15-424: Foundations of Cyber-Physical Systems

Lecture Notes on Differential Equations & Differential Invariants

André Platzer

Carnegie Mellon University
Lecture 10

1 Introduction

Lecture 5 on Dynamical Systems & Dynamic Axioms gave us a first simple proof principle for differential equations if we find a representable solution of the differential equation. The axiom [′] replaces properties of differential equations with suitably quantified properties of solutions, with a universal quantifier over all durations of the solution. Yet, that does not work for all differential equations, because only some of them have explicit closed-form solutions, and, of those, only very few have solutions that are simple enough to be quantified over without leaving the decidable parts of the resulting arithmetic.

Lecture 2 on Differential Equations & Domains allows many more differential equations to be part of CPS models than just the ones that happen to have simple solutions. In fact, in a certain sense, most of the interesting differential equations do not possess useful closed-form solutions. Today's lecture reinvestigates the way we prove properties of differential equations from a much more fundamental perspective, which will lead to a way of proving properties of CPS with more general differential equations.

More details can be found in [Pla10a, Pla10b, Chapter 3.5] and also [Pla12b]. Differential invariants were originally conceived in 2008 [Pla10a, Pla08] and later used for an automatic proof procedure for hybrid systems [PC08].

2 Global Descriptive Power of Local Differential Equations

Differential equations let physics evolve continuously for longer periods of time. They describe such global behavior locally.

Note 1 (Local descriptions of global behavior by differential equations)**.** *The key principle behind the descriptive power of differential equations is that they describe the evolution of a continuous process over time using only a local description of the direction into which the system evolves at any point in space. The solution of a differential equation is a global description of how the system evolves, while the differential equation itself is a local characterization.*

This difference between local description and global behavior can be exploited for proofs.

The semantics of a differential equation was described in Lecture 2 as:

$$\rho(x' = \theta \,\&\, H) = \{(\varphi(0), \varphi(r)) \;:\; \varphi(t) \models x' = \theta \text{ and } \varphi(t) \models H \text{ for all } 0 \leq t \leq r \text{ for a solution } \varphi : [0, r] \to \mathcal{S} \text{ of any duration } r\}$$

The solution φ describes the global behavior of the system, which is specified locally by the right-hand side θ of the differential equation.

Lecture 2 has shown a number of examples illustrating the descriptive power of differential equations. That is, examples in which the solution was very complicated even though the differential equation was rather simple. This is a strong property of differential equations: they can describe even complicated processes in simple ways. Yet, that representational advantage of differential equations does not carry over into the verification when verification is stuck with proving properties of differential equations only by way of their solutions, which, by the very nature of differential equations, are more complicated.

This lecture, thus, investigates ways of proving properties of differential equations using the differential equations themselves, not their solutions. This technique is called *differential invariants* [Pla10a, Pla12b].

3 Differential Equations vs. Loops

A programmatic way of developing an intuition for differential invariants leads through a comparison of differential equations with loops [Pla12a]. This perhaps surprising relation can be made completely rigorous and is at the heart of a deep connection equating discrete and continuous dynamics proof-theoretically [Pla12a]. We will stay at the surface of this connection but still leverage the relation of differential equations to loops for our intuition.

To get started with relating differential equations to loops, compare

$$x' = \theta \qquad \text{vs.} \qquad (x' = \theta)^*$$

How does the differential equation $x' = \theta$ compare to the same differential equation in a loop $(x' = \theta)^*$ instead? Unlike the differential equation $x' = \theta$, the repeated differential equation $(x' = \theta)^*$ can run the differential equation $x' = \theta$ repeatedly. Albeit, on second

thought, does that get the repetitive differential equation $(x' = \theta)^*$ to any more states than where the differential equation $x' = \theta$ could evolve to?

Not really, because chaining lots of solutions of differential equations from a repetitive differential equation $(x' = \theta)^*$ together will give a single solution for the same differential equation $x' = \theta$ that we could have followed just once all the way.[1]

Note 2 (Looping differential equations). *$(x' = \theta)^*$ is equivalent to $x' = \theta$, i.e. both have the same transition semantics. Differential equations "are their own loop".*[2]

In light of Note 2, differential equations look somewhat like loops. Like nondeterministic repetitions, differential equations might stop right away. Like nondeterministic repetitions, differential equations could evolve for longer or shorter durations. Like in nondeterministic repetitions, the outcome of the evolution of the system so far determines what happens next. And, in fact, in a deeper sense, differential equations actually really do correspond to loops [Pla12a].

