $\mu_{\mathcal{A}} : C_{\mathcal{A}} \to C_{\mathrm{F}}$ (Corollary A.3) is given by (13), that is, for $\tau = \mathbf{x}^{\boldsymbol{\alpha}}$, $\mu_{\mathcal{A}}(f)(\tau) = \frac{1}{\alpha!}\frac{\partial f}{\partial \tau}(0)$. In other words, $\mu_{\mathcal{A}}$ maps the analytic function f into the CFPS obtained from the Taylor expansion of f from 0. Now fix a coherent Σ. Let $\psi : U \to \mathcal{A}$ be a solution of $\mathbf{iP} = (\Sigma,\rho)$, in the classical sense, and assume it analytic. This means, letting the homomorphic extension $\mathcal{P} \to \mathcal{A}$ of ψ be still be denoted by ψ, that

(a) $\psi(u_\tau)(0) = \rho(u_\tau)$ for each $u_\tau \in \mathcal{P}\mathrm{a}(\Sigma)$; and,

(b) $\psi(u_\tau) = \psi(F)$ for each $u_\tau = F$ in Σ^∞.

We want to show that for each $E \in \mathcal{P}$ the Taylor expansion of $\psi(E)$, seen as a CFPS, coincides with $\phi_{\mathbf{iP}}(E)$, the unique solution obtained from Theorem 2.5: formally, that $\mu_{\mathcal{A}}(\psi(E)) = \phi_{\mathbf{iP}}(E)$. This is a consequence of the following lemma.

▶ **Lemma A.12.** *Let Σ be coherent. Then any analytic solution ψ is a coalgebra morphism $C_{\mathbf{iP}} \to C_{\mathcal{A}}$.*

Proof. First, by repeating verbatim the proof of Lemma A.5, we check that

$$\text{whenever } E =_\Sigma F \text{ then } \psi(E) = \psi(F). \tag{20}$$

Indeed, if $E \to_\Sigma F$, this is a consequence of property (b) above of the definition of solution (in the classical sense), and the fact that ψ is a homomorphic extension from U to $\mathcal{P}$; the proof for the general case follows from this fact and from the definition of $=_\Sigma$. Second, we will exploit the following fact:

$$\text{whenever } F \in \mathcal{P}_0(\Sigma) \text{ then } \psi(F)(0) = \rho(F). \tag{21}$$

This is shown by an induction on F, where the base case $F = u_\tau$ relies on the above definition of solution, part (a). We can now repeat basically the same arguments of the uniqueness part of Theorem 2.5, as follows. Let $E \in \mathcal{P}$. There are two steps in the proof.

- $\psi(E)(0) = \psi(S\,E)(0) = \rho(S\,E) = o_\rho(E)$, where the first equality follows from (20) and the second one from (21).
- For each x, $\frac{\partial\psi(E)}{\partial x} = \psi(\delta_\Sigma(E,x))$. First, we note that $\frac{\partial\psi(E)}{\partial x} = \psi(D_x E)$. This is proven by induction on the size of E: in the base case when $E = u_\tau$, just use the fact that, by the above definition of solution (in the analytic sense), part (b), $\frac{\partial\psi(u_\tau)}{\partial x} = \frac{\partial}{\partial x}\frac{\partial\psi(u)}{\partial\tau} = \frac{\partial\psi(u)}{\partial\tau x} = \psi(u_{\tau x}) = \psi(D_x u_\tau)$; in the induction step, use the fact that ψ is a homomorphism over $\mathcal{P}$, and the differentiation rules of D_x and $\frac{\partial}{\partial x}$ for sum and product. Now applying (20), we get $\psi(D_x E) = \psi(S\,D_x E) = \psi(\delta_\Sigma(E,x))$, which is the wanted equality.

◀

▶ **Proposition A.13** (conservative extension for pure systems)**.** *Let Σ be a coherent system and ρ an initial data specification for Σ. Let ψ be an analytic solution of (Σ,ρ). Then $\mu_{\mathcal{A}} \circ \psi = \phi_{(\Sigma,\rho)}$.*

Proof. From the lemma just proven, and since the composition of two coalgebra morphisms is a coalgebra morphism, we have that $\mu_{\mathcal{A}} \circ \psi : C_{\mathbf{iP}} \to C_{\mathrm{F}}$ is a coalgebra morphism. By the uniqueness of such morphism (Corollary A.3), we have $\mu_{\mathcal{A}} \circ \psi = \phi_{(\Sigma,\rho)}$, which is the wanted claim. ◀

Stratified systems In what follows, we let $\mathcal{A}_k$ ($k \ge 0$) denote the set of k-arguments analytic functions defined in a neighborhood of $0 \in \mathbb{R}^k$. For $f \in \mathcal{A}_n$, let $X = \{x_1,...,x_n\}$ represent the arguments of f, and let $Y \subseteq X$: we let $f_Y \in \mathcal{A}_{|Y|}$ denote the function obtained from f by fixing to 0 the arguments not in Y.

