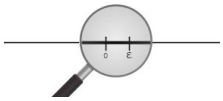
15-824 TERM PAPER
NON-STANDARD SEMANTICS FOR DIFFERENTIAL DYNAMIC LOGIC

NOAH ABOU EL WAFA



Abstract. Differential dynamic logic provides a remarkable framework for verifying hybrid systems. Yet proving properties of differential equations remains challenging. This is expected, as it necessitates syntactic reasoning about continuous phenomena. We present a new approach to this problem by viewing the continuous evolution of a physical proccess as a discrete process on an infinitely small scale. For this purpose we introduce a modification $\mathsf{d}\mathcal{L}^*$ of differential dynamic logic that incorporates infinitesimals. We also present a proof system for $\mathsf{d}\mathcal{L}^*$ that allows for natural inductive proving of differential invariants. Properties proved this way in $\mathsf{d}\mathcal{L}^*$ can be transferred to $\mathsf{d}\mathcal{L}$ seamlessly and we show that we may view $\mathsf{d}\mathcal{L}^*$ as a useful tool for proving $\mathsf{d}\mathcal{L}$-formulas.

1. Introduction

The main tools to prove properties of differential equations in differential dynamic logic are differential invariants. Similar to loop invariants, which are properties that remain true for an arbitrary number of repetitions of a program, differential invariants are properties that remain true for an arbitrary evolution of a differential equation. While loop invariants can be proved naturally by induction, the lack of a natural notion of a single step or iteration seems to preclude a similar, simple approach to proving differential invariants. To make this kind of reasoning possible additional complexity is required. This becomes apparent when considering the axiom

$$(\overleftrightarrow{\Delta})\quad [x' = f(x)F] \leftrightarrow \forall t \geqslant 0 \exists \varepsilon < 0 \exists h_0 > 0 \forall 0 < h < h_0 [(x := x + hf(x))^*](t \geqslant 0 \rightarrow \neg U_\varepsilon(\neg F))$$

proved sound in [Pla12]. (Here $U_\varepsilon(F)$ is shorthand for $\exists y\ (x - y < \varepsilon \wedge y - x < \varepsilon \wedge F(y))$.) Axiom $\overleftrightarrow{\Delta}$ is theoretically very interesting, as it allows us to translate properties of a differential equation into a completely discrete statement about loops. Unfortunately the large number of quantifier alterations appears to make proving properties of differential equations by such a discretization impractical.

For this reason another approach is commonly taken when proving properties of continuous programs in differential dynamic logic. Instead of reasoning locally by discretization, one works on a larger scale by averaging. By the typical mean value arguments, one can deduce properties of a function from properties of its derivatives. An example of a proof rule of this form would be

$$\vdash x > 0 \rightarrow [x' := f(x)]x > 0 \quad \Rightarrow \quad \vdash x > 0 \rightarrow [x' = f(x)]x > 0.$$

However such arguments do not suffice to prove every differential invariant one may be interested in. For example the last proof rule is not helpful whenever $f(x) = -x$. However it is shown in [PT20] that any analytic differential invariant can be proved in $\mathsf{d}\mathcal{L}$, by adding more differential equations (so-called differential ghosts).

In this paper we present a complementary approach to the one outlined in the last paragraph. Instead of looking at the differential equation on a larger scale, we look much closer. By introducing infinitesimals (numbers that are smaller than every real) into our formal language, we can talk about what is happening on a small scale. This way we can discretize a differential equation without accumulating a significant error. In fact the error will stay infinitesimal. A

discretization is naturally amenable to syntactic reasoning and will allow us to prove the soundness of a proof rule for differential invariants, which turns a differential invariant into an entirely inductive problem without adding quantifiers. Interestingly proving differential invariants of a certain kind in the non-standard setting again requires the use of differential ghosts. However these are of a very different nature to those introduced to prove differential invariants in $\mathsf{d}\mathcal{L}$.

There has been some interesting prior work done on proving properties of differential equations with tools of non-standard analysis. See in particular [Ben97], which contains examples of properties of differential equations studied using tools of non-standard analysis. To the extent of our knowledge, few attempts have been made to apply non-standard analysis to the verification of hybrid systems. A notable exception being [SH11], which contains a development of a Hoare-style logic, using non-standard analysis to verify hybrid systems. The approach taken here is different in that it builds on differential dynamic logic. This has the advantage of close integration, that will be made explicit in the transfer principle and yields alternative ways to (purely syntactically) prove $\mathsf{d}\mathcal{L}$ formulas by using the proof calculus for the non-standard semantics. Moreover [SH11] does not enable us to prove differential invariants that refer to the real part of a hyperreal. This difficulty can be overcome in our context by introducing an error tracker.

Section 2 contains a brief primer on non-standard analysis. This is followed by the definition of the non-standard semantics for differential dynamic logic and an investigation of the relationship of those semantics to the semantics of $\mathsf{d}\mathcal{L}$ in Section 3. Following this, in Section 4, we initiate the development of a proof calculus for the non-standard semantics. This is continued in Section 5, in which we investigate a way to prove differential invariants in the non-standard setting.

Our presentation of non-standard analysis will be entirely self-contained, with the exception that we do not give a proof of Łoś's theorem.

2. A non-standard model of Real Arithmetic

There are various ways of constructing models of real arithmetic that include infinitesimals. Following Robinson's original approach we introduce them as an ultrapower of the reals. For background on non-standard analysis we refer to [Kei]. Recall that an ultrafilter on $\mathbb{N}$ is a set $U \subseteq \mathcal{P}(\mathbb{N})$ of subsets of $\mathbb{N}$ with the following four properties

(1) $\varnothing \notin U$,
(2) for all $A, B \subseteq \mathbb{N}$ if $A \in U$ and $A \subseteq B$ then $B \in U$,
(3) for all $A, B \subseteq \mathbb{N}$ if $A \in U$ and $B \in U$ then $A \cap B \in U$,
(4) for all $A \subseteq \mathbb{N}$ either $A \in U$ or $\mathbb{N} \backslash A \in U$.

An ultrafilter is non-principal, if it does not contain singletons. It is a consequence of the axiom of choice, that non-principal ultrafilter on $\mathbb{N}$ exist.

We fix $\mathcal{L} = (+, -, \cdot, ^{-1}, 0, 1, <, \{q\}_{q\in\mathbb{Q}})$ to be the first order language of ordered fields together with constant symbols for every rational. Let $\mathbb{R}$ denote the real field as an $\mathcal{L}$-structure. Also once and for all fix a non-principal ultrafilter U on $\mathbb{N}$. We write $\mathbb{R}^{\mathbb{N}}$ for the set of functions from $\mathbb{N}$ to $\mathbb{R}$. The ultrafilter U induces an equivalence relation on $\mathbb{R}^{\mathbb{N}}$ defined by

$$f \sim g \quad \Leftrightarrow \quad \{m \in \mathbb{N} : f(m) = g(m)\} \in U.$$

For $f \in \mathbb{R}^{\mathbb{N}}$ let $[f]$ denote the corresponding equivalence class of f. Recall that the ultrapower of $\mathbb{R}$ by U is an $\mathcal{L}$-structure on the set of equivalence classes

$$\mathbb{R}^{\mathbb{N}}/U = \{[f] : f \in \mathbb{R}^{\mathbb{N}}\}.$$

For brevity we denote this $\mathcal{L}$-structure by $\mathbb{R}^*$. The only property of this structure that we shall need is Łoś's theorem, which states that

$$\mathbb{R}^* \vDash \varphi(f_0, \ldots, f_n) \quad \Leftrightarrow \quad \{m \in \mathbb{N} : \mathbb{R} \vDash \varphi(f_0(m), \ldots, f_n(m))\} \in U$$

for any $\mathcal{L}$-formula $\varphi(x_0, \ldots, x_n)$ and elements $f_0, \ldots, f_n \in \mathbb{R}^*$.

To each $x \in \mathbb{R}$ we associate the constant map $c_x : m \mapsto x$. Then Łoś's theorem shows that the map $\iota : \mathbb{R} \to \mathbb{R}^*$ defined by $x \mapsto [c_x]$ is an elementary embedding. Hence $\mathbb{R}^*$ has the same first-order theory as $\mathbb{R}$. In other words $\mathbb{R}^*$ is a real closed field. We informally identify $\mathbb{R}$ with its image under ι and treat it as a subfield of $\mathbb{R}$. Elements of $\mathbb{R}^*$ are sometimes called

hyperreals. Elements of the set $\mathbb{N}^* = \{[f] \in \mathbb{R}^* : f \in \mathbb{N}^{\mathbb{N}}\}$ are called hypernaturals. Note that the hypernaturals are unbounded in $\mathbb{R}^*$: For any $x \in \mathbb{R}^*$ let $n(m)$ be some natural number greater than $x(m)$. Then $n \in \mathbb{N}^*$ and $x < n$.

