Formal Verification of Station Keeping Maneuvers for a Planar Autonomous Hybrid System

Benjamin Martin
LIX, Ecole Polytechnique, CNRS
Université Paris-Saclay, 91120 Palaiseau, France
bmartin@lix.polytechnique.fr

Khalil Ghorbal
INRIA, Rennes, France
khalil.ghorbal@inria.fr

Eric Goubault
LIX, Ecole Polytechnique, CNRS
Université Paris-Saclay, 91120 Palaiseau, France
goubault@lix.polytechnique.fr

Sylvie Putot
LIX, Ecole Polytechnique, CNRS
Université Paris-Saclay, 91120 Palaiseau, France
putot@lix.polytechnique.fr

We formally verify a hybrid control law designed to perform a *station keeping* maneuver for a planar vehicle. Such maneuver requires that the vehicle reaches a neighborhood of its station in finite time and remains in it while waiting for further instructions. We model the dynamics as well as the control law as a hybrid program and formally verify both the reachability and safety properties involved. We highlight in particular the automated generation of invariant regions which turns out to be crucial in performing such verification. We use the theorem prover `Keymaera X` to discharge some of the generated proof obligations.

1 Introduction

Formal hybrid modelling languages such as hybrid automata [1] or hybrid programs [12] offer a convenient way to describe a wide variety of hybrid systems. In this paper, we consider a piecewise continuous system where the continuous dynamics are subject to discrete switching. The plant part is modeled as a Dubins vehicle, that is a vehicle describing planar circular curves at a constant speed. The heading of the vehicle respects a hybrid control law, here taken from [8], designed for station keeping maneuver. This means that the vehicle is expected to reach a neighborhood of its station in finite time and remains in it as long as it is not asked to do differently. Possible applications for such maneuvers are e.g. an autonomous sailboat that needs to anchor around a GPS position waiting to be picked up, or an autonomous drone that needs to keep a station at a given position while waiting for further instructions.

Our goal is to formally verify the given control law while investigating to which extent such verification could be automated. In particular, we use recent symbolic computation techniques to automatically generate algebraic and semialgebraic invariant regions [11, 7]. Such regions are then exploited to formally verify the reachability and safety properties of the station keeping maneuver using the hybrid theorem prover `Keymaera X` [4]. We compare our findings with the results from [8] where another proof is conducted by means of guaranteed numerical methods.

L. Bulwahn, M. Kamali, S. Linker (Eds.): First Workshop on Formal Verification of Autonomous Vehicles (FVAV 2017).
EPTCS 257, 2017, pp. 91–104, doi:10.4204/EPTCS.257.9

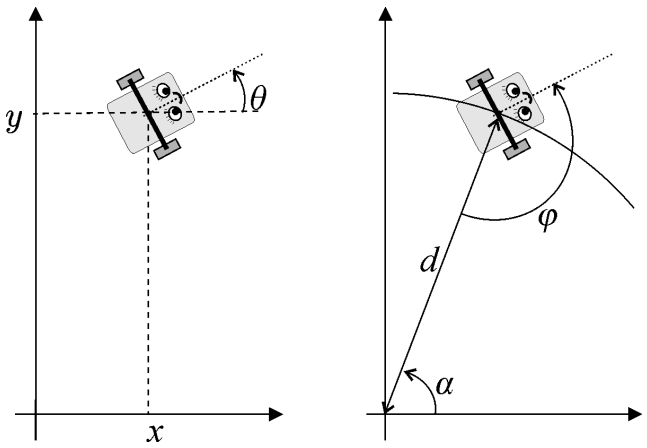

Figure 1: Cartesian coordinates (left) and polar coordinates (right). Courtesy to [8].

2 The Station Keeping Maneuver

The Dubins vehicle in Cartesian coordinates is described by the following system:

$$\begin{cases} \dot{x} &= \cos(\theta) \\ \dot{y} &= \sin(\theta) \\ \dot{\theta} &= u \end{cases}, \tag{1}$$

where (x,y,θ) defines the pose of the vehicle composed of its position in the plane (x,y) as well as its heading angle θ. The vehicle is always moving at a constant speed (here fixed to 1). The variable u encodes an input control that affects directly the heading's angular velocity. Following [8], we consider the above plant model in the polar coordinates (d,φ,α) depicted in Figure 1. The above ODE then becomes:

$$\begin{cases} \dot{d} &= -\cos(\varphi) \\ \dot{\varphi} &= \frac{\sin(\varphi)}{d} + u \\ \dot{\alpha} &= -\frac{\sin(\varphi)}{d} \end{cases}, \tag{2}$$

where radius d is a positive real satisfying $d^2 = x^2 + y^2$ and the heading angle θ is linearly related to φ and α: $\varphi - \theta + \alpha = \pi$. The angle φ can be understood as a bearing which measures the angle of the head of the vehicle with respect to a given position of the plan (here the origin).

The polar coordinates have in fact numerous advantages over the Cartesian coordinates. On one hand, one gets a decoupling of the state variables for free since the derivatives of φ and d are independent of α calling for a model reduction where only the states (d,φ) are considered. On the other hand, since φ appears only as a direct argument of the sine and cosine functions, one can restrict φ to $[0,2\pi)$ with no loss of generality: the vector field is invariant under the action of the transitive additive group that takes φ to $\varphi + 2k\pi$. Recall that the polar coordinates transformation presents a singularity at the origin $(x,y) = (0,0)$, where d vanishes. We will go back to this issue later. First, we recall the piecewise control law for u proposed in [8].

$$u = \begin{cases} 1 & \text{if } \cos(\varphi) \leq \frac{\sqrt{2}}{2} \\ -\sin(\varphi) & \text{otherwise} \end{cases}, \tag{3}$$

The intuition is that the vehicle constantly turns left when it is not pointing towards the origin (modelled by $\cos(\varphi) \leq \frac{\sqrt{2}}{2}$), otherwise the input is proportional to the bearing φ so as to push it towards 0, in which case the vehicle is moving towards the origin.

