

Control Barrier Function-Based Safety Filters: Characterization of Undesired Equilibria, Unbounded Trajectories, and Limit Cycles*

Pol Mestres^{1*}, Yiting Chen², Emiliano Dall’Anese², Jorge Cortés¹

¹Department of Mechanical and Aerospace Engineering, UC San Diego.

²Department of Electrical and Computer Engineering, Boston University.

*Corresponding author(s). E-mail(s): pomestre@ucsd.edu;

Contributing authors: yich4684@bu.edu; edallane@bu.edu;

cortes@ucsd.edu;

Abstract

This paper focuses on “safety filters” designed based on Control Barrier Functions (CBFs): these are modifications of a nominal stabilizing controller typically utilized in safety-critical control applications to render a given subset of states forward invariant. The paper investigates the dynamical properties of the closed-loop systems, with a focus on characterizing undesirable behaviors that may emerge due to the use of CBF-based filters. These undesirable behaviors include unbounded trajectories, limit cycles, and undesired equilibria, which can be locally stable and even form a continuum. Our analysis offer the following contributions: (i) conditions under which trajectories remain bounded and (ii) conditions under which limit cycles do not exist; (iii) we show that undesired equilibria can be characterized by solving an algebraic equation, and (iv) we provide examples that show that asymptotically stable undesired equilibria can exist for a large class of nominal controllers and design parameters of the safety filter (even for convex safe sets). Further, for the specific class of planar systems, (v) we provide explicit formulas for the total number of undesired equilibria and the proportion of saddle points and asymptotically stable equilibria, and (vi) in the case of linear planar systems, we present an exhaustive analysis of their global stability properties. Examples throughout the paper illustrate the results.

Keywords: Safety-critical control, control barrier functions, stabilization, safety filters

1 Introduction

Cyber-physical systems and autonomous systems – from individual and connected self-driving vehicles to complex infrastructures such as smart grids and transportation networks – must comply with safety and operational constraints, while ensuring a desired level of performance and efficiency. Control Barrier Functions (CBFs) have emerged as a widely used tool for the design of architectural control frameworks that guarantee the forward invariance of a given set of desirable, “safe” states (termed in what follows as *safe set* of the system) [2–4]. In this context, CBF theory is a workhorse for the design of *safety filters*, which are utilized to (minimally) adjust the input provided by a nominal controller (typically designed to achieve properties such as stability, optimality, or robustness) to ensure forward invariance of the safe set. However, recent works [1, 5–9] have shown that when the nominal controller is augmented with a safety filter, the resulting closed-loop system may not preserve desired properties such as stability and robustness, and it may in fact lead to undesirable behaviors such that the emergence of undesired equilibria. This paper provides a detailed characterization of the undesired equilibria and delves deep into these undesirable behaviors, presenting new findings that demonstrate how safety filters can give rise to unbounded trajectories and limit cycles.

Literature review. CBFs [2–4] are a well-established tool to design controllers that render a given set forward invariant. Several works have combined CBFs with other control-theoretic tools, such as control Lyapunov functions (CLFs) [10], in order to obtain controllers with both safety and stability guarantees [6, 7, 11], and robustness guarantees [12, 13]; CBF-based controllers with input constraints have also been developed in, e.g., [14]. Within this line of works, this paper focuses on *safety filters* [15]. Safety filters yield a controller that minimally modifies the nominal one while ensuring forward invariance of a safe set. A critical research question is whether the closed-loop system under the safety filter retains the stability and robustness properties of the dynamical system with the nominal controller only. A first set of results about the dynamical properties of the closed-loop system with safety filters was provided in [16], which estimated the region of attraction of the desirable equilibrium (taken to be the origin w.l.o.g.). However, questions remain on the asymptotic behavior of the trajectories with initial condition outside the estimated region of attraction. The emergence of undesired equilibria due to CBF-based controllers which are asymptotically stable was noted in [5], and since then they have been studied profusely [6–9]; here, in the controller and filter design, the CBF is assumed to be given throughout the analysis. In previous work [17], we have shown that the undesired equilibria that emerge in safety filters and their stability properties, whatever they may be, are independent of the choice of CBF, for a broad class of CBFs (related by an appropriate equivalence relation). However, [17] does not study what the actual dynamical properties are, which is precisely the focus of this paper. Our recent work [1] provides a characterization of undesired equilibria and, for the special case of linear planar systems and

* A preliminary version of this paper appeared as [1] at the IEEE Conference on Decision and Control.

ellipsoidal obstacles, finds such undesired equilibria explicitly and characterizes their stability properties. The results show that for this class of systems and safe sets, if the system is underactuated there always exists a single undesired equilibrium, whereas if the system is fully actuated, the number and stability properties of undesired equilibria depend on the choice of nominal controller. However, the results of [1] are limited to linear and planar dynamics, and to safe sets that are the complement of an ellipsoid.

Statement of contributions. We study the dynamical properties of closed-loop systems under CBF-based safety filters, paying special attention to the emergence of undesired behaviors. The contributions are summarized as follows:

- (i) We characterize the undesired equilibria that emerge in closed-loop systems with general control-affine dynamical systems, a stabilizing, locally-Lipschitz nominal controller, and a CBF-based safety filter. We show that finding the undesired equilibria is equivalent to solving an algebraic equation.
- (ii) We provide an example showing that, in general, the set of undesired equilibria can be a continuum. This motivates our next contribution, which consists in providing conditions under which the equilibria are isolated points.
- (iii) We show that, in general, the trajectories of the closed-loop system can be unbounded. We then show how, by appropriately selecting some of the parameters of the safety filter, and under mild assumptions, one can ensure that the trajectories of the closed-loop system remain bounded.
- (iv) In the case of planar systems, we show that by suitably tuning the parameters of the safety filter, the closed-loop system does not contain any limit cycles. This implies that all trajectories of the closed-loop system either converge to the origin or to an undesired equilibrium. Therefore, the solutions of the algebraic equation for the undesired equilibria define all the possible limits of trajectories of the closed-loop system. Since solving this algebraic equation for general systems is complicated, we also provide qualitative results regarding the structure of the set of undesired equilibria. We show that if the safe set is bounded, the number of undesired equilibria is even, and half of them are saddle points, whereas if the unsafe set is bounded, the number of undesired equilibria is odd, equal to $2l - 1$ with $l \in \mathbb{Z}_{>0}$, and l of them are saddle points.
- (v) We illustrate the existence of undesired equilibria and their stability properties for linear planar systems in a variety of different cases. For underactuated systems and safe sets that are parametrizable in polar coordinates, we show that no undesired equilibria exist. We provide different examples in which asymptotically stable undesired equilibria exist, including a fully actuated system with a convex safe set, an underactuated system with a safe set not parameterizable in polar coordinates, and an underactuated and fully actuated systems with a bounded unsafe set. We also provide an example with nonconvex unsafe set where asymptotically stable undesired equilibria exist for any choice of stabilizing nominal controller. Finally, for the special case where the unsafe set is an ellipse, we provide analytical expressions for the undesired equilibria and their stability properties. We show that if the system is underactuated, there exists exactly one undesired equilibria, which is a saddle point, whereas if the system is fully actuated, the behavior is much richer and includes a variety of different cases.

Our contributions highlight the intricate relationship between the system dynamics, the geometry of the safe set, and the existence of undesired equilibria and their stability properties. They also serve as a cautionary note to practitioners, for whom we provide a variety of methods to tune (when possible) their controllers to avoid this plethora of undesirable behaviors.

2 Preliminaries

In this section we introduce the notation and provide preliminaries on CBFs and safety filters. The reader familiar with these contents can safely skip this section.

2.1 Notation

We denote by $\mathbb{Z}_{>0}$, $\mathbb{Z}_{\geq 0}$ and \mathbb{R} the set of positive integers, nonnegative integers, and real numbers, respectively. We use bold symbols to represent vectors and non-bold symbols to represent scalar quantities. Let $n \in \mathbb{Z}_{>0}$; $\mathbf{0}_n$ represents the n -dimensional zero vector, $\mathbf{0}_n$ the $n \times n$ -dimensional zero matrix and \mathbf{I}_n the n -dimensional identity matrix. We also write $[n] = \{1, \dots, n\}$. Given $\mathbf{x} \in \mathbb{R}^n$, $\|\mathbf{x}\|$ denotes its Euclidean norm. Given a matrix $G \in \mathbb{R}^{n \times n}$, $\|\mathbf{x}\|_G = \sqrt{\mathbf{x}^T G \mathbf{x}}$ and $\det(G)$ denotes its determinant. A function $\beta : \mathbb{R} \rightarrow \mathbb{R}$ is of extended class \mathcal{K}_∞ if $\beta(0) = 0$, β is strictly increasing and $\lim_{s \rightarrow \pm\infty} \beta(s) = \pm\infty$. Given a set $S \subset \mathbb{R}^n$, we denote by $\text{Int}(S)$ and ∂S the interior and boundary of S , respectively. For a continuously differentiable function $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $\nabla h(\mathbf{x})$ denotes its gradient at \mathbf{x} . Consider the system $\dot{\mathbf{x}} = f(\mathbf{x})$, with $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ locally Lipschitz. Then, for any initial condition $\mathbf{x}_0 \in \mathbb{R}^n$ at time t_0 , there exists a maximal interval of existence $[t_0, t_1)$ such that $\mathbf{x}(t; \mathbf{x}_0)$ is the unique solution to $\dot{\mathbf{x}} = f(\mathbf{x})$ on $[t_0, t_1)$, cf. [10]. For f continuously differentiable and \mathbf{x}_* an equilibrium point of f (i.e., $f(\mathbf{x}_*) = \mathbf{0}_n$), \mathbf{x}_* is *degenerate* if the Jacobian of f evaluated at \mathbf{x}_* has at least one eigenvalue with real part equal to zero; otherwise, we refer to \mathbf{x}_* as *hyperbolic*. Given a hyperbolic equilibrium point with $k \in \mathbb{Z}_{>0}$ eigenvalues with negative real part, the Stable Manifold Theorem [18, Section 2.7] ensures that there exists an invariant k -dimensional manifold S for which all trajectories with initial conditions lying on S converge to \mathbf{x}_* . The global stable manifold at \mathbf{x}_* is defined as $W_s(\mathbf{x}_*) = \bigcup_{\{t \leq 0, \mathbf{x}_0 \in S\}} \mathbf{x}(t; \mathbf{x}_0)$. An equilibrium point \mathbf{x}_* is *isolated* if there exists an open neighborhood \mathcal{U} of \mathbf{x}_* such that \mathbf{x}_* is the only equilibrium point in \mathcal{U} .

