From Zonotopes to Proof Certificates

A Formal Pipeline for Safe Control Envelopes

Jonathan Hellwig, Lukas Schäfer, Long Qian, Matthias Althoff and André Platzer

iFM 2025

November 28, 2025

Introduction

Cyber-Physical Systems (CPS)

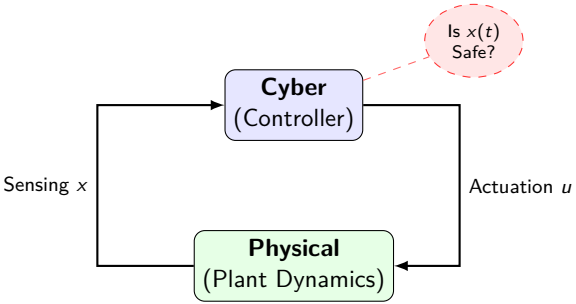
Systems where software meets physics:

- **Examples:** Autonomous vehicles, surgical robots, power grids.
- **Challenge:** Logic interacts with continuous physical dynamics.

The Safety Critical Problem

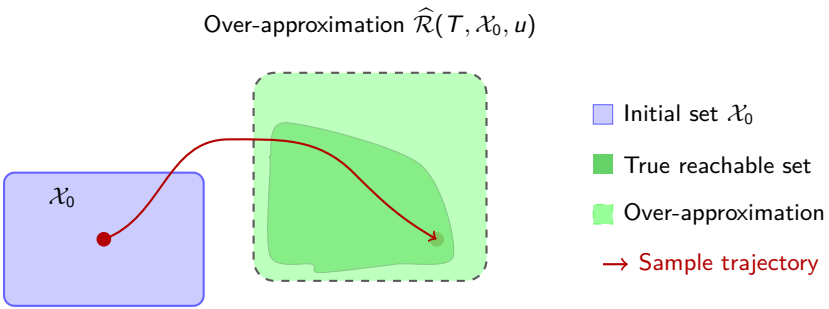
How do we guarantee safety when the physical environment is continuous?

Solution: Reachability Analysis.



Continuous Interaction

Reachability Analysis



Problem Statement

For the time-triggered control system:

$$x'(t) = f(x(t), u_{\lfloor t/\Delta t \rfloor}), \quad x_0 \in \mathcal{X}_0, \quad u_k \in \mathcal{U}$$

Do we have $\forall T \geq 0 : \mathcal{R}([0, T], \mathcal{X}_0, u) \subseteq \text{Safe}$?

Robust control invariant

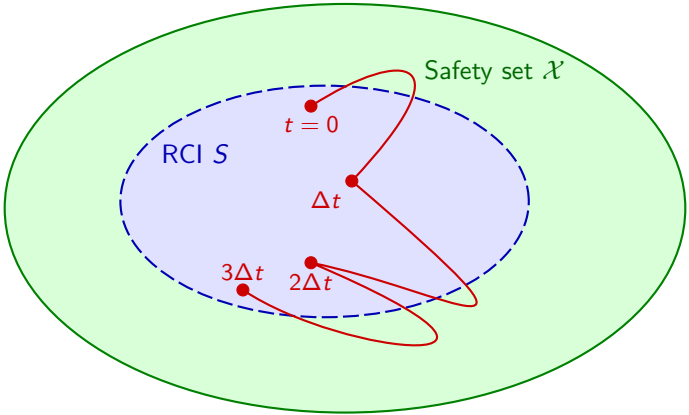
Challenge: The Time Horizon

- Reachability analysis computes tight over-approximations for **finite time** $T \geq 0$:

$$\widehat{\mathcal{R}}([0, T], \mathcal{X}_0, u)$$

- As $T \to \infty$, these sets generally **diverge** or become too conservative.
- **Question:** How can we guarantee safety for **unbounded** time?
- **Answer:** We need an inductive argument.

Robust control invariants



- Safety set $\mathcal{X}$ is never violated.
- At $k\Delta t$ there exists a new control input to keep the system in S.

Robust control invariant sets

Control Envelope $\mathcal{E}$

A **control envelope** $\mathcal{E} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ generalizes a controller to a set of **admissible inputs**. At any state x, any input $u \in \mathcal{E}(x)$ is valid.

