

Extending Taylor Approximation to Hybrid Automata with Integrals

Ruggero Lanotte¹ and Simone Tini¹

University of Insubria (IT)

Abstract. In [10,11] we proposed a technique to approximate Hybrid Automata (HA) with Polynomial HA. The idea was to replace functions appearing in formulae with their Taylor Polynomials. Here we extend this technique to HA with formulae admitting integral functions. We prove that we get over-approximations of the original HA. We study the conditions ensuring that: 1. the “distance” between the formulae of the original HA and its approximation get close to 0 when increasing the degree of the Taylor polynomial (syntactical approximation), 2. the “distance” between the configurations reached in n steps by the two HA get close to 0 when increasing the degree of the Taylor polynomial (semantic approximation).

1 Introduction

Hybrid automata [1, 2] (HA, for short) are a widely studied model for *hybrid systems* [13], i.e. systems with discrete and continuous state changes. HA extend classic finite state machines with continuously evolving *variables*, and exhibit two kinds of state changes: discrete jump transitions, occurring instantaneously, and continuous flow transitions, occurring while time elapses. The two kinds of transitions are guarded by *jump conditions* and *activity functions*, resp., which are formulae expressing constraints on the source and target value of the variables. Extensions to HA are considered to deal with particular scopes. As an example in [12, 8] HA are extended with data structures to face with safety and security problems. But most of hybrid system applications is modelling and verifying systems where digital computational processes interact with analog physical ones. In this setting, integrals have several applications. In physics and engineering, where hybrid systems are widely used, we mention: work and impulse, electromagnetism, first moment and center of mass, application in fluid mechanics.

As an example, the HA in Fig. 1 models a controller of a tank. The controller continuously senses the level of water and fills or empties it, aiming to keep the level between m and M litres ($M > m$). The water level, represented by variable x , varies with time depending on input/output flows. When the controller fills the tank (state *in*), the flow rate depends on time y , and is $1 - \cos(y^2)$ litres/second. Thus, after a time t the water level is increased by $\int_0^t 1 - \cos(y^2) dy$, as modeled by activity function ϕ_{in} . When the controller empties the tank (state *out*), the flow rate at time y is y^2 . Thus, after a time t the water level is decreased by $\int_0^y y^2 dt = \frac{1}{3}t^3$, as modeled by activity function ϕ_{out} .

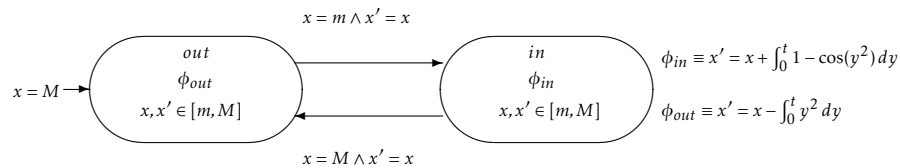


Fig. 1. The tank controller

In this example it is relevant to solve integrals by finding their antiderivatives. Unfortunately, it is well known that the integration problem is "difficult", and in many cases impossible. For instance the antiderivative is non elementary for the filling flow function $1 - \cos(y^2)$ we consider. Indeed the antiderivatives cannot be expressed by an algebraic expression of rationals, exponentials, logarithms, absolute values and trigonometric functions. A classical practical example of non elementary antiderivative function is given by the Gauss integral error $\int_a^b e^{-x^2} dx$. Therefore this problem cannot be considered a problem with restricted impact. Moreover in [14] it is showed that the integration problem is undecidable for functions with a non elementary antiderivative.

Our work is inspired by the necessity of using integrals in modeling real problems with HA, meanwhile dealing with the problem of managing and solving integrals, which is in general hard and even impossible for non elementary antiderivatives.

HA are usually used to prove *safety properties* (i.e. properties requiring that a given set of *bad configurations* cannot be reached). The decidability of *reachability* problem (i.e. whether or not a given configuration can be reached) becomes determinant. Unfortunately, for most classes of HA, reachability is undecidable [7] and the introduction of integrals complicates this analysis. However, for some classes of HA, computing the successors (or predecessors) of configurations sets is reasonably efficient, and, therefore, reachability *in a limited number of steps* is decidable. For instance for *Polynomial HA* computing the successors of configuration sets is decidable [15]. A methodology proposed in [6] fills this gap: the idea is to over-approximate an HA H with another HA H' s.t. computing the successors of configuration sets for H' is decidable and the computations of H' are a superset of the set of all the possible computations of H . Hence, if we prove that a *bad* configuration cannot be reached by H' then we can infer that it cannot be reached by the original H . In [6] it is required that the approximation H' is in the class of the *Linear HA*, for which the successors of configuration sets are computable.

The notion of approximation is then strengthened in [6] by ϵ -approximation, which is motivated by the need to limit the *error* introduced by the approximation. In [6] an asymptotically complete approximation operator, called *rationaly rectangular phase-portrait approximation*, is given which approximates any jump condition or activity function by a predicate satisfied by all points lying in a space consisting of a products of intervals with rational endpoints.

In [10, 11] over-approximations are based on replacing functions over variables with their Taylor polynomial. Since Taylor polynomials allow us to ap-

proximate functions and integrals, in the present paper we extend our technique in [10, 11] to over-approximate HA with integrals. In detail, given any HA H and $k \in \mathbb{N}$, $A(H, k)$ is the set of the Polynomial HA (for which successors of configuration sets is decidable) that are obtained by replacing in jump conditions and activity functions of H each integral $\int_{I_1}^u f(\vec{x})dx$ with a polynomial based on Taylor polynomial theory. The resulting polynomial HA over-approximates the original one according to [6]. We study the conditions ensuring that our approximation is asymptotically complete, in the sense that, for each $\epsilon > 0$ there exists some k_0 s.t., for all $k > k_0$, $A(H, k)$ contains only ϵ -approximations for H . This analysis of the error is *syntactic*, meaning that it does not consider the behaviour of H and its approximation. We consider also *semantic* analysis of the error and study conditions ensuring that, when k tends to the infinity, the behaviour of any $H_k \in A(H, k)$ gets close to the behaviour of H .