Combining the control law (3) together with the plant in polar coordinates (2), one gets a switched system where two different dynamics, implied by two different controls, can be applied depending on the state of the system.

In the following sections we prove that (i) the vehicle reaches in a finite time a position at a reasonably short distance from a beacon positioned at the center of the coordinate system, and (ii) stays in that region for an indefinite time. To do so, we first perform a qualitative analysis of the two continuous dynamics obtained exhibiting interesting invariant regions.

3 Generating Positive Algebraic Invariants

For convenience, we recall the formal definition of positive invariant sets. Let $\phi(x_0,.) : \mathbb{R} \to \mathbb{R}^n$ denote the solution of the initial value problem $\dot{x} = f(x)$ for a given ODE f and initial value x_0. Let $I \subset \mathbb{R}$ denote the maximal interval on which $\phi(x_0,.)$ is defined. Recall that I need not be the entire real line and that, depending on f, the solution may be defined on a bounded interval. When I is bounded, the system exhibits a finite time blow-up problem [2], that is in general at least one variable diverges in finite time. Such problems are intimately related to the singularities of the solutions and are often hard to detect and characterize. We will carefully discuss and analyze such issues for our case study. Notice, however, that to the best of our knowledge there are currently no automated methods to detect whether I is bounded or not where non-linear dynamics are involved (when f is linear, $I = \mathbb{R}$).

Definition 1. *A set $S \subseteq \mathbb{R}^n$ is* positive invariant *for f if and only if for any $x_0 \in S$, the corresponding solution $\phi(x_0,\cdot)$ satisfies $\phi(x_0,t) \in S$ for all $t \in [0,+\infty) \cap I$, that is for all non-negative time t as long as the solution is defined.*

In order to verify that a set S is positive invariant for an ODE, it is enough to show that the flow f is entering, constant or inner tangential on the boundaries of S. When S is semi-algebraic (that is defined by boolean combinations of polynomial equalities and inequalities), this can be done by checking the sign of the Lie derivatives (and, if necessary, higher-order Lie derivatives) of the active boundaries of S [9].

Recently in [11, 5, 9, 14, 7], many effective methods for constructing algebraic and semi-algebraic positive invariant sets have been proposed. Those methods apply, however, only to polynomial vector fields. We thus start by transforming the system (2) into a polynomial differential system. This could be done classically by adding fresh variables corresponding to the transcendental functions. The so obtained dynamics is a sound approximation (abstraction) of the original dynamics that one could refine by re-introducing the links between the extra variables and the hidden transcendental functions they represent [3]. For our dynamics, one obtains the following algebraic system :

$$\left\{\begin{array}{lcl} \dot{g} &=& -(he+u)h \\ \dot{h} &=& (he+u)g \\ \dot{e} &=& ge^2 \\ \dot{d} &=& -g \\ \dot{\varphi} &=& he+u \end{array}\right. , \tag{4}$$

where the variable g encodes cosine function $\cos(\varphi)$, h the sine function $\sin(\varphi)$ and e the inverse $\frac{1}{d}$. For the abstraction to be precise enough, one has to respect those functions initially, for instance if the initial

Control	Darboux Polynomial	Cofactor
$u = 1$	e	ge
	$1+2eh$	$2ge$
$u = -h$	e	ge
	h	$(e-1)g$

Table 1: Darboux polynomials for Eq. (4).

values of g and h are fixed, then the initial value of ϕ is entirely determined. Likewise, if d is fixed initially, then the initial value of e is fixed and is equal to the inverse of d. In this case study, the control law u, as well as the switching conditions (cf. (3)), are expressible directly with the extra variables g, h, and e. Therefore, the control law as well as the plant could be rewritten entirely algebraically.

Darboux polynomials and positive invariants Darboux polynomials, and more generally the Darboux criterion, are fundamental building blocks for the qualitative analysis of ODE and hence invariant generation. They are also at the heart of symbolic integration methods [6, 10].

Definition 2. *Let $\dot{x} = f(x)$ denote an ODE. A polynomial p is* Darboux *for f if and only if*

$$\dot{p} = cp$$

where $c \in \mathbb{R}[x]$ is a polynomial, called cofactor and where $\dot{p}$ denotes the time derivative of the polynomial p with respect to f.

The polynomial $\dot{p}$ is also known as the *Lie derivative* of p with respect to f and is formally defined as

$$\langle \nabla p, f \rangle,$$

where ∇p is the gradient of p and $\langle ., . \rangle$ is the standard inner product on $\mathbb{R}^n$.

Searching for Darboux polynomials, up to a given fixed degree, can be performed algorithmically [11, 5] by deriving all the constraints that the unknown coefficients of a polynomial (or template) have to satisfy and then solve the so obtained system.

We were able to exploit such techniques to automatically generate Darboux polynomials of Eq. (4). For a better performance, we restricted ourselves to the three dimensional dynamics in (g,h,e) since they define a closed form ODE on their own. Table 1 summarizes our findings depending on the selected control. From there, we recovered other Darboux polynomials for the five dimensional system by exploiting the algebraic invariant $de = 1$ known to be satisfied by construction.

From Table 1, in the constant control mode ($u = 1$), we observe that the cofactors of the two Darboux polynomials are the same up to a multiplication by the integer 2. This suggests the following rational invariant function for this mode:

$$V_{cst} := \frac{1+2eh}{e^2}, \tag{5}$$

obtained by first matching the two cofactors by raising the power of the Darboux polynomial e to match the multiplicative integer 2 (since the cofactor of e^2 is twice the cofactor of e) and then dividing the two Darboux polynomials with the same cofactor, namely $1+2eh$ and e^2. Such relation between Darboux polynomials and rational invariant functions is well known in the literature [6]. One can indeed easily check that the Lie derivative of V_{cst} vanishes for all t, that is $\dot{V}_{cst} = 0$, as long as the control input remains equal to 1. Moreover, since $de = 1$ by construction of e, we have a polynomial equivalent formula for V_{cst}, namely $d^2 + 2dh$.