We call an element $x \in \mathbb{R}^*$ infinitesimal, if $|x| < \frac{1}{n}$ for all $n \in \mathbb{N}$. By construction $\mathbb{R}^*$ contains positive infinitesimals. An example of such an element is $\varepsilon = [m \mapsto \frac{1}{m}] \in U$. For every $n \in \mathbb{N}$ the set

$$\{m \in \mathbb{N} : \mathbb{R} \vDash 0 < \frac{1}{m} < \frac{1}{n}\}$$

is cofinite and therefore belongs to U. Hence by Łoś's theorem $\mathbb{R}^* \vDash 0 < \varepsilon < \frac{1}{n}$. We fix this ε for the rest of the proposal. So $\mathbb{R}^*$ is, unlike $\mathbb{R}$, a non-Archimedean field. Recall that an ordered field is Archimedean if for all positive elements x there is $n \in \mathbb{N}$ such that

$$\underbrace{x + \ldots + x}_{n \text{ times}} > 1.$$

Since $\mathbb{R}^*$ is a field containing infinitely small elements, it must also contain infinitely large ones. For example $\frac{1}{\varepsilon}$ is greater than any standard natural number. We call elements of $\mathbb{R}^*$ with this property infinite. Note that x is infinite iff $\frac{1}{x}$ is infinitesimal. A hyperreal $x \in \mathbb{R}^*$ is finite if it is not infinite, equivalently if there is some standard natural number $n \in \mathbb{N}$ such that $|x| < n$. (In the literature these are sometimes also called unlimited and limited respectively.)

An important observation is that the product of a finite hyperreal x and an infinitesimal y is again infinitesimal: As x is finite there is $m \in \mathbb{N}$ such that $|x| < m$. Now $|y| < \frac{1}{n \cdot m}$ for any $n \in \mathbb{N}$, since y is infinitesimal. Hence $|x \cdot y| < m \cdot \frac{1}{n \cdot m} = \frac{1}{n}$ for all $n \in \mathbb{N}$.

Although $\mathbb{R}^*$ contains many more reals than $\mathbb{R}$, each new finite element has a closest unique standard real. This is made precise in the next proposition:

Proposition 1 (Standard Part Principle)**.** *For any finite hyperreal $x \in \mathbb{R}^*$ there is a unique $\widehat{x} \in \mathbb{R}$ such that $x - \widehat{x}$ is infinitesimal. We call $\widehat{x}$ the standard part of x. We write $\widehat{x} = \infty$, if $x \in \mathbb{R}^*$ is infinite.*

Proof. Assume without loss of generality that $x > 0$ and let $\widehat{x} = \sup\{y \in \mathbb{R} : y < x\}$. We show that $x - \widehat{x}$ is infinitesimal. Suppose it were not, then there is some $n \in \mathbb{N}$ such that $x - \widehat{x} > \frac{1}{n}$. But then $x + \frac{1}{n} \in \mathbb{R}$ is less than x and greater than $\widehat{x}$ contradicting the choice of $\widehat{x}$ as the supremum.

For uniqueness note that $a, b \in \mathbb{R}$ are such that $x - a$ and $x - b$ are infinitesimal. Then $(x - a) - (x - b) = a - b$ is infinitesimal as the difference of infinitesimals. But $a - b \in \mathbb{R}$. So $a = b$, since 0 is the only infinitesimal in $\mathbb{R}$. □

Observe that $\widehat{x + y} = \widehat{x} + \widehat{y}$ whenever x and y are finite, as the sum of infinitesimals is infinitesimal. Note also that x is infinitesimal iff $\widehat{x} = 0$. We say two finite hyperreals x and y are infinitely close if $x - y$ is infinitesimal and write $x \approx y$. Then $x \approx y$ is equivalent to $\widehat{x} = \widehat{y}$.

Notationally we write $[a, b]^*$ for the set $\{c \in \mathbb{R}^* : a \leqslant c \leqslant b\}$ to distinguish it from the usual interval $[a, b] = \{c \in \mathbb{R} : a \leqslant c \leqslant b\}$.

In many ways the hypernaturals behave similar to the natural numbers. An important difference is that, while every subset of $\mathbb{N}$ is well-ordered, there are subsets of $\mathbb{N}^*$, for example $\mathbb{N}^* \backslash \mathbb{N}$, that are not well-ordered. This can be remedied by restricting our attention to a particular class of subsets of $\mathbb{R}^*$, the so-called internal sets. And indeed every internal subset of $\mathbb{N}^*$ is well-ordered. We call a set $X \subseteq \mathbb{R}^*$ internal if there is a sequence $X_m \subseteq \mathbb{R}$ such that

$$x \in X \quad \Leftrightarrow \quad \{m \in \mathbb{N} : x(m) \in X_m\} \in U.$$

It is easy to see that the intersection and union of two internal sets is internal and the complement of an internal set is internal. As announced we have:

Proposition 2. *Internal subsets of $\mathbb{N}^*$ are well-ordered.*

Proof. Suppose $X_m \in \mathbb{R}$ is a sequence of sets witnessing the assumption that X is internal. We can without loss of generality assume that all X_m are non-empty. Let $f(m) = \min X_m$ and note that $[f] \in X$. It follows from Łoś's theorem that $[f]$ is the minimal element of X. □

This immediately gives us the following important property of internal sets.

Proposition 3 (Internal Induction)**.** *Suppose $X \subseteq \mathbb{N}^*$ is internal and inductive. That is $0 \in X$ and $x + 1 \in X$, whenever $x \in X$. Then $X = \mathbb{N}^*$.*

We remark that one cannot dispense with the assumption that X is internal. For example the set $\mathbb{N}$ has the two inductive properties, but is distinct from $\mathbb{N}^*$. So in particular it follows from the proposition that neither $\mathbb{N}$ nor $\mathbb{N}\backslash\mathbb{N}^*$ are internal.

Proof. Towards a contradiction assume that $Y = \mathbb{N}^*\backslash X$ is non-empty. Note that Y is also internal. By the previous proposition Y is well-ordered. Let $[f]$ be the minimal element of Y. Then $[f] \neq 0$. So let $g(m) = \min\{f(m) - 1, 0\}$ and note that by Łoś's theorem, $[g] + 1 = [f]$. Now $[g] \in X$, since $[f]$ was chosen minimal. This is a contradiction to X being inductive. □

As an example application of the internal induction principle, we note a simple fact that we will repeatedly use.

Proposition 4. *Let $a, b \in \mathbb{R}^*$ with $a \geqslant 0$. If $b - a \geqslant \varepsilon$ then there is $n \in \mathbb{N}^*$ such that $n \cdot \varepsilon \in (a, b]$.*

Proof. If $a = 0$, it is immediate that $n = 1$ is as needed. Suppose then that $a > 0$ and that there is no such $n \in \mathbb{N}^*$. Consider the set

$$X = \{n \in \mathbb{N}^* : n \cdot \varepsilon \leqslant a\}.$$

We show that X is internal and inductive. Clearly $0 \in X$ as $a \geqslant 0$. Also $(n + 1) \cdot \varepsilon \leqslant a + \varepsilon \leqslant b$ for $n \in X$. So $n + 1 \in X$ since $n + 1 \notin (a, b]$. The sequence $X_m = \{n \in \mathbb{N} : n \cdot \varepsilon(m) \leqslant a(m)\}$ witnesses that X is internal, since for any $n \in \mathbb{R}$ by Łoś's theorem:

$$n \in X \Leftrightarrow \{m \in \mathbb{N} : n(m) \in \mathbb{N} \wedge n(m) \cdot \varepsilon(m) \leqslant a(m)\} \in U \Leftrightarrow \{m \in \mathbb{N} : n(m) \in A_m\} \in U.$$

Finally by the internal induction principle $X = \mathbb{N}^*$. This is a contradiction, since $\mathbb{N}^*$ is unbounded in $\mathbb{R}^*$ and X is bounded by $\frac{a}{\varepsilon}$. □

We now consider solutions to differential equations in the non-standard setting. Suppose $f(x, \delta, y_0, \ldots, y_k)$ is an $\mathcal{L}$-term, with only $x, \delta, y_0, \ldots, y_k$ occuring in f. For notational clarity we assume $k = 0$. (In the following, if not explicitly specified otherwise, f will always be assumed to be such a term.) Let $f^{\mathbb{R}}$ be the interpretation of f in $\mathbb{R}$, that is $f^{\mathbb{R}}(x) = y$ iff $\mathbb{R} \vDash f(x) = y$. (We will drop the superscript, when it is clear that we are referring to the interpretation of f in $\mathbb{R}$.) Similarly by f^* denote the interpretation in $\mathbb{R}^*$ so $f^*(x) = y$ iff $\mathbb{R}^* \vDash f(x) = y$. Consider the (symbolic) ordinary differential equation $x' = f(x, \delta, y)$. We define an infinitesimal analogue to the Euler approximation as follows: Begin by defining the stepping function

$$A_{\delta,y}(x) = x + \delta \cdot f^{\mathbb{R}}(x, \delta, y).$$

The operation $A_{\delta,y}$ is a real Euler step of length δ along the differential equation $x' = f(x)$. A solution to a differential equation will be a walk along the differential equation, that is a sequence of steps following f. This makes sense in the non-standard context, where we may choose an infinitesimal step size and thereby compute a solution up to an infinitesimal error. However if we choose an infinitesimal step size, we will need many steps. Since even making an infinitesimal step for every natural number will not get us very far. Let us first however take finitely many steps. Define the k-time repeated application A^k of A by $A^0 = \mathrm{id}$ and $A^{k+1} = A \circ A^k$.