Let us fix a coherent stratified system H and and an initial data specification ρ for H. Let $\psi : U \to \mathcal{A}$ be an analytic solution of (H,ρ), in the classical sense. This means, letting the homomorphic extension $\mathcal{P} \to \mathcal{A}$ of ψ be still be denoted by ψ, that for each $\Gamma_i = (\Sigma_i, X_i) \in \overline{H}$ and for each $u_\tau = F$ in Γ_i^∞:

$$\psi(u_\tau)_{X_i} = \psi(F)_{X_i}\,. \tag{22}$$

▶ **Theorem A.14** (conservative extension for stratified systems)**.** *Let H be a coherent stratified system and ρ an initial data specification for H. Let ψ be an analytic solution of (H,ρ). Then $\mu_{\mathcal{A}} \circ \psi = \phi_{(H,\rho)}$.*

Proof. Let $\overline{H} = \{\Gamma_0,...,\Gamma_k\}$, with $\Gamma_i = (\Sigma_i, X_i)$. For each $i = 0,...,k$, we let $\mu_i : \mathcal{A}_{|X_i|} \to \mathbb{R}[[X_i]]$ denote the final morphism into $\mathbb{R}[[X_i]]$ obtained by turning $\mathcal{A}_{|X_i|}$ into a coalgebra with inputs in X_i and outputs in $\mathbb{R}$ (see previous paragraph). In particular, $\mu_{\mathcal{A}} = \mu_1$. Now, let $\mathbf{iP}_i = (\Gamma_i, \rho_i)$ be the same sequence of IVPs defined in the proof of Theorem A.10, and $\phi_{\mathbf{iP}_i}$ be the corresponding unique solutions. Let $\psi_i : \mathcal{P}_{\Gamma_i} \to \mathcal{A}_{|X_i|}$ be defined as $\psi_i(E) \triangleq \psi(E)_{X_i}$. We now show that for each $i = 0, ..., k$, ψ_i is an analytic solution — in the classical sense, defined by (a), (b) in the previous paragraph — of $\mathbf{iP}_i$. From this, by invoking Proposition A.13 we will have, for each i

$$\phi_{\mathbf{iP}_i} = \mu_i \circ \psi_i\,. \tag{23}$$

From this the thesis will follow by considering $i = 1$, as by Theorem A.10, $\phi_{(H,\rho)} = \phi_{\mathbf{iP}_1}$. We proceed now to actually show that ψ_i is an analytic solution of $\mathbf{iP}_i$. In fact, condition (b) coincides with (22), so we have to check only condition (a). We proceed by induction on a fixed linear order compatible with $\prec$. In the base case, we have $\mathcal{P}\mathrm{a}(\Gamma_i) = \emptyset$, hence condition (a) holds vacuously. In the induction step, consider any $u_\tau \in \mathcal{P}\mathrm{a}(\Gamma_i)$. By definition of ρ_i (cf. proof of Theorem A.10), $\rho_i(u_\tau) = \phi_{\mathbf{iP}_j}(u_\tau)(\epsilon)$, for the unique j such that $u_\tau \in \mathcal{P}\mathrm{r}(\Gamma_j)$; clearly $j \prec i$. By induction hypothesis, and (23), $\phi_{\mathbf{iP}_j}(u_\tau)(\epsilon) = \psi_j(u_\tau)(0)$. Now, denoting by $0_i, 0_n$ and 0_j the zero's in $\mathbb{R}^{|X_j|}, \mathbb{R}^n$ and $\mathbb{R}^{|X_i|}$, respectively, we have by definition of ψ_k: $\psi_j(u_\tau)(0_j) = \psi(u_\tau)(0_n) = \psi_i(u_\tau)(0_i)$. To sum up, $\rho_i(u_\tau) = \psi_i(u_\tau)(0)$, hence (a) is proven. ◂

▸ Remark A.15 (equational reasoning on analytic solutions). Consider a coherent H, with the additional property that for each initial data specification ρ there exists a unique analytic solution, say $\psi_{(H,\rho)}$, around 0. Then Theorem A.14 ensures that, in terms of valid polynomial equalities, considering analytic solutions or CFPSs makes no difference at all. More precisely, letting $\mathrm{sp}^{\mathcal{A}}_H(P) \triangleq \{E \in \mathcal{P} : \psi_{(H,\rho)}(E) = 0 \text{ for each } \rho \in \mathbf{V}(X \cup P)\}$, for such a H we have that $\mathrm{sp}^{\mathcal{A}}_H(P) = \mathrm{sp}_H(P)$.