2.2 Control barrier functions and safety filters

Consider a control-affine system of the form

$$\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}, \quad (1)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ are locally Lipschitz, $\mathbf{x} \in \mathbb{R}^n$ is the state, and $\mathbf{u} \in \mathbb{R}^m$ is the input. Suppose that a locally Lipschitz nominal controller $k : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is designed so that the system $\dot{\mathbf{x}} = \tilde{f}(\mathbf{x}) := f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$ has a unique equilibrium, which furthermore is globally asymptotically stable. In the remainder of the paper, we

assume without loss of generality that such equilibrium is the origin. In the following, we introduce the notion of safety filter and filtered system. To this end, we first recall the definition of CBF.

Definition 2.1 (Control Barrier Function). *Let $\mathcal{C} \subset \mathbb{R}^n$ be a subset of \mathbb{R}^n . Let $h : \mathbb{R}^n \rightarrow \mathbb{R}$ be a continuously differentiable function such that*

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}, \quad (2a)$$

$$\partial\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\}. \quad (2b)$$

The function h is a **CBF** of \mathcal{C} for the system (1) if there exists an extended class \mathcal{K}_∞ function α such that, for all $\mathbf{x} \in \mathcal{C}$, there exists $\mathbf{u} \in \mathbb{R}^m$ satisfying $\nabla h(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}) + \alpha(h(\mathbf{x})) \geq 0$. If this inequality holds strictly, we refer to h as a strict CBF.

The set \mathcal{C} corresponds to the set of *safe* states. With this definition of CBF, hereafter we refer to the *filtered* system as:

$$\dot{\mathbf{x}} = \tilde{f}(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x}), \quad (3)$$

where the map $v : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is defined as the unique solution to the following optimization problem:

$$\begin{aligned} v(\mathbf{x}) &= \arg \min_{\boldsymbol{\theta} \in \mathbb{R}^m} \|\boldsymbol{\theta}\|_{G(\mathbf{x})}^2 \\ \text{s.t. } &\nabla h(\mathbf{x})^\top(\tilde{f}(\mathbf{x}) + g(\mathbf{x})\boldsymbol{\theta}) + \alpha(h(\mathbf{x})) \geq 0, \end{aligned} \quad (4)$$

with $G : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times m}$ continuously differentiable and positive definite for all $\mathbf{x} \in \mathbb{R}^n$. Note that (4) is a linearly-constrained convex quadratic program (QP), parametrized by the state vector \mathbf{x} . We assume the following.

Assumption 1 (Origin in the interior of \mathcal{C}). *The origin is in the interior of the safe set, i.e., $h(\mathbf{0}_n) > 0$.*

We note that if h is a strict CBF, (4) is strictly feasible for all $\mathbf{x} \in \mathcal{C}$. Moreover, since the objective function in (4) is strongly convex, and the constraints are affine, $v(\mathbf{x})$ is unique and well-defined for all $\mathbf{x} \in \mathcal{C}$. Furthermore, if h is a strict CBF, using arguments similar to [19, Lemma III.2], one can show that $v(\mathbf{x})$ is locally Lipschitz. If h is a strict CBF, it also follows that $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}) \neq \mathbf{0}_n$ for all $\mathbf{x} \in \partial\mathcal{C}$ and therefore from [3, Thm. 2], it follows that the set \mathcal{C} is forward invariant for the system (3). Because of this feature, v is typically referred to as the *safety filter* associated with k .

3 Problem Statement

We consider a control-affine system as in (1) and a safe set $\mathcal{C} \subset \mathbb{R}^n$ defined as the 0-superlevel set of a differentiable function $h : \mathbb{R}^n \rightarrow \mathbb{R}$. We assume that h is a strict CBF of \mathcal{C} ; this function, along with the associated extended class \mathcal{K}_∞ function α in Definition 2.1, are assumed to be given. In our setup, we consider the general setting where: (i) a given controller $k : \mathbb{R}^n \rightarrow \mathbb{R}^m$ globally asymptotically stabilizes

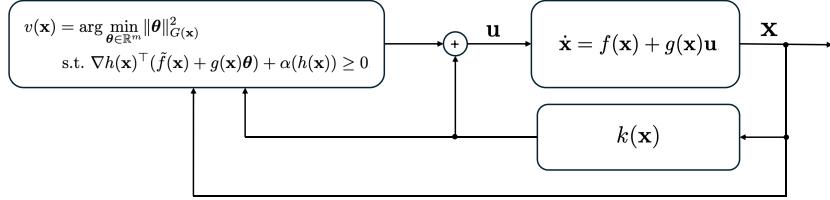


Fig. 1 Closed-loop system that is the subject of this paper. A nominal controller k is used as input to the safety filter, that finds the controller closest to k that satisfies the CBF condition.

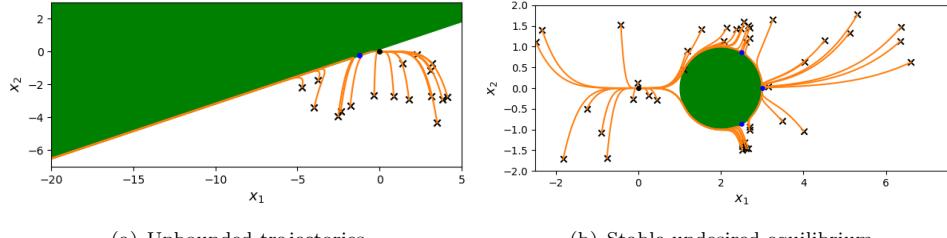


Fig. 2 Control-affine systems with a safety filter with (a) half-plane and (b) circular obstacles. The plots show the trajectories from random initial conditions, the undesired equilibria (colored in blue), and the desired equilibrium (the origin, colored in black).

the origin; (ii) a safety filter is added as in (3) to render \mathcal{C} forward invariant; and, (iii) Assumption 1 holds. Figure 1 illustrates the setup considered in this paper.

Even though the origin is globally asymptotically stable under the nominal controller k , and v is designed to minimally modify k while ensuring that \mathcal{C} is forward invariant, studying the dynamical behavior of (3) is challenging. As noted in, e.g., [16], the filtered system does not necessarily inherit the global asymptotic stability properties of the controller k , and can even have undesired equilibria [1, 5–8] (cf. Figure 2). Most of these works focus on studying conditions under which such undesired equilibria exist or can be confined to specific regions of interest, but do not study other dynamical properties of the closed-loop system, such as boundedness of trajectories, existence of limit cycles, or regions of attraction. In practical applications, these properties (e.g., ensuring that trajectories are bounded, unexpected limit cycles do not arise, or the region of attraction of undesired equilibria is small) is critical to guarantee a desirable performance of the closed-loop system. With this premise, the goal of this paper is as follows:

Problem 1. *Given the system (1) and the safety filter $v(\mathbf{x})$, characterize the boundedness of trajectories, the region of attraction of the origin, existence of undesired equilibria and their regions of attraction, and the existence of limit cycles for the dynamics (3).*

4 Undesired Equilibria, Bounded Trajectories, and Limit Cycles

In this section, we study the dynamical properties of (3), including undesired equilibria, (un)boundedness of trajectories, and limit cycles. We start by providing a precise expression for the unique optimal solution $v(\mathbf{x})$ of the problem (4). For brevity, define $\eta(\mathbf{x}) := \nabla h(\mathbf{x})^T \tilde{f}(\mathbf{x}) + \alpha(h(\mathbf{x}))$. Then, for any $\mathbf{x} \in \mathcal{C}$

$$v(\mathbf{x}) = \begin{cases} \mathbf{0}_m, & \text{if } \eta(\mathbf{x}) \geq 0, \\ \bar{u}(\mathbf{x}), & \text{if } \eta(\mathbf{x}) < 0, \end{cases} \quad (5)$$

where $\bar{u}(\mathbf{x}) := -\frac{\eta(\mathbf{x})G(\mathbf{x})^{-1}g(\mathbf{x})^\top \nabla h(\mathbf{x})}{\|g(\mathbf{x})^\top \nabla h(\mathbf{x})\|_{G^{-1}(\mathbf{x})}^2}$. We note that since h is a strict CBF, $g(\mathbf{x})^\top \nabla h(\mathbf{x}) \neq \mathbf{0}_m$ if $\eta(\mathbf{x}) < 0$ and $\mathbf{x} \in \mathcal{C}$. This implies that \bar{u} (and hence v) is well defined on \mathcal{C} .