Recall that we have fixed a particular infinitesimal $\varepsilon > 0$. We now extend the path following the differential equation with fixed step size ε by defining $W^n : (\mathbb{R}^*)^2 \mapsto \mathbb{R}^*$ for a hypernatural $n \in \mathbb{N}^*$ by

$$W^n(x_0, y) = z \quad \Leftrightarrow \quad \{m \in \mathbb{N} : A^{n(m)}_{\varepsilon(m),y(m)}(x_0(m)) = z(m)\} \in U.$$

This is our notion of a solution to the differential equation with initial value x_0. In contrast to the real flow, this is sometimes referred to as a hyperreal walk or a hyperfinite Euler approximation. The purpose of this paper is to see how we can exploit the more discrete flavour of the hyperreal walk in syntactic reasoning about properties of differential equations.

Observe that the hyperreal walk is always defined. In particular our definition did not rely on any smoothness properties of the function $f^{\mathbb{R}}$. It is important to note however that it is only

meaningful when it is finite. There are interesting results asserting finiteness of the hyperreal walk that are parallel to the Peano Existence theorem. See for example [Kei].

We will repeteadly use that $W^0(x_0, y) = x_0$ is the identity and

$$W^{n+1}(x_0, y) = W^n(x_0, y) + \varepsilon f^*(W^n(x_0, y), \varepsilon, y),$$

for $n \in \mathbb{N}$, which follows using Łoś's theorem. Sometimes we will drop the parameters ε and y entirely to aid readability and write $W^{n+1}(x_0) = W^n(x_0) + \varepsilon f^*(W_n(x_0))$, when there is no chance of ambiguity. It is important that in the above definition we allow the differential equation f to depend on the step size. This will later allow us to consider differential equations tracking the (infinitesimal) error.

The hyperreal walk is supposed to model a solution of a differential equation up to an infinitesimal displacement. We want to make this precise by relating the hyperreal walk to the standard notion of such a solution. Our next proposition says that as long as a hyperreal walk stays finite, it is an ordinary solution to the initial value problem defined by f up to an infinitesimal error.

Proposition 5. *Suppose $y_0 \in \mathbb{R}$ and there is $N \in \mathbb{N}^*$ such that $T = N\varepsilon$ and $W^n(y_0)$ are finite for all $n \leqslant N$. Then*

$$y : [0, \widehat{T}] \to \mathbb{R}, \quad y(\widehat{n\varepsilon}) \mapsto \widehat{W^n(y_0)}$$

is well-defined with $y(0) = y_0$ and $y'(t) = f(y(t))$.

A proof of this result can be found in [Kei] (Theorem 14.1) using the non-standard definition of derivative and integral. That those coincide with the standard notions is proved for example in [Ben97] (Corollary 6.2).

We will later also require a kind of converse result to the last proposition. The next lemma provides such a result by telling us that if there is a solution to a differential equation, then it almost agrees with the hyperreal walk.

Lemma 6. *Suppose $N \in \mathbb{N}^*$ such that $T = N\varepsilon$ is finite and there is $y : [0, \widehat{T}] \to \mathbb{R}$ satisfiying $y(0) = y_0$ and $y'(t) = f(y(t))$. Then for all $n \leqslant N$ the hyperreal walk $W^n(y_0)$ is finite and*

$$W^n(y_0) \approx y(\widehat{n\varepsilon}).$$

The proof is similar to the estimation of the global error of an Euler approximation. It can be found in the appendix.

3. Non-Standard Semantics

As mentioned above we now want to use the framework of non-standard analysis to redefine the semantics for $\mathsf{d}\mathcal{L}$ in the hope that the more discrete nature of hyperfinite walks lends itself to useful and natural axioms to prove properties of differential equations. To achieve this we will need to be able to refer to the step ε in our formal language. Moreover it will later be useful to be able to formally refer to the standard parts of a hyperreal. So we slightly extend the syntax of differential dynamic logic to accommodate reasoning about infinitesimals. For reference we restate first the syntax for differential dynamic logic $\mathsf{d}\mathcal{L}$ as introduced in [Pla12]:

Terms:	$e ::= x \mid x' \mid c \mid e + \tilde{e} \mid e - \tilde{e} \mid e \cdot \tilde{e} \mid e/\tilde{e}$
Formulas:	$P, Q ::= e \geqslant \tilde{e} \mid \neg P \mid P \wedge Q \mid \exists x P \mid [\alpha]P$
Hybrid Programs:	$\alpha, \beta ::= x := e \mid ?Q \mid x' = f(x) \;\&\; Q \mid \alpha \cup \beta \mid \alpha; \beta \mid \alpha^*$

We define syntax for an extension $\mathsf{d}\mathcal{L}^*$ in exactly the same way only adding to the definition of terms another constant symbol ε and another unary function symbol $\widehat{\ }$. We assume that in ODE programs f is a $\widehat{\ }$-free term. However ε may occur in f. Let $\mathcal{L}^*$ be the (first-order) expansion of $\mathcal{L}$ by adding the new symbols ε and $\widehat{\ }$. Then the modality free $\mathsf{d}\mathcal{L}$-formulas are exactly the first-order $\mathcal{L}$-formulas and the modality free $\mathsf{d}\mathcal{L}^*$-formulas are exactly the first order $\mathcal{L}^*$-formulas.

The notions $\vee$, $\rightarrow$, $\leftrightarrow$, $=$, $\forall$ and $\langle\rangle$ are used as shorthands, using their usual definitions. Additionally we write $e \approx k$ to abbreviate $\widehat{e} = \widehat{k}$. We will also abbreviate by $\mathrm{fin}(x)$ the formula $\widehat{x} = 0 \wedge x \neq 0$ asserting that x is finite.

Recall that a state ω is a map from the set of variables $\mathcal{V}$ to $\mathbb{R}$. Similarly a hyperstate ω is a map from the set of variables $\mathcal{V}$ to $\mathbb{R}^*$. It is occassionally useful to replace the value of a variable x by some other value a. We write $\omega\frac{a}{x}$ for the state that is obtained from ω in this way.

We define the semantics for $\mathsf{d}\mathcal{L}^*$ as follows: For a hyperstate and a $\mathsf{d}\mathcal{L}$-term e, let $\omega_*[\![e]\!]$ be the usual interpretation of e in $\mathbb{R}^*$ where $\omega_*[\![\varepsilon]\!] = \varepsilon$ and $\omega_*[\![\widehat{e}]\!] = \widehat{\omega_*[\![e]\!]}$, if this is not ∞ and $\omega_*[\![\widehat{e}]\!] = 0$ otherwise. By recursion on the complexity of the formula we define the semantics of $\mathsf{d}\mathcal{L}^*$-formulas P as a satisfaction relation $\omega \vDash_* P$ (for hyperstates ω) and the semantics of a hybrid program α as a transition relation $[\![P]\!]_* \subseteq \mathcal{S}^* \times \mathcal{S}^*$. (We also write $[\![P]\!]_*$ for the set of hyperstates ω with $\omega \vDash P$.) The definitions are for the most part identical to the definitions for $\mathsf{d}\mathcal{L}$. Indeed for formulas the definition is the same as in [Pla18] verbatim. Similarly for hybrid programs the definitions remain almost entirely unchanged. We remark only on the cases of loops and differential equations in hybrid programs.