With this rough relation in mind, let's advance the dictionary translating differential equation phenomena into loop phenomena and back. The local description of a differential equation as a relation $x' = \theta$ of the state to its derivative corresponds to the local description of a loop by a repetition operator α^*. The global behavior of a solution of a differential equation $x' = \theta$ corresponds to the full execution of a system that performs a repetition in a loop α^*. We also say that the local relation $x' = \theta$ is the generator of the global system solution and that the loop body α is the generator of the global behavior of repetition of the loop, because both local generators tell us everything about the system by way of their global interpretation as either differential or repetitive effect. Proving a property of a differential equation in terms of its solution corresponds to proving a property of a loop by unwinding it (infinitely long) by axiom $[^{*n}]$ from Lecture 5 on Dynamical Systems & Dynamic Axioms.

Now Lecture 7 on Control Loops & Invariants made the case that unwinding the iterations of a loop can be a rather tedious way of proving properties about the loop, because there is no good way of ever stopping to unwind, unless a counterexample can be found after a finite number of unwindings. Lecture 7 introduced induction with invariants instead to prove properties of loops, by, essentially, cutting the loop open and arguing that the generic state after any run of the loop body has the same characterization as the generic state before. After all these analogous correspondences between loops and differential equations, the obvious question is what the differential equation analogue proof concept would be that corresponds to proofs by induction for loops, which is the premier technique for proving loops.

Induction can be defined for differential equations using what is called *differential invariants* [Pla10a, Pla12b]. The have a similar principle as the proof rules for induction for loops. Differential invariants prove properties of the solution of the differential

[1] This is related to classical results about the continuation of solutions, e.g., [Pla10b, Proposition B.1].

[2] Beware not to confuse this with the case for differential equations with evolution domain constraints, which is subtly different.

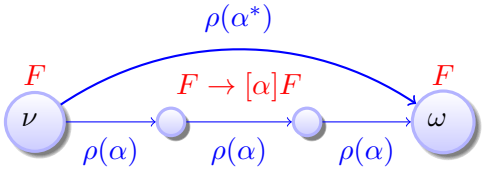
equation using only its local generator: the right-hand side of the differential equation.

Note 3 (Correspondence map between loops and differential equations).

loop α^*	***differential equation*** $x' = \theta$
can skip over	*can evolve for duration 0*
repeat any number $n \in \mathbb{N}$ *of times*	*evolve for any duration* $0 \leq r \in \mathbb{R}$
effect depends on previous iteration	*effect depends on past solution*
local generator α	*local generator* $x' = \theta$
full execution trace	*global solution* φ
proof by unwinding iterations $[^{*n}]$	*proof by solution* $[']$
proof by induction with invariant ind	*proofs by differential invariants*

Recall from Lecture 7:

$$\rho(\alpha^*) = \bigcup_{n\in\mathbb{N}} \rho(\alpha^n) \qquad \text{with} \quad \alpha^{n+1} \equiv \alpha^n;\alpha \text{ and } \alpha^0 \equiv ?true$$



$$(ind)\ \frac{\Gamma \vdash F, \Delta \quad F \vdash [\alpha]F \quad F \vdash \psi}{\Gamma \vdash [\alpha^*]\psi, \Delta}$$

4 Intuition of Differential Invariants

Just as inductive invariants are the premier technique for proving properties of loops, differential invariants [Pla10a, Pla12b] provide the primary inductive technique we use for proving properties of differential equations (without having to solve them).

The core principle behind loop induction is that the induction step investigates the local generator α ands shows that it never changes the truth-value of the invariant F (also see the core induction proof rule *ind* from Lecture 7). Let us try to establish the same inductive principle, just for differential equations.

What does the local generator of a differential equation $x' = \theta$ tell us about the evolution of a system? And how does it relate to the truth of a formula F all along the solution of that differential equation? That is, to the truth of the dL formula $[x' = \theta]F$ expressing that all runs of $x' = \theta$ lead to states satisfying F. Fig. 1 depicts an example of a vector field for a differential equation, a global solution (in red), and an unsafe region $\neg F$ (shown in blue). The safe region F is the complement of the blue unsafe region $\neg F$.

One way of proving that $[x' = \theta]F$ is true in a state ν would be to compute a solution from that state ν, check every point in time along the solution to see if it is in the safe region F or the unsafe region $\neg F$. Unfortunately, these are uncountably infinitely

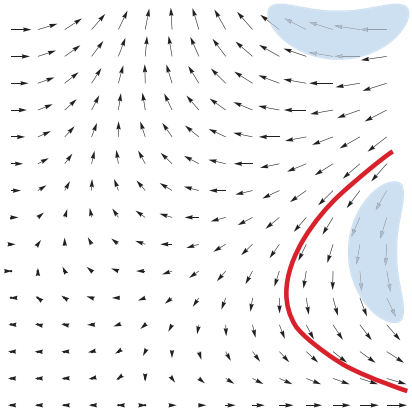


Figure 1: Vector field and one solution of a differential equation that does not enter the blue regions

many points in time to check. Furthermore, that only considers a single initial sate ν, so proving validity of a formula would require considering every of the uncountably infinitely many possible initial states and computing a solution in each of them. That is why this naïve approach would not compute.