Unfortunately, not all systems of PDEs posses an analytic solution, even when confining to the polynomial format as we do — in stark contrast with the case of ODEs. The following example of a linear PDE system is drawn from [14]

$$u_{xx} = u_{xy} + u_{yy} + v$$
$$v_{yy} = v_{xy} + v_{yy} + u$$

with the initial conditions $u(0,y) = u_x(0,y) = \exp(y)$ and $v(x,0) = v_y(x,0) = \exp(x)$. The initial conditions can be easily recast into polynomial form as follows: $u_y(0,y) = u(0,y)$ and $u_{xy}(0,y) = u_x(0,y)$ (similarly for v), with the initial data specified by $P = \{u-1, u_x-1, v-1, v_x-1\}$. This results in a stratified system $H = \{(\Sigma_1, \{x,y\}), (\Sigma_2, \{y\}), (\Sigma_3, \{x\})\}$ that is coherent w.r.t. to the ranking considered in [14]: $u \prec v \prec u_y \prec u_x \prec v_x \prec v_y \prec \cdots$. As a consequence, H has a unique CFPS solution for each initial data specification over $\mathcal{P}\mathrm{a}(H) = \{u, u_x, v, v_y\}$. Lemaire [14] shows however that H has no analytic solution. Informally, the reason is that its Taylor coefficients grow too fast as the order of the derivatives grows.

Syntactic formats that guarantee existence and uniqueness of analytic solutions of PDEs IVPs are known: for instance, one has the Cauchy-Kovalevskaya format [17, Ch.2.6], generalized by the Riquier format [21], further generalized by Rust et al. [25].

A.4 Proofs of Section 4

The next lemma says that the normal form functions S_{Σ_1} and S_H preserve the sum and product operations on polynomials defined in (3). In what follows, we shall abbreviate S_{Σ_1} as S_1.

▸ **Lemma A.16.** *Let H be an autonomous and coherent stratified system. Then for each $E, F \in \mathcal{P}^{\mathrm{a}}$, we have $S_H(E+F) = S_H E + S_H F$ and $S_H(E \cdot F) = (S_H E) \cdot (S_H F)$. The same holds true for S_1.*

Proof. Let us consider the statement for S_H. We only consider the sum, as the product is similar. Fix an arbitrary initial data specification ρ for H and denote by $\phi_{\mathbf{iP}}$ the unique solution of $\mathbf{iP} = (H,\rho)$

(Theorem 3.5). We have:

$$\begin{aligned}
(S_H(E+F))(\rho) &= \phi_{\mathbf{iP}}(S_H(E+F))(\epsilon) && (24)\\
&= \phi_{\mathbf{iP}}(E+F)(\epsilon) && (25)\\
&= \phi_{\mathbf{iP}}(E)(\epsilon)+\phi_{\mathbf{iP}}(F)(\epsilon) && (26)\\
&= \phi_{\mathbf{iP}}(S_H E)(\epsilon)+\phi_{\mathbf{iP}}(S_H F)(\epsilon) && (27)\\
&= (S_H E)(\rho)+(S_H F)(\rho) && (28)\\
&= (S_H E+S_H F)(\rho) && (29)
\end{aligned}$$

where: (24) and (28) follow from (6) with $\tau=\epsilon$; (25) and (27) follow from Lemma A.11; (26) follows because ϕ is a homomorphism; (29) follows by definition of homomorphic extension of ρ to $\mathcal{P}_0(H)$. In other words, we have shown that $(S_H(E+F))(\rho)-(S_H E+S_H F)(\rho) = (S_H(E+F)-(S_H E+S_H F))(\rho) = 0$, for arbitrary $\rho\in\mathbb{R}^k$ ($k=|\mathcal{P}\mathrm{a}(H)|$). Since[11] $(S_H(E+F)-(S_H E+S_H F))\in\mathcal{P}_0^{\mathrm{a}}(H)=\mathbb{R}[\mathcal{P}\mathrm{a}(H)]$, we can conclude that $S_H(E+F)-(S_H E+S_H F)=0$, that is $S_H(E+F)=S_H E+S_H F$. The proof for $S_1=S_{\Sigma_1}$ is similar. ◀

We need need two 'substitution lemmas' for templates, also to effectively compute (9). These prove the equalities in (8).

▶ **Lemma A.17.** *Let H be a an autonomous and coherent stratified system. Let π a polynomial template, $v\in\mathbb{R}^s$.*

1. $\delta_{\Sigma_1}(\pi[v],x)=\delta_{\Sigma_1}(\pi,x)[v]$ *for any $x\in X$;*
2. $S_H(\pi[v])=(S_H\pi)[v]$.