The semantics of a repetition are defined exactly the same way as for differential dynamic logic. To be precise $(\omega_0,\ \omega_n) \in [\![\alpha^*]\!]_*$ exactly if there are $\omega_1, \ldots, \omega_{n-1}$ such that $(\omega_i, \omega_{i+1}) \in [\![\alpha]\!]_*$ for all i. Note that we do not allow hypernatural repetitions. This is because discrete programs are only supposed to run finitely many times. Although it may be theoretically interesting to consider longer repetitions, they do not faithfully model cyber-physical systems.

For differential equations we make the following definition:

Definition 7. For a continuous program $(\omega, \nu) \in [\![x' = f(x)\ \&\ Q]\!]_*$ exactly if there is $r \in \mathbb{R}^*$ finite and a map

$$\varphi : [0, r]^* \to \mathcal{S}^*, \quad t \mapsto \varphi_t$$

such that $\varphi_0 = \omega$ except at x', $\varphi_r = \nu$ and $\varphi \vDash_* x' = f(x)\ \&\ Q$.

Here $\varphi \vDash_* x' = f(x)\ \&\ Q$ abbreviates the statement that $\varphi_z(y) = \varphi_0(y)$ for all $y \in \mathcal{V}\backslash\{x, x'\}$ whenever $z \in [0, r]^*$ and for all $n \in \mathbb{N}^*$ with $n \cdot \varepsilon \leqslant r$:

(1) $\varphi_{n\cdot\varepsilon}(x) = W^n(\varphi_0(x))$, where W^n is the hyperreal walk along f,
(2) $z \mapsto \varphi_z(x)$ is constant on $[n \cdot \varepsilon, (n+1) \cdot \varepsilon)^*$ and
(3) $\varphi_{n\cdot\varepsilon} \in [\![x' = f(x) \wedge Q]\!]_*$.

Before we investigate these semantics further, we remark that those are meaningful and interesting semantics. The main reason for the investigation of differential dynamic logic is after all to model and verify cyber-physical systems. So it is crucial that satisfaction of a $\mathsf{d}\mathcal{L}$ formula can expresses an interesting and important fact about a model of a hybrid system. In the case of the usual semantics of a $\mathsf{d}\mathcal{L}$-formula this is immediately apparent from the definition of the semantics. For example validity of the formula

$$x = 0 \wedge t = 0 \to [x' = 1, t' = 1\ \&\ t < 1]x < 10$$

(in $\mathsf{d}\mathcal{L}$) can be interpreted to assert that a car travelling at a velocity of 1 m/s will not crash into an obstacle 10 meters away within the next second. However it is perhaps not immediately obvious that these non-standard semantics have the same virtue. A good case can be made that the non-standard semantics do describe, for example, continuous motion meaningfully. In fact this non-standard approach is close to Leibniz's early development of calculus. We will go into no detail and will instead make the fact that the two semantics really describe the same reality mathematical and precise in the transfer principle below.

For first-order $\mathsf{d}\mathcal{L}^*$-formulas, i.e. those which are modality free, satisfaction as defined above agrees with the standard definition of satisfaction in the first-order structure $\mathbb{R}^*$. More explicitly, for any $\mathcal{L}^*$-formula P, the semantics are defined so that

$$\omega \in [\![P]\!]_* \quad \Leftrightarrow \quad \mathbb{R}^* \vDash P[\omega],$$

where $\mathbb{R}^*$ is an ultrapower. This is analogous to the equivalence

$$\omega \in [\![Q]\!] \quad \Leftrightarrow \quad \mathbb{R} \vDash Q[\omega]$$

for $\mathcal{L}$-formulas Q in $\mathsf{d}\mathcal{L}$.

There is a close relationship between satisfaction in $\mathbb{R}$ and $\mathbb{R}^*$ encapsulated in the next proposition. This is a first approximation to a transfer principle between $\mathsf{d}\mathcal{L}$ and $\mathsf{d}\mathcal{L}^*$.

Proposition 8 (Weak Transfer Principle)**.** *For an $\mathcal{L}$-formula P and a standard state $\omega \in \mathcal{S}$:*

$$\omega \in [\![P]\!] \quad \Leftrightarrow \quad \omega \in [\![P]\!]_* $$

Proof. We have

$$\omega \in [\![P]\!] \quad \Leftrightarrow \quad \mathbb{R} \vDash P[\omega] \quad \Leftrightarrow \quad \{m \in \mathbb{N} : \mathbb{R} \vDash P[\omega]\} \in U \quad \Leftrightarrow \quad \mathbb{R}^* \vDash P[\omega] \quad \Leftrightarrow \quad \omega \in [\![P]\!]_*$$

where the penultimate equivalence is by Łoś's theorem. □

We now investigate how the non-standard semantics for $\mathsf{d}\mathcal{L}^*$-formulas relate to the non-standard semantics. It is to be expected that there is indeed a form of the transfer principle, since after all both semantics are designed to describe the same reality. Naively one may hope that this takes the form of the previous proposition simply extended to all $\mathsf{d}\mathcal{L}$-formulas. We shall later see that this can not work. Intuitively the reason is that on an ε scale the solution φ to a differential equation, is linear. So a differential invariant describing quadratic behaviour, while on a large scale accurate, fails in the infinitesimal. For example we will later see that the formula

$$v = a \cdot t \wedge x = \frac{a}{2} \cdot t^2 \to [x' = v, v' = a, t' = 1 \;\&\; x, y, t < 10](v = a \cdot t \wedge x = \frac{a}{2} \cdot t^2)$$

does hold in a state where $a \neq 0$ with respect to the non-standard semantics. However it is valid with respect to the standard semantics, as is easily seen using the differential equation solution axiom ['] for $\mathsf{d}\mathcal{L}$. (See [Pla18].) This can be seen to indicate that we are looking too closely when working in the hyperreals. We will return to this phenomenon later.

The reason for considering non-standard semantics was to see how they could be used in syntactic reasoning about differential equations. Consequently we focus our attention on how first-order properties of differential equations can be transferred between $\mathsf{d}\mathcal{L}$ and $\mathsf{d}\mathcal{L}^*$. That is we concentrate on FOD formulas, i.e. formulas of the form $[x' = f(x)]P$ where P is an $\mathcal{L}$-formula. Above we remarked that a differential walk is always close enough to a solution to a differential equation in the usual sense and vice versa. This allows us to prove the following transfer principle for FOD formulas.

Proposition 9 (FOD Transfer Principle)**.** *For $\mathcal{L}$-formulas $p(x), q(x)$ and a standard state $\omega \in \mathcal{S}$:*

$$\omega \in [\![[x' = f(x) \;\&\; q(x)]p(x)]\!] \quad \Leftrightarrow \quad \omega \in [\![[x' = f(x) \;\&\; q(\widehat{x}) \wedge \mathrm{fin}(x)]p(\widehat{x})]\!]_*.$$

Proof. For the forward direction consider $\omega \in [\![[x' = f(x) \;\&\; q(x)]p(x)]\!]$. Suppose $\varphi : [0, r]^* \to \mathcal{S}^*$ witnesses that

$$(\omega, \varphi_r) \in [\![x' = f(x) \;\&\; q(\widehat{x}) \wedge \mathrm{fin}(x)]\!]_*.$$

Pick $N \in \mathbb{N}^*$ such that $N\varepsilon \in (r - \varepsilon, r]$. As r is finite $N\varepsilon$ is also finite. Moreover $\varphi_{n\varepsilon} \in [\![q(\widehat{x}) \wedge \mathrm{fin}(x)]\!]_*$ for all $n \leqslant N$ by the evolution domain constraint. Hence $W^n(\omega(x))$ is finite for all $n \leqslant N$. Setting $t = \widehat{N\varepsilon}$ by Proposition 5

$$y : [0, t] \to \mathbb{R}, \quad y(\widehat{n\varepsilon}) \mapsto \widehat{W^n(\omega(x))}$$

is a solution to the differential equation $x' = f(x)$ with initial value $\omega(x)$. Define a function $\psi : [0, t] \to \mathcal{S}$ such that $\psi_t(x) = y(t)$ and $\psi_t(x') = \psi_t[\![f(x)]\!]$. Setting $\psi_t = \omega$ everywhere else, it follows that $\psi \vDash x' = f(x)$. Moverover $y(\widehat{n\varepsilon}) = \widehat{\varphi_{n\varepsilon}(x)}$ and hence $\psi_{\widehat{n\varepsilon}} \in [\![q(x)]\!]_*$ whenever $\widehat{n\varepsilon} \in [0, t]$. So by the weak transfer principle $\psi \vDash x' = f(x) \;\&\; q(x)$. Now $\psi_t \in [\![p(x)]\!]$ using $\omega \in [\![[x' = f(x) \;\&\; q(x)]p(x)]\!]$. Again using the weak transfer principle and $\psi_t(x) = \widehat{\varphi_r}(x)$ we conclude that $\varphi_r \in [\![p(\widehat{x})]\!]_*$.