A similar idea can still be made to work when the symbolic initial-value problem can be solved with a symbolic initial value x and a quantifier for time can be used, which is what the solution axiom [′] does. Yet, even that only works when a solution to the symbolic initial-value problem can be computed and the arithmetic resulting from the quantifier for time can be decided. For polynomial solutions, this works, for example. But polynomial come from very simple systems (called nilpotent linear differential equation systems).

Reexamining the illustration in Fig. 1, we suggest an entirely different way of checking whether the system could ever lead to an unsafe state in $\neg F$ when following the differential equation $x' = \theta$. The intuition is the following. If there were a vector in Fig. 1 that points from a safe state in F to an unsafe state $\neg F$ (in the blue region), then following that vector could get the system into an unsafe $\neg F$. If, instead, all vectors point from safe states to safe states in F, then, intuitively, following such a chain of vectors will only lead from safe states to safe states. So if the system also started in a safe state, it would stay safe.

Let us make this intuition rigorous to obtain a sound proof principle.

5 Deriving Differential Invariants

How can the intuition about directions of evolution of a logical formula F with respect to a differential equation $x' = \theta$ be made rigorous? We develop this step by step.

As an example, consider a conjecture about the rotational dynamics where d and e

represent the direction of a vector rotating clockwise in a circle of radius r (Fig. 2):

$$d^2 + e^2 = r^2 \to [d' = e, e' = -d]d^2 + e^2 = r^2 \tag{1}$$

The conjectured dℒ formula (1) is valid, because, indeed, if the vector (d, e) is initially at

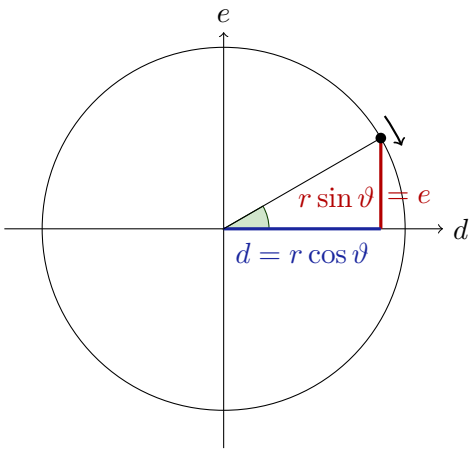


Figure 2: One scenario for the rotational dynamics and relationship of vector (d, e) to radius r and angle ϑ

distance r from the origin (0,0), then it will always be when rotating around the origin, which is what the dynamics does. That is, the point (d, e) will always remain on the circle of radius r. But how can we prove that? In this particular case, we could possibly investigate solutions, which are trigonometric functions (although the ones shown in Fig. 2 are not the only solution). With those solutions, we could perhaps find an argument why they stay at distance r from the origin. But the resulting arithmetic will be unnecessarily difficult and, after all, the argument for why the simple dℒ formula (1) is valid should be easy. And it is, after we have discovered the right proof principle as this lecture will do.

First, what is the direction into which a continuous dynamical system evolves? The direction is exactly described by the differential equation, because the differential equation describes in which direction the state evolves at every point in space. So the direction into which a continuous system obeying $x' = \theta$ follows from state ν is exactly described by the time-derivative of the state being the term θ, i.e. $[\![\theta]\!]_\nu$. Recall that term θ can mention x and other variables so its value $[\![\theta]\!]_\nu$ depends on the state ν.

Note 4. *Proving dℒ formula $[x' = \theta]F$ does not require us to answer where the system evolves to but how the evolution of the system relates to formula F and the set of states ν in which F evaluates to true.*

The logical formula F is built from atomic formulas that are comparisons of (polynomial or rational) terms. Let η denote such a (polynomial) term in the variable (vector)

x. The semantics of a polynomial term η in a state ν is the real number $[\![\eta]\!]_\nu$ that it evaluates to. In which direction does the value of η evolve when following the differential equation $x' = \theta$ for some time? That depends both on the term η that is being evaluated and on the differential equation $x' = \theta$ that describes the evolution of x.

Directions of evolutions are described by derivatives, after all the differential equation $x' = \theta$ describes that the time-derivative of x is θ. Let's derive some term η of interest and see what that tells us about how η evolves over time. How can we derive η? The term η could be built from any of the operators discussed in Lecture 2, to which we now add division for rational terms to make it more interesting. Let Σ denote the set of all variables. *Terms* θ are defined by the grammar (where θ, η are terms, x a variable, and r a rational number constant):

$$\theta, \eta ::= x \mid r \mid \theta + \eta \mid \theta - \eta \mid \theta \cdot \eta \mid \theta/\eta$$

It is, of course, important to take care that division θ/η only makes sense in a context where the divisor η is guaranteed not to be zero in order to avoid undefinedness. Thus, we only allow division to be used in a context where the divisor is ensured not to be zero.