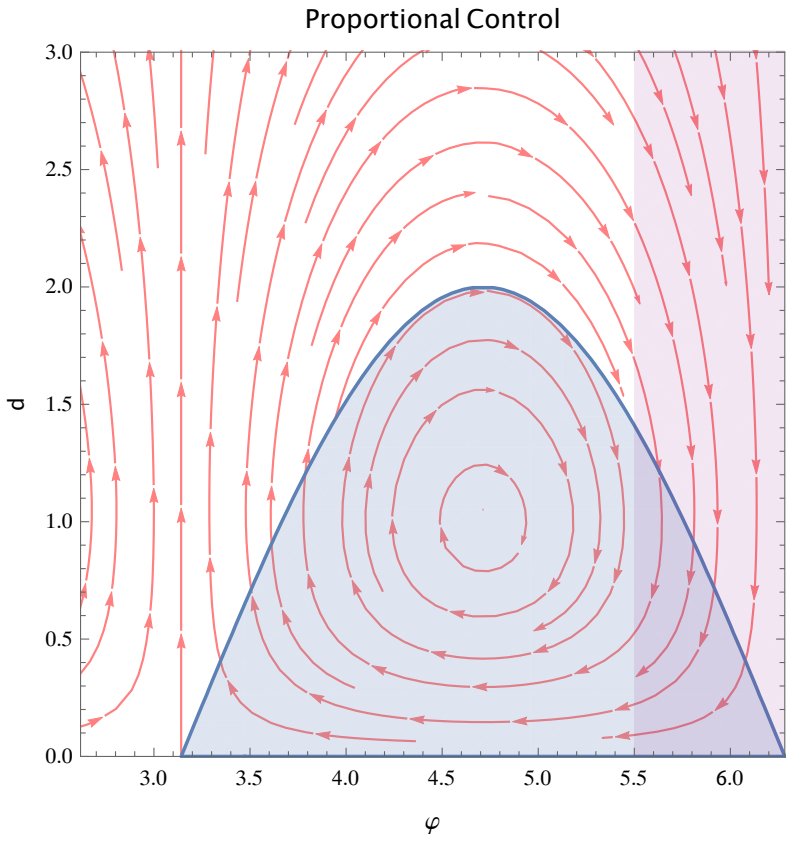

Figure 2: The set $V_{cst} \leq 0$ is depicted in blue.

For the proportional mode, when $u = -h$, it turns out that the time derivative of V_{cst} keeps a constant sign:

$$\dot{V}_{cst} = -\frac{2g(h+1)}{e} \leq 0 \ .$$

This is because $h+1 \geq 0$ (recall that h is defined as a sine function) and $g > \frac{\sqrt{2}}{2}$ by definition of g and the considered control law (3).

Figure 2 shows the region $V_{cst} \leq 0$ (depicted in blue). Observe in particular that it contains the equilibrium of the system $(d, \varphi) = (1, \frac{3\pi}{2})$ and that d is upper bounded by 2, meaning that the vehicle is at a fairly close distance to the origin of the Cartesian coordinate system. Thus, the set $V_{cst} \leq 0$ seems to be a good station keeping candidate for the switched system.

In the next sections, we formally prove that $V_{cst} \leq 0$ is an invariant set for the switched system and more importantly that it is reachable from any initial condition of the vehicle provided that initially $d > 0$ and $\varphi \neq 0$.

4 Safety Analysis

In this section, we will first model the switched system as a hybrid program, then prove the invariance of $V_{cst} \leq 0$.

Definition 3 (Hybrid Program Model)**.**

$$\alpha := \left\{ \left\{ \text{Plant}_{|u=1} \ \& \ d > 0 \wedge 2g \leq \sqrt{2} \right\} \cup \left\{ \text{Plant}_{|u=-h} \ \& \ d > 0 \wedge 2g > \sqrt{2} \right\} \right\}^*$$

where the Plant *dynamics are defined as in* (4)*.*

The hybrid program α in Definition 3 shows a piece-wise continuous system that models the behavior of the vehicle when the control law is applied. The entire feedback loop runs for any non-negative number of iterations, modeled by the star $\{\}^*$. The loop is made of an non-deterministic choice (modelled by the operator $\cup$) between the dynamics induced by the two possible controls ($\text{Plant}_{|u=1}$ and $\text{Plant}_{|u=-h}$). Any dynamics are applied as long as the states remains within the evolution domain given by the conditions after the & symbol. Here, the different conditions on g imposes to follow the control law given by (3). The condition $d > 0$, present in both evolution domains, ensures that the polar coordinates rewriting is valid.

In the sequel, we will be using $\Delta := \left(g^2 + h^2 = 1 \wedge ed = 1 \wedge d > 0\right)$ to encode the fact that the initial value of the variables (g, h, φ, d, e) is coherent, that is once g and h are fixed such that $g^2 + h^2 = 1$, $\varphi \in [0, 2\pi[$ is known and is such that $\cos(\varphi) = g$ and $\sin(\varphi) = h$, although its value is not given explicitly. The variable e is entirely determined via the equation $de = 1$ as soon as a positive d is chosen. We are now ready to formally state the positive invariance of the region $V_{cst} \leq 0$.