2 Hybrid Automata

In this section we recall the formalism of Hybrid Automata (see, e.g., [13]).

A *vector* of dimension n over a set U is a tuple $\vec{u} = (u_1, \dots, u_n)$ in U^n . By \vec{u}_i we denote the i^{th} element u_i . We denote by $\vec{u} \oplus (u)$ the vector $(u_1, \dots, u_n, u) \in U^{n+1}$. Then, for $\vec{u} = (u_1, \dots, u_n)$ and $\vec{v} = (v_1, \dots, v_m)$, we denote by $\vec{u} \oplus \vec{v}$ the vector $(u_1, \dots, u_n, v_1, \dots, v_m)$ in U^{n+m} . A *space* over U^n is a set of vectors in U^n .

We assume a finite set of *real variables* X ranged over by x, y, z, w, \dots . Each $x \in X$ can assume values in $\text{Dom}(x) \subseteq \mathbb{R}$. An *evaluation* over X is a mapping $v: X \rightarrow \mathbb{R}$ s.t. $v(x) \in \text{Dom}(x)$ for $x \in X$. For an evaluation v , a variable y and a real $c \in \text{Dom}(x)$, the evaluation $v[y := c]$ is defined by $v[y := c](x) = v(x)$, for $x \neq y$, and $v[y := c](y) = c$. For vectors $\vec{x} = (x_1, \dots, x_n)$ over X^n and $\vec{u} = (u_1, \dots, u_n)$ with $u_i \in \text{Dom}(x_i)$, we write $v[\vec{x} := \vec{u}]$ for $v[x_1 := u_1] \dots [x_n := u_n]$. Then, $v(x_1, \dots, x_m)$ denotes the vector $(v(x_1), \dots, v(x_m)) \in \mathbb{R}^m$. We denote by \vec{X} the vector $(x_1, \dots, x_{|X|})$ over $X^{|X|}$. Finally, we write $[\vec{X} := \vec{u}]$ to denote the evaluation v s.t. $v(\vec{X}) = \vec{u}$.

We assume a set of *function symbols* F , together with an *arity* mapping $r: F \rightarrow \mathbb{N}$ that assigns to each $f \in F$ its rank $r(f)$. If $r(f) = 0$ then f is called a *constant*. We assume a unique *interpretation* I associating to each function symbol $f \in F$ a continuous function $I(f): \text{Dom}(f) \rightarrow \mathbb{R}$ s.t. $\text{Dom}(f) \subseteq \mathbb{R}^{r(f)}$. Being I unique, sometimes with abuse of notation we use f for $I(f)$. In order to build polynomials with rational coefficients, we require that F contains the constant symbol q with $I(q) = q$ for all $q \in \mathbb{Q}$, and symbols $+$, \cdot , $-$ denoting, resp., binary summation, binary multiplication and unary negation over reals.

Definition 1. *The set $\Phi(F, X)$ of the formulae over F and X is the least set s.t.:*

- $\Phi(F, X)$ contains all basic formulae of the form

$$\int_{I_1}^{u_1} \left(\dots \left(\int_{I_n}^{u_n} f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)})) dw_{i_n} \right) \dots \right) dw_{i_1} \sim ax, \text{ where:}$$

- $n \geq 0$ and, whenever $n > 0$, then $l_1, u_1, \dots, l_n, u_n \in X \cup \mathbb{Q}$;
 - $w_1, \dots, w_{r(f)} \in X \setminus \{l_1, u_1, \dots, l_n, u_n\}$ and $\{i_1, \dots, i_n\} \subseteq \{1, \dots, r(f)\}$;
 - $g_1, \dots, g_{r(f)} \in F$ are polynomial functions s.t. $\text{Dom}(w_i) \subseteq \text{Dom}(g_i)$;
 - $f \in F$ with $\text{Dom}(f) \subseteq g_1(\text{Dom}(w_1)) \times \dots \times g_{r(f)}(\text{Dom}(w_{r(f)}))$;
 - \sim is a comparison operator in $\{<, \leq, =, \geq, >\}$;
 - $x \in X$ and $a \in \{0, 1\}$;
 - $[\min(l, r), \max(l, r)] \subseteq \text{Dom}(w_{i_j})$ for $l \in \text{Dom}(l_j), r \in \text{Dom}(u_j), j = 1, \dots, n$.
- $\neg\phi$ is in $\Phi(X, F)$ whenever ϕ is in $\Phi(X, F)$;
- $\phi_1 \vee \phi_2$ and $\phi_1 \wedge \phi_2$ are in $\Phi(X, F)$ whenever both ϕ_1 and ϕ_2 are in $\Phi(X, F)$;
- $\forall y. \phi$ and $\exists y. \phi$ are in $\Phi(X, F)$ whenever $y \in X$ and ϕ is in $\Phi(X, F)$.

The subset of polynomial formulae is obtained by restricting to (i) those $f \in F$ that are polynomial functions s.t. $\text{Dom}(f)$ is a product of intervals with bounds in $\mathbb{Q} \cup \{\pm\infty\}$, (ii) those variables $x \in X$ s.t. $\text{Dom}(x)$ is an interval with bounds in $\mathbb{Q} \cup \{\pm\infty\}$.