Example: An interval around a nominal controller $-Kx$:

$$\mathcal{E}(x) = \{u \in \mathbb{R} \mid -Kx - \delta \leq u \leq -Kx + \delta\}$$

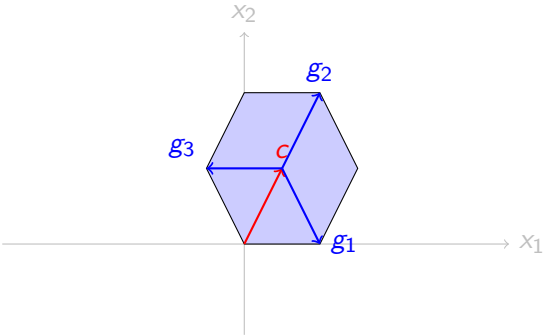
Robust Control Invariant (RCI) Set

A set $S \subseteq \mathbb{R}^n$ is an RCI set if there exists an envelope $\mathcal{E}$ such that:

1. **One-step invariance:** $\mathcal{R}(\Delta t, S, \mathcal{E}) \subseteq S$
2. **One-step safety:** $\mathcal{R}([0, \Delta t], S, \mathcal{E}) \subseteq \mathcal{X}$

How do we find such sets S and envelopes $\mathcal{E}$ for real systems?

Zonotopes



$$Z = \{c + \sum_{i=1}^{m} g_i \lambda_i \mid \lambda \in [-1, 1]^m\}$$

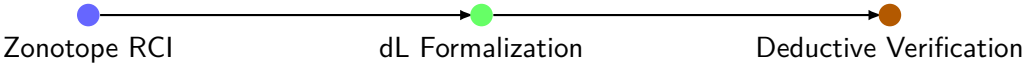
Zonotopes are

- Convex symmetric, bounded polytope.
- Closed under linear transformations and set addition.
- Used to efficiently compute

$$\mathcal{R}([0, \Delta t], S, \mathcal{E}), \quad \mathcal{R}(\Delta t, S, \mathcal{E})$$

  in a tool like CORA.
- Can be used to find RCI sets and control envelopes via convex optimization.

Contribution



Goal: CORA Tool outputs → independent proof check with Differential Dynamic Logic (dL).
Technical challenges:

- Complex set representations of zonotopes in KeYmaera X.
- We cannot reuse the same method for both one-step safety and invariance.

System model in dL

Hybrid program for time-triggered control system:

$$\begin{aligned}\text{init} &::= u := *; ?E(x,u)\\ \text{ctrl} &::= \mathit{if}(t = \Delta t)\{u := *; ?E(x,u); t := 0\}\\ \text{plant} &::= x' = f(x), t' = 1 \& t \leq \Delta t\end{aligned}$$

Safety Specification

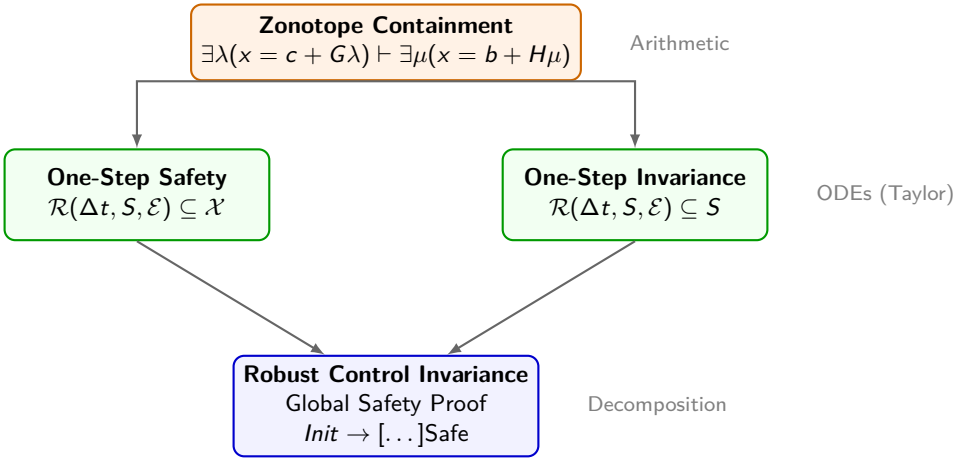
Goal: Prove the safety of the time-triggered control system:

$$X_0 \wedge t = 0 \rightarrow [\text{init}; (\text{ctrl}; \text{plant})^*]\,\text{Safe}$$

Hoare tripel:

$$\{X_0 \wedge t = 0\}\ \text{init}; (\text{ctrl}; \text{plant})^*\ \{\text{Safe}\}$$

Proof Architecture



Formalizing robust control invariant sets in dL

Robust control invariant set

Recall the properties:

- **One-step safety**: $\mathcal{R}([0, \Delta t], S, \mathcal{E}) \subseteq \mathcal{X}$,
- **One-step invariance**: $\mathcal{R}(\Delta t, S, \mathcal{E}) \subseteq S$.