If η is a sum $a + b$, its derivative is the derivative of a plus the derivative of b. If η is a product $a \cdot b$, its derivative is the derivative of a times b plus a times the derivative of b. The derivative of a rational number constant $r \in \mathbb{Q}$ is zero.[3] The other operators are similar, leaving only the case of a single variable x. What is its derivative?

Before you read on, see if you can find the answer for yourself.

[3] Of course, the derivative of real number constants $r \in \mathbb{R}$ is also zero, but only rational number constants are allowed in the first-order logic of real arithmetic, more precisely, of real-closed fields.

The exact value of the derivative of x certainly depends on the state and on the evolution of the system. So for now, we just define the derivative of a variable x to be the symbol x' and consider what to do with it later.

Definition 1 (Derivation)**.** The operator $(\cdot)'$ that is defined as follows on terms is called *syntactic (total) derivation*:

$$(r)' = 0 \qquad \text{for numbers } r \in \mathbb{Q} \tag{2a}$$
$$(x)' = x' \qquad \text{for variable } x \in \Sigma \tag{2b}$$
$$(a+b)' = (a)' + (b)' \tag{2c}$$
$$(a-b)' = (a)' - (b)' \tag{2d}$$
$$(a \cdot b)' = (a)' \cdot b + a \cdot (b)' \tag{2e}$$
$$(a/b)' = ((a)' \cdot b - a \cdot (b)')/b^2 \tag{2f}$$

Even though the following names are not crucial for the understanding of this course, let's briefly align Def. 1 with the algebraic structures from differential algebra [Kol72]. Case (2a) defines number symbols as *differential constants*, which do not change during continuous evolution. Their total derivative is zero. Equation (2c) and the *Leibniz* or *product rule* (2e) are defining conditions for *derivation operators on rings*. The derivative of a sum is the sum of the derivatives (additivity or a homomorphic property with respect to addition, i.e. the operator $(\cdot)'$ applied to a sum equals the sum of the operator applied to each summand) according to equation (2c). Furthermore, the derivative of a product is the derivative of one factor times the other factor plus the one factor times the derivative of the other factor as in (2e). Equation (2d) is a derived rule for subtraction according to $a - b = a + (-1) \cdot b$ and again expresses a homomorphic property, now with respect to subtraction. In addition, equation (2b) uniquely defines operator $(\cdot)'$ on the *differential polynomial algebra* spanned by the *differential indeterminates* $x \in \Sigma$. It says that we understand the differential symbol x' as the derivative of the symbol x for all state variables $x \in \Sigma$. Equation (2f) canonically extends $(\cdot)'$ to the *differential field of quotients* by the usual *quotient rule*. As the base field $\mathbb{R}$ has no zero divisors[4], the right-hand side of (2f) is defined whenever the original division a/b can be carried out, which, as we assumed, is guarded by $b \neq 0$.

The derivative of a division a/b uses a division, which is where we need to make sure not to accidentally divide by zero. Yet, in the definition of $(a/b)'$, the division is by b^2 which has the same roots that b has. So $b = 0 \leftrightarrow b^2 = 0$ is valid for any term b. Hence, in any context in which a/b was defined, its derivative $(a/b)'$ will also be.

Which of the terms should we derive when trying to prove (1)? Since that is not necessarily clear so far, let's turn the formula (1) around and consider the following equivalent dℒ formula instead, which only has a single nontrivial term to worry about:

$$d^2 + e^2 - r^2 = 0 \rightarrow [d' = e, e' = -d]d^2 + e^2 - r^2 = 0 \tag{3}$$

[4]In this setting, $\mathbb{R}$ have no zero divisors, because the formula $ab = 0 \rightarrow a = 0 \vee b = 0$ is valid, i.e. a product is zero only if a factor is zero.

Derivation of the relevant term $d^2 + e^2 - r^2$ in the postcondition of (3) gives

$$(d^2 + e^2 - r^2)' = 2dd' + 2ee' - 2rr' \tag{4}$$

Def. 1 makes it possible to derive polynomial and rational terms. Deriving them with the total derivative operator $(\cdot)'$ does not result in a term over the signature of the original variables in Σ, but, instead, a differential term, i.e. a term over the extended signature $\Sigma \cup \Sigma'$, where $\Sigma' \overset{\text{def}}{=} \{x' : x \in \Sigma\}$ is the set of all differential symbols x' for variables $x \in \Sigma$. In particular, the total derivative $(\eta)'$ of a polynomial term η is not a polynomial term, but may mention differential symbols such as x'. All syntactic elements of those differential terms are easy to interpret based on the semantics of terms defined in Lecture 2, except for the differential symbols. What is the meaning of a differential symbol x'?