4.1 Characterization of undesired equilibria

Our first result leverages expression (5) to provide a necessary and sufficient condition for the existence of undesired equilibria of the filtered system (3).

Lemma 4.1. (Conditions for undesired equilibria): *Consider system (3). Let h be a strict CBF and suppose Assumption 1 holds. Let $\mathbf{x}_* \in \mathbb{R}^n$ be such that $\tilde{f}(\mathbf{x}_*) \neq \mathbf{0}_n$. Then, \mathbf{x}_* is an equilibrium of (3) if and only if there exists $\delta < 0$ such that*

$$h(\mathbf{x}_*) = 0 \text{ and} \quad (6a)$$

$$\tilde{f}(\mathbf{x}_*) = \delta g(\mathbf{x}_*)G(\mathbf{x}_*)^{-1}g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*). \quad (6b)$$

Moreover, \mathbf{x}_* is an equilibrium of (3) independently of the choice of h and α .

Proof. Note that if x_* and $\delta < 0$ satisfy (6), then by multiplying (6b) by $\nabla h(\mathbf{x}_*)^\top$ we get $\|g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*)\|_{G^{-1}(\mathbf{x}_*)}^2 \delta = \nabla h(\mathbf{x}_*)^\top \tilde{f}(\mathbf{x}_*)$. Now, if $g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*) = 0$, it would follow that $\nabla h(\mathbf{x}_*)^\top \tilde{f}(\mathbf{x}_*) = 0$. Since $h(\mathbf{x}_*) = 0$, this would imply that $\nabla h(\mathbf{x}_*)^\top (\tilde{f}(\mathbf{x}_*) + g(\mathbf{x}_*)\mathbf{u}) + \alpha(h(\mathbf{x}_*)) = \nabla h(\mathbf{x}_*)^\top (f(\mathbf{x}_*) + g(\mathbf{x}_*)\mathbf{u}) + \alpha(h(\mathbf{x}_*)) = 0$, for all $\mathbf{u} \in \mathbb{R}^m$, which contradicts the assumption that h is a strict CBF. Therefore, $g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*) \neq 0$ and $\delta = \frac{\nabla h(\mathbf{x}_*)^\top \tilde{f}(\mathbf{x}_*)}{\|g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*)\|_{G^{-1}(\mathbf{x}_*)}^2}$. Then, it follows that $\eta(\mathbf{x}_*) < 0$ and $\tilde{f}(\mathbf{x}_*) + g(\mathbf{x}_*)\bar{u}(\mathbf{x}_*) = \mathbf{0}_n$. Therefore, \mathbf{x}_* is an equilibrium of (3).

Conversely, if \mathbf{x}_* is an equilibrium of (3), then $\tilde{f}(\mathbf{x}_*) + g(\mathbf{x}_*)\bar{u}(\mathbf{x}_*) = \mathbf{0}_n$. It follows that $0 = \nabla h(\mathbf{x}_*)^\top (\tilde{f}(\mathbf{x}_*) + g(\mathbf{x}_*)\bar{u}(\mathbf{x}_*)) = -\alpha(h(\mathbf{x}_*))$. Since $\alpha(\cdot)$ is an extended class \mathcal{K}_∞ function, it must hold that $h(\mathbf{x}_*) = 0$. Note also that $\eta(\mathbf{x}_*) < 0$, since otherwise $\bar{u}(\mathbf{x}_*) = \mathbf{0}_m$ and hence $\tilde{f}(\mathbf{x}_*) + g(\mathbf{x}_*)\bar{u}(\mathbf{x}_*) = \tilde{f}(\mathbf{x}_*) = f(\mathbf{x}_*) + g(\mathbf{x}_*)k(\mathbf{x}_*) = \mathbf{0}_n$, which can only hold if $\mathbf{x}_* = \mathbf{0}_n$, contradicting Assumption 1. Hence, $\eta(\mathbf{x}_*) < 0$ and one has that $f(\mathbf{x}_*) = \frac{\eta(\mathbf{x}_*)}{\|g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*)\|_{G^{-1}(\mathbf{x}_*)}^2} g(\mathbf{x}_*)G(\mathbf{x}_*)^{-1}g(\mathbf{x}_*)^\top \nabla h(\mathbf{x}_*)$, implying (6b) with $\delta < 0$. Finally, the fact that \mathbf{x}_* is an equilibrium of (3) independently of h and α follows from [17, Corollary 4.5]. \square

Hereafter, given a solution $(\mathbf{x}_*, \delta_{\mathbf{x}_*})$ to (6), we refer to $\delta_{\mathbf{x}_*}$ as the *indicator* of \mathbf{x}_* . Lemma 4.1 characterizes the undesired equilibria of closed-loop systems obtained from safety filters. We note that related results exist in the literature: [6, Theorem 2] characterizes the undesired equilibria for a CBF-based control design that includes CLF constraints and [7, Proposition 5.1] characterizes undesired equilibria for a design that includes such CLF constraints as a penalty term in the objective function. However, both of these designs can introduce undesired equilibria in the interior of the safe set, whereas as shown in Lemma 4.1, (3) can only introduce undesired equilibria in the boundary of the safe set.

Based on Lemma 4.1, we define the sets

$$\begin{aligned}\mathcal{E} &:= \{\mathbf{x} \in \mathbb{R}^n : \exists \delta \in \mathbb{R} \text{ s.t. } (\mathbf{x}, \delta) \text{ solves (6)}\}, \\ \hat{\mathcal{E}} &:= \{\mathbf{x} \in \mathbb{R}^n : \exists \delta < 0 \text{ s.t. } (\mathbf{x}, \delta) \text{ solves (6)}\}.\end{aligned}$$

We refer to \mathcal{E} and $\hat{\mathcal{E}}$ as the sets of *potential undesired equilibria* and of *undesired equilibria* of (3), respectively. Note that $\hat{\mathcal{E}} \subset \mathcal{E}$. By Lemma 4.1, determining the equilibrium points of system (3) is equivalent to solving (6) and checking the sign of δ . Under appropriate conditions, [17, Proposition 10] provides an explicit expression for the Jacobian of $\tilde{f}(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x})$ at $\mathbf{x} \in \hat{\mathcal{E}}$ and shows that one of its eigenvalues is $-\alpha'(0)$ (provided that α is differentiable), and the rest of the eigenvalues are independent of α . If $\alpha'(0) > 0$, it follows that the Jacobian evaluated at $\mathbf{x} \in \hat{\mathcal{E}}$ always has a negative eigenvalue.

Even if we are able to identify the set of undesired equilibria, characterizing the dynamical properties of the closed-loop system (3) is still challenging. In the following, we provide a variety of results regarding the boundedness of trajectories, the existence of limit cycles, and the stability properties of undesired equilibria. We also investigate the possible asymptotic behaviors of trajectories of the closed-loop system (3).

4.2 Boundedness of trajectories

The following result states a set of general conditions under which the trajectories of (3) are bounded.

Proposition 4.2. (Conditions for boundedness of trajectories): *Consider system (3) and suppose that h is a strict CBF and Assumption 1 holds. Let $\partial\mathcal{C}$ be bounded. Moreover, assume that the extended class \mathcal{K}_∞ function $\alpha(\cdot)$ is linear, i.e., $\alpha(z) = az$ for some $a > 0$. Then, for any compact set $\Phi \subset \mathbb{R}^n$ with $\partial\Phi \subset \mathcal{C}$, there exists $\tilde{a}_\Phi > 0$ and a compact set $\tilde{\Phi}$ containing Φ such that, by taking $a > \tilde{a}_\Phi$, $\tilde{\Phi}$ is forward invariant under (3). As a consequence, any trajectory of (3) with initial condition in $\Phi \cap \mathcal{C}$ is bounded (because it remains in $\tilde{\Phi}$ at all times).*

Proof. Since $\partial\mathcal{C}$ is bounded, either \mathcal{C} is bounded or $\mathbb{R}^n \setminus \mathcal{C}$ is bounded.

Case 1: \mathcal{C} bounded. In this case, \mathcal{C} is actually compact, and the result holds because \mathcal{C} is forward invariant under (3): for any compact Φ with $\partial\Phi \subset \mathcal{C}$, necessarily $\Phi \subset \mathcal{C}$ and we can take $\tilde{\Phi} = \mathcal{C}$ and any extended class \mathcal{K}_∞ function.