Proof. Let $\pi=\sum_i \ell_i\gamma_i$, for distinct monomials $\gamma_i\in\mathcal{D}^{\otimes}$. Facts (1) and (2) easily follow from the distributivity properties of S_H and S_1 (Lemma A.16). As an example, for (1) we have

$$\begin{aligned}
\delta_{\Sigma_1}(\pi[v],x) &= \delta_{\Sigma_1}(\sum_i \ell_i[v]\gamma_i,x)\\
&= S_1\sum_i \ell_i[v]D_x\gamma_i\\
&= \sum_i \ell_i[v]S_1 D_x\gamma_i\\
&= \sum_i \ell_i[v]\delta_{\Sigma_1}(\gamma_i,x)\\
&= (\sum_i \ell_i\delta_{\Sigma_1}(\gamma_i,x))[v]\\
&= \delta_{\Sigma_1}(\pi,x)[v] && (30)
\end{aligned}$$

The proof for (2) is similar. ◀

We finally arrive at the proof of the stabilization property stated in (11).

▶ **Lemma A.18** (property (11))**.** *Let $\mathrm{Post}_H(P_0,\pi)=(V_m,J_m)$, under the hypotheses of Theorem 4.3. Then for each $j\geq 1$, one has $V_m=V_{m+j}$ and $J_m=J_{m+j}$.*

[11] Here we are crucially using the autonomicity assumptiom.

Proof. We proceed by induction on j. The base case $j = 1$ follows from the definition of m. Assuming by induction hypothesis that $V_m = \cdots = V_{m+j}$ and that $J_m = \cdots = J_{m+j}$, we prove now that $V_m = V_{m+j+1}$ and that $J_m = J_{m+j+1}$. The key to the proof is the following fact

$$(S_H\pi_{\tau x})[v] \in J_m, \ \ \forall\ |\tau| = m + j,\ x \in X \text{ and } v \in V_m\,. \tag{31}$$

From this fact the thesis will follow, as we show below.

1. $V_m = V_{m+j+1}$. To see this, observe that for each $v \in V_{m+j} = V_m$ (the equality here follows from the induction hypothesis), it follows from (31) and the definition of J_m that $(S_H\pi_{\tau x})[v]$ can be written as a finite sum of the form $\sum_l h_l \cdot (S_H\pi_{\tau_l})[w_l]$, with $0 \le |\tau_l| \le m$ and $w_l \in V_m$. For each $0 \le |\tau_l| \le m$, $(S_H\pi_{\tau_l})[w_l] \in I_0$ by assumption, from which it easily follows that also $(S_H\pi_{\tau x})[v] = \sum_l h_l \cdot (S_H\pi_{\tau_l})[w_l] \in I_0$. Since fact holds for each τ of size m and $x \in X$, hence for each τ of size $m + 1$, it shows that $v \in V_{m+j+1}$, proving that $V_{m+j+1} \supseteq V_{m+j} = V_m$. The reverse inclusion is obvious.

2. $J_m = J_{m+j+1}$. As a consequence of $V_{m+j+1} = V_{m+j} (= V_m)$ (the previous point), we can write

$$\begin{aligned} J_{m+j+1} &= \Big\langle \bigcup_{|\tau|\le m+j} (S_H\pi_\tau)[V_{m+j}] \ \cup \bigcup_{|\xi|=m+j+1} (S_H\pi_\xi)[V_{m+j}] \Big\rangle \\ &= \Big\langle J_{m+j} \ \cup \bigcup_{|\xi|=m+j+1} (S_H\pi_\xi)[V_{m+j}] \Big\rangle \\ &= \Big\langle J_m \ \cup \bigcup_{|\xi|=m+j+1} (S_H\pi_\xi)[V_m] \Big\rangle \end{aligned}$$

where the last step follows by induction hypothesis. From (31), we have that for $|\xi| = m + j + 1$, $(S_H\pi_\xi)[V_m] \subseteq J_m$, which implies the thesis for this case, as $\langle J_m \rangle = J_m$.

We prove now (31). In this proof, we shall make use of the following equalities satisfied by S_H. For each $E \in \mathcal{P}$ and $x \in X$

$$S_H D_x S_H E = S_H D_x E \tag{32}$$

$$S_H S_1 E = S_H E\,. \tag{33}$$

The proof of (32) is essentially identical to that of Lemma A.4 (induction on the rank of the leading derivative in E) and is omitted. Concerning (33), note that $E =_{\Sigma_1} S_1 E$ implies $E =_H S_1 E$, which in turn implies $E =_H S_H S_1 E$, that is $S_H E = S_H S_1 E$. Let us now proceed to the proof of (31). Fix any $v \in V_m$. First, note that for $|\tau| = m + j$ and $x \in X$, by definition $\pi_{\tau x}[v] = \delta_{\Sigma_1}(\pi_\tau[v], x) = S_1 D_x(\pi_\tau[v])$ (where in the first step we have used Lemma A.17; here $S_1 = S_{\Sigma_1}$). Now consider $S_H\pi_\tau$: by induction hypothesis, $(S_H\pi_\tau)[V_m] = (S_H\pi_\tau)[V_{m+j}] \subseteq J_{m+j} = J_m$, hence $(S_H\pi_\tau)[v]$ can be written as a finite sum