Theorem 1 ($V_{cst} \leq 0$ is a Positive Invariant)**.**

$$V_{cst} \leq 0 \ \wedge \ \Delta \rightarrow \ [\alpha] \, V_{cst} \leq 0$$

The box modality around α means that for all runs of the hybrid program, the following post-condition must be true. When Theorem 1 is written as a hybrid program in `Keymaera X`[1], we use the polynomial form of $V_{cst} = d^2 + 2dh$, which is a valid rewriting of the rational form if $de - 1 = 0$ is an invariant of the hybrid system. If one assumes that $de - 1 = 0$ holds initially (as this is the case here) then it is possible to prove that it holds for all time. Indeed the polynomial $de - 1$ is a Darboux polynomial for the dynamics defined in (4) for all inputs u. This fact cannot be proved currently within the theorem prover `Keymaera X` as a proof rule based on the Darboux criterion is not yet available. Therefore, it is currently necessary to add the conditions $de - 1 = 0$ into the evolution domain in order to complete the proof in `Keymaera X`. The proof itself then is mostly based on the differential invariant (DI) proof rule [12], which is essentially a conservative lifting of barrier certificates [13] to the boolean connectives. For convenience, we give in Eq. (6) the conditions required for barrier certificates, or likewise the premises of the proof rule DI for the simple case of region of the form $p \leq 0$.

$$(DI) \frac{\forall x. ([x' := f(x)] \dot{p} \leq 0)}{p \leq 0 \rightarrow [\dot{x} = f(x) \& H] p \leq 0} \tag{6}$$

In other words, the invariance of $p \leq 0$ can be deduced if its time (Lie) derivative is negative. The reason (DI) succeeds is due to the fact that when $g > \frac{\sqrt{2}}{2}$ and the proportional control is applied, one gets $\dot{V}_{cst} = -2gd(h+1)$, which is negative since g is positive, d is positive, and h, as a cosine, is greater or equal to -1. The set $V_{cst} \leq 0$ is thus not only a positive invariant when the constant control is applied but also for the proportional control when applied according to the control law Eq. (3). The output obtained with `Keymaera X` is shown on Fig. 3.

[1] Source files for this hybrid program are available at the following link `http://ben-martin.fr/publications`

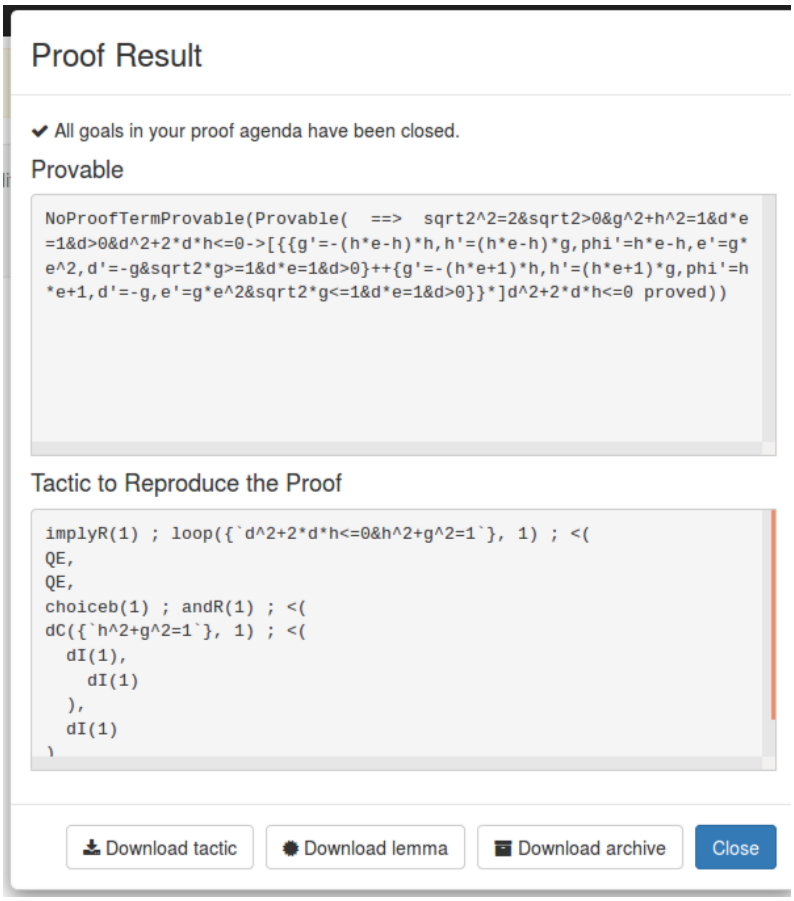

Figure 3: `Keymaera X`: last window showing the proof of Theorem 1 is concluded.

5 Reachability Analysis

We prove in this section that the region $V_{cst} \leq 0$ is reachable from (almost) anywhere in the phase space. We use the same hybrid program model given in Definition 3. We first discuss the special case $\varphi = 0$ (Section 5.1), and then the generic case where $\varphi > 0$ (Section 5.2).

5.1 The Special Case ($\varphi = 0$)

When $\varphi = 0$ initially, then g and h have to be instantiated to 1 and 0 respectively. This configuration sets the control input u to $-h$ and one can see from Table 1 that h is a Darboux polynomial and therefore $h = 0$ is an invariant equation as long as the system evolves following the Plant equations (cf. (4)) with the domain $d > 0$. In this particular case, and since $h = 0$ is invariant, the set of equations simplifies to

$$\begin{cases} \dot{g} &= 0 \\ \dot{h} &= 0 \\ \dot{e} &= e^2 \\ \dot{d} &= -1 \\ \dot{\varphi} &= 0 \end{cases}, \tag{7}$$

meaning that the bearing angle φ, as well as its cosine g, will remain constants. Subsequently, this also means that the control input u will also remain fixed to $-h$, that is 0. The vehicle shows here an interesting behavior as the time derivative for d is strictly decreasing: the vehicle is heading straight to the origin $(x,y) = (0,0)$. However, because of the evolution domain constraint $d > 0$, the dynamics will be followed as long as this constraint is not violated. But then, by design, the system is forced again to execute the differential equation which only makes d closer and closer to 0 (and thus e closer to infinity because of $ed = 1$ is an invariant). We have here in fact a finite time explosion problem: if the system starts at a distance d_0 from the origin, with $\varphi_0 = 0$, the dynamics are only defined for $t \in [0, d_0[$ and the maximal interval of definition is upper bounded. At $t = d_0$, the model hits the singularity of the polar coordinates transformation and it is no longer valid as is. A careful analysis shows that, right after the singularity, d remains an infinitesimal (and is thus continuous as one expects) and φ is discontinuous as it jumps from 0 to π switching the control input from $-h$ to 1. This discontinuity comes in fact from switching the direction of the radius vector (of magnitude d) and does affect neither the position nor the heading of the vehicle. The vehicle follows therefore a new trajectory with d very small but positive and $\varphi = \pi$. This new initial position is part of the generic case discussed in the next section.