Before you read on, see if you can find the answer for yourself.

6 The Meaning of Prime

The meaning $[\![x]\!]_\nu$ of a variable symbol x is defined by the state ν. The meaning of a differential symbol x' cannot be defined in a state ν, because derivatives do not even exist in isolated points. Along a (differentiable) continuous evolution $\varphi : [0, r] \to \mathcal{S}$ of a system, however, we can make sense of what x' means. At any point in time $\zeta \in [0, r]$ along such a continuous evolution φ, the differential symbol x' can be taken to mean the time-derivative of the value $[\![x]\!]_{\varphi(\zeta)}$ of x at ζ [Pla10a]:

Definition 2 (Differentially augmented state in differential state flow)**.** The value of x' at time $\zeta \in [0, r]$ of a differentiable function $\varphi : [0, r] \to \mathcal{S}$ of some duration $r \in \mathbb{R}$ is defined as:

$$[\![x']\!]_{\varphi(\zeta)} = \frac{\mathsf{d}\varphi(t)(x)}{\mathsf{d}t}(\zeta)$$

Intuitively, $[\![x']\!]_{\varphi(\zeta)}$ is determined by considering how the value $\varphi(\zeta)(x) = [\![x]\!]_{\varphi(\zeta)}$ of x changes along the function φ when we change time ζ "only a little bit". Visually, it corresponds to the slope of the tangent at time ζ; see Fig. 3.

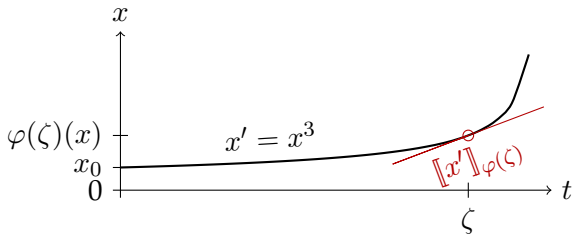


Figure 3: Differential state flow

Yet, what is the right-hand side in Def. 2, i.e. the time-derivative of the value of x along φ at time ζ? For differentiable φ, that is always defined, but that does not mean it would be computable. If, however, the continuous evolution φ is generated by a differential equation $x' = \theta$, i.e. φ solves $x' = \theta$, then $[\![x']\!]_{\varphi(\zeta)}$ can be described easily in terms of that differential equation, because at any time $\zeta \in [0, r]$ the time-derivative of the value of x is $[\![\theta]\!]_{\varphi(\zeta)}$, by definition of what it means for φ to be a solution of $x' = \theta$ (cf. Lecture 2).

Now Def. 1 defines how to derive a term η syntactically and Def. 2 defines how to interpret the differential symbols that occur in the total derivative $(\eta)'$. When interpreting all differential symbols as defined in Def. 2 for an evolution φ that follows the differential equation $x' = \theta$, this defines a value for the derivative $(\eta)'$ of any term η along that function φ. What does this value mean? How does it relate to how the value of η changes over time?

Before you read on, see if you can find the answer for yourself.

When interpreting differential symbols by derivatives along a function φ, the value of $(\eta)'$ at any time ζ coincides with the analytic time-derivative of the value of η at ζ.

The following central lemma, which is the differential counterpart of the substitution lemma, establishes the connection between syntactic derivation of terms and semantic differentiation as an analytic operation to obtain analytic derivatives of valuations along differential state flows. It will allow us to draw analytic conclusions about the behaviour of a system along differential equations from the truth of purely algebraic formulas obtained by syntactic derivation. In a nutshell, the following lemma shows that, along a flow, analytic derivatives of valuations coincide with valuations of syntactic derivations.

Lemma 3 (Derivation lemma)**.** *Let $\varphi : [0,r] \to \mathcal{S}$ be a differentiable function of duration $r > 0$. Then for all terms η that are defined all along φ and all times $\zeta \in [0,r]$:*

$$\frac{\mathsf{d}\,[\![\eta]\!]_{\varphi(t)}}{\mathsf{d}t}(\zeta) = [\![(\eta)']\!]_{\varphi(\zeta)}$$

where differential symbols are interpreted according to Def. 2. In particular, $[\![\eta]\!]_{\varphi(\zeta)}$ is continuously differentiable.

Proof. The proof is an inductive consequence of the correspondence of the semantics of differential symbols and analytic derivatives along a flow (Def. 2). It uses the assumption that φ remains within the domain of definition of η and is continuously differentiable in all variables of η. In particular, all denominators are nonzero during φ.