Theorem (Robust control invariance)

$$\frac{\begin{array}{l} E(x,u), t = 0 \vdash [\mathsf{plant}]X(x) \\ E(x,u), t = 0 \vdash [\mathsf{plant}](t = \Delta t \rightarrow \exists u E(x,u)) \end{array}}{X_0(x), t = 0 \vdash [\mathsf{init}; (\mathsf{ctrl}; \mathsf{plant})^*]X(x)}$$

Taylor models

Idea: Given the ODE

$$x(t)' = f(x(t)), x_0 \in X_0, t \in [0, \Delta t],$$

Approximate the solution with provable error bounds:

$$x(t) \in p(t) + I(t), \quad t \in [0, \Delta t].$$

Example

$$p_{x_1}(t, \lambda) = \lambda_1 + t\lambda_2 + \frac{t^2}{2}\lambda_3, \quad p_{x_2}(t, \lambda) = \lambda_2 + t\lambda_3, \quad p_u(t, \lambda) = \lambda_3.$$

Interval error bounds:

$$\underline{I}_{x_1}(t) = -101020 \cdot 10^{-11} t,\ \overline{I}_{x_1}(t) = 101020 \cdot 10^{-11} t,$$
$$\underline{I}_{x_2}(t) = -10^{-6} t,\ \overline{I}_{x_2}(t) = 10^{-6} t,$$
$$\underline{I}_{u}(t) = -10^{-6} t,\ \overline{I}_{u}(t) = 10^{-6} t.$$

Provable Taylor models in dL

Theorem

$$\frac{\begin{array}{l}\exists\lambda\exists t(\mathit{TM}_{p,\mathbf{I}}(x,\lambda,t)\wedge 0\le t\le \Delta t\wedge \|\lambda\|_\infty\le 1)\ \vdash P(x)\\ X_0(x)\vdash \exists\lambda\,(\mathit{TM}_{p,\mathbf{I}}(x,\lambda,0)\wedge \|\lambda\|_\infty\le 1)\\ \mathit{TM}_{p,\mathbf{I}}(x,\lambda,t),\ 0\le t\le \Delta t,\ \|\lambda\|_\infty\le 1\vdash \partial_t \mathit{TM}_{p,\mathbf{I}}(x,\lambda,t)\end{array}}{X_0(x),\ t=0\vdash [\mathsf{plant}]P(x)}$$

Key idea: Prove that the Taylor model over-approximates the ODE solution provably in dL.

One step invariance

One-step invariance: $\mathcal{R}(\Delta t, S, \mathcal{E}) \subseteq S$.

Theorem (One-step invariance)

$$\frac{\begin{array}{r}\langle b, H\rangle(x,u) \vdash \langle c_x, G_x\rangle(x,u) \\ \langle c, G\rangle(x,u) \vdash \exists\lambda\,(TM_{p,\mathbf{I}}(x,u,\lambda,0) \wedge \|\lambda\|_\infty \le 1) \\ TM_{p,\mathbf{I}}(x,u,\lambda,t),\ 0 \le t \le \Delta t,\ \|\lambda\|_\infty \le 1 \vdash \partial_t TM_{p,\mathbf{I}}(x,u,\lambda,t)\end{array}}{\langle c, G\rangle(x,u),\ t = 0 \vdash [\text{plant}]\,(t = \Delta t \to \exists u\,\langle c, G\rangle(x,u))}$$

Guarantees that at time Δt, there exists an admissible control to stay in S.

One-step safety in dL

One-step safety: $\mathcal{R}([0, \Delta t], S, \mathcal{E}) \subseteq \mathcal{X}$

Theorem (One-step safety)

$$\frac{\begin{array}{rl} \langle b, H\rangle(x) \vdash & X(x) \\ \langle c, G\rangle(x, u) \vdash & \exists \lambda \, (TM_{p,\mathbf{I}}(x, u, \lambda, 0) \wedge \|\lambda\|_\infty \leq 1) \\ TM_{p,\mathbf{I}}(x, u, \lambda, t),\ 0 \leq t \leq \Delta t,\ \|\lambda\|_\infty \leq 1 \vdash & \partial_t TM_{p,\mathbf{I}}(x, u, \lambda, t) \end{array}}{\langle c, G\rangle(x, u),\ t = 0 \vdash [\text{plant}]\, X(x)}$$

Guarantees $X(x)$ holds throughout $[0, \Delta t]$ under the plant dynamics.

Zonotope Containment as a Formula

The Verification Goal: To prove $\mathcal{Z}_1 \subseteq \mathcal{Z}_2$, we must verify that every x in $\mathcal{Z}_1$ exists in $\mathcal{Z}_2$:

$$\forall x \left(\underbrace{\exists \lambda \, (x = c + G\lambda \wedge \dots)}_{x \in \mathcal{Z}_1} \longrightarrow \underbrace{\exists \mu \, (x = b + H\mu \wedge \dots)}_{x \in \mathcal{Z}_2} \right)$$

This formula can be rewritten in prenex normal form as:

$$\forall x \, \forall \lambda \, \exists \mu \, (\dots)$$

Why is this hard?