5.2 The Generic Case ($\varphi > 0$)

The phase space is now restricted to $(0 < \varphi < 2\pi)$. We assume that in this case all trajectories are defined for all $t \geq 0$. Our reachability analysis exploits the recent invariant-based liveness proof rule, (SP), introduced by Sogokon and Jackson in [15, Proposition 10]. The idea is to use special invariant sets, so called *staging sets*, that contains the initial set and from which the system can only escape to go to the target set one wants to prove reachable. To further prove that the system will eventually leave the staging set in finite time, a progress function must be also supplied. These two ingredients are sufficient to prove that *any* trajectory starting in the staging set will eventually reach the target set in finite time. The (SP) rule is defined as follows: $X_T \subset \mathbb{R}^n$ is the target set we want to prove reachable in finite time and $X_0 \subset \mathbb{R}^n$ is the initial set. The premises of the proof rule are sufficient to prove its conclusion under the assumption that the solution is defined for as long as needed to reach the target set.

$$\vdash \exists \varepsilon > 0.\ \forall x.\ S \rightarrow (p \geq 0 \wedge \dot{p} \leq -\varepsilon)$$

$$(\mathrm{SP})\frac{X_0 \wedge \neg X_T \vdash S \qquad \vdash S \rightarrow [\dot{x} = f(x)\ \&\ \neg(H \wedge X_T)]S \qquad X_0 \vee S \vdash H}{\vdash X_0 \rightarrow \langle \dot{x} = f(x)\ \&\ H \rangle X_T} \tag{8}$$

The diamond modality around the hybrid program means that at least one run leads to satisfying the post-conditions (liveliness). The proof rule has four premises and relies essentially on a strictly decreasing progress function p within an invariant set S. The progress of p is proven using the positive real number ε. The intuition of the proof rule (SP) is that if there exists a bounded from below and decreasing function along the trajectories in an invariant region with respect to certain evolution domain constraint, then the flow cannot stay indefinitely inside and must eventually exit the region described by those constraints.

The aim of this section is to prove that the region defined by $V_{cst} \leq 0$ is reachable from any generic initial position (i.e. excluding the case $\varphi = 0$). Doing so with a single application of the (SP) rule can be cumbersome. Therefore, we suggest to build a chain of staging sets that prove reachability of intermediate regions, the chain leading to the reachability of $V_{cst} \leq 0$. In what follows, we partition the

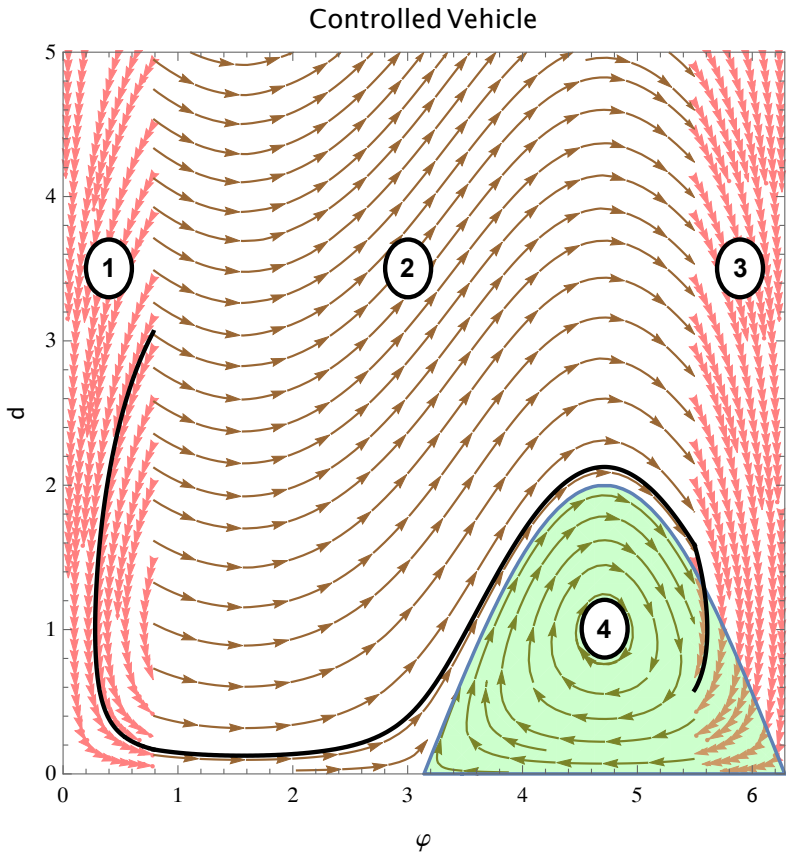

Figure 4: Phase space of the controlled Dubins vehicle in polar coordinates. The black curve shows a possible trajectory of the vehicle until reaching $V_{cst} \leq 0$ (region ④). When the vehicle is in regions ① or ③, $u = 1$; In region ② $u = -h$.

phase space into 4 regions as shown in Fig. 4:

$$\begin{aligned} ① &:= 0 < \varphi < \frac{\pi}{4} \wedge d > 0 \\ ② &:= \frac{\pi}{4} \leq \varphi \leq \frac{7\pi}{4} \wedge d > 0 \wedge V_{cst} > 0 \\ ③ &:= \frac{7\pi}{4} < \varphi < 2\pi \wedge d > 0 \wedge V_{cst} > 0 \\ ④ &:= V_{cst} \leq 0 \end{aligned} \tag{9}$$

This partition has been found manually by separating the regions with different control and the region found safe in the previous section. We prove sequentially that ② can be reached from ①, ③ can be reached from ② and that ④ can be reached from ③.