$\sum_l h_l \cdot (S_H\pi_{\tau_l}[w_l])$, with $0 \le |\tau_l| \le m$ and $w_l \in V_m$ and $h_l \in \mathcal{P}_0(H)$. Summing up, we have:

$$
\begin{aligned}
(S_H\pi_{\tau x})[v] &= S_H S_1 D_x(\pi_\tau[v]) \\
&= S_H D_x(\pi_\tau[v]) && (34)\\
&= S_H D_x S_H(\pi_\tau[v]) && (35)\\
&= S_H D_x \sum_l h_l \cdot S_H\pi_{\tau_l}[w_l] && (36)\\
&= S_H \sum_l (D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot D_x S_H(\pi_{\tau_l}[w_l]) && (37)\\
&= \sum_l S_H(D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot S_H D_x S_H(\pi_{\tau_l}[w_l]) && (38)\\
&= \sum_l S_H(D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot S_H D_x(\pi_{\tau_l}[w_l]) && (39)\\
&= \sum_l S_H(D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot S_H S_1 D_x(\pi_{\tau_l}[w_l]) && (40)\\
&= \sum_l S_H(D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot S_H \delta_1(\pi_{\tau_l}[w_l], x) && (41)\\
&= \sum_l S_H(D_x h_l) \cdot S_H\pi_{\tau_l}[w_l] + h_l \cdot S_H\pi_{\tau_l x}[w_l] && (42)
\end{aligned}
$$

where:

- (34) follows from (33);
- (35) follows from (32);
- (36) follows from the equality for $S_H(\pi_\tau[v]) = (S_H\pi)[v]$ (here we use Lemma A.17) proven above;
- (37) follows from distributing D_x over sum and products, and applying the rules for total derivatives;
- (38) follows from distributing S_H (Lemma A.16) over sums and products, and further noting that $S_H h_l = h_l$, as $h_l \in \mathcal{P}_0(H)$;
- (39) follows again from (32);
- (40) follows again from (33);
- (41) follows from the definition of δ_1;
- (42) follows from Lemma A.17.

Now, for each $w_l \in V_m = V_{m+1}$, the term $S_H\pi_{\tau_l x}[w_l]$, with $0 \le |\tau_l x| \le m+1$, is by definition in $J_{m+1} = J_m$. Thus (42) proves that $S_H\pi_{\tau x}[v] \in J_m$, as required. ◀

A.5 Computational details for the Post algorithm in Section 4

We refer the reader to [10, Ch.3,Sect.1,Th.2] for the definitions of Gröbner basis G, of reduction mod G, as well as of the technical notion of elimination order; the lexicographic order is one such order. See [3, Lemma 3] for a proof of the following lemma.

▶ **Lemma A.19.** *Let $\mathbf{z} = \{z_1, ..., z_k\}$ and $\mathbf{a} = \{a_1, ..., a_s\}$ be disjoint sets of indeterminates. Let $G \subseteq \mathbb{R}[\mathbf{z}]$ be a Gröbner basis in $\mathbb{R}[\mathbf{a} \cup \mathbf{z}]$ w.r.t. a monomial elimination order for the a_i s in $\mathbf{a}$. Consider $p \in \mathrm{Lin}(\mathbf{a})[\mathbf{z}]$, seen as a polynomial in $\mathbb{R}[\mathbf{a} \cup \mathbf{z}]$, and $r = p \bmod G$. Then r is linear in $\mathbf{a}$. Moreover, for each $v \in \mathbb{R}^s$, $p[v] \bmod G = r[v]$.*

For $\pi \in \mathrm{Lin}[\mathbf{a}][\mathbb{R}]$, let coeff($\pi$) be the set of coefficients (linear expressions) of π. Recall that for a Gröbner basis G and a polynomial E, $E \bmod G$ denotes the remainder of the division of E by G. Here we use the fact that $G \subseteq \mathcal{P}_0^{\mathrm{a}}(H)$ is also a Gröbner over the larger polynomial ring $\mathbb{R}[\{a_1, ..., a_s\} \cup \mathcal{P}\mathrm{a}(H)]$, which contains also all templates, once an elimination monomial order (e.g. lexicographic) for the a_i s is fixed.