General-purpose Quantifier Elimination (QE) is **prohibitive**:

- **Structure:** The $\forall x \forall \lambda \exists \mu$ alternation is computationally expensive.
- **Complexity:** Doubly exponential in the number of variables.

Efficient Zonotope Containment

Theorem (Witness-Based Containment)

Arithmetic Check

$$\vdash HH^+ = I$$
$$\vdash b - c = H\beta$$
$$\vdash \|H\Gamma - G\|_\infty \leq \varepsilon$$
$$\vdash \|(\Gamma, \beta)\|_\infty \leq 1 - \varepsilon\|H^+\|_\infty$$

$$\frac{}{\exists\lambda(x = c + G\lambda) \vdash \exists\mu(x = b + H\mu)}\ \text{(Witness Rule)}$$

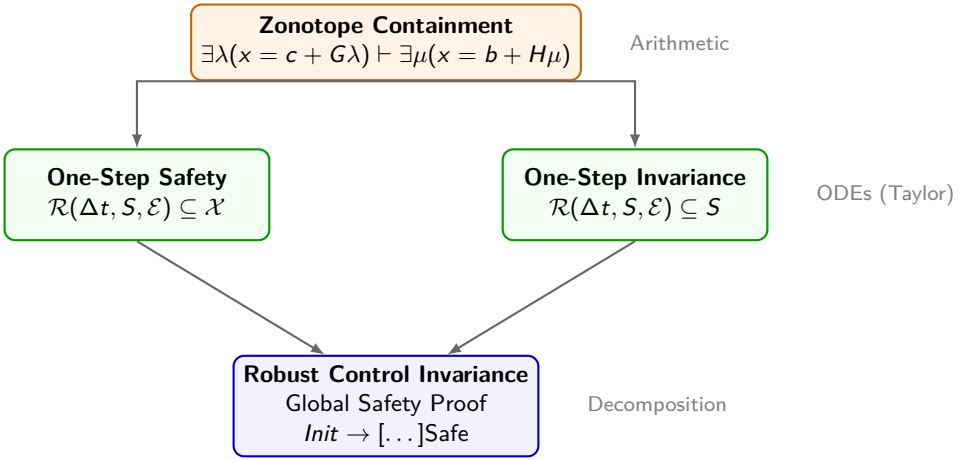
1. Witness Generation (Untrusted)

- **Goal:** Find witnesses Γ (matrix) and β (vector).
- **Method:** Solved via external Linear Programming / Convex Optimization.

2. Formal Verification (KeYmaera X)

The prover checks the purely rational arithmetic obligations.

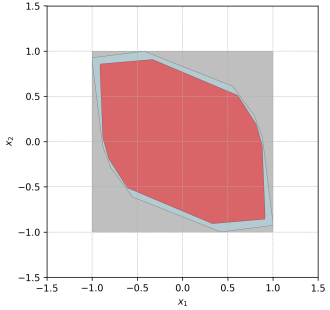
Proof Architecture



Experimental Results

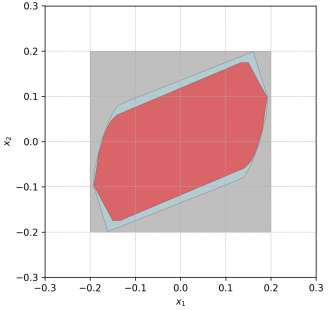
Safety Set $\mathcal{X}$ | Computed RCI S | Reachable Set (Taylor)

Double Integrator



$$x_1' = x_2 + w_1$$
$$x_2' = \frac{1}{m}u + w_2$$

Jet Engine



$$x_1' = -x_2 - \frac{3}{2}x_1^2 - \frac{1}{2}x_1^3 + w$$
$$x_2' = u$$

Conclusion & Outlook

Summary of Contributions

- **Formalization:** Defined zonotope-based reachability analysis within Differential Dynamic Logic (dL).
- **Theory:** Established sufficient conditions to verify Robust Control Invariant (RCI) sets using external certificates.
- **Validation:** Successfully verified nonlinear case studies (Jet Engine, Double Integrator).

Limitations & Future Work

- **Scalability:** Currently a proof-of-concept; need to optimize for higher-dimensional systems.
- **Automation:** Eliminate manual artifact transfer by creating an automated pipeline from CORA to KeYmaera X and vice versa.

Contact: jonathan.hellwig@kit.edu