For the converse direction consider $\omega \in [\![[x' = f(x) \;\&\; q(\widehat{x}) \wedge \mathrm{fin}(x)]p(\widehat{x})]\!]_*$. Set $x_0 = \omega(x)$ and suppose $\varphi : [0, r] \to \mathcal{S}$ witnesses that $(\omega, \varphi_r) \in [\![x' = f(x) \;\&\; q(x)]\!]_*$. Now set $N = \frac{r}{\varepsilon}$. Then by Lemma 6 the hyperreal walk W^n is finite and $W^n(x_0) \approx \varphi_{\widehat{n\varepsilon}}(x)$ for any $n \leqslant N$. Define a function $\psi : [0, N\varepsilon] \to \mathcal{S}$ with $\psi_{n\varepsilon}(x) = W^n(x_0)$ and $\psi_t(x') = \psi_t[\![f(x)]\!]_*$ such that that $\psi \vDash_* x' = f(x)$. Note that $\widehat{\psi_{n\varepsilon}(x)} = \varphi_{\widehat{n\varepsilon}}(x)$. Now using that $\varphi_{\widehat{n\varepsilon}} \in [\![q(x)]\!]$ it follows that $\psi_{n\varepsilon} \in [\![q(\widehat{x})]\!]_*$. So by the condition on ω also $\psi_{n\varepsilon} \in [\![p(\widehat{x})]\!]_*$ and consequently $\varphi_r \in [\![p(x)]\!]$. □

4. A Proof Calculus for the Non-Standard Semantics

We may take the transfer principle in the last section to justify our interest in the non-standard semantics and a corresponding proof calculus. Being able to syntactically derive a certain property of a differential equation in such a proof system will easily translate into a property in $\mathsf{d}\mathcal{L}$. The remainder of this paper contains an initial development of such a proof system $\vdash_*$ exploiting the presence of infinitesimals. In light of the transfer principle we are mostly interested in proving $\mathsf{d}\mathcal{L}^*$ formulas of the form:

$$r \wedge \widehat{x} = x \wedge \ldots \wedge \widehat{y} = y \to [x' = f(x) \,\&\, q(\widehat{x}) \wedge \mathrm{fin}(x)]p(\widehat{x}).$$

(We will call such formulas transferable.) Towards this goal we develop a proof system $\vdash_*$ similar to the proof calculus for $\mathsf{d}\mathcal{L}$ in [Pla18]. As a base we take standard sequent calculus for first order logic and add axioms to deal with the added complexity of hybrid programs. First we assume that $\vdash_* P$, whenever P is a first order $\mathcal{L}^*$-formula with $\mathbb{R}^* \models P$. We add also add the axiom [:=] and proof rules G and $M[\cdot]$ as defined in [Pla18]. The soundness proofs for those axioms are identical to those for $\mathsf{d}\mathcal{L}$. The non-standard semantics only diverge from the standard semantics, when it comes to differential equations. We introduce three useful axioms for differential equations here and turn to proving differential invariants in the next section.

The first axiom tells us that any differential equation can run for at least 0 time units:

$$[x' = f(x) \,\&\, q(x)]p(x) \to (q(x) \to p(x)). \tag{DX}$$

Soundness of (DX) is immediate. Secondly we prove a version of the differential ghost axiom which allows us to introduce an additional differential equation.

Proposition 10. *For $\mathsf{d}\mathcal{L}^*$ formulas P, Q and a hat-free $\mathcal{L}^*$-term g the formula*

$$[x' = f(x) \,\&\, q(x)]p(x) \leftrightarrow \exists y[x' = f(x), y' = g(\vec{x}) \,\&\, q(x)]p(x) \tag{DG}$$

is valid with respect to $\models_*$.

Proof. Soundness of the backward direction is immediate, since y does not occur in P and Q. Soundness of the forward implication holds, because in the non-standard setting there always is a hyperreal walk along $y' = f(\vec{x})$ of arbitrary duration. (However the walk need not be finite.) □

The dG proof rule from [Pla18] follows immediately from this axiom. Finally, although we will not use it here, we also give an explicit proof of the differential cut axiom from [Pla18], which allows us to accumulate properties of differential equations when proving them one at a time.

Proposition 11. *For $\mathsf{d}\mathcal{L}^*$-formulas P, Q, C the formula*

$$[x' = f(x) \,\&\, Q]C \to ([x' = f(x) \,\&\, Q]P \leftrightarrow [x' = f(x) \,\&\, Q \wedge C]P) \tag{DC}$$

is valid with respect to $\models_*$.

Proof. Consider a hyperstate $\omega \in [\![[x' = f(x) \,\&\, Q]C]\!]_*$. The forward direction is immediate, since any $\varphi \models_* x' = f(x) \,\&\, Q \wedge C$ naturally also satisfies $\varphi \models_* x' = f(x) \,\&\, Q$.

Conversely suppose $\omega \in [\![[x' = f(x) \,\&\, Q \wedge C]P]\!]_*$ and consider a witness $\varphi \models_* x' = f(x) \,\&\, Q$ to $(\omega, \varphi_r) \in [\![x' = f(x) \,\&\, Q]\!]_*$. Because any restriction of φ also satisfies this, it follows from the first assumption on ω that $\varphi_t \in [\![C]\!]_*$ for all $t \leqslant r$. Hence also $\varphi \models_* x' = f(x) \,\&\, Q \wedge C$, which implies $\varphi_r \in [\![P]\!]_*$ as needed. □

The dC proof rule then follows as in [Pla18]. In the following we assume that (DG), (DC) and (DX) are part of the proof system $\vdash_*$.

Note that we have not mentioned any axioms for loops. The reason for this is that we view the non-standard semantics as a useful way of proving $\mathsf{d}\mathcal{L}$ properties of differential equation, by the transfer principle. They do not seem to have the potential to improve syntactic reasoning about loops over $\mathsf{d}\mathcal{L}$. Still a useful proof theory for transferable formulas could be a valuable tool to prove interesting $\mathsf{d}\mathcal{L}$-properties of differential equations.

There is a theoretical limit to a proof calculus for transferable formulas. It is shown in [Pla12] that $\mathsf{d}\mathcal{L}$ is decidable relative to an oracle for FOD-tautologies. And Proposition 9 says that any FOD-formula can be (computably) translated into a transferable formula. Hence by

incompleteness of $\mathsf{d}\mathcal{L}$ [Pla08], no recursive axiom system for $\vdash_*$ can prove every transferable $\mathsf{d}\mathcal{L}^*$-formula.

5. A Differential Invariant Proof Rule

The advantages of the non-standard approach become apparent when proving properties of differential equations. In this section we exploit the presence of infinitesimal and the discrete nature of the hyperreal walk to prove soundness of an axiom turning a differential invariant into an inductive property. We first make a very simple observation that we will use repeatedly.

Lemma 12. *Suppose $\varphi : [0,r]^* \to \mathcal{S}^*$ and $\varphi \vDash_* x' = f(x)\ \&\ Q$, then*

$$\varphi_{(k+1)\cdot\varepsilon}(x) = (\varphi_{k\cdot\varepsilon})_*[\![x + \varepsilon \cdot f(x)]\!]$$

for any $k \in \mathbb{N}^*$ such that $(n+1)\cdot\varepsilon \leqslant r$.

Proof. This follows from the definition of the semantics for continuous programs and $W^{n+1}(x_0) = W^n(x_0) + \varepsilon f^*(W_n(x_0))$. □

Now we can turn a property of a differential invariant into an inductive property:

Proposition 13. *For a modality free $\mathsf{d}\mathcal{L}$-formula $p(x)$ a $\mathsf{d}\mathcal{L}^*$-formula $q(x)$, the $\mathsf{d}\mathcal{L}^*$-formula*

$$\begin{aligned}(\mathrm{DI}^*)\qquad & [x' = f(x)\ \&\ q(x)]p(x)\\ \leftrightarrow\quad & \big((q(x) \to p(x)) \wedge [x' = f(x)\ \&\ q(x)]\,(p(x) \to [x := x + \varepsilon\cdot f(x)](q(x) \to p(x))\big)\end{aligned}$$

is valid with respect to $\vDash_*$.