In [11] we restricted to basic formulae of Def. 1 with $n = 0$, i.e. general continuous function without integrals. The definition of basic formulae could appear restrictive at first glance. We argue that Def. 1 gives us expressiveness and flexibility by some examples:

1. By existential quantification, arbitrary expressions be compared. For instance, $e^{x+\sin y} \leq x/(y^2 + 1)$ is expressed by $\exists z. (e^{x+\sin y} \leq z \wedge x/(y^2 + 1) = z)$.
2. By existential quantification, we give to the user as much freedom as possible in choosing the functions to be approximated. For instance, for $f, g \in F$, we can rewrite a formula $h(\vec{x}) \sim ax$ with $h = f \circ g$ by $\exists y. g(\vec{x}) = y \wedge f(y) \sim ax$. In the first case the function $f \circ g$ is approximated, in the second case f and g are approximated separately, e.g. in order to approximate the exponential and the sin separately, the formula $\exists z. (e^{x+\sin y} \leq z \wedge x/(y^2 + 1) = z)$ in item 1 can be rewritten as $\exists z_1. \exists z_2. (e^{z_1} \leq z_2 \wedge x + \sin y = z_1 \wedge x/(y^2 + 1) = z_2)$.
3. Also arbitrary expressions dealing with integrals can be compared. For instance, the expression $h(x) + \int_x^y f(z)dz = \int_x^y g(z)dz$ can be expressed by $\exists w_1 \exists w_2. \int_x^y f(z)dz = w_1 \wedge \int_x^y g(z)dz = w_2 \wedge h(x) + w_1 = w_2$.
4. Expressions with integrals can be arguments of functions. For instance, $\cos\left(\int_x^y f(z)dz\right) > 0$ can be expressed by $\exists w. \cos(w) > 0 \wedge \int_x^y f(z)dz = w$, and $\int_0^5 f\left(x, \int_0^3 g(y)dy\right)dx \leq 7$ by $\exists z. \int_0^3 g(y)dy = z \wedge \int_0^5 f(x, z)dx \leq 7$.
5. We can deal also with general bounds for integrals. For instance the formula $\int_4^{e^x} f(y)dy \leq x$ can be expressed by $\exists z. \int_4^z f(y)dy \leq x \wedge e^x = z$.

We write $v \models \phi$ to denote that *the evaluation v satisfies the formula ϕ* . Relation \models is defined inductively as follows:

$$- v \models \int_{l_1}^{u_1} \left(\dots \left(\int_{l_n}^{u_n} f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)})) dw_{i_n} \right) \dots \right) dw_{i_1} \sim ax \text{ iff}$$

$$\int_{v(l_1)}^{v(u_1)} \left(\dots \left(\int_{v(l_n)}^{v(u_n)} I(f)(I(g_1)(e_1), \dots, I(g_{r(f)})(e_{r(f)})) dw_{i_n} \right) \dots \right) dw_{i_1} \sim I(a)v(x)$$

where either $e_j = w_j$, if $j \in \{i_1, \dots, i_n\}$, or $e_j = v(w_j)$, otherwise.

- $v \models \neg\phi$ iff $v \not\models \phi$ (namely $v \models \phi$ does not hold).
- $v \models \phi_1 \wedge \phi_2$ (resp. $v \models \phi_1 \vee \phi_2$) iff $v \models \phi_1$ and $v \models \phi_2$ (resp. $v \models \phi_1$ or $v \models \phi_2$).
- $v \models \forall y. \phi$ (resp. $v \models \exists y. \phi$) iff $v[y := c] \models \phi$ for all (resp. for some) $c \in \text{Dom}(y)$.

For a formula $\phi \in \Phi(F, X)$, let $\llbracket \phi \rrbracket$ denote the set $\{v: X \rightarrow \mathbb{R} \mid v \models \phi\}$ of the evaluations satisfying ϕ . Two formulae ϕ_1, ϕ_2 are *equivalent* iff $\llbracket \phi_1 \rrbracket = \llbracket \phi_2 \rrbracket$.

Definition 2. The subset of the normal forms in $\Phi(F, X)$ contains the formulae of the form $Q_1 y_1. \dots Q_m y_m. \phi$, where: (i) $Q_i \in \{\forall, \exists\}$ for $i = 1, \dots, m$; (ii) ϕ contains neither quantifiers nor negations; (iii) ϕ contains only relations in $\{<, \leq\}$; (iv) all basic formulae in ϕ are of the following form, for $n \leq r(f)$ and $z_1, \dots, z_n \in X$:

$$\int_0^{z_1} \left(\dots \left(\int_0^{z_n} f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)})) dw_n \right) \dots \right) dw_1 \sim ax.$$

Proposition 1. Given any formula $\phi \in \Phi(X, F)$, there exists a normal form equivalent to ϕ that can be constructed from ϕ .

E.g. $\int_4^z e^y dy \leq w$ is equivalent to the normal form $\exists x. \int_0^x e^{y+4} dy \leq w \wedge z-4 = x$.

Definition 3. An Hybrid Automaton (HA for short) H over X and F is a tuple of the form $H = \langle \phi_{init}, Q, q_0, T, Act \rangle$, where:

- $\phi_{init} \in \Phi(F, X)$ is the initial condition.
- Q is a finite set of states, and $q_0 \in Q$ is the initial state.
- $T \subseteq Q \times \Phi(F, \{x_1, \dots, x_{|X|}, x'_1, \dots, x'_{|X|}\}) \times Q$ is a finite set of transitions. Variables $x'_1, \dots, x'_{|X|}$ represent the values taken by $x_1, \dots, x_{|X|}$ after the firing of a transition.
- $Act: Q \rightarrow \Phi(F, \{x_1, \dots, x_{|X|}, t, x'_1, \dots, x'_{|X|}\})$ is the activity function assigning to each state q a formula $Act(q)$. Variable t represents time elapsing.¹

Then, H is a Polynomial Hybrid Automaton (PHA for short) iff $\phi_{init}, Act(q)$ for all states q and ϕ for each transition (q, ϕ, q') are all polynomial formulae.

Example 1. The tank controller represented in Fig. 1 has two states: in state *in* the controller fills the tank, in state *out* the controller empties the tank. The jump condition $x = m \wedge x' = x$ (resp. $x = M \wedge x' = x$) ensures that the jump from *out* to *in* (resp. from *in* to *out*) happens when the level of the water is m (resp. M), and the firing of the transition does not cause any change in the water level.

In state *in*, the water flow rate at time y is $1 - \cos(y^2)$. Hence, staying in *in* for t units of time causes a water level growing of $\int_0^t 1 - \cos(y^2) dy$. This is modelled by the activity function ϕ_{in} , which can be written in normal form in several ways. Let ϕ' be the formula $x \leq M \wedge x \geq m \wedge x' \leq M \wedge x' \geq m$, or $x, x' \in [m, M]$ for short. Given the functions f, g s.t. $f(y) = 1 - \cos(y^2)$ and $g(y) = -\cos(y^2)$, we can write ϕ_{in} in the two following ways, which are semantically equivalent:

$$\phi_{in}^1 \equiv \phi' \wedge x' - x = z \wedge \int_0^t f(y) dy = z \quad \phi_{in}^2 \equiv \phi' \wedge x' - x - t = z \wedge \int_0^t g(y) dy = z$$

¹ Note that invariants can be expressed by means of universal quantifiers (see [11]).