Case 2: $\mathbb{R}^n \setminus \mathcal{C}$ is bounded. Since the system $\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$ renders the origin globally asymptotically stable, by [20, Theorem 4.17], there exists a radially

unbounded Lyapunov function $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$. Let Γ be a Lyapunov sublevel set of V containing $(\mathbb{R}^n \setminus \mathcal{C}) \cup \Phi$ and such that $h(\mathbf{x}) > 0$ for all $\mathbf{x} \in \partial\Gamma$. Note that such Γ exists because $\mathbb{R}^n \setminus \mathcal{C}$ is bounded and V is radially unbounded. Now, since $h(\mathbf{x}) > 0$ for all $\mathbf{x} \in \partial\Gamma$, there exists \tilde{a}_Φ sufficiently large such that $\eta(\mathbf{x}) = \nabla h(\mathbf{x})^T(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})) + \tilde{a}_\Phi h(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \partial\Gamma$. Let $\tilde{\Phi} = \Gamma$. Since Γ is a Lyapunov sublevel set, this implies that Γ is forward invariant for $\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$. By taking $\alpha(z) = az$ with $a > \tilde{a}_\Phi$, and since $\nabla h(\mathbf{x})^T(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})) + \alpha(h(\mathbf{x})) \geq 0$ for all $\mathbf{x} \in \partial\Gamma$, the safety filter is inactive in $\partial\Gamma$ and therefore Γ is also forward invariant under (3). \square

Next we show that if the assumptions of Proposition 4.2 do not hold, the trajectories of (3) might not be bounded (as illustrated in Figure 2(a)). The following result provides technical conditions under which for linear systems and affine CBFs, (3) has unbounded trajectories.

Proposition 4.3. (Unbounded trajectories): *Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ and consider the LTI system $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$, the function $h : \mathbb{R}^n \rightarrow \mathbb{R}$ given by $h(\mathbf{x}) = \mathbf{a}^T \mathbf{x} - b$ and the set $\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$. Let h be a strict CBF and suppose Assumption 1 holds. Further assume that there exists $\mathbf{c} \in \mathbb{R}^n$, $\zeta_1 > 0$, and $\zeta_2 \geq 0$ satisfying $\mathbf{c} \neq \mathbf{0}_n$, $\mathbf{c}^T B = \mathbf{0}_m^T$, and $\mathbf{c}^T A = \zeta_1 \mathbf{c}^T + \zeta_2 \mathbf{a}^T$. Then, for any locally Lipschitz controller $\hat{\mathbf{u}} : \mathbb{R}^n \rightarrow \mathbb{R}^m$, and any initial condition \mathbf{x}_0 in $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} \geq b, \zeta_1 \mathbf{c}^T \mathbf{x} + \zeta_2 b > 0\}$, the solution of $\dot{\mathbf{x}} = A\mathbf{x} + B\hat{\mathbf{u}}(\mathbf{x})$ with initial condition at \mathbf{x}_0 satisfies $\lim_{t \rightarrow +\infty} \|\mathbf{x}(t; \mathbf{x}_0)\| = +\infty$.*

Proof. We note that

$$\frac{d}{dt}(\mathbf{c}^T \mathbf{x}) = \mathbf{c}^T(A\mathbf{x} + B\hat{\mathbf{u}}(\mathbf{x})) = (\zeta_1 \mathbf{c} + \zeta_2 \mathbf{a})^T \mathbf{x} \geq \zeta_1 \mathbf{c}^T \mathbf{x} + \zeta_2 b.$$

This implies that the set $\{\mathbf{x} \in \mathbb{R}^n : \zeta_1 \mathbf{c}^T \mathbf{x} + \zeta_2 b \geq 0\}$ is forward invariant. Additionally, if $\zeta_1 \mathbf{c}^T \mathbf{x}_0 + \zeta_2 b > 0$, then $\zeta_1 \mathbf{c}^T \mathbf{x}(t; \mathbf{x}_0) + \zeta_2 b \geq \zeta_1 \mathbf{c}^T \mathbf{x}_0 + \zeta_2 b$ and therefore $\mathbf{c}^T \mathbf{x}(t; \mathbf{x}_0) \geq \mathbf{c}^T \mathbf{x}_0$ for all $t \geq 0$. It follows that

$$\mathbf{c}^T \mathbf{x}(t; \mathbf{x}_0) = \mathbf{c}^T \mathbf{x}_0 + \int_0^{+\infty} \frac{d}{dt} \mathbf{c}^T \mathbf{x}(t; \mathbf{x}_0) dt \geq \mathbf{c}^T \mathbf{x}_0 + \int_0^{+\infty} (\zeta_1 \mathbf{c}^T \mathbf{x}_0 + \zeta_2 b) dt = +\infty,$$

which implies that $\lim_{t \rightarrow +\infty} \|\mathbf{x}(t; \mathbf{x}_0)\| = +\infty$. \square

Remark 4.4. (Underactuated systems always have unbounded solutions for some safe set): In the setting of Proposition 4.3, if $m < n$, $\ker(B) \neq 0$ and therefore there exists $\mathbf{c} \neq \mathbf{0}_n$ such that $\mathbf{c}^T B = \mathbf{0}_m^T$. Then, for any $\zeta_1 > 0$ and $\zeta_2 > 0$, by letting $\mathbf{a} = \frac{1}{\zeta_2}(A^T \mathbf{c} - \zeta_1 \mathbf{c})$ we satisfy the hypothesis of Proposition 4.3. Hence, for any underactuated linear system, there exists a safe set for which any controller induces unbounded solutions. On the other hand, for fully actuated systems, there does not exist $\mathbf{c} \in \mathbb{R}^n$ satisfying $\mathbf{c}^T B \neq \mathbf{0}_m^T$ and therefore the conditions of Proposition 4.3 are never met. \square

Proposition 4.3 and Remark 4.4 show that in general, one cannot guarantee that the trajectories of (3) are bounded.

4.3 Limit cycles

Here we turn our attention to limit cycles. The following result ensures that, by taking a linear extended class \mathcal{K}_∞ function $\alpha(x) = ax$, $a > 0$, with sufficiently large slope, closed-loop planar systems (3) do not have limit cycles.

Proposition 4.5. (No limit cycles in planar systems): *Consider system (3) with $n = 2$. Let h be a strict CBF with a linear extended class \mathcal{K}_∞ function $\alpha(\cdot)$, i.e., $\alpha(z) = az$ for some $a > 0$. Suppose Assumption 1 holds and that $\mathbb{R}^2 \setminus \mathcal{C}$ is comprised of a finite number of connected components. Then, there exists $\hat{a} > 0$ sufficiently large such that, by taking $a > \hat{a}$, the closed-loop system does not contain any limit cycles in \mathcal{C} . Moreover, all bounded trajectories with initial condition in \mathcal{C} converge to an equilibrium point.*

Proof. We start by noting that, if there exists a limit cycle with at least one point in \mathcal{C} , then it is contained in \mathcal{C} (because the set is forward invariant under (3)). Let γ be a limit cycle contained in \mathcal{C} . Note that γ corresponds to a maximal trajectory of (3) and cannot contain an equilibrium point. We distinguish four different cases:

- (i) If γ does not encircle the origin and does not encircle any connected component of $\mathbb{R}^n \setminus \mathcal{C}$, we reach a contradiction because limit cycles must encircle an equilibrium point [21, Corollary 6.26], and all equilibrium points other than the origin of (3) are in $\partial\mathcal{C}$ by Lemma 4.1.
- (ii) Suppose that γ encircles the origin, but does not encircle any connected component of $\mathbb{R}^2 \setminus \mathcal{C}$. Note that since the origin is globally asymptotically stable for the nominal system, and the origin is contained in $\text{Int}(\mathcal{C})$, there exists $\mathbf{q}_1 \in \partial\mathcal{C}$ such that the solution $\mathbf{y}(t; \mathbf{q}_1)$ of the nominal system satisfies $\mathbf{y}(t; \mathbf{q}_1) \in \text{Int}(\mathcal{C})$ for all $t > 0$. Since $\mathbf{q}_1 \in \partial\mathcal{C}$, $\nabla h(\mathbf{x})^T(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x}))|_{\mathbf{x}=\mathbf{q}_1} > 0$ and since ∇h , f , g and k are continuous there exists a neighborhood $\mathcal{N}_{\mathbf{q}_1}$ of \mathbf{q}_1 such that $\nabla h(\mathbf{x})^T(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x}))|_{\mathbf{x}=\mathbf{z}} > 0$ for all $\mathbf{z} \in \mathcal{N}_{\mathbf{q}_1}$. Now, note that there exists $d_1 > 0$ such that $h(\mathbf{y}(t; \mathbf{q}_1)) > d_1$ for all $t > 0$ such that $\mathbf{y}(t; \mathbf{q}_1) \notin \mathcal{N}_{\mathbf{q}_1}$. This means that there exists $a_1 > 0$ such that $\nabla h(\mathbf{y}(t; \mathbf{q}_1))^T(f(\mathbf{y}(t; \mathbf{q}_1)) + g(\mathbf{y}(t; \mathbf{q}_1))k(\mathbf{y}(t; \mathbf{q}_1))) + a_1 h(\mathbf{y}(t; \mathbf{q}_1)) > 0$ for all $t \geq 0$. Therefore, $\mathbf{y}(t; \mathbf{q}_1)$ is also a trajectory of (3) by taking α to be an extended class \mathcal{K}_∞ function with slope greater than a_1 , which means that it intersects with γ , violating the existence and uniqueness of solutions of (3). Hence, by taking $a > a_1$, we ensure such γ cannot exist.
- (iii) Suppose that γ encircles one or more connected components in $\mathbb{R}^2 \setminus \mathcal{C}$, but not the origin, cf. Figure 3. Let S (resp., \bar{S}) be the union of the connected components encircled (resp., not encircled) by γ . Since the origin is globally asymptotically stable for the nominal system, there exists \mathbf{q}_2 in the boundary of S so that the solution $\mathbf{y}(t; \mathbf{q}_2)$ of the nominal system satisfies one of the following:
 - (a) there exists $t_1 < 0$ with $\tilde{\mathbf{q}}_2 = \mathbf{y}(t_1; \mathbf{q}_2) \in \partial\bar{S}$ and $\mathbf{y}(t; \mathbf{q}_2) \in \text{Int}(\mathcal{C})$ for all $t \in (t_1, 0)$. In this case, there exists t'_1 such that $\mathbf{y}(t'_1; \mathbf{q}_2) \in \gamma$, and there exists $a_{S,2} > 0$ such that $\{\mathbf{y}(t; \mathbf{q}_2)\}_{t \in [t_1, t'_1]}$ is a trajectory of (3) by taking a linear extended class \mathcal{K}_∞ function with slope greater than $a_{S,2}$. This violates the existence and uniqueness of solutions of (3), because $\{\mathbf{y}(t; \mathbf{q}_2)\}_{t \in [t_1, t'_1]}$ is a solution and intersects γ .