- If η is a variable x, the conjecture holds immediately by Def. 2:

  $$\frac{\mathsf{d}\,[\![x]\!]_{\varphi(t)}}{\mathsf{d}t}(\zeta) = \frac{\mathsf{d}\,\varphi(t)(x)}{\mathsf{d}t}(\zeta) = [\![(x)']\!]_{\varphi(\zeta)}.$$

  The derivative exists, because φ is assumed to be differentiable.

- If η is of the form $a+b$, the desired result can be obtained by using the properties of analytic derivatives, synctactic derivations (Def. 1), and valuation of terms (Lecture 2):

  $$\begin{aligned}
  &\frac{\mathsf{d}}{\mathsf{d}t}([\![a+b]\!]_{\varphi(t)})(\zeta) && \\
  =\ &\frac{\mathsf{d}}{\mathsf{d}t}([\![a]\!]_{\varphi(t)} + [\![b]\!]_{\varphi(t)})(\zeta) && [\![\cdot]\!]_\nu \text{ homomorphic for } + \\
  =\ &\frac{\mathsf{d}}{\mathsf{d}t}([\![a]\!]_{\varphi(t)})(\zeta) + \frac{\mathsf{d}}{\mathsf{d}t}([\![b]\!]_{\varphi(t)})(\zeta) && \frac{\mathsf{d}}{\mathsf{d}t} \text{ is a (linear) derivation} \\
  =\ &[\![(a)']\!]_{\varphi(\zeta)} + [\![(b)']\!]_{\varphi(\zeta)} && \text{by induction hypothesis} \\
  =\ &[\![(a)' + (b)']\!]_{\varphi(\zeta)} && [\![\cdot]\!]_\nu \text{ homomorphic for } + \\
  =\ &[\![(a+b)']\!]_{\varphi(\zeta)} && (\cdot)' \text{ is a syntactic derivation}
  \end{aligned}$$

- The case where η is of the form $a \cdot b$ or $a - b$ is similar, using Leibniz product rule (2e) or subtractivity (2d) of Def. 1, respectively.
- The case where η is of the form a/b uses (2f) of Def. 1 and further depends on the assumption that $b \neq 0$ along φ. This holds as the value of η is assumed to be defined all along state flow φ.
- The values of numbers $r \in \mathbb{Q}$ do not change during a state flow (in fact, they are not affected by the state at all); hence their derivative is $(r)' = 0$. □

Lemma 3 shows that the value of the total derivative of a term coincides with the analytic derivative of the term, provided that differential symbols are interpreted according to Def. 2. Along a differential equation $x' = \theta$, the differential symbols have a simple interpretation, the interpretation determined by the differential equation. Putting these thoughts together leads to replacing differential symbols with the corresponding right-hand sides of their respective differential equations. That is, replacing left-hand sides of differential equations with their right-hand sides.

Note 8. *The direction into which the value of a term η evolves as the system follows as differential equation $x' = \theta$ depends on the term η and the differential equation $x' = \theta$ that locally describes the evolution of x.*

The substitution property can be lifted to differential equations, i.e., differential equations can be used for equivalent substitutions along differential state flows respecting the corresponding differential constraints. In a nutshell, the following lemma can be used to substitute right-hand sides of differential equations for the left-hand side derivatives for flows along which these differential equations hold. For comparison, the classical substitution property says that equals can be substituted for equals, i.e., left-hand sides of equations can be substituted by right-hand sides of equations within formulas in which the equations hold.

Lemma 4 (Differential substitution property for terms)**.** *If $\varphi : [0, r] \to \mathcal{S}$ solves the differential equation $x' = \theta$, i.e. $\varphi \models x' = \theta$, then $\varphi \models (\eta)' = (\eta)'^{\theta}_{x'}$ for all terms η, i.e.:*

$$[\![(\eta)']\!]_{\varphi(\zeta)} = [\![(\eta)'^{\theta}_{x'}]\!]_{\varphi(\zeta)} \quad \text{for all } \zeta \in [0, r]$$

Proof. The proof is a simple inductive consequence of Lemma 3 using that $[\![x']\!]_{\varphi(\zeta)} = [\![\theta]\!]_{\varphi(\zeta)}$ at each time ζ in the domain of φ. □

The operation mapping term η to $(\eta)'^{\theta}_{x'}$ is called *Lie-derivative* of η with respect to $x' = \theta$.

Differential substitution of the differential equation $d' = e, e' = -d$ from (3) into (4) results in

$$(d^2 + e^2 - r^2)'^{e\ -d}_{d'\ e'} = (2dd' + 2ee' - 2rr')^{e\ -d}_{d'\ e'} = 2de + 2e(-d) + 2rr'$$

Oops, that did not make all differential symbols disappear, because r' is still around, since r did not have a differential equation in (3). Stepping back, what we mean by a differential equation like $d' = e, e' = -d$ that does not mention r' is that r is not supposed to change. If r is supposed to change during a continuous evolution, there has to be a differential equation for r.