Lemma 1.

$$\left(\frac{\sqrt{2}}{2} < g \wedge h > 0 \wedge \Delta\right) \rightarrow \langle\, \text{Plant}_{|u=-h} \,\rangle \left(g = \frac{\sqrt{2}}{2} \wedge h = \frac{\sqrt{2}}{2}\right)$$

Proof. Apply the proof rule (SP) with the progress function $p := d$, $\varepsilon := \frac{\sqrt{2}}{2}$ and the invariant set $S := X_0$.

All premises can be checked automatically. The first, second and fourth premises can be discharged using a quantifier elimination procedure over the reals (CAD for instance). The most involved premise is the third one where one has to prove the invariance of the staging set S. As shown in [9] this can be also reduced to a universal quantifier elimination problem and can thus be discharged using CAD. □

Prop. 1 exhibits a particular run of the hybrid program α in which it reaches the bearing angle $\varphi = \frac{\pi}{4}$ from any coherent initial position satisfying $\varphi \in (0, \frac{\pi}{4})$.

Proposition 1.

$$\left(0 < \varphi < \frac{\pi}{4} \wedge \Delta\right) \rightarrow \langle\alpha\rangle \left(\varphi = \frac{\pi}{4}\right)$$

Proof. When $\varphi \in (0, \frac{\pi}{4})$ initially then $g > \frac{\sqrt{2}}{2}$ and the control input u is set to $-h$. According to Lem. 1, and since $d > 0$ is a positive invariant, $\varphi = \frac{\pi}{4}$ is reachable by continuously following the dynamics while fixing u to $-h$. This is allowed by the control law since, as long as the system evolves within the staging set, $g > \frac{\sqrt{2}}{2}$ is satisfied. □

Similarly, we prove that region ③ is reachable from ② by exhibiting a run of α that reaches the region ②.

Lemma 2.

$$\left(g \leq \frac{\sqrt{2}}{2} \wedge V_{cst} > 0 \wedge \Delta\right) \rightarrow \langle\, \text{Plant}_{|u=1} \,\rangle \left(\frac{\sqrt{2}}{2} < g \wedge h < 0\right)$$

Proof. Apply the proof rule (SP) with the progress function $p := -\varphi + \frac{7\pi}{4}$, $\varepsilon := \frac{1}{2}$ and the invariant set $S := X_0$. To prove that $S \rightarrow p \geq 0$, we need to use the fact that $(g,h) = (\cos(\varphi), \sin(\varphi))$. The proof is not involved but has to be done manually as it requires manipulating transcendental functions. □

Proposition 2.

$$\left(\frac{\pi}{4} \leq \varphi \leq \frac{7\pi}{4} \wedge V_{cst} > 0 \wedge \Delta\right) \rightarrow \langle\alpha\rangle \left(\frac{7\pi}{4} < \varphi < 2\pi\right)$$

Finally, we prove the reachability of the invariant set ④ from ③.

Lemma 3.

$$\left(\frac{\sqrt{2}}{2} < g \wedge h < 0 \wedge V_{cst} > 0 \wedge \Delta\right) \rightarrow \langle\, \text{Plant}_{|u=-h} \,\rangle \ (V_{cst} \leq 0)$$

Proof. Apply the proof rule (SP) with the progress function $p := d$, $\varepsilon := \frac{\sqrt{2}}{2}$ and the invariant set $S := X_0$. □

Proposition 3.

$$\left(\frac{7\pi}{4} < \varphi < 2\pi \wedge V_{cst} > 0 \wedge \Delta\right) \rightarrow \langle\alpha\rangle \ (V_{cst} \leq 0)$$

Combined together, Propositions 1, 2 and 3 give:

Theorem 2 (Reachability of V_{cst})**.** *The region $V_{cst} \leq 0$ is reachable from any coherent generic position:*

$$(0 < \varphi < 2\pi \wedge \Delta) \rightarrow \langle\alpha\rangle \ (V_{cst} \leq 0)$$

The proof rule (SP) is not yet available in `Keymaera X`. All the premises can be however discharged using a quantifier elimination procedure, including the invariance of the staging set S. The latter is a direct consequence of the decidability of the invariance of semialgebraic sets [9]. We used our proper implementation of the procedure described in [9] together with the `Reduce` procedure in Wolfram Mathematica to discharge all the premises of the (SP) proof rule, except for Lem. 2 where transcendental functions are involved.

Remark 1. *Theorem 2 tells nothing about the time required to reach the target set, only that this time is finite. The rule (SP) embeds, however, a bounded progress function which can be used to determine an upper bound on the time required to reach the target region. For example starting at (φ_0, d_0) in region ① ($d > 0$), and given the upper bound on the decrease of the progress function d, we can conclude that the region ② is entered in at most at $\sqrt{2}d_0$ seconds. The quality of this upper bound depends on the positive bound used for the progress function and can thus be arbitrarily large (but always finite). Notice that a finer analysis for the time spent at a given traversed region would benefit from a lower bound of the progress function, although such bound is not required for proving the reachability itself.*

Remark 2. *There exists in fact an attractor for the switched system that is inside $V_{cst} \leq 0$, namely $V_{cst} \leq -\frac{1}{2}$. However, proving its reachability is much more involved than proving the reachability of $V_{cst} \leq 0$, because a trajectory may loop for some time before reaching it. It also features a sliding mode at the boundary $\varphi = \frac{7\pi}{4}$ for d in $\left[\frac{1}{\sqrt{2}}, 1\right]$. In fact the only entry to the attractor is the point $\left(\frac{7\pi}{4}, \frac{1}{\sqrt{2}}\right)$ which is reached whenever the system enters its sliding mode. Notice also that when the system loops around this attractor, one has to consider roots of the Lambert W function. We do not carry such proof in `Keymaera X`, however, since $V_{cst} \leq 0$ is sufficient to prove that the maneuver reaches a region close to the origin. (see Fig. 5).*