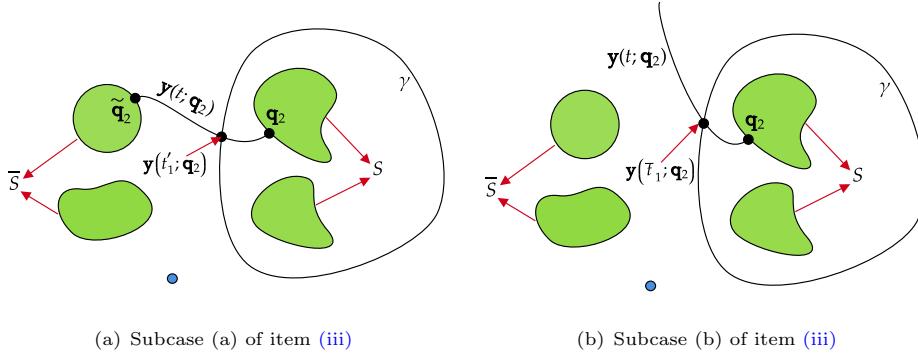


Fig. 3 Sketch of the setting considered in item (iii) of the proof of Proposition 4.5. The connected components comprising $\mathbb{R}^2 \setminus \mathcal{C}$ are depicted in green, whereas the origin is represented by the blue dot.

- (b) $\mathbf{y}(t; \mathbf{q}_2) \in \mathcal{C}$ for all $t < 0$. By Lemma 8.1, $\lim_{t \rightarrow -\infty} \|\mathbf{y}(t; \mathbf{q}_2)\| = \infty$. This means that there exists $\bar{t}_1 < 0$ such that $\mathbf{y}(\bar{t}_1; \mathbf{q}_2) \in \gamma$; This means that there exists $\tilde{a}_{S,2} > 0$ such that $\{\mathbf{y}(t; \mathbf{q}_2)\}_{t \in [\bar{t}_1, \frac{\bar{t}_1}{2}]}$ is a trajectory of (3) by taking a linear extended class \mathcal{K}_∞ function with slope greater than $\tilde{a}_{S,2}$. However, this solution intersects with γ , violating the existence and uniqueness of solutions of (3). Hence by taking a linear extended class \mathcal{K}_∞ function with slope greater than $a_{S,2}$ and $\tilde{a}_{S,2}$, we ensure such γ cannot exist.
- (iv) Suppose that γ encircles one or more connected components in $\mathbb{R}^2 \setminus \mathcal{C}$ and the origin. Let S' (resp., \bar{S}') be the subset of the connected components of $\mathbb{R}^2 \setminus \mathcal{C}$ encircled (resp., not encircled) by γ . Again, since the origin is globally asymptotically stable for the nominal system, there exists \mathbf{q}_3 in the boundary of S' so that the solution $\mathbf{y}(t; \mathbf{q}_3)$ of the nominal system satisfies one of the following:
 - (a) there exists $t_2 < 0$ with $\mathbf{y}(t_2; \mathbf{q}_3) \in \partial \bar{S}'$ and $\mathbf{y}(t; \mathbf{q}_3) \in \text{Int}(\mathcal{C})$ for all $t \in (t_2, 0)$;
 - (b) $\mathbf{y}(t; \mathbf{q}_3) \in \mathcal{C}$ for all $t < 0$.
 By following an argument analogous to case (iii), there exists $\check{a}_{S'} > 0$ sufficiently large such by taking a linear extended class \mathcal{K}_∞ function with slope greater than $\check{a}_{S'}$, we ensure such γ cannot exist.

Note that the values of $a_{S,2}$ and $\tilde{a}_{S,2}$ defined in (iii) depend on the set of connected components of $\mathbb{R}^2 \setminus \mathcal{C}$ encircled by the limit cycle. Since there is a finite number of bounded connected components, there exists a^* such that $a^* > a_{S,2}$ and $a^* > \tilde{a}_{S,2}$ for all possible sets of connected components S of $\mathbb{R}^2 \setminus \mathcal{C}$. Similarly, the value $\check{a}_{S'}$ defined in (iv) depends on S' , but there exists \check{a}^* such that $\check{a}^* > \check{a}_{S'}$ for all possible S' . By taking $a > \hat{a} := \max\{a^*, \check{a}^*\}$, it follows that the closed-loop system does not contain any limit cycles in \mathcal{C} . Finally, since no limit cycles exist in \mathcal{C} for $a > \hat{a}$, by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1] all bounded trajectories with initial condition in \mathcal{C} converge to an equilibrium point. \square

The following example shows that limit cycles can exist for systems of the form (3) in dimension $n \geq 3$.

Example 4.6. (Existence of limit cycles in higher dimensions): Consider the safe set $\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x} - [x_c, 0, 0]^T\|^2 - r^2 \geq 0\}$. Let $B \in \mathbb{R}^{3 \times 3}$ invertible, $G(\mathbf{x}) = B^\top B$, $K = \mathbf{0}_3$ and nominal controller $k(\mathbf{x}) = K\mathbf{x} \equiv \mathbf{0}_3$. Next, let $0 < p_1 < \frac{r}{r+x_c}p_2$, $p_2 > 0$, $p_3 > 0$, and define

$$A := \begin{bmatrix} -p_1 & 0 & 0 \\ 0 & -p_2 & p_3 \\ 0 & -p_3 & -p_2 \end{bmatrix}.$$

Consider the closed-loop system (3) obtained with $f(\mathbf{x}) = A\mathbf{x}$, $g(\mathbf{x}) = B$, $h(\mathbf{x}) = \|\mathbf{x} - [x_c, 0, 0]^T\|^2 - r^2$ and $k(\mathbf{x}) \equiv 0$.

Define also $\hat{p} := \frac{x_c p_1}{p_2 - p_1}$ and $\hat{q} := \sqrt{r^2 - \hat{p}^2}$, then $0 < \hat{p} < r$ and $p_1 = \frac{\hat{p}}{\hat{p} + x_c}p_2$. Consider

$$\mathbf{x}_0 := \begin{bmatrix} \hat{p} \\ \hat{q} \\ \hat{q} \end{bmatrix}, \quad \hat{\mathbf{x}}(t; \mathbf{x}_0) := \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \end{bmatrix} = \begin{bmatrix} x_c + \hat{p} \\ \hat{q} \sin(p_3 t) \\ \hat{q} \cos(p_3 t) \end{bmatrix}.$$

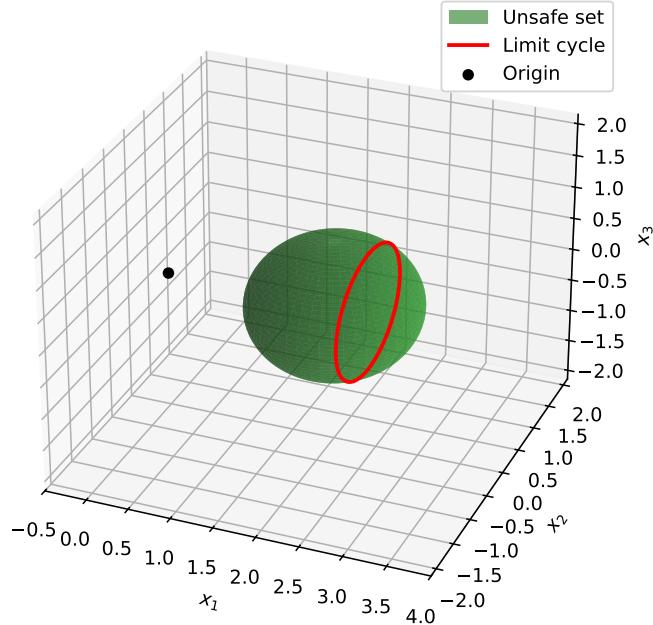


Fig. 4 Depiction of the setting considered in Example 4.6, with parameters $x_c = 2$, $r = 1$, $p_1 = 1$, $p_2 = 6$, $p_3 = 1$. The obstacle is depicted in green, the limit cycle $\hat{\mathbf{x}}$ in red, and the origin in black.