Proof. For the forward direction we consider a hyperstate

$$\omega \in [\![[x' = f(x)\ \&\ q(x)]p(x)]\!]_* \tag{2}$$

and prove the two conjuncts separately. To see that $\omega \in [\![q(x) \to p(x)]\!]_*$ assume $\omega \in [\![q(x)]\!]_*$. Now let $\varphi : [0,0] \to \mathcal{S}^*$ be defined by $\varphi_0 = \omega\frac{\omega_*[\![f(x)]\!]}{x'}$. Then $\varphi \vDash_* x' = f(x)\ \&\ q(x)$ using $\omega \in [\![q(x)]\!]_*$ and $W^0(x,y) = x$. This implies that $(\omega,\varphi_0) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$. Hence $\varphi_0 \in [\![p(x)]\!]_*$. As x' may not occur freely in $p(x)$ and φ_0 agrees with ω everywhere else, we conclude $\omega \in [\![p(x)]\!]_*$.

For the second conjunct we consider $(\omega,\nu) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$ with $\nu \in [\![p(x)]\!]_*$ and show

$$\nu \in [\![[x := x + \varepsilon\cdot f(x)](q(x) \to p(x)]\!]_*.$$

So let η be a state with $(\nu,\eta) \in [\![x := x + \varepsilon\cdot f(x)]\!]_*$. If $\eta \notin [\![q(x)]\!]_*$ there is nothing to show. Otherwise pick a witness $\varphi : [0,r]^* \to \mathcal{S}^*$ to $(\omega,\nu) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$. We define an extension $\psi : [0, r+\varepsilon]^* \to \mathcal{S}^*$ as follows: Pick the $n \in \mathbb{N}^*$ such that $n\cdot\varepsilon \in (r, r+\varepsilon]$. For $t < n\cdot\varepsilon$ set $\psi_t = \varphi_r$. For $t \geqslant n\cdot\varepsilon$ define

$$\psi_t(x) = \varphi(r)[\![x + \varepsilon f(x)]\!]_*$$

and $\psi(x') = \varphi(x)\frac{\tilde{\varphi}_t(x)}{x}[\![f(x)]\!]_*$. Finally set $\psi_t(y) = \varphi_t(y)$ for all $y \in \mathcal{V}\backslash\{x,x'\}$ whenever $t \geqslant n\cdot\varepsilon$. Note that by the last lemma $\eta(x) = \psi_{r+\varepsilon}(x)$. So since $q(x)$ may not mention x' and η agrees with $\psi_{r+\varepsilon}$ everywhere else, we see $\psi_{r+\varepsilon} \in [\![q(x)]\!]_*$. It is now easy to check that $\psi \vDash_* x' = f(x)\ \&\ q(x)$. This shows that ψ witnesses $(\omega,\psi_{r+\varepsilon}) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$. So assumption (2) above yields $\psi_{r+\varepsilon} \in [\![p(x)]\!]_*$. Again using that $\psi_{r+\varepsilon}$ and η agree everywhere but on x' we conclude $\eta \in [\![p(x)]\!]_*$.

Now for the backward direction consider a hyperstate $\omega \in [\![q(x) \to p(x)]\!]_*$ with

$$\omega \in [\![[x' = f(x)\ \&\ q(x)]\,(p(x) \to [x := x + \varepsilon\cdot f(x)](q(x) \to p(x)))]\!]_*. \tag{3}$$

We need to show that $\nu \in [\![p(x)]\!]_*$ whenever $(\omega,\nu) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$. So consider a witness $\varphi : [0,r]^* \to \mathcal{S}^*$ to $(\omega,\nu) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$ and pick some $n \in \mathbb{N}^*$ with $n\cdot\varepsilon \in (r-\varepsilon, r]$. By the definition of the semantics for continuous programs, $\nu = \varphi_r = \varphi_{n\cdot\varepsilon}$. We shall show $\varphi_{n\cdot\varepsilon} \in [\![p(x)]\!]_*$. To this end consider

$$X = \{k \in \mathbb{N}^* : k\cdot\varepsilon > r \vee \varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_*\}.$$

We are done if we show $X = \mathbb{N}^*$ by internal induction.

To see that X is inductive, we observe first that $\varphi_t \in [\![q(x)]\!]_*$ for all $t \in [0,r]^*$ as $\varphi \vDash_* x' = f(x)\ \&\ q(x)$. Then $\varphi_0 \in [\![p(x)]\!]_*$ since $\omega \in [\![q(x) \to p(x)]\!]_*$ and ω agrees with φ_0 on all possibly

free variables in $q(x)$ and $p(x)$. Hence $0 \in X$. Now suppose $k \in X$ and $(k+1)\cdot\varepsilon < r$. We need to show $\varphi_{(k+1)\cdot\varepsilon} \in [\![p(x)]\!]_*$. Observe that $\varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_*$ as $k \in X$. Moreover the restriction of φ to $[0, k\cdot\varepsilon]^*$ witnesses that $(\omega, \varphi_{k\cdot\varepsilon}) \in [\![x' = f(x)\ \&\ q(x)]\!]_*$. Hence by (3):

$$\varphi_{k\cdot\varepsilon} \in [\![[x := x + \varepsilon\cdot f(x)](q(x) \to p(x))]\!]_*.$$

Since now $\varphi_{k\cdot\varepsilon}$ and $\varphi_{(k+1)\cdot\varepsilon}$ agree everywhere, but possibly on x and x' and

$$\varphi_{(k+1)\cdot\varepsilon}(x) = (\varphi_{k\cdot\varepsilon})_*[\![x + \varepsilon\cdot f(x)]\!]$$

we see $\varphi_{(k+1)\cdot\varepsilon} \in [\![q(x) \to p(x)]\!]_*$. So since $\varphi_{(k+1)\cdot\varepsilon} \in [\![q(x)]\!]_*$ we conclude that $k+1 \in X$.

It remains to show that X is internal. It is easy to see that $Y = \{k \in \mathbb{N}^* : k\cdot\varepsilon > r\}$ is internal as witnessed by $Y_m = \{k \in \mathbb{N} : k\cdot\varepsilon(m) > r(m)\}$. So since the union of internal sets is internal it suffices to show that $Z = \{k \in \mathbb{N}^* : \varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_*\}$ is internal. For the following suppose $\{x, y_0, \dots, y_\ell\}$ contains all the free variables of p. Say $a_0 = \varphi_0(y_0), \dots, a_\ell = \varphi_0(y_\ell)$. We claim that internality of Z is witnessed by

$$X_m = \{k \in \mathbb{N} : \mathbb{R} \vDash p(A^k_{\varepsilon(m)}(\varphi_0(x)(m), a_0, \dots, a_\ell))\}.$$

In other words $k \in Z$ iff $\{m \in \mathbb{N} : k(m) \in X_m\} \in U$.

Fix $k \in \mathbb{N}^*$ and write $b = \varphi_{k\cdot\varepsilon}$. Since φ_0 and $\varphi_{k\cdot\varepsilon}$ agree except at x and x' and p does not mention x' we see that

$$\varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_* \quad\Leftrightarrow\quad \varphi_0\frac{b}{x} \in [\![p(x)]\!]_* \quad\Leftrightarrow\quad \mathbb{R}^* \vDash p[b, a_0, \dots, a_\ell]$$

By Łoś's theorem

$$\varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_* \quad\Leftrightarrow\quad \{m \in \mathbb{R} : \mathbb{R} \vDash p(b(m), a_0(m), \dots, a_\ell(m))\} \in U.$$

Note that this step crucially uses that p is an $\mathcal{L}$-formula. That is p does not contain modalities, hats, ε or the predicate fin. Write also $c = \varphi_0(x)$ and observe that $b = W^k(c, a_0, \dots, a_\ell)$ as $\varphi \vDash x' = f(x)$. Recall that the definition of the hyperreal walk says that

$$\{m \in \mathbb{N} : A^{k(m)}_{\varepsilon(m), y(m)}(c(0)) = b(m)\} \in U.$$

This combined with the last equivalence implies

$$\varphi_{k\cdot\varepsilon} \in [\![p(x)]\!]_* \quad\Leftrightarrow\quad \{m \in \mathbb{R} : \mathbb{R} \vDash p(A^{k(m)}_{\varepsilon(m), y(m)}(c(0)), a_0(m), \dots, a_\ell(m))\} \in U,$$

where we use that U is closed under intersections and supersets. Hence $k \in Z$ iff $\{m \in \mathbb{N} : k(m) \in X_m\} \in U$, as required. □

The formula proved valid is useful in proving differential invariants in $\mathsf{d}\mathcal{L}^*$. Before we turn to some example applications of this proposition, we (syntactically) derive a consequence that is more convenient to use:

Corollary 14. *For a modality free $\mathsf{d}\mathcal{L}$-formula $p(x)$ the following are equivalent*

(1) $\vdash_* p(x) \to [x' = f(x)\ \&\ q(x)]p(x)$
(2) $\vdash_* p(x) \wedge q(x) \to [x := x + \varepsilon f(x)](q(x) \to p(x))$

Proof. To see that (1) implies (2) cut in $p(x) \to [x' = f(x)\ \&\ q(x)]p(x)$ to reduce (2) to

$$[x' = f(x)\ \&\ q(x)]p(x) \vdash p(x) \wedge q(x) \to [x := x + \varepsilon f(x)](q(x) \to (x)).$$

Then use the equivalence in DI* on the antecedent and apply DX to complete the proof.