Note 10 (Explicit change). *Hybrid programs are* explicit change*: nothing changes unless an assignment or differential equation specifies how (compare the semantics from Lecture 3). In particular, if a differential equation (system) $x' = \theta$ does not mention z', then z does not change during $x' = \theta$, so the original system $x' = \theta$ and $x' = \theta, z' = 0$ are equivalent.*

We will often assume $z' = 0$ without further notice for variables z that do not change during a differential equation.

Since (3) does not have a differential equation for r, Note 10 implies that its differential equation $d' = e, e' = -d$ is equivalent to $d' = e, e' = -d, r' = 0$. Hence, when adding zero derivatives for all unchanged variables, differential substitution of the differential equation $d' = e, e' = -d$ along with the explicit-change assumption $r' = 0$ into (4) gives

$$(d^2 + e^2 - r^2)'^{\,e\ -d\ 0}_{d'\ e'\ r'} = (2dd' + 2ee' - 2rr')^{\,e\ -d\ 0}_{d'\ e'\ r'} = 2de + 2e(-d) \tag{5}$$

This is good news, because the last part of (5) is a standard term of first-order logic of real arithmetic, because it no longer has any differential symbols. So we can make sense of $2de+2e(-d)$ and, by Lemma 4, its value along a solution of $d' = e, e' = -d$ is the same as that of the derivative $(d^2 + e^2 - r^2)'$, which, by Lemma 3 is the same as the value of the time-derivative of the original term $d^2 + e^2 - r^2$ along such a solution. Simple arithmetic shows that the term $2de + 2e(-d)$ in (5) is 0. Consequently, by Lemma 3 and Lemma 4, the time-derivative of the term $d^2 + e^2 - r^2$ in the postcondition of (3) is 0 along any solution φ of its differential equation:

$$\begin{aligned}\frac{\mathsf{d}[\![d^2 + e^2 - r^2]\!]_{\varphi(t)}}{\mathsf{d}t}(\zeta) &\overset{\text{Lem3}}{=} [\![(d^2 + e^2 - r^2)']\!]_{\varphi(\zeta)}\\ &\overset{\text{Lem4}}{=} [\![(d^2 + e^2 - r^2)'^{\,e\ -d\ 0}_{d'\ e'\ r'}]\!]_{\varphi(\zeta)}\\ &\overset{(5)}{=} [\![2de + 2e(-d)]\!]_{\varphi(\zeta)} = 0\end{aligned}$$

for all times ζ. That means that the value of $d^2 + e^2 - r^2$ never changes during the rotation, and, hence (3) is valid, because $d^2 + e^2 - r^2$ stays 0 if it was 0 in the beginning, which is what (3) assumes.

7 Differential Invariant Terms

In order to be able to use the above reasoning as part of a sequent proof, we need to capture arguments like these in a proof rule, preferably one that is more general than

this particular argument. The argument is not specific to the term $d^2 + e^2 - r^2$ but works for any other term η and for any differential equation $x' = \theta$. This would give us a soundness proof for the following proof rule.

Lemma 5 (Differential invariant terms)**.** *The following special case of the differential invariants proof rule is sound, i.e. if its premise is valid then so is its conclusion:*

$$(DI_{=0})\ \frac{\vdash \eta'^{\theta}_{x'} = 0}{\eta = 0 \vdash [x' = \theta]\eta = 0}$$

Proof. Assume the premise $\eta'^{\theta}_{x'} = 0$ to be valid, i.e. true in all states. In order to prove that the conclusion $\eta = 0 \vdash [x' = \theta]\eta = 0$ is valid, consider any state ν. Assume that $\nu \models \eta = 0$, as there is otherwise nothing to show (sequent is trivially *true* since antecedent evaluates to *false*). If $\zeta \in [0, r]$ is any time during any solution $\varphi : [0, r] \to \mathcal{S}$ of any duration $r \in \mathbb{R}$ of $x' = \theta$ beginning in initial state $\varphi(0) = \nu$, then

$$\frac{\mathsf{d}[\![\eta]\!]_{\varphi(t)}}{\mathsf{d}t}(\zeta) \overset{\text{Lem3}}{=} [\![(\eta)']\!]_{\varphi(\zeta)} \overset{\text{Lem4}}{=} [\![(\eta)'^{\theta}_{x'}]\!]_{\varphi(\zeta)} \overset{\text{premise}}{=} 0$$

By antecedent, $\nu \models \eta = 0$, i.e. $[\![\eta]\!]_\nu = 0$, in the initial state $\nu = \varphi(0)$.