However, when non-polynomial functions are approximated by their Taylor polynomials, in the former case we approximate f and in the latter g .

In state *out*, the water flow rate at time y is y^2 . Hence, staying in state *out* for t units of time causes a water level decrement of $\int_0^t y^2 dy$. This is modelled by the activity function $\phi_{out} \equiv \phi' \wedge x - x' = z \wedge \int_0^t y^2 dy = z$. Obviously $\int_0^t y^2 dy = \frac{t^3}{3}$. Therefore in this case no approximation is necessary.

A *configuration* of an HA H is a pair (q, \vec{u}) , with $q \in Q$ and $\vec{u} = (u_1, \dots, u_{|X|})$ a vector in $\mathbb{R}^{|X|}$ representing that each variable x_i has value u_i . H can evolve from (q, \vec{u}) to (q', \vec{u}') , written $(q, \vec{u}) \rightarrow (q', \vec{u}')$, by an activity or transition step. An *activity step* describes the evolution from (q, \vec{u}) due to remaining in q and passing of time. In δ time units, $Act(q)$ takes H to a new evaluation of the variables:

$$\text{if } \delta \geq 0 \text{ and } [\vec{X} := \vec{u}, t := \delta, \vec{X}' := \vec{u}'] \models Act(q), \text{ then } (q, \vec{u}) \rightarrow (q, \vec{u}').$$

A *transition step* describes the evolution from (q, \vec{u}) due to a transition from q :

$$\text{if } (q, \phi, q') \in T \text{ and } [\vec{X} := \vec{u}, \vec{X}' := \vec{u}'] \models \phi, \text{ then } (q, \vec{u}) \rightarrow (q', \vec{u}').$$

A *run* is a sequence of (activity and transition) steps $(q_0, \vec{u}_0) \rightarrow \dots \rightarrow (q_i, \vec{u}_i) \dots$ with q_0 the initial state and $[\vec{X} := \vec{u}_0] \in \llbracket \phi_{init} \rrbracket$. A configuration (q, \vec{u}) is *reachable in n steps* iff there is a run $(q_0, \vec{u}_0) \rightarrow \dots \rightarrow (q_n, \vec{u}_n) \dots$ s.t. $q_n = q$ and $\vec{u}_n = \vec{u}$. A configuration is *reachable* iff it is reachable in n steps for some $n \geq 0$.

A *region* R of a HA H is a set of configurations. The region reachable by H from a region R is denoted $Post(R, H)$. Formally: $Post(R, H) = \{(q', \vec{u}') \mid \exists (q, \vec{u}) \in R, \text{ such that } (q, \vec{u}) \rightarrow (q', \vec{u}')\}$. Let $Post^n(H)$ denote either the region $\{(q_0, \vec{u}_0) \mid [\vec{X} := \vec{u}_0] \in \llbracket \phi_{init} \rrbracket\}$, if $n = 0$, or the region $Post(Post^{n-1}(H), H)$, if $n > 0$. Moreover, let $Post(H)$ denote the region $\bigcup_{n \in \mathbb{N}} Post^n(H)$. The following result is folklore.

Theorem 1. *For each $n \in \mathbb{N}$, a configuration (q, \vec{u}) is reachable in n steps iff $(q, \vec{u}) \in Post^n(H)$. Hence (q, \vec{u}) is reachable iff $(q, \vec{u}) \in Post(H)$.*

The following result follows from Tarski's results [15] and from the fact that the antiderivative of a polynomial is a polynomial.

Theorem 2. *If H is polynomial and $n \in \mathbb{N}$ then $(q, \vec{u}) \in Post^n(H)$ is decidable.*

3 Taylor Approximation

The i^{th} derivative of $f \in F$ wrt. coordinate j^{th} is denoted $D_j^i f$. Let C^k denote the set of the functions that are derivable k times, namely $f \in C^k$ iff $D_1^{j_1} \dots D_{r(f)}^{j_{r(f)}} f$ exists whenever $j_1 + \dots + j_{r(f)} = k$.

Definition 4. *Assume a function $f \in C^k$ and a vector $\vec{v} \in Dom(f)$. The polynomial of Taylor of degree k for f wrt. \vec{v} is defined by*

$$P^k(f, \vec{w}, \vec{v}) = \sum_{j_1 + \dots + j_{r(f)} \leq k} \frac{(D_1^{j_1} \dots D_{r(f)}^{j_{r(f)}} f)(\vec{v}) \cdot (w_1 - v_1)^{j_1} \dots (w_{r(f)} - v_{r(f)})^{j_{r(f)}}}{j_1! \dots j_{r(f)}!}$$

where \vec{w} is the vector of variables $(w_1, \dots, w_{r(f)})$. For $\vec{u} \in \text{Dom}(f)$, the value $r^k(f, \vec{u}, \vec{v})$ defined by $r^k(f, \vec{u}, \vec{v}) = f(\vec{u}) - P^k(f, \vec{u}, \vec{v})$ is called the remainder.

The intuition is that $P^k(f, \vec{w}, \vec{v})$ is a polynomial that approximates $f(\vec{w})$, and the error of the approximation in $\vec{u} \in \text{Dom}(f)$ is given by $r^k(f, \vec{u}, \vec{v})$. This error is quantified by the following result, known as Lagrange Remainder Theorem.