6 Related Work

An interval-based numerical approach has been proposed in [8] to validate the controller. The idea is to construct a discrete abstraction of the state space of the system and then to build a graph where nodes correspond to discrete regions and transitions between nodes mean that the system can potentially move from one region to the other. To actually build such transitions, the author considered guaranteed numerical tests based on interval analysis. This final graph was then used to show that the vehicle will be eventually trapped in a limited region of the state space. By construction, such a region is an over-approximation of the actual attractor of the system. This region is not precisely given in [8] but it can be approximated by the following set:

$$
\begin{aligned}
&(0 \leq d \leq 2) \vee \\
&\left(\left(0 \leq \varphi \leq \frac{\pi}{6} \vee \frac{7\pi}{4} \leq \varphi \leq 2\pi\right) \wedge 0 \leq d \leq 7.6\right) \vee \\
&\left(\frac{\pi}{2} \leq \varphi \leq \frac{7\pi}{4} \wedge 2 \leq d \leq 7.6 \wedge \frac{112}{\pi}\varphi - 25d - 6 \geq 0\right).
\end{aligned}
$$

In comparison, the zero level set of V_{cst} describes more accurately the behavior of the system, entailing a sharper analysis. Fig. 6 depicts the attractor from [8] in comparison to the ones found in this paper. It is worth noting that the method from [8] does not require the algebraic rewriting. It works directly with the original system in polar coordinates and uses guaranteed numerical computations to find the invariant set. As in our approach, however, finding invariant candidates was essentially manual.

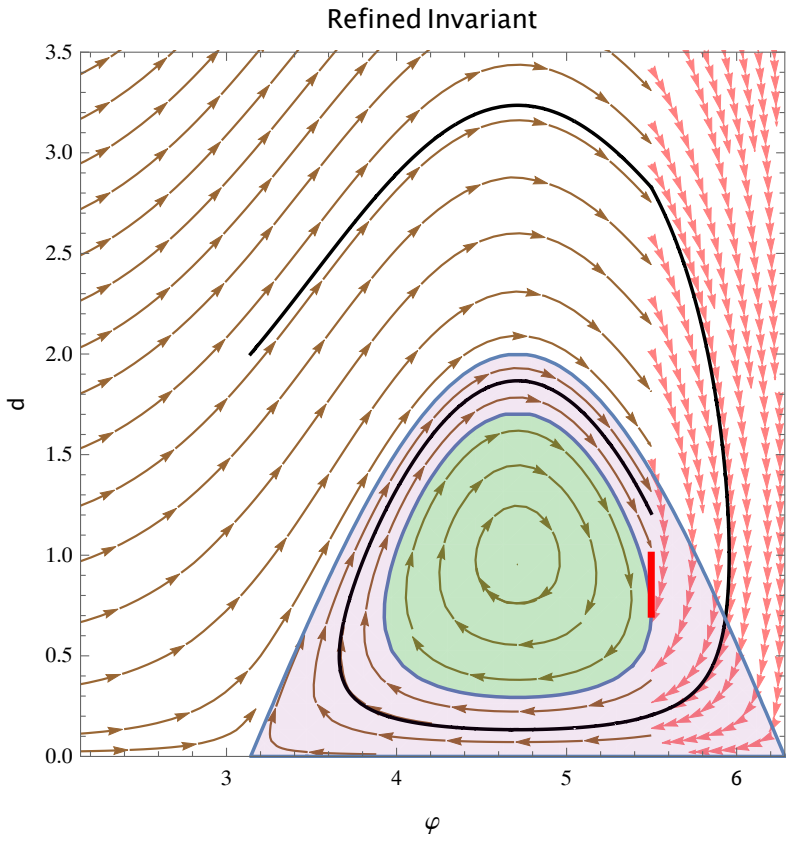

Figure 5: The refined invariant of the controlled vehicle. The red segment shows the sliding mode region before entering the attractor. The black curve shows a possible looping trajectory around the attractor. The region $V_{cst} \leq 0$ is shown in purple for convenience.

7 Discussion and Conclusion

We have developed a formal proof for the safety and liveness of an autonomous switched system, corresponding to a planar Dubins car whose goal is to perform a station keeping maneuver around the origin. This proof can be synthesized in three steps. First, we have used recent algebraic methods to derive algebraic invariant properties for the switched system. In particular, an algebraic invariant region, corresponding to a station keeping behavior, has been identified and formally checked with the hybrid system theorem prover `Keymaera X`. Second, the unbounded time reachability analysis of the system has been performed. To do so, the phase space was partitioned into subregions, each equipped with a progress function to prove that the system eventually leaves the region. To complete the proof, we exploited the invariants we generated to show that a region is only exited at certain locations.

Although this case study appears simple, this formal proof relies on non-trivial elements in particular for the reachability analysis. We believe, however, that the current state-of-the-art techniques and tools are mature enough to handle such the proof obligations for such case study (up to properly implementing the used proof rules in a theorem prover like `Keymaera X`). Some questions remain about how far can the application of these tools be automatized. For example, we were required to decompose the state space for the proof of reachability. The decomposition is obtained by hand from the autonomous system