▶ **Lemma A.20.** *Under the hypotheses of Theorem 4.3, let $G \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ be a Gröbner basis of I_0. Then $V_i = \mathrm{span}\,(\cup_{|\tau|\leq i}\mathrm{coeff}((S_H\pi_\tau) \bmod G))$. As a consequence $J_i = \langle\ \cup_{|\tau|\leq i}(S_H\pi_\tau)[B_i]\ \rangle$, where B_i is a basis of V_i.*

Proof. Let $\mathbf{z} = \mathcal{P}\mathrm{a}(H)$. Let $G \subseteq \mathcal{P}^{\mathrm{a}}_0(H)$ be the given Gröbner basis of I_0: G can also be considered as a Gröbner basis in the larger ring $\mathbb{R}[\mathbf{a} \cup \mathbf{z}]$, w.r.t. some elimination order for the parameters a_i s in $\mathbf{a}$. Fix any $\tau \in X^{\otimes}$. Applying Lemma A.19 with $p = S_H\pi_\tau$, we have that for each $v \in \mathbb{R}^s$: $(S_H\pi_\tau)[v] \in I_0$ iff $r^{(\tau)}[v] = 0$, where $r^{(\tau)} \triangleq S_H\pi_\tau \bmod G$; this is true iff $v \in \mathrm{span}(\mathrm{coeff}(r^{(\tau)}))$. Hence, by definition (9), $v \in V_i$ iff $v \in \mathrm{span}(\mathrm{coeff}(r^{(\tau)}))$ for each $|\tau| \leq i$. This is in turn equivalent to $v \in \mathrm{span}(\cup_{|\tau|\leq i}\mathrm{coeff}(r^{(\tau)}))$, which is the first part of the statement. The last part follows because, for any template π, vector space $V \subseteq \mathbb{R}^s$ and basis B of V, one has $\langle\ \pi[V]\ \rangle = \langle\ \pi[B]\ \rangle$. ◀

▶ Remark A.21 (on completeness). Completeness (equality) in part (a) of Theorem 4.3 is only guaranteed if P_0 is chosen such that $I_0 = \mathbf{I}(\mathbf{V}(P))$, otherwise $\pi[V_m]$ is just a postcondition. When $I_0 = \mathbf{I}(\mathbf{V}(P))$, I_0 is said to be a *real radical* of P. Computing real radicals is a computationally hard problem, in the general case. For a number of special cases relevant to our goals, fortunately, the real radical is trivial. For instance, if P only contains elements of the form $d - e$, for d an indeterminate and e an indeterminate or a constant, then $\langle\ P\ \rangle = \mathbf{I}(\mathbf{V}(P))$, so that $\langle\ P\ \rangle$ is a real radical. Also note that the completeness in part (b) of Theorem 4.3 does *not* depend on having a real radical at hand. See [3] for further discussion on the real radical problem.

A.6 Additional examples 1: boundary problems

A *boundary problem* prescribes the form of the solution at some specified curve, rather than an initial condition like an IVP. Any scalar, first order boundary problem can be transformed into an IVP via a suitable change of coordinates, hence becoming amenable to analysis with our algorithm. One can exploit the *method of characteristics* [11, Ch.3] as a systematic recipe for carrying out this transformation. The resulting technique is illustrated via the following example.

Consider the PDE $u_x^2 + u_y^2 = 1$ (the *Eikonal* equation), with the boundary condition $u_{|C} = 0$, where C is the unit circle centered at the origin. According to the method of characteristics, one can transform a boundary problem into a *family* of hopefully simpler ODE IVPs. For our purposes, we need not worry about the details of this transformation (see [15, Ch.2] for a detailed derivation). It suffices to know it results in the following ODE IVPs, depending on a parameter $r \in \mathbb{R}$. Here s is the only independent variable, while x, y, z, p, q are the dependent variables.

$$
\begin{array}{lll}
\frac{dx}{ds}(s;r) \;=\; 2p & \frac{dy}{ds}(s;r) \;=\; 2q & \frac{dz}{ds}(s;r) \;=\; 2p^2 + 2q^2 \\
\frac{dp}{ds}(s;r) \;=\; 0 & \frac{dq}{ds}(s;r) \;=\; 0 & \\
x(0;r) \;=\; \cos(r) & y(0;r) \;=\; \sin(r) & z(0;r) \;=\; 0 \\
p(0;r) \;=\; \cos(r) & q(0;r) \;=\; \sin(r)\,. &
\end{array}
$$

According to the theory of ODEs, for each r the above IVP has a unique solution in a neighborhood of $s = 0$. The union of the solutions' trajectories $(x(s;r), y(s;r), z(s;r))$ represents the solution u of the original problem, in the sense that for each r, and for each s in a neighborhood of 0

$$z(s;r) \quad = \quad u(x(s;r), y(s;r))\,.$$

As $(x(0;r), y(0;r))$ represents a parametrization of the circle C depending on $r \in \mathbb{R}$, the above formula says that we can represent the solution u via z at least locally, that is near the boundary C. Also note that $z(0;r) = 0$, as required by the boundary condition. At this stage, to obtain an explicit formula for

u, the method of characteristics prescribes to try the following: (1) solve the given IVPs, obtaining formulae for x, y, z as functions of (s, r); (2) invert the functions x and y, that is express (s, r) in terms of (x, y). This way one can rewrite $z(s; r) = u(x(s; r), y(s; r))$ as a function of x and y alone.