Theorem 3 (Lagrange). For a function $f \in C^{k+1}$, a convex set $S \subseteq \text{Dom}(f)$ and two vectors \vec{u}, \vec{v} in S , there exists a vector \vec{z} on the segment linking \vec{u} and \vec{v} s.t.:

$$r^k(f, \vec{u}, \vec{v}) = \sum_{j_1 + \dots + j_{r(f)} = k+1} \frac{(D_1^{j_1} \dots D_{r(f)}^{j_{r(f)}} f)(\vec{z}) \cdot (u_1 - v_1)^{j_1} \dots (u_{r(f)} - v_{r(f)})^{j_{r(f)}}}{j_1! \cdot \dots \cdot j_{r(f)}!}.$$

Our aim is to give an upper bound to $|r^k(f, \vec{u}, \vec{v})|$, under suitable hypothesis.

Definition 5. A function $f \in C^{k+1}$ is analytic in $S \subseteq \text{Dom}(f)$ if there are two constants C, L s.t., for all $j_1, \dots, j_{r(f)}$ with $j_1 + \dots + j_{r(f)} \leq k+1$ and $\vec{z} \in S$, we have

$$|(D_1^{j_1} \dots D_{r(f)}^{j_{r(f)}} f)(\vec{z})| \leq L \cdot C^{j_1 + \dots + j_{r(f)}}.$$

Then, f is analytic if f is analytic in $\text{Dom}(f)$ and $\text{Dom}(f)$ is convex.

Example 2. Trigonometric functions are analytic. For instance, for the function $\sin(x)$ it is sufficient to take the constants $L = C = 1$. Exponential and logarithmic functions are analytic in finite intervals. For instance, for function e^{2x} and interval $[0, 10]$, it is sufficient to take the constants $C = 2$ and $L = e^{20}$.

Let us assume an analytic function $f \in C^{k+1}$. Then, for \hat{C} and \hat{L} the minimal values satisfying the condition of Def. 5, for any k we denote with $C(f, k)$ the value $\hat{L} \cdot \hat{C}^{k+1}$. Moreover, let $R^k(f, \vec{w}, \vec{v})$ denote the polynomial over \vec{w} defined by

$$R^k(f, \vec{w}, \vec{v}) = \frac{C(f, k) \cdot (r(f))^{k+1} \cdot \prod_{j=1}^{r(f)} ((w_j - v_j)^2)^{\lfloor \frac{k+1}{2} \rfloor + 1}}{\lfloor \frac{k+1}{r(f)} \rfloor!}$$

By definition, $R^k(f, \vec{u}, \vec{v})$ is an upper bound to $|r^k(f, \vec{u}, \vec{v})|$ for all $\vec{u} \in \text{Dom}(f)$. Moreover, $R^k(f, \vec{u}, \vec{v})$ gets close to 0 when k tends to the infinity. Formally:

Proposition 2 ([11]). Let $f \in F$ be analytic. Then, for all $\vec{u}, \vec{v} \in \text{Dom}(f)$ we have: (1) $|r^k(f, \vec{u}, \vec{v})| \leq R^k(f, \vec{u}, \vec{v})$, and (2) $\lim_{k \rightarrow \infty} R^k(f, \vec{u}, \vec{v}) = 0$.

From $|r^k(f, \vec{u}, \vec{v})| \leq R^k(f, \vec{u}, \vec{v})$, $f(\vec{u}) = r^k(f, \vec{u}, \vec{v}) + P^k(f, \vec{u}, \vec{v})$ and monotonicity of the integral we get the following result.

Proposition 3. Let $f \in F$ be analytic and $\vec{v} \in \text{Dom}(f)$. Then for all vectors $\vec{e} = (g_1(w_1), \dots, g_n(w_n)) \oplus (g_{n+1}(c_{n+1}), \dots, g_{r(f)}(c_{r(f)}))$ with $c_{n+1}, \dots, c_{r(f)} \in \mathbb{R}$, and $r_1, \dots, r_n \in \mathbb{R}$ s.t. $g([0, r_1]) \times \dots \times g([0, r_n]) \times \{g_{n+1}(c_{n+1})\} \times \dots \times \{g_{r(f)}(c_{r(f)})\} \subseteq \text{Dom}(f)$, we have

$$\int_0^{r_1} \left(\dots \left(\int_0^{r_n} f(\vec{e}) dw_n \right) \dots \right) dw_1 \geq \int_0^{r_1} \left(\dots \left(\int_0^{r_n} (P^k(f, \vec{e}, \vec{v}) - R^k(f, \vec{e}, \vec{v})) dw_n \right) \dots \right) dw_1.$$

If we replace $f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)}))$ with $P^k(f, (g_1(w_1), \dots, g_{r(f)}(w_{r(f)}), \vec{v}) - R^k(f, (g_1(w_1), \dots, g_{r(f)}(w_{r(f)}), \vec{v}))$ in a basic formula, by Prop. 3 we get a less demanding formula, provided the operator \sim is in $\{<, \leq\}$, like in normal forms.

4 Approximation of Hybrid Automata

Approximations of HA are obtained by weakening formulae [6].

Definition 6 ([6]). An HA H' is an approximation of an HA H if H' is obtained from H by replacing each formula ϕ in H with a formula ϕ' s.t. $\llbracket \phi \rrbracket \subseteq \llbracket \phi' \rrbracket$.

We aim to give a notion of approximation for HA respecting Def. 6. We start with a notion of approximation for normal forms inspired by Prop. 3.

Definition 7. For a normal form $\phi \in \Phi(X, F)$ and $k \in \mathbb{N}$, if each $f \in F \setminus \{+, \cdot, -\}$ that appears in ϕ is derivable $k+1$ times and is analytic, then the approximation of ϕ of degree k is the set of formulae denoted $\mathbf{A}(\phi, k)$ defined inductively wrt. ϕ as follows:

1. If $\phi \equiv \int_0^{z_1} \left(\dots \left(\int_0^{z_n} f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)})) \, dw_n \right) \dots \right) dw_1 \sim ax$, then either $\mathbf{A}(\phi, k)$ is the singleton $\{\phi\}$, if f is a polynomial, or $\mathbf{A}(\phi, k)$ contains all the formulae

$$\phi_{k, \vec{v}} \equiv \int_0^{z_1} \left(\dots \left(\int_0^{z_n} \left(P^k(f, \overrightarrow{g(w)}, \vec{v}) - R^k(f, \overrightarrow{g(w)}, \vec{v}) \right) dw_n \right) \dots \right) dw_1 \sim ax$$

with $\overrightarrow{g(w)} = (g_1(w_1), \dots, g_{r(f)}(w_{r(f)}))$ and $\vec{v} \in \text{Dom}(f)$;