Note that for any t ,

$$\nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))^T (A + BK) \hat{\mathbf{x}}(t; \mathbf{x}_0) = -2 \begin{bmatrix} \hat{p} \\ \hat{q} \sin(p_3 t) \\ \hat{q} \cos(p_3 t) \end{bmatrix}^T \begin{bmatrix} \frac{\hat{p}}{\hat{p} + x_c} p_2 & 0 & 0 \\ 0 & p_2 & -p_3 \\ 0 & p_3 & p_2 \end{bmatrix} \hat{\mathbf{x}}(t; \mathbf{x}_0) = -2r^2 p_2 < 0.$$

Moreover, let

$$\kappa(t) = \frac{\nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))^T (A + BK) \hat{\mathbf{x}}(t; \mathbf{x}_0) + \alpha(h(\hat{\mathbf{x}}(t; \mathbf{x}_0)))}{\nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))^T BG^{-1}B^T \nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))}$$

and note that since $K = \mathbb{0}_3$,

$$\begin{aligned} & (A + BK) \hat{\mathbf{x}}(t; \mathbf{x}_0) - \kappa(t) BG^{-1}B^T \nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0)) \\ &= A \hat{\mathbf{x}}(t; \mathbf{x}_0) - \frac{\nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))^T A \hat{\mathbf{x}}(t; \mathbf{x}_0)}{\nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))^T \nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0))} \nabla h(\hat{\mathbf{x}}(t; \mathbf{x}_0)) \\ &= \begin{bmatrix} -\hat{p}p_2 \\ -p_2 \hat{q} \sin(p_3 t) + p_3 \hat{q} \cos(p_3 t) \\ -p_2 \hat{q} \sin(p_3 t) - p_3 \hat{q} \cos(p_3 t) \end{bmatrix} + p_2 \begin{bmatrix} \hat{p}_1 \\ \hat{q} \sin(p_3 t) \\ \hat{q} \cos(p_3 t) \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ p_3 \hat{q} \cos(p_3 t) \\ -p_3 \hat{q} \sin(p_3 t) \end{bmatrix} = \frac{d}{dt} \begin{bmatrix} x_c + \hat{p} \\ \hat{q} \sin(p_3 t) \\ \hat{q} \cos(p_3 t) \end{bmatrix} = \frac{d}{dt} \begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \end{bmatrix} \end{aligned}$$

Hence $\hat{\mathbf{x}}(t; \mathbf{x}_0)$ is a valid trajectory of the closed-loop system (3) and it is a limit cycle. Note also that since $\hat{\mathbf{x}}(t) \in \partial \mathcal{C}$ for all $t \geq 0$, by [17, Corollary 4.5] this trajectory exists for any choice of h and α . Figure 4 depicts the limit cycle $\hat{\mathbf{x}}$. \triangle

4.4 Structure of the set of undesired equilibria

In this section, we investigate the set of undesired equilibria of (3). The following example shows that in general, this set can be a continuum.

Example 4.7. (Continuum of undesired equilibria): Let $n = 2, m = 1$ and consider (3) with

$$f(\mathbf{x}) = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \mathbf{x}, \quad g(\mathbf{x}) = B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad h(\mathbf{x}) = \begin{cases} (x_1 + 1)^2 + x_2, & x_1 \geq -1, \\ x_2, & -2 < x_1 < -1, \\ (x_1 + 2)^2 + x_2, & x_1 \leq -2, \end{cases}$$

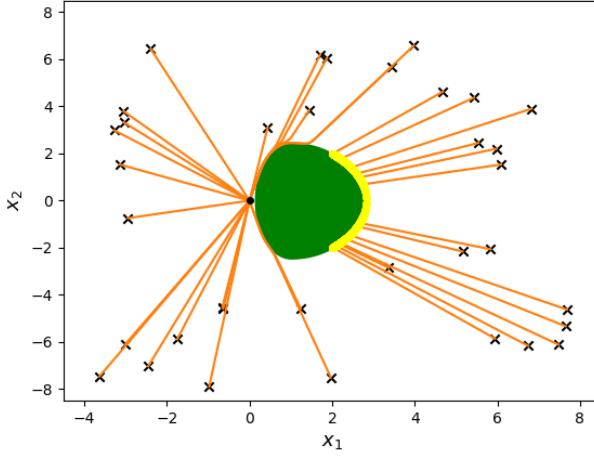


Fig. 5 Plot of different trajectories of (8). Trajectories are depicted in orange. The unsafe set is colored in green. Black crosses denote initial conditions, the black dot denotes the origin, and the yellow region denotes a continuum of undesired equilibria.

and $\alpha(s) = 10s$. Note that h is continuously differentiable. Moreover,

$$\nabla h(\mathbf{x})^T B = \begin{cases} 2 \begin{bmatrix} x_1 + 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0, & x_1 \geq -1, \\ 2 \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0, & -2 < x_1 < -1, \\ 2 \begin{bmatrix} x_1 + 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq 0, & x_1 \leq -2, \end{cases}$$

Therefore, h is a strict CBF. Next, we show that the set $\{t(1,0) : -2 \leq t \leq -1\}$ is contained in the set of undesired equilibria for any linear stabilizing controller $k(\mathbf{x}) = k_1 x_1 + k_2 x_2$, $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ and extended class \mathcal{K}_∞ function α . Since k is a stabilizing controller, it follows that $f(\mathbf{x}) + Bk(\mathbf{x})$ is Hurwitz and therefore $-2+k_2 < 0$, $1+k_1 > 0$. For any $\sigma \in [-2, -1]$, the point $\mathbf{x}_\sigma = (\sigma, 0) \in \mathbb{R}^2$ satisfies (6) with associated indicator equal to $\frac{G(\mathbf{x}_\sigma)}{(1+k_1)\sigma} < 0$. Hence, the set $\{\sigma(1,0) : -2 \leq \sigma \leq -1\}$ is contained in the set of undesired equilibria for any linear stabilizing controller $k(\mathbf{x}) = k_1 x_1 + k_2 x_2$, $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ and extended class \mathcal{K}_∞ function α . Figure 5 shows some of the trajectories for the corresponding closed-loop system (3). Since the undesired equilibria are not isolated, the study of their stability properties requires using the notion of *semistability* [23]. \triangle

The following result provides conditions under which a continuum of undesired equilibria of (3) does not exist.

Lemma 4.8. (Sufficient conditions for isolated equilibria): *An undesired equilibrium \mathbf{x}_* is isolated if the Jacobian of (3) evaluated at \mathbf{x}_* does not have imaginary eigenvalues. If $\partial\mathcal{C}$ is bounded, then each undesired equilibria of (3) is isolated if and only if $|\hat{\mathcal{E}}| < \infty$.*

Proof. Given an undesired equilibrium \mathbf{x}_* , if the Jacobian of (3) evaluated at \mathbf{x}_* does not have imaginary eigenvalue, then there exists a neighborhood of \mathbf{x}_* such that the linearization of (3) around \mathbf{x}_* does not contain any equilibrium point other than itself. By the Hartman-Grobman Theorem [18, Section 2.8], there also exists a neighborhood of \mathbf{x}_* for which (3) does not contain any undesired equilibrium and hence \mathbf{x}_* is isolated.

Consider the case when $\partial\mathcal{C}$ is bounded. Clearly, if $|\hat{\mathcal{E}}| < \infty$ (i.e., the number of undesired equilibria is finite), then each of the undesired equilibria is isolated. Conversely, if the number of undesired equilibria is infinite, consider an infinite sequence of undesired equilibria $\{\mathbf{x}_{*,i}\}_{i=1}^{+\infty}$. Since $\partial\mathcal{C}$ is compact, there exists a convergent subsequence $\{\mathbf{x}_{*,i_k}\}_{k=1}^{+\infty}$ such that $\lim_{k \rightarrow \infty} \mathbf{x}_{*,i_k} = \mathbf{q}_*$. Since (3) is continuous under the assumption that h is a strict CBF (cf. [19, Lemma III.2]), $0 = \lim_{k \rightarrow \infty} f(\mathbf{x}_{*,i_k}) + g(\mathbf{x}_{*,i_k})v(\mathbf{x}_{*,i_k}) = f(\mathbf{x}_*) + g(\mathbf{q}_*)v(\mathbf{x}_*)$ and hence \mathbf{x}_* is an equilibrium, which is non-isolated. \square

While Lemma 4.8 provides sufficient conditions for the existence of isolated undesired equilibria, finding their number is, in general, challenging. However, this is possible for the special case of planar systems, as we show next.

Proposition 4.9. (Number and stability properties of undesired equilibria for a planar system with bounded obstacle): *Let h be a strict CBF and suppose Assumption 1 holds. Let $\mathbb{R}^2 \setminus \mathcal{C}$ be a bounded connected set. Consider (3) with $n = 2$ and assume its undesired equilibria $\hat{\mathcal{E}} = \{\mathbf{x}_*^{(i)}\}_{i=1}^k \subset \partial\mathcal{C}$ are either asymptotically stable or saddle points. Then, k is odd, and $\frac{k+1}{2}$ equilibria are saddle points and $\frac{k-1}{2}$ are asymptotically stable.*

Proof. Let $\mathcal{L} \subset [k]$ be the index set of undesired equilibria that are saddle points. We show that $|\mathcal{L}| = \frac{k+1}{2}$. Figure 6 serves as visual aid for the different elements employed in the proof. Let Φ be a compact set containing the origin and $\mathbb{R}^2 \setminus \mathcal{C}$ in its interior. This implies that $\partial\Phi \subset \mathcal{C}$. By Propositions 4.2 and 4.5, there exists $a_\Phi > 0$ and a compact set $\tilde{\Phi}$ containing Φ such that (3), with extended class \mathcal{K}_∞ function with slope greater than a_Φ , makes $\tilde{\Phi}$ forward invariant and $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycles. Since [17, Proposition 11] ensures that the stability properties of undesired equilibria are independent of α , we can assume without loss of generality that α takes this form. Now, for $j \in \mathcal{L}$, let γ_j be a subset of the one-dimensional local stable manifold of (3) at $\mathbf{x}_*^{(j)}$ such that:

- (i) γ_j corresponds to a maximal trajectory of (3);
- (ii) there exists $T > 0$ with $\gamma_j(t) \in \text{Int}(\mathcal{C})$ for all $t > T$.