The converse implication is immediate from DI* concluding with G. □

Unlike the differential invariant rule dI for $\mathsf{d}\mathcal{L}$ this equivalence gives us not only a straightforward way to prove a differential invariant but also one to disprove it. This is similar to the results in [PT20].

In many ways using the last corollary is a very natural way of proving a differential invariant. Consider the formula $x > 0 \to [x' = -x]x > 0$. In $\mathsf{d}\mathcal{L}$ this invariant would be proved using differential ghosts. By the last corollary however, for $\mathsf{d}\mathcal{L}^*$, it suffices to show

$$\vdash_* x > 0 \to [x := x - \varepsilon\cdot x]x > 0.$$

But this is immediate from

$$\mathbb{R}^* \models x > 0 \to x - \varepsilon x > 0.$$

Very similarly using Corollary 14 provability of

$$x > 0 \to [x' = -x^2 \ \& \ \mathrm{fin}(x)]x > 0$$

follows from $\mathbb{R}^* \models x > 0 \wedge \mathrm{fin}(x) \to x - \varepsilon x^2 > 0$. Note that $x - \varepsilon x^2$ is equivalent to $\frac{1}{x} > \varepsilon$, which follows from x being finite and positive.

Above we mentioned a formula that, while valid in $\mathsf{d}\mathcal{L}$, can not be expected to be valid in the $\models_*$ semantics. We will slightly modify the formula and consider instead:

$$\mathrm{fin}(a) \wedge v = a \cdot t \wedge x = \frac{a}{2} \cdot t^2 \to [x' = v, v' = a, t' = 1 \ \& \ x, v, t < 10]v = a \cdot t \wedge x = \frac{a}{2} \cdot t^2.$$

Using the previous corollary it is not difficult to see that provability of the formula in $\vdash_*$ is equivalent to $\vdash_* \frac{a}{2} \cdot \varepsilon^2 = 0$. Since $\vdash_*$ is a sound calculus $\omega \models_* P$ would imply that $\omega \models_* \frac{a}{2} \cdot \varepsilon^2 = 0$. However this is simply not true if $\omega(a) \neq 0$.

This is an interesting limiting example and points to some of the challenges that the non-standard approach faces. To overcome it one would like to replace equality with almost equality. That is we could instead try to prove the invariant $v = a \cdot t \wedge x \approx \frac{a}{2} \cdot t^2$. Indeed by the transfer principle for differential equations this is a valid differential invariant in the non-standard semantics. So isolating an axiom that applies to prove this would be an important step. We remark that DI* does not apply immediately. After all our proposition claiming soundness of DI* relied on the assumption that the formula $p(x)$ was an $\mathcal{L}$-formula. In particular no hat and no $\approx$ can occur in $p(x)$.

To remedy this we need a quantitative bound on the error. Unfortunately we can not expect to find such a bound that works for evolutions of arbitrary duration. Instead we introduce a differential equation that models the bound, which may change, but only at an infinitesimal rate. It is curios that although the non-standard proof calculus for differential invariants is conceptually very different from the standard calculus, proving the necessary differential invariants again appears to necessitate an increase in the dimension of the differential equation. However the kind of differential ghost measuring the error is again unlike those used in [PT20] to prove all kinds of differential invariants.

The next proposition shows that for a certain kind of error evolution the error remains infinitesimal for any finite time.

Proposition 15. *For any $m \in \mathbb{N}\backslash\{0\}$ and any $\mathcal{L}$-term g the $\mathsf{d}\mathcal{L}^*$-formula*

$$\rho \approx 0 \wedge \mathrm{fin}(g) \to [\rho' = g \cdot \varepsilon^m]\rho \approx 0 \tag{V}$$

is valid with respect to $\models_*$.

To be able to state this proposition it was vital that we allowed the constant symbol ε to occur on the right hand side of a differential equation.

Proof. Consider a hyperstate ω such that $\omega(\rho) \approx 0$ and $a = \omega(g)$ is finite. Suppose ν is another state such that $(\omega, \nu) \in [\![\rho' = g \cdot \varepsilon^m]\!]_*$. Let r be finite and $\varphi : [0, r]^* \to \mathcal{S}^*$ be a witness such that $\varphi \models_* \rho' = g \cdot \varepsilon$. Pick the $n \in \mathbb{N}$ such that $n\varepsilon \in (r - \varepsilon, r]$. Then

$$\nu(\rho) = \varphi_{n \cdot \varepsilon}(\rho) = W^n(\varphi(0)(\rho)) = \varphi_0(\rho) + an\varepsilon\varepsilon^m,$$

where the last equality follows from the definition of W^n and Łoś's theorem. Note that $n\varepsilon$ is finite, as r is finite. Hence $an\varepsilon\varepsilon^m$ is infinitesimal as the product of a finite and an infinitesimal hyperreal. Thus

$$\widehat{\nu(\rho)} = \widehat{\omega(\rho)} + \widehat{an\varepsilon\varepsilon^m} = 0$$

and $\nu \in [\![\rho \approx 0]\!]_*$ as required. □

Let us now assume that V is part of the calculus $\vdash_*$. We can now finally prove the invariant

$$\mathrm{fin}(a) \wedge v = a \cdot t \wedge x \approx \frac{a}{2} \cdot t^2 \to [x' = v, v' = a, t' = 1 \ \& \ x, v, t < 10](v = a \cdot t \wedge x \approx \frac{a}{2} \cdot t^2)$$

in the $\vdash_*$-calculus. Using (V), (DG) and $M[\cdot]$ this invariant follows from

$$\vdash_* v = a\cdot t \wedge \rho = x - \frac{a}{2}\cdot t^2 \to [x' = v, v' = a, t' = 1, \rho' = -\frac{a}{2}\varepsilon](v = a\cdot t \wedge \rho = x - \frac{a}{2}t^2).$$

Now DI* is applicable and proving the last formula is equivalent to proving

$$\vdash_* v = a\cdot t \wedge \rho = x - \frac{a}{2}\cdot t^2 \to \left((v+\varepsilon a) = a\cdot(t+\varepsilon) \wedge (\rho - \frac{a}{2}\varepsilon^2) = (x+\varepsilon v) - \frac{a}{2}(t+\varepsilon)^2\right),$$

which is easily verified to hold in $\mathbb{R}^*$.

We observe that proving the differential invariants in the first two examples was very natural and intuitive. Unlike in the standard $\mathsf{d}\mathcal{L}$-case the proofs of these formulas did not rely on ghosts, which introduce some conceptual complexity, but instead turned the problem into a simple verification. However if we want to prove a corresponding transferable formula we would again need to introduce an error tracking differential equation similarly to what we did in the last example. The process of finding the correct growth rate of the error appears to be an entirely mechanical process, as we only need to compute the error a single infinitesimal step causes. This may indicate that any transferrable differential invariant can (recursively) be translated into an equivalent formula which is equivalent to a first order problem by Corollary 14. (Perhaps by using the bounds obtained in the proof of Lemma 6?) If this succeeds, it would allow us to reduce any $\mathsf{d}\mathcal{L}$ differential invariance question to a first-order $\mathbb{R}^*$-tautology. It would be interesting to see what results in this direction, which would be analogous to those of [PT20], can be obtained in this non-standard context.

If the goal in the last paragraph can be achieved, it would naturally raise the question of whether there are (usable) decision procedures for $\mathbb{R}^*$-tautologies in $\mathcal{L}^*$ that can be used to decide differential invariants efficiently. There appears to be interesting related material in [DH95], [CD83] and [dMP13].

Appendix A. Proof of the Approximation Lemma

We give a proof of the approximation lemma from Section 2. This is the non-standard version of the error bound on the Euler discretization. (See for example Theorem 3 of [Pla12].) The proof is very similar.