2. If $\phi \equiv \phi^1 \wedge \phi^2$ then $\mathbf{A}(\phi, k) = \{\phi_k^1 \wedge \phi_k^2 \mid \phi_k^1 \in \mathbf{A}(\phi^1, k) \text{ and } \phi_k^2 \in \mathbf{A}(\phi^2, k)\}$;
3. If $\phi \equiv \phi^1 \vee \phi^2$ then $\mathbf{A}(\phi, k) = \{\phi_k^1 \vee \phi_k^2 \mid \phi_k^1 \in \mathbf{A}(\phi^1, k) \text{ and } \phi_k^2 \in \mathbf{A}(\phi^2, k)\}$;
4. If $\phi \equiv \exists y. \phi'$ then $\mathbf{A}(\phi, k) = \{\exists y. \phi'_k \mid \phi'_k \in \mathbf{A}(\phi', k)\}$;
5. If $\phi \equiv \forall y. \phi'$ then $\mathbf{A}(\phi, k) = \{\forall y. \phi'_k \mid \phi'_k \in \mathbf{A}(\phi', k)\}$.

Let us prove that all formulae in $\mathbf{A}(\phi, k)$ are less demanding than ϕ .

Theorem 4. For a normal form ϕ and $k \in \mathbb{N}$ s.t. $\mathbf{A}(\phi, k)$ is defined, then $\llbracket \phi \rrbracket \subseteq \llbracket \phi' \rrbracket$ for all $\phi' \in \mathbf{A}(\phi, k)$.

Proof (sketch). By structural induction over ϕ . The proof of the base case follows from Prop. 3, the inductive steps are standard.

From the approximation of normal forms we get an approximation of HA.

Definition 8. Assume an HA H s.t. $\mathbf{A}(\phi, k)$ is defined for each formula ϕ in H . The approximation of degree k for H is the set of the PHA denoted $\mathbf{A}(H, k)$ that are obtained from H by replacing each formula ϕ in H with some formula in $\mathbf{A}(\phi, k)$.

An immediate corollary of Thm. 4 states that Def. 8 respects Def. 6.

Corollary 1. Given any HA H and $k \in \mathbb{N}$, all PHA in $\mathbf{A}(H, k)$ are approximations of H according to Def. 6.

Example 3. Let us consider the tank controller H of Ex. 1 where $\phi_{in} \equiv \phi_{in}^2$. The set $\mathbf{A}(H, 4)$ contains the automaton obtained from H by approximating function g in ϕ_{in} by choosing the real 0 as vector \vec{v} . (ϕ_{out} does not change since all functions are polynomial). Since $D_w^k(-\cos(w)) = -D_w^k \cos(w) = -\cos(w + k \cdot \frac{\pi}{2})$, it holds that $P^4(\cos, y^2, 0) \equiv -\cos(0) - \cos(\frac{\pi}{2}) \cdot (y^2)^1 - \cos(2 \cdot \frac{\pi}{2}) \cdot \frac{(y^2)^2}{2!} - \cos(3 \cdot \frac{\pi}{2}) \cdot \frac{(y^2)^3}{3!} - \cos(4 \cdot \frac{\pi}{2}) \cdot \frac{(y^2)^4}{4!} = -1 + \frac{y^4}{2} - \frac{y^8}{24}$. Moreover, $R^4(\cos, y^2, 0) = C(\cos, 4) \cdot \frac{(y^2)^{6+1}}{120}$. Now, $C(\cos, 4) = \max\{|\cos(w + 4 \cdot \frac{\pi}{2})| : w \in \text{Dom}(g)\} = 1$, therefore we have that

$$(\phi_{in}^2)_{4,0} \equiv \phi' \wedge x' - x - t = z \wedge \int_0^t -1 + \frac{y^4}{2} - \frac{y^8}{24} + \frac{y^{12} + 1}{120} dy = z.$$

The behaviour of any PHA H_k in $\mathbf{A}(H, k)$ approximates the behaviour of H , meaning that any configuration reachable by H is reachable also by H_k , in the same number of steps.

Theorem 5. *Given any HA H and $k, n \in \mathbb{N}$, if $\mathbf{A}(H, k)$ is defined, then, for all PHA $H_k \in \mathbf{A}(H, k)$ it holds that $\text{Post}^n(H) \subseteq \text{Post}^n(H_k)$.*

Proof (sketch). By Thm. 4 and the monotonicity of Post.

Thm. 5 gives us a sound method for showing that H cannot reach some *bad* configuration (q, \vec{u}) in n steps. In fact, by Thm. 2 it is computable if (q, \vec{u}) can be reached in n steps by a PHA H_k in $\mathbf{A}(H, k)$. By Thm. 5 if (q, \vec{u}) cannot be reached in n steps by H_k then it cannot be reached in n steps by H as well.

5 Analysis of the Error

In order to limit the error introduced by the approximation, Def. 6 is strengthened in [6] by the notion of ϵ -approximation, which requires that any vector satisfying a formula ϕ' of the approximation H' must be “close” to at least one vector satisfying the corresponding formula ϕ in the original HA H . We reformulate this notion in terms of a notion of neighborhood of a space in \mathbb{R}^n .

Given two vectors $\vec{u} = (u_1, \dots, u_n)$ and $\vec{v} = (v_1, \dots, v_n)$ in \mathbb{R}^n , let $d(\vec{u}, \vec{v})$ denote their *distance* $d(\vec{u}, \vec{v}) = \sqrt{(u_1 - v_1)^2 + \dots + (u_n - v_n)^2}$.

Given a vector \vec{v} and a real $\epsilon > 0$, let $\mathbf{N}(\vec{v}, \epsilon)$ denote the space of vectors $\{\vec{u} \mid d(\vec{v}, \vec{u}) \leq \epsilon\}$. Then, for a space $S \subseteq \mathbb{R}^n$ and a real $\epsilon \geq 0$, the *neighborhood of ray* ϵ of space S is the set of spaces $\mathbf{N}(S, \epsilon) = \{S' \supseteq S \mid \forall \vec{v}' \in S' \exists \vec{v} \in S \text{ s.t. } d(\vec{v}, \vec{v}') \leq \epsilon\}$.