Note that such γ_j always exists because the stable manifold is a union of trajectories and the stable manifold of (3) at $\mathbf{x}_*^{(j)}$ is tangent to $\nabla h(\mathbf{x}_*^{(j)})$. The fact that the stable manifold of (3) at $\mathbf{x}_*^{(j)}$ is tangent to $\nabla h(\mathbf{x}_*^{(j)})$ follows from the fact that $\nabla h(\mathbf{x}_*^{(j)})$ is an eigenvector of the Jacobian of (3) evaluated at $\mathbf{x}_*^{(j)}$ with negative eigenvalue (cf. [24, Proposition 6.2]), and the Stable Manifold Theorem [18, Section 2]. By Lemma 8.3,

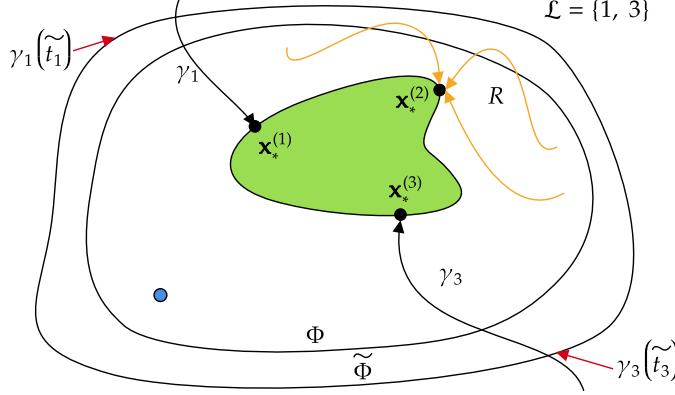


Fig. 6 Sketch of the setting considered in the proof of Proposition 4.9. The unsafe set depicted in green, whereas the origin is represented by the blue dot.

since $\tilde{\Phi}$ is forward invariant and $\tilde{\Phi} \cap \mathcal{C}$ does not contain limit cycles, there exists at least one $j \in \mathcal{L}$ and $\tilde{t}_j \in \mathbb{R}$ such that $\gamma_j(\tilde{t}_j) \in \partial\tilde{\Phi}$ and $\gamma_j(t) \notin \tilde{\Phi}$ for all $t < \tilde{t}_j$.

Moreover by Lemma 8.2(i), γ_j cannot be tangent to \mathcal{C} for any $j \in \mathcal{L}$. Therefore $\{\gamma_j\}_{j \in \mathcal{L}}$ divide $\tilde{\Phi} \cap \mathcal{C}$ into $|\mathcal{L}|$ regions (note that this would not be the case if there was not at least one $j \in \mathcal{L}$ and $\tilde{t}_j \in \mathbb{R}$ such that $\gamma_j(\tilde{t}_j) \in \partial\tilde{\Phi}$ and $\gamma_j(t) \notin \tilde{\Phi}$ for all $t < \tilde{t}_j$), each of which is an open connected set. Let R be any one of such regions. We now show that $\text{Cl}(R)$ contains exactly one asymptotically stable equilibrium point.

Indeed, first suppose that there is no asymptotically stable equilibrium point in $\text{Cl}(R)$. Let $\tilde{\gamma}$ be any trajectory with initial condition in $\text{Cl}(R)$. By the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], since $\tilde{\Phi}$ is forward invariant, $\tilde{\gamma}$ converges to either a limit cycle or an equilibrium point. It cannot converge to a limit cycle because $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycles. Moreover, if \bar{x}_* and \tilde{x}_* are the equilibrium points in $\{\mathbf{x}_*^{(j)}\}_{j=1}^k$ whose stable manifolds define the boundary of R , $\tilde{\gamma}$ cannot converge to an equilibrium other than \bar{x}_* or \tilde{x}_* , since otherwise $\tilde{\gamma}$ would intersect the stable manifolds of \bar{x}_* or \tilde{x}_* , which would contradict the uniqueness of solutions of (3). But it can also not converge to \bar{x}_* or \tilde{x}_* : if, for example, $\tilde{\gamma}$ converged to \bar{x}_* , there would be two different trajectories converging to \bar{x}_* with different tangent vectors at \bar{x}_* , which would contradict the fact that \bar{x}_* is a saddle point. Therefore, $\text{Cl}(R)$ contains at least one asymptotically stable equilibrium.

Next suppose that there are multiple asymptotically stable undesired equilibria in $\text{Cl}(R)$. By [20, Theorem 8.1], the boundary of the regions of attraction of asymptotically stable equilibria is formed by trajectories. Let $\tilde{\gamma}$ be one such trajectory with initial condition in R . Note that since R is contained in $\tilde{\Phi}$ and $\tilde{\Phi}$ is forward invariant, $\tilde{\gamma}$ is bounded. By the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], this trajectory must converge to an equilibrium point or a limit cycle. It can not converge to a limit cycle because $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycles. It can not converge to an equilibrium point because it is not in any region of attraction of an asymptotically stable equilibrium, there are no saddle points in R , and it can not converge to \bar{x}_* (resp. \tilde{x}_*), because otherwise \bar{x}_* (resp. \tilde{x}_*) would have two trajectories converging

to it with different tangent vectors, which would contradict the fact that $\bar{\mathbf{x}}_*$ (resp. $\tilde{\mathbf{x}}_*$) is a saddle point.

Therefore, we conclude that in each of the $|\mathcal{L}|$ regions formed by $\{\gamma_j\}_{j \in \mathcal{L}}$, there is exactly one asymptotically stable equilibrium in their boundary. Since there are $k - |\mathcal{L}|$ other undesired equilibria, and since the origin is asymptotically stable, this means that $|\mathcal{L}| = k - |\mathcal{L}| + 1$. Hence, $|\mathcal{L}| = \frac{k+1}{2}$. Note also that by [17, Proposition 10], the stability properties of undesired equilibria are independent of the choice of extended class \mathcal{K}_∞ function α . Hence, even though in our arguments we have chosen a specific extended class \mathcal{K}_∞ function α , the statement holds for any such α . \square

Under the assumptions of Proposition 4.9, since $\mathbb{R}^2 \setminus \mathcal{C}$ is bounded, [25, Proposition 3] implies that there does not exist a safe globally asymptotically stabilizing controller. If no limit cycles exist (for example, under the conditions of Proposition 4.5), by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], this implies that there must exist at least one undesired equilibrium, and hence $k \geq 1$. Note also that as shown in [17, Proposition 6.2], under a large class of CBFs, the stability properties of the undesired equilibria remain the same.

Next we give a result similar to Proposition 4.9 for the case when the safe set \mathcal{C} is compact and connected.

Proposition 4.10. (Number of undesired equilibria for compact connected safe set): *Let h be a strict CBF and suppose Assumption 1 holds. Let \mathcal{C} be a compact connected set. Consider (3) with $n = 2$ and assume its undesired equilibria $\hat{\mathcal{E}} = \{\mathbf{x}_*^{(i)}\}_{i=1}^k \subset \partial\mathcal{C}$ are either asymptotically stable or saddle points. Then, k is even, and $\frac{k}{2}$ equilibria are saddle points and $\frac{k}{2}$ equilibria are asymptotically stable.*

Proof. The proof follows a similar argument to that of Proposition 4.9. Let $\mathcal{L} \subset [k]$ be the set of indices of undesired saddle points. We show that $|\mathcal{L}| = \frac{k}{2}$. For $j \in \mathcal{L}$, let γ_j be a subset of the one-dimensional local stable manifold of (3) at $\mathbf{x}_*^{(j)}$ such that

- (i) γ_j corresponds to a maximal trajectory of (3);
- (ii) there exists $T > 0$ with $\gamma_j(t) \in \text{Int}(\mathcal{C})$ for all $t > T$.

Note that for each $j \in \mathcal{L}$, either there exists $\tilde{t}_j \in \mathbb{R}$ such that $\gamma_j(\tilde{t}_j) \in \partial\mathcal{C}$ and $\gamma_j(t) \notin \mathcal{C}$ for $t < \tilde{t}_j$ or $\lim_{t \rightarrow -\infty} \gamma_j(t) = \mathbf{x}'_*$, with \mathbf{x}'_* an undesired equilibrium that is a saddle point. Indeed, otherwise, since \mathcal{C} is bounded, by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1] $\lim_{t \rightarrow -\infty} \gamma_j(t)$ would converge to a limit cycle or another equilibrium point. However, by Proposition 4.5, there exists $\hat{a} > 0$ such that (3) with linear extended class \mathcal{K}_∞ function with slope greater than \hat{a} does not have any limit cycles in \mathcal{C} . Since [17, Proposition 11] ensures that the stability properties of undesired equilibria are independent of α , we can assume without loss of generality that $\lim_{t \rightarrow -\infty} \gamma_j(t)$ does not converge to a limit cycle. Moreover, $\lim_{t \rightarrow -\infty} \gamma_j(t)$ cannot converge to an asymptotically stable equilibrium, since it belongs to the one-dimensional local stable manifold of $\mathbf{x}_*^{(j)}$.