If the duration of φ is $r = 0$, we have $\varphi(0) \models \eta = 0$ immediately, because $\nu \models \eta = 0$. For duration $r > 0$, we show that $\eta = 0$ holds all along the flow φ, i.e., $\varphi(\zeta) \models \eta = 0$ for all $\zeta \in [0, r]$.

Suppose there was a $\zeta \in [0, r]$ with $\varphi(\zeta) \models \eta \neq 0$, which will lead to a contradiction. The function $h : [0, r] \to \mathbb{R}$ defined as $h(t) = [\![\eta]\!]_{\varphi(t)}$ satisfies the relation $h(0) = 0 \neq h(\zeta)$, because $h(0) = [\![\eta]\!]_{\varphi(0)} = [\![\eta]\!]_\nu$ and $\nu \models \eta = 0$ by antecedent of the conclusion. By Lemma 3, h is continuous on $[0, r]$ and differentiable at every $\xi \in (0, r)$. By mean value theorem, there is a $\xi \in (0, \zeta)$ such that $\frac{\mathsf{d}h(t)}{\mathsf{d}t}(\xi) \cdot (\zeta - 0) = h(\zeta) - h(0) \neq 0$. In particular, we can conclude that $\frac{\mathsf{d}h(t)}{\mathsf{d}t}(\xi) \neq 0$. Now Lemma 3 implies that $\frac{\mathsf{d}h(t)}{\mathsf{d}t}(\xi) = [\![(\eta)']\!]_{\varphi(\xi)} \neq 0$. This, however, is a contradiction, because the premise implies that the formula $(\eta)' = 0$ is true in all states along φ, including $\varphi(\xi) \models (\eta)' = 0$, which contradicts $[\![(\eta)']\!] \neq 0$. ∎

This proof rule enables us to prove (3) easily in dℒ's sequent calculus:

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\vdash 2de + 2e(-d) - 0 = 0}\,\mathbb{R}}{\vdash (2dd' + 2ee' - 2rr' = 0)^{e\ -d\ -0}_{d'\ e'\ r'}}}{d^2 + e^2 - r^2 = 0 \vdash [d' = e, e' = -d]d^2 + e^2 - r^2 = 0}\,DI_{=0}}{\vdash d^2 + e^2 - r^2 = 0 \to [d' = e, e' = -d]d^2 + e^2 - r^2 = 0}\,{\to}\mathrm{r}$$

The line proof step that This is an exciting development, because, thanks to differential invariants, the property (3) of a differential equation with a nontrivial solution has a very simple proof that we can easily check.

8 Summary

This lecture showed one simple special form of differential invariants: the form where the differential invariants are terms whose value always stays 0 along all solutions of a differential equation. The next lecture will investigate more general forms of differential invariants and more advanced proof principles for differential equations.

The most important insight of today's lecture was that complicated behavior of systems defined in terms of real analytic properties and semantics can be captured by purely syntactical proof principles using derivations. The derivation lemma proved that the values of syntactic derivations coincides with the analytic derivatives of the values. The differential substitution lemma allowed us the intuitive operation of substituting differential equations into terms. Proving properties of differential equations using these simple proof principles is much more civilized and effective than working with solutions of differential equations. The proofs are also computationally easier, because the proof arguments are local.

Exercises

Exercise 1. What happens in the proof of Lemma 5 if there is no solution φ? Show that this is not a counterexample to proof rule $\text{DI}_{=0}$, but that the rule is sound in that case.

References

[Kol72] Ellis Robert Kolchin. *Differential Algebra and Algebraic Groups*. Academic Press, New York, 1972.

[PC08] André Platzer and Edmund M. Clarke. Computing differential invariants of hybrid systems as fixedpoints. In Aarti Gupta and Sharad Malik, editors, *CAV*, volume 5123 of *LNCS*, pages 176–189. Springer, 2008. doi:10.1007/978-3-540-70545-1_17.

[Pla08] André Platzer. *Differential Dynamic Logics: Automated Theorem Proving for Hybrid Systems*. PhD thesis, Department of Computing Science, University of Oldenburg, Dec 2008. Appeared with Springer.

[Pla10a] André Platzer. Differential-algebraic dynamic logic for differential-algebraic programs. *J. Log. Comput.*, 20(1):309–352, 2010. doi:10.1093/logcom/exn070.

[Pla10b] André Platzer. *Logical Analysis of Hybrid Systems: Proving Theorems for Complex Dynamics*. Springer, Heidelberg, 2010. doi:10.1007/978-3-642-14509-4.

[Pla12a] André Platzer. The complete proof theory of hybrid systems. In *LICS*, pages 541–550. IEEE, 2012. doi:10.1109/LICS.2012.64.

[Pla12b] André Platzer. The structure of differential invariants and differential cut elimination. *Logical Methods in Computer Science*, 8(4):1–38, 2012. doi:10.2168/LMCS-8(4:16)2012.