Definition 9. *A formula $\phi' \in \Phi(F, X)$ is an ϵ -approximation of a formula $\phi \in \Phi(F, X)$ iff $\{v(X) \mid v \in \llbracket \phi' \rrbracket\} \in \mathbf{N}(\{v(X) \mid v \in \llbracket \phi \rrbracket\}, \epsilon)$.*

Definition 10 ([6]). *An HA H' is an ϵ -approximation of an HA H if H' is obtained from H by replacing each formula ϕ in H with a formula ϕ' s.t. ϕ' is an ϵ -approximation of ϕ .*

Our aim is to study the conditions over formulae in H ensuring that, for any $\epsilon > 0$, there exists some $k_0 \in \mathbb{N}$ s.t. for all $k > k_0$ we have that the set $A(H, k)$ contains only ϵ -approximations for H . In [11] we argued that formulae of the form $f(\vec{x}) \sim c$ with $\sim \in \{<, >\}$ should be avoided, since they describe open sets. In [11] we argued also that we can manage only formulae constraining variables within bounded intervals, thus avoiding variables that can tend to the infinity.

Definition 11. A normal form $\phi \in \Phi(F, X)$ is bounded iff for any variable x in ϕ we have that $\text{Dom}(x) = [l_x, r_x]$, for suitable rationals $l_x, r_x \in \mathbb{Q}$, and for each function f in ϕ we have that $\text{Dom}(f) = [l_1, r_1] \times \dots \times [l_{r(f)}, r_{r(f)}]$, for suitable rationals $l_1, r_1, \dots, l_{r(f)}, r_{r(f)} \in \mathbb{Q}$.

5.1 Syntactical Analysis of the Error

First of all let us give the intuition why for bounded normal formulae with comparison operator \leq we have that and for all $\epsilon > 0$ there exists some k_0 s.t. for all $k > k_0$, $A(\phi, k)$ contains only ϵ -approximations of ϕ .

Consider a normal form $\phi \equiv \int_0^d f(x, y) dx \leq 0$. All formulae in $A(\phi, k)$ are of the form $\int_0^d (P(f, (x, y), (c_x, c_y)) - R^k(f, (x, y), (c_x, c_y))) dx \leq 0$ for a vector $(c_x, c_y) \in \text{Dom}(f)$. Since ϕ is bounded, we can split $\text{Dom}(\int_0^d f(x, y) dx)$ (which is a function over variable y) in m closed intervals S_1, \dots, S_m of size strictly $< \epsilon$. Let $i_1, \dots, i_l \in \{1, \dots, m\}$ be the indexes s.t. no evaluation in $\llbracket \phi \rrbracket$ maps y to $S_{i_1} \cup \dots \cup S_{i_l}$, namely there is no $u \in S_{i_1} \cup \dots \cup S_{i_l}$ satisfying $\int_0^d f(x, u) dx \leq 0$. It is enough to show that no evaluation v_k in any $\llbracket \phi_k \rrbracket$ with $\phi_k \in A(\phi, k)$ maps y to $S_{i_1} \cup \dots \cup S_{i_l}$. In fact, if $v_k(y) \in S_j$ with $j \notin \{i_1, \dots, i_l\}$, by the definition of j_1, \dots, j_l there is some $v \in \llbracket \phi \rrbracket$ with $v(y) \in S_j$ and, since the size of S_j is bounded by ϵ , we infer $v_k(y) - v(y) < \epsilon$.

Hence the target is to show that there exists some k_0 s.t. for all $k > k_0$ we have that for all $u \in S_{i_1} \cup \dots \cup S_{i_l}$ the following inequality holds:

$$\int_0^d (P(f, (x, u), (c_x, c_y)) - R^k(f, (x, u), (c_x, c_y))) dx > 0. \quad (1)$$

Since $S_{i_1} \cup \dots \cup S_{i_l}$ is a closed set, $\int_0^d f(x, u) dx$ is a continuous function (which follows by the continuity of f), and the comparison symbol \leq guarantees that $\int_0^d f(x, u) dx$ is strictly positive in $S_{i_1} \cup \dots \cup S_{i_l}$, we can define $\vartheta = \min\{\int_0^d f(x, u) dx \mid u \in S_{i_1} \cup \dots \cup S_{i_l}\}$. Since $[0, d] \times S_{i_1} \cup \dots \cup S_{i_l}$ is a closed set, we can define $e_k = \max\{R^k(f, (u', u), (c_x, c_y)) \mid u' \in [0, d] \wedge u \in S_{i_1} \cup \dots \cup S_{i_l}\}$. By Prop. 2.2 we can find a k_0 s.t. for all $k > k_0$, $e_k < \vartheta/(2 \cdot d)$. Assume $k > k_0$. We show Eq. 1 by

$$\begin{aligned}
& \int_0^d (P(f, (x, u), (c_x, c_y)) - R^k(f, (x, u), (c_x, c_y))) dx \\
& \geq \int_0^d (P(f, (x, u), (c_x, c_y)) - e_k) dx \\
& = \int_0^d (P(f, (x, u), (c_x, c_y)) + r^k(f, (x, u), (c_x, c_y)) - r^k(f, (x, u), (c_x, c_y)) - e_k) dx \\
& = \int_0^d f(x, u) - r^k(f, (x, u), (c_x, c_y)) - e_k dx \geq \int_0^d f(x, u) - e_k - e_k dx \\
& > \int_0^d f(x, u) - \frac{\vartheta}{d} dx = \int_0^d f(x, u) dx - \vartheta \geq \int_0^d f(x, u) dx - \int_0^d f(x, u) dx = 0
\end{aligned}$$

with the first inequality by the definition of e_k and the monotonicity of the integral, the second by $|r^k(f, (x, u), (c_x, c_y))| \leq R^k(f, (x, u), (c_x, c_y))$ and the definition of e_k , the third by $e_k < \frac{\vartheta}{2d}$, and the last inequality by the definition of ϑ .