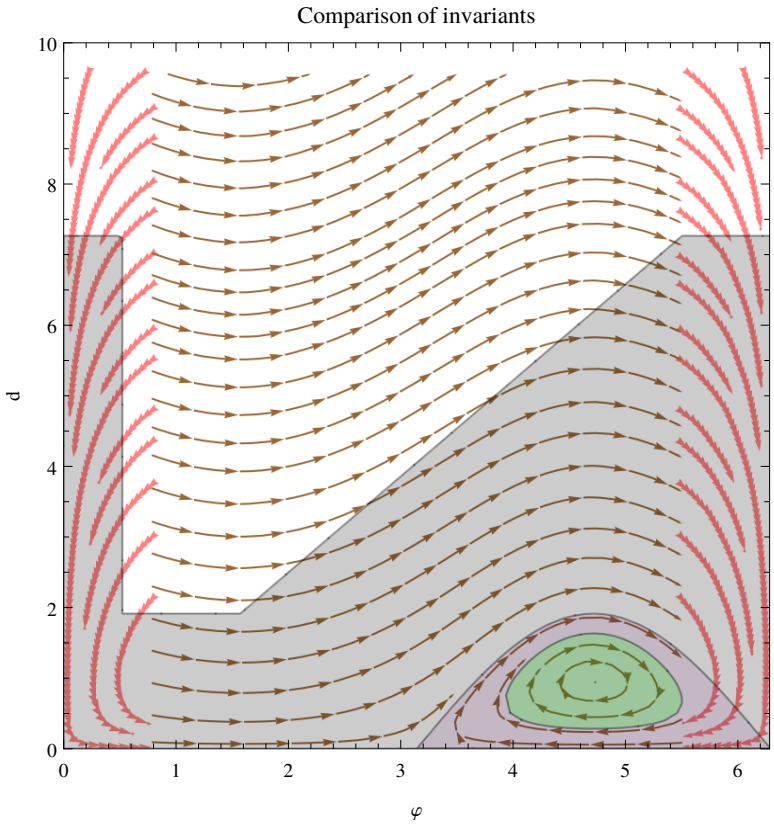

Figure 6: In gray the invariant region obtained in [8], in purple $V_{cst} \leq 0$ and in green $V_{cst} \leq -\frac{1}{2}$.

itself and the invariant found for the safety proof. We can think that for other applications, such a decomposition could be partially inferred automatically in a similar manner.

Another challenging future work avenue we are keen to investigate is the interaction between the formal proof process and the design of the control law. For instance, how can one derive a feedback for the designer when the proof fails ? This becomes more intricate when the proof strategy proceeds by sufficient conditions not by equivalences, as for reachability. A promising direction would be to direct the design so as to ease the generation of suitable differential variants/invariants sets (e.g. using Darboux polynomials), like for control Lyapunov functions.

Acknowlegments

This work is partially supported by the DGA / MRIS under the project “Sureté de Fonctionnement des Systèmes Robotiques Complexes” and by the Academic and Research Chair “Engineering of Complex Industrial Systems”, sponsored by Thales, Naval Group, Dassault Aviation and DGA.

References

[1] Rajeev Alur, Costas Courcoubetis, Thomas A Henzinger & Pei-Hsin Ho (1993): *Hybrid automata: An algorithmic approach to the specification and verification of hybrid systems*. In: *Hybrid systems*, Springer,

pp. 209–229, doi:10.1007/3-540-57318-6_30.

[2] John M. Ball (1978): *Finite Time Blow-Up in Nonlinear Problems*. In Michael G. Crandall, editor: *Nonlinear Evolution Equations*, Academic Press, pp. 189 – 205, doi:10.1016/B978-0-12-195250-1.50015-1.

[3] David C. Carothers, G. Edgar Parker, James S. Sochacki & Paul G. Warne (2005): *Some properties of solutions to polynomial systems of differential equations*. Electronic Journal of Differential Equations 2005(40), pp. 1–17.

[4] Nathan Fulton, Stefan Mitsch, Jan-David Quesel, Marcus Völp & André Platzer (2015): *KeYmaera X: An Axiomatic Tactical Theorem Prover for Hybrid Systems*. In Amy P. Felty & Aart Middeldorp, editors: *CADE*, *LNCS* 9195, Springer, pp. 527–538, doi:10.1007/978-3-319-21401-6_36.

[5] Khalil Ghorbal & André Platzer (2014): *Characterizing Algebraic Invariants by Differential Radical Invariants*. In: *TACAS*, Springer, pp. 279–294, doi:10.1007/978-3-642-54862-8_19.

[6] Alain Goriely (2001): *Integrability and Nonintegrability of Dynamical Systems*. Advanced Series in Nonlinear Dynamics, World Scientific, doi:10.1142/3846.

[7] Eric Goubault, Jacques-Henri Jourdan, Sylvie Putot & Sriram Sankaranarayanan (2014): *Finding non-polynomial positive invariants and Lyapunov functions for polynomial systems through Darboux polynomials*. In: *2014 American Control Conference*, pp. 3571–3578, doi:10.1109/ACC.2014.6859330.

[8] Luc Jaulin (2013): *Outer Approximation of Attractors Using an Interval Quantization*. Reliable Computing 19, pp. 261–273.

[9] Jiang Liu, Naijun Zhan & Hengjun Zhao (2011): *Computing semi-algebraic invariants for polynomial dynamical systems*. In: *EMSOFT*, ACM, pp. 97–106, doi:10.1145/2038642.2038659.

[10] Yiu-Kwong Man (1993): *Computing Closed Form Solutions of First Order ODEs Using the Prelle-Singer Procedure*. J. Symb. Comput. 16(5), pp. 423–443, doi:10.1006/jsco.1993.1057.

[11] Nadir Matringe, Arnaldo Vieira Moura & Rachid Rebiha (2010): *Generating Invariants for Non-linear Hybrid Systems by Linear Algebraic Methods*. In: *SAS*, *LNCS* 6337, Springer, pp. 373–389, doi:10.1007/978-3-642-15769-1_23.

[12] André Platzer (2008): *Differential Dynamic Logic for Hybrid Systems*. J. Autom. Reasoning 41(2), pp. 143–189, doi:10.1007/s10817-008-9103-8.

[13] Stephen Prajna & Ali Jadbabaie (2004): *Safety Verification using Barrier Certificates*. In: *HSCC*, *LNCS* 2993, Springer, pp. 477–492, doi:10.1007/978-3-540-24743-2_32.

[14] Sriram Sankaranarayanan (2010): *Automatic invariant generation for hybrid systems using ideal fixed points*. In: *HSCC*, pp. 221–230, doi:10.1145/1755952.1755984.

[15] Andrew Sogokon & Paul B. Jackson (2015): *Direct Formal Verification of Liveness Properties in Continuous and Hybrid Dynamical Systems*. In: *FM*, *LNCS* 9109, Springer, pp. 514–531, doi:10.1007/978-3-319-19249-9_32.