Proof of Lemma 6. Consider the function $\check{y} : [0, N\varepsilon] \to \mathbb{R}^*$ such that $v, w \in \mathbb{R}^*$ that

$$\mathbb{R}^* \vDash \check{y}(v) = w \quad \Leftrightarrow \quad \{m \in \mathbb{N} : y(v(m)) = w(m)\} \in U.$$

For the purpose of this proof we consider an expansion $\mathcal{L}'$ of the language $\mathcal{L}$ of ordered fields by a function symbol x. In $\mathbb{R}$ this symbol will be interpreted by the function y. (We extend y to $\mathbb{R}$ by setting it to 0 outside of its domain.) In $\mathbb{R}^*$ we interpret y by $\check{y}$. Note that $\mathbb{R}^*$ is still an ultrapower of $\mathbb{R}$ in the expanded language and we can apply Łoś's theorem in the expansion. In particular we view $\mathbb{R}$ as an elementary $\mathcal{L}'$-substructure of $\mathbb{R}^*$. We begin by transferring some first-order properties between y and $\check{y}$.

Claim 1: Suppose $n \leqslant N$ then $\check{y}(n\varepsilon) \approx y(\widehat{n\varepsilon})$.
Consider $m \in \mathbb{N}$ arbitrary and fix $r = \widehat{n\varepsilon}$. By continuity of y there is $k \in \mathbb{N}$ such that

$$\mathbb{R} \vDash \forall z \left(|z - r| < \frac{1}{k} \to |x(z) - x(r)| < \frac{1}{m}\right).$$

Reading this in $\mathbb{R}^*$ and instantiating $z = n\varepsilon$ we see $|\check{y}(n\varepsilon) - \check{y}(r)| < \frac{1}{m}$. As m was arbitrary $\check{y}(n\varepsilon) \approx \check{y}(r) = y(\widehat{n\varepsilon})$.

Claim 2: Suppose $K = \max_{\xi\in[0,\widehat{T}]} |y''(\xi)|$ then whenever $n + 1 \leqslant N$:

$$|\check{y}(n\varepsilon + \varepsilon) - W^{n+1}(y_0)| \leqslant |\check{y}(n\varepsilon) + \varepsilon f(\check{y}(n\varepsilon)) - W^{n+1}(y_0)| + \frac{\varepsilon^2}{2}K.$$

The first-order Taylor approximation with Lagrange remainder to y yields

$$\mathbb{R} \vDash \forall h \forall a \forall x \left(0 \leqslant a < a + h \leqslant \widehat{T} \to |x(a+h) - c| \leqslant |x(a) + hf(x(a)) - c| + \frac{h^2}{2}K\right).$$

The claim follows by tansferring to $\mathbb{R}^*$ and instantiating $h = \varepsilon$, $a = n\varepsilon$ and $c = W^{n+1}(y_0)$.

Claim 3: There is $L \in \mathbb{R}$ such that f^* is L-Lipschitz on

$$C^* := \{v \in \mathbb{R}^* : \exists t \in [0, \widehat{T}]\ |v - \check{y}(t)| \leqslant 1\}.$$

We first note that $C = \{v \in \mathbb{R} : \exists t \in [0, \widehat{T}] : |v - \check{y}(t)| \leqslant 1\}$ is a compact subset of $\mathbb{R}$ as the continuous image of the compact set $y([0, \widehat{T}]) \times B_1(0)$ under the addition map. Since f is continuously differentiable, it is Lipschitz continuous on the compact set C. Fix a Lipschitz constant $L \in \mathbb{R}$ for f. Now that $\mathbb{R} \vDash \forall a \forall b \exists t_0, t_1 \in [0, \widehat{T}]\ (|a - x(t_0)| \leqslant 1 \wedge |a - x(t_1)| \leqslant 1 \to |f(a) - f(b)| \leqslant L|a - b|)$ transferring to $\mathbb{R}^*$ we see that L is a Lipschitz constant for f^* on C^*.

We now turn to the proof that $W^n(y_0) \approx y(\widehat{n\varepsilon})$. Define $H_n \in \mathbb{N}^*$ for all $n \leqslant N$ by

$$H_n(m) = \sum_{i=0}^{n(m)} (1 + L\varepsilon(m))^i \frac{\varepsilon(m)}{2} K.$$

We will show that $|\check{y}(n\varepsilon) - W^n(y_0)| \leqslant H_n\varepsilon$ for all $n \leqslant N$. We proceed by internal induction. It is easy to see that the relevant set is internal using the definitions of $H_n, \check{y}$ and W_n. The induction start for $n = 0$ is immediate. So suppose the inequality holds for $n \leqslant N - 1$ and compute

$$\begin{aligned}
|\check{y}(n\varepsilon + \varepsilon) - W^{n+1}(y_0)| &\leqslant |\check{y}(n\varepsilon) + \varepsilon f^*(\check{y}(n\varepsilon)) - W^{n+1}(y_0)| + \frac{\varepsilon^2}{2} K \\
&\leqslant |\check{y}(n\varepsilon) + \varepsilon f^*(\check{y}(n\varepsilon)) - W^n(y_0) - \varepsilon f^*(W^n(y_0))| + \frac{\varepsilon^2}{2} K \\
&\leqslant |\check{y}(n\varepsilon) - W^n(y_0)| + \varepsilon |f^*(\check{y}(n\varepsilon)) - f^*(W^n(y_0))| + \frac{\varepsilon^2}{2} K \\
&\leqslant (1 + L\varepsilon)|\check{y}(n\varepsilon) - W^n(y_0)| + \frac{\varepsilon^2}{2} K \\
&\leqslant (1 + L\varepsilon) H_n \varepsilon + \frac{\varepsilon^2}{2} K \\
&= H_{n+1}\varepsilon \leqslant H_N \varepsilon.
\end{aligned}$$

Next we check that H_N is finite. Using $(1 + x) \leqslant e^x$ for positive x we check that

$$H_N(m) \leqslant \frac{K}{2} \sum_{i=0}^{N(m)} e^{L\varepsilon(n)i} \varepsilon(m) \leqslant \frac{K}{2} \int_0^{\widehat{T}} e^{Lx} dx = K \frac{e^{L\widehat{T}} - 1}{2L} = K'$$

for U-almost all m, where the second inequality uses that the sum is a lower Riemann sum of the integrand for U-almost all m. Hence $H_N \leqslant K'$, in particular it is finite. Thus $|\check{y}(n\varepsilon) - W^n(y_0)| \leqslant H_N\varepsilon \approx 0$.

Finally we conclude from the last paragraph and Claim 1 that

$$|y(\widehat{n\varepsilon}) - W^n(y_0)| \leqslant |y(\widehat{n\varepsilon}) - \check{y}(n\varepsilon)| + |\check{y}(n\varepsilon) - W^n(y_0)| \approx 0.$$

Hence $W^n(y_0) \approx y(\widehat{n\varepsilon})$ as required. □

References

[BBCP12] Albert Benveniste, Timothy Bourke, Benoît Caillaud, and Marc Pouzet. Non-standard semantics of hybrid systems modelers. *Journal of Computer and Systems Sciences International*, 78(3):877–910, 2012.

[Ben97] E. Benoit. Applications of nonstandard analysis in ordinary differential equations. In L.O. Arkeryd, N.J. Cutland, and C.W. Henson, editors, *Nonstandard Analysis*, volume 493. Springer, 1997.

[Cav16] Evan Cavallo. Differential equations via temporal logic and infinitesimals. 15-824 Term Paper, Spring 2016.

[CD83] Gregory Cherlin and Max Dickmann. Real closed rings ii. model theory. *Annals of Pure and Applied Logic*, 1983.

[DH95] Lou Van Den Dries and Adam H.Lewenberg. T-convexity and tame extensions. *Journal of Symbolic Logic*, 1995.

[dMP13] Leonardo de Moura and Grant Olney Passmore. Computation in real closed infinitesimal and transcendental extensions of the rationals. In *Automated Deduction – CADE-24. CADE 2013. Lecture Notes in Computer Science*, 2013.

[Kei] H. Jerome Keisler. Foundations of infinitesimal calculus.

[Pla08] André Platzer. Differential dynamic logic for hybrid systems. *J. Autom. Reas.*, 41(2):143–189, 2008.

[Pla12] André Platzer. The complete proof theory of hybrid systems. In *LICS '12: Proceedings of the 2012 27th Annual IEEE/ACM Symposium on Logic in Computer Science*, 2012.

[Pla18] André Platzer. *Logical Foundations of Cyber-Physical Systems*. Springer, 2018.

[PT20] André Platzer and Yong Kiam Tan. Differential equation invariance axiomatization. *J. ACM*, 67(1):6:1–6:66, 2020.

[SH11] Kohei Suenaga and Ichiro Hasuo. Programming with infinitesimals: A while-language for hybrid system modeling. In *Lecture Notes in Computer Science*. Springer, 2011.