One can avoid to carry out steps (1) and (2) explicitly by exploiting the Post algorithm. In fact, seeing r as an independent *variable*, rather than as a parameter, one can turn the above family of ODE IVPs into a AFP, coherent stratified system H of PDEs for the functions $x(s, r), y(s, r), \ldots$: say $H = \{(\Sigma_1, \{s, r\}), (\Sigma_2, \{r\})\}$, for the obvious choices of Σ_1 and Σ_2. Now, one can use Post to systematically search for all valid polynomial relations linking x, y, z. If the resulting polynomial system can be solved for z, obtaining say $z = f(x, y)$, one can deduce $u(x, y) = f(x, y)$, at least for (x, y) sufficiently near to the boundary[12] C. In the present case, we run $\text{Post}_H(P, \pi)$ with $P = \{x - 1, y, z, p - 1, q\}$ (encoding initial values for x, y, z, p, q) and π the complete template of total degree 2 over the variables $\{x, y, z\}$, which has 10 parameters. We get stabilization at $m = 5$ (after about 5s), obtaining a 1-parameter result template π', where $\pi'[1] = x^2 + y^2 - z^2 - 2z - 1 = x^2 + y^2 - (z+1)^2$. Therefore $x^2 + y^2 = (z+1)^2$ is the only valid polynomial relation of degree ≤ 2 for this system. Solving for z, we obtain $z = \pm\sqrt{x^2 + y^2} - 1$. The function involving the negative square root does not satisfy the boundary condition, so we deduce that $u = z = \sqrt{x^2 + y^2} - 1$ is the solution of the original problem.

A.7 Additional examples 2: conservation laws

Conservation laws may provide important qualitative insight about a system and are also crucial in applications. The following definitions are rephrased in our notation and specialized to the polynomial case we consider from [17, Ch.4,Sect.3]. Given a stratified system H, a (polynomial) *conservation law* for H is a n-tuple $\mathbf{C} = (C_1, \ldots, C_n) \in \mathcal{P}^n$ such that the equation

$$\nabla \mathbf{C} \stackrel{\triangle}{=} D_{x_1} C_1 + \cdots + D_{x_n} C_n = 0 \tag{43}$$

is valid under all solutions of H; equivalently, such that $\nabla \mathbf{C} \in \mathrm{sp}_H(\emptyset)$. This can be generalized to $\nabla \mathbf{C} \in \mathrm{sp}_H(P)$, for any given $P \subseteq \mathcal{P}_0(H)$ defining a set of initial data specifications. In this context, the expressions C_i are called *conserved currents*. For $n = 1$ and $X = \{t\}$, (43) expresses a first integral of motion of the system. For $n = 2$ and $X = \{t, x\}$, (43) expresses, informally, that variations in the *density* $\rho \stackrel{\triangle}{=} C_1$, are compensated by variations in the spatial *flux* $\phi \stackrel{\triangle}{=} C_2$. See [17, Ch.4,Prop.4.20]. The literature on conservation laws often confines to the the special case $H = \{(\Sigma, X)\}$ and $P = \emptyset$. We will call such laws *global* for Σ.

Since an equation $\nabla \mathbf{C} = 0$ is a particular polynomial invariant of the system, in principle we can apply Post to the systematic search of polynomial conservation laws for a given IVP. We demonstrate this application on the following IVP for the wave equation in one spatial dimension:

$$u_{tt} = u_{xx} \qquad u_t(0, x) = 0 \qquad u(0, x) = A\sin(x) + B\cos(x) \tag{44}$$

for arbitrary real constants A, B. More specifically, the one above is a Cauchy problem. This problem is coded up as an AFP, coherent stratified system $H = \{(\Sigma_1, \{t, x\}), (\Sigma_2, \{x\})\}$, where the auxiliary variables v, w represent generic sinusoids $A\sin(x) + B\cos(x)$ and $A\cos(x) + B\sin(x)$, respectively:

$$\Sigma_1 = \{u_{tt} = u_{xx}\,, v_t = 0\,, w_t = 0\} \qquad \Sigma_2 = \{u_t = 0, u_x = w, v_x = w\,, w_x = -v\}.$$

For this example we fix, somewhat arbitrarily, a subset $S = \{u, u_t, u_x, u_{tx}, u_{xx}\} \subseteq \mathcal{D}$, and look for all polynomial conservation laws of degree ≤ 2 that can be built out of S. To this end, we first build π_1 and π_2, two complete polynomial templates of degree 2 on the indeterminates in S, using two