Theorem 6. *Given any bounded normal form $\phi \in \Phi(F, X)$ s.t. each subformula*

$$\int_0^{z_1} \left(\dots \left(\int_0^{z_n} f(g_1(w_1), \dots, g_{r(f)}(w_{r(f)})) dw_n \right) \dots \right) dw_1 \sim ax$$

in ϕ is such that \sim is \leq , then, for each $\epsilon > 0$, there exists some k_0 s.t. for each $k > k_0$, the set $\mathbf{A}(\phi, k)$ contains only ϵ -approximations for ϕ .

The result above can be immediately extended to automata.

Corollary 2. *Given any HA H s.t. each formula in H satisfies the hypothesis of Thm. 6, then, for each $\epsilon > 0$, there exists some k_0 s.t. for each $k > k_0$ the set $\mathbf{A}(H, k)$ contains only ϵ -approximations for H .*

5.2 Semantical Analysis of the Error

Our aim is to measure how close the behaviors of the PHA in $\mathbf{A}(H, k)$ and the behavior of H are.

Definition 12. *Let $\epsilon \geq 0$. The neighborhood of ray ϵ of a region R is the set of regions $N(R, \epsilon) = \{R' \supseteq R \mid \forall (q', \vec{u}') \in R'. \exists (q, \vec{u}) \in R. q = q' \text{ and } d(\vec{u}, \vec{u}') \leq \epsilon\}$.*

Under the hypothesis of Thm. 6, for all $n \in \mathbb{N}$, if k tends to the infinity, then the behavior of length at most n of each PHA $H_k \in \mathbf{A}(H, k)$ gets close to the behavior of H , in the sense that $\text{Post}^n(H_k)$ is in a neighborhood of $\text{Post}^n(H)$ of ray arbitrarily small. This comes from the fact that $\text{Post}^n(H_k)$ can be expressed by means of a formula by using existential quantifications.

Theorem 7. *Consider an HA H s.t. each formula in H satisfies the hypothesis of Thm. 6. For each $\epsilon > 0$ and $n \in \mathbb{N}$, there exists some k_0 s.t., for all $k > k_0$, we have $\text{Post}^n(H_k) \in N(\text{Post}^n(H), \epsilon)$ for all $H_k \in \mathbf{A}(H, k)$.*

6 Conclusion and Future Works

In this paper we have defined syntactical over-approximations for Hybrid Automata enriched with integrals. The approximation is based on Taylor polynomials. We have also studied their syntactical and semantical convergence w.r.t. the original specifications.

As future work we will also study under-approximations based on the same technique. The idea is to define the *under-approximation* of degree k of a formula ϕ by using the polynomial which approximates the remainder to increasing the Taylor polynomial. Moreover we can extend our work with function variables by following the theory developed in [4, 5]. Finally, our results can be used to study cyber physical attacks ([9]) by using tools like as Ariadne ([3]) based on Taylor theory.

References

1. R. Alur, C. Courcoubetis, N. Halbwachs, T. A. Henzinger, P. H. Ho, X. Nicollin, A. Olivero, J. Sifakis, S. Yovine. The Algorithmic Analysis of Hybrid Systems. *Theor. Comput. Sci.* 138(1) (1995) 3–34.
2. R. Alur, T. A. Henzinger, P. H. Ho. Automatic Symbolic Verification of Embedded Systems. *IEEE Trans. Software Eng.* 22(6) (1996) 181–201.
3. A. Balluchi, A. Casagrande, P. Collins, A. Ferrari, T. Villa, A. L. Sangiovanni-Vincentelli. Ariadne, a Framework for Reachability Analysis of Hybrid Automata. *Proc. Int. Symp. on Mathematical Theory of Networks and Systems*, 2006.
4. V. Castiglioni, R. Lanotte, S. Tini. A Function Elimination Method for Checking Satisfiability of Arithmetical Logics. *Proc. of the 23th International Workshop CS&P 2014, CEUR Workshop Proceedings 1269*, pp. 46–57 (2014).
5. V. Castiglioni, R. Lanotte, S. Tini. A Function Elimination Method for Checking Satisfiability of Arithmetical Logics. *Fundamenta Informaticae* 143: 51–71 (2016).
6. T. A. Henzinger, P. H. Ho, H. Wong-Toi. Algorithmic Analysis of Nonlinear Hybrid Systems. *IEEE Trans. Automat. Contr.* 43(4) (1998) 540–554.
7. T. A. Henzinger, P. W. Kopke, A. Puri, P. Varaiya. What’s Decidable About Hybrid Automata? *J. Comput. Syst. Sci.* 57(1) (1998) 94–124.
8. R. Lanotte. Expressive Power of Hybrid Systems with Real Variables, Integer Variables and Arrays. *J. Autom. Lang. Comb.* 12 (3): 373–405 (2007).
9. R. Lanotte, M. Merro, R. Muradore, L. Viganó. A Formal Approach to Cyber-Physical Attacks. Submitted for publication.
10. R. Lanotte, S. Tini. Taylor Approximation for Hybrid Systems. *Proc. Hybrid Systems: Computation and Control, LNCS 3114, Springer, Berlin, 1999*, pp. 402–416.
11. R. Lanotte, S. Tini. Taylor Approximation for Hybrid Systems. *Information and Computation* 205(11): 1575–1607 (2007).
12. R. Lanotte, A. Maggiolo-Schettini, S. Tini. Information flow in hybrid systems. *ACM Trans. Embedded Comput. Syst.* 3 (4): 760–799 (2004).
13. A. Pnueli and J. Sifakis (Eds.), Special Issue on Hybrid Systems, *Theor. Comput. Sci.* 138(1) (1995).
14. D. Richardson. Some Undecidable Problems Involving Elementary Functions of a Real Variable. *J. Symbolic Logic* 33 (1968), no. 4, 514–520.
15. A. Tarski. A Decision Method for Elementary Algebra and Geometry. University of California Press, Berkeley, California, 1951.