[12] Technically, under mild conditions [15, Ch.2] that are satisfied in the present example, the function $G(s, r) \stackrel{\triangle}{=} (x(s, r), y(s, r))$ is locally invertible around $s = 0$. Therefore, for each (x_0, y_0) sufficiently near to the boundary C and for $(s_0, r_0) = G^{-1}(x_0, y_0)$, we have: $u(x_0, y_0) = u(G(s_0, r_0)) = z(s_0, r_0) = f(G(s_0, r_0)) = f(G(G^{-1}(x_0, y_0))) = f(x_0, y_0)$.

disjoint sets of template parameters. Then let $\pi \triangleq D_t\pi_1 + D_x\pi_2$ represent a template for divergences. As there are no constraints on the initial data ($P = \emptyset$), we run $\textsc{Post}(\emptyset, \pi)$, obtaining an output (V, J), after 4 iterations and about 7s. By theorem Theorem 4.3(a), $\pi[V] \subseteq \mathrm{sp}_H(\emptyset)$, and since $\pi[V] = (D_t\pi_1 + D_x\pi_2)[V] = D_t(\pi_1[V]) + D_x(\pi_2[V]) = \{D_t\pi_1[v] + D_x\pi_2[v] : v \in V\}$, we have found the set of all polynomial conservation laws of H of the desired type. From V and π_1, π_2, we can also recover explicitly the vector space of conserved density-flux pairs (ρ, ϕ):

$$(\pi_1, \pi_2)[V] \triangleq \{(\pi_1[v], \pi_2[v]) : v \in V\}.$$

A basis for $(\pi_1, \pi_2)[V]$ can be easily built out of the result template returned by $\textsc{Post}$. We report below the density-flux pairs of just two nontrivial[13] conservation laws in the basis we computed[14].

$$\rho_1 = -2u_xu_t \qquad \phi_1 = u_x^2 + u_t^2 \qquad\qquad \rho_2 = uu_{tx} - u_xu_t \qquad \phi_2 = u^2/2 + u_x^2/2\,.$$

Importantly, the found conservation laws $(\pi_1, \pi_2)[V]$ are not necessarily valid for different IVPs of the same PDE. In particular, while $(\pi_1, \pi_2)[V]$ includes *all* global conservation laws of the considered type for the wave equation, this inclusion is in general strict. For instance, only the leftmost law above is global for the wave equation. Indeed, if we change the first initial condition in (44) to $u_t(0, x) = C\exp(-x^2)$ (C constant), and repeat the experiment, we end up with a different set of laws, not including e.g. the rightmost law above.

▶ Remark A.22 (global vs. IVP conservation laws). Methods to search for *global* conservation laws have traditionally been linked to symmetries of the system, on account of a celebrated theorem by Emmy Noether [17, Ch.4,Sect.4]. Alternative, direct methods exist that are more widely applicable, like those centered on characteristics [17, Ch.4]. In our context, let us see the given PDE equations as a set of polynomials, $\Sigma = \{u_{\tau_1} - E_1, ..., u_{\tau_k} - E_k\} \subseteq \mathbb{R}[X \cup D]$, for $D \subseteq \mathcal{D}$, with $N \triangleq |X \cup D| < +\infty$. Under suitable technical conditions on Σ (*nondegeneracy*, [17, Ch.4]), the variety $\mathbf{V}(\Sigma) \subseteq \mathbb{R}^N$ coincides with the union of the graphs of the analytic solutions of Σ (and their derivatives present in D). Then $\nabla\mathbf{C}$, or more generally any polynomial $G \in \mathbb{R}[X \cup D]$, vanishes on the solutions of Σ if and only if $G \in \mathbf{I}(\mathbf{V}(\Sigma))$. Under the mentioned technical condition, one can assume $G \in \langle\, \Sigma \,\rangle$. The polynomial coefficients Q_j s.t. $G = \sum_j Q_jE_j$ ($E_j \in \Sigma$) are known as *characteristics*. Characteristics that yield conservation laws can be searched quite effectively by analytical or algebraic means. Unfortunately, it is not obvious how to extend this approach to IVPs. In fact, the subset of the solutions satisfying the given initial conditions, represented in terms of their graphs, may have a complicated geometry, with no algebraic description. Even in cases where such descriptions exist, it is unclear how to build them systematically. This explains why, when searching for IVPs conservation laws, one may have to resort to methods that are more “brute force" in spirit, like the one outlined above.

[13] A polynomial conservation law $\mathbf{C} = (C_1, ..., C_n)$ is *trivial* if it is a linear combination of laws satisfying either of these two conditions: (a) for each i, $D_{x_i}C_i \in \mathrm{sp}_H(P)$; or, (b) $\nabla\mathbf{C} = 0$ as a polynomial in $\mathcal{P}$. See [17, Ch.4,Sect.4].

[14] Code for this example available at `https://github.com/micheleatunifi/PDEPY/blob/master/PDEv2.py`. The concrete form of the returned basis depends on the underlying platform.