Now we note that, since for all $j \in \mathcal{L}$, either there exists \tilde{t}_j such that $\gamma_j(\tilde{t}_j) \in \partial\mathcal{C}$ and $\gamma_j(t) \notin \mathcal{C}$ for $t < \tilde{t}_j$ or $\lim_{t \rightarrow -\infty} \gamma_j(t) = \mathbf{x}'_*$, with \mathbf{x}'_* an undesired equilibrium that

is a saddle point, the trajectories $\{\gamma_j\}_{j \in \mathcal{L}}$ divide \mathcal{C} into $|\mathcal{L}| + 1$ connected sets in \mathcal{C} . By an argument analogous to the one in the proof of Proposition 4.9, in each of those sets there must exist exactly one asymptotically stable equilibrium. Since the origin is asymptotically stable under (3), this implies that the number of undesired equilibria that are saddle points is equal to the number of undesired equilibria that are asymptotically stable, proving $|\mathcal{L}| = \frac{k}{2}$. \square

Note that by [17, Proposition 11], the Jacobian of (3) evaluated at an undesired equilibrium has at least one negative eigenvalue. Therefore, the assumption in Propositions 4.9 and 4.10 that all undesired equilibria are either asymptotically stable or saddle points is satisfied if at any undesired equilibrium the other eigenvalue of the Jacobian is nonzero. If the other eigenvalue is zero and the equilibrium is degenerate, the point is asymptotically stable if the trajectories in its central manifold converge to it and it is a saddle point if the trajectories in its central manifold diverge from it.

We also note that Propositions 4.9 and 4.10 provide information about the number and the stability properties of undesired equilibria even in the case where the algebraic equations (6) defining the undesired equilibria are difficult to solve. We also point out that both results require $\partial\mathcal{C}$ to be bounded and hence, by Lemma 4.8, the assumption that the number of equilibria is finite is equivalent to each of them being isolated.

We finalize this section by introducing a class of safe sets that do not introduce undesired equilibria.

Proposition 4.11. (Class of safe sets with global asymptotic stability of origin): *Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a global Lyapunov function for the nominal system $\dot{\mathbf{x}} = f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})$. Let $c > 0$ and suppose $h(\mathbf{x}) := c - V(\mathbf{x})$ is a strict CBF of $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$. Then, the origin is asymptotically stable with region of attraction containing \mathcal{C} . In particular, (3) does not contain any undesired equilibria.*

Proof. Let us first show that the safety filter is inactive at all points in \mathcal{C} . Indeed, for all $\mathbf{x} \in \mathcal{C}$, $\eta(\mathbf{x}) = \nabla h(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})) + \alpha(h(\mathbf{x})) = -\nabla V(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x})) + \alpha(h(\mathbf{x})) \geq 0$. Therefore, $v(\mathbf{x}) = \mathbf{0}_m$ for all $\mathbf{x} \in \mathcal{C}$. This implies that for all $\mathbf{x} \in \mathcal{C}$, $\nabla V(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})k(\mathbf{x}) + g(\mathbf{x})v(\mathbf{x})) < 0$ for all $\mathbf{x} \in \mathcal{C} \setminus \{\mathbf{0}_n\}$, and therefore all trajectories with initial condition in $\mathcal{C} \setminus \{\mathbf{0}_n\}$ converge to the origin, i.e., the origin is asymptotically stable with region of attraction containing \mathcal{C} . In particular, this implies that no undesired equilibria exist (i.e., $\hat{\mathcal{E}} = \emptyset$), since otherwise trajectories with initial condition in such undesired equilibria do not converge to the origin. \square

5 Dynamical Properties of Safety Filters for Linear Planar Systems

In this section, we focus on linear planar systems. Due to their simpler structure, solving (6) leads to additional results and insights, compared to the general treatment presented in the previous section. Consider the LTI planar system

$$\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}, \quad (7)$$

where $\mathbf{x} = [x_1, x_2]^\top \in \mathbb{R}^2$, $\mathbf{u} \in \mathbb{R}^m$, with $m \in \{1, 2\}$, $A \in \mathbb{R}^{2 \times 2}$, and $B \in \mathbb{R}^{2 \times m}$ having full column rank. We make the following assumption.

Assumption 2 (Stabilizability). *The system (7) is stabilizable. Moreover, $\mathbf{u} = -K\mathbf{x}$, $K \in \mathbb{R}^{2 \times m}$, is a stabilizing controller such that $\tilde{A} = A - BK$ is Hurwitz.* \square

In this setup, the system (3) is then customized as follows.

$$\dot{\mathbf{x}} = F(\mathbf{x}) := (A - BK)\mathbf{x} + Bv(\mathbf{x}), \quad (8)$$

where the safety filter is given by

$$v(\mathbf{x}) = \begin{cases} 0, & \text{if } \eta(\mathbf{x}) \geq 0, \\ -\frac{\eta(\mathbf{x})G(\mathbf{x})^{-1}B^T \nabla h(\mathbf{x})}{\|B^T \nabla h(\mathbf{x})\|_{G^{-1}(\mathbf{x})}^2}, & \text{if } \eta(\mathbf{x}) < 0, \end{cases} \quad (9)$$

with $\eta(\mathbf{x}) := \nabla h(\mathbf{x})^T (A - BK)\mathbf{x} + \alpha(h(\mathbf{x}))$.

As shown in [17, Proposition 6.2], the stability properties of the undesired equilibria in the different results of this section hold for a large class of choices of the CBF h and the function α .

5.1 Bounded safe set

Here we discuss various results and examples for the case where the safe set is compact and contains the origin. We start by showing that for linear, planar underactuated systems and safe sets that are parametrizable in polar coordinates by a continuously differentiable function, the system (8) does not have undesired equilibria.

Proposition 5.1. (No undesired equilibria for underactuated planar systems and safe sets parametrizable in polar coordinates): *Consider (7) with $m = 1$ (i.e., the system is underactuated), and suppose that Assumptions 1 and 2 hold. Let $r : \mathbb{R} \rightarrow \mathbb{R}_{>0}$ be a continuously differentiable, 2π -periodic function, and such that*

$$\partial\mathcal{C} = \{(r(\theta) \cos(\theta), r(\theta) \sin(\theta)) : \theta \in [0, 2\pi]\}, \quad (10)$$

and let h be a strict CBF of \mathcal{C} . Then, (8) does not have any undesired equilibria for any $K \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}_{>0}$.

Proof. For convenience, denote

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 & k_2 \end{bmatrix},$$

in (7) and let $\beta = a_{11}b_2 - b_1a_{21}$, $\gamma = a_{22}b_1 - b_2a_{12}$. Further define $R : \mathbb{R}^2 \setminus \{\mathbf{0}_2\} \rightarrow \mathbb{R}$ as

$$R(\mathbf{y}) = \begin{cases} r\left(\frac{\pi}{2}\right)^2 & \text{if } y_1 = 0, y_2 > 0, \\ r\left(\frac{3\pi}{2}\right)^2 & \text{if } y_1 = 0, y_2 < 0, \\ r(\arctan(\frac{y_2}{y_1}))^2 & \text{otherwise,} \end{cases}$$

We recall the following three facts:

- (i) By [17, Corollary 4.5], undesired equilibria are independent of the choice of CBF;
- (ii) By Lemma 8.4, since h is a strict CBF of \mathcal{C} , any CBF of \mathcal{C} is a strict CBF of \mathcal{C} ;
- (iii) By [3, Theorem 3], since \mathcal{C} is safe, any continuously differentiable function whose 0-superlevel is \mathcal{C} is a CBF of \mathcal{C} .

These facts imply that, without loss of generality, we can assume that h is such that in a neighborhood \mathcal{N}_c of $\partial\mathcal{C}$, $h(\mathbf{y}) = -\|\mathbf{y}\|^2 + R(\mathbf{y})$ for all $\mathbf{y} \in \mathcal{N}_c$, and that such h is a strict CBF. Note that for all $\mathbf{x} \in \partial\mathcal{C}$ with $x_1 \neq 0$, we have

$$\nabla h(\mathbf{x}) = \begin{pmatrix} -2x_1 + 2r(\arctan(\frac{x_2}{x_1}))r'(\arctan(\frac{x_2}{x_1}))\frac{-x_2}{x_1^2+x_2^2} \\ -2x_2 + 2r(\arctan(\frac{x_2}{x_1}))r'(\arctan(\frac{x_2}{x_1}))\frac{x_1}{x_1^2+x_2^2} \end{pmatrix}. \quad (11)$$

Since ∇h is continuous, if $x_1 = 0$ and $x_2 > 0$, we have

$$\nabla h(0, x_2) = \begin{pmatrix} -2x_1 + 2r(\arctan(\frac{\pi}{2}))r'(\arctan(\frac{\pi}{2}))\frac{-1}{x_2} \\ -2x_2 \end{pmatrix}, \quad (12)$$

whereas if $x_1 = 0$ and $x_2 < 0$,

$$\nabla h(0, x_2) = \begin{pmatrix} -2x_1 + 2r(\arctan(\frac{3\pi}{2}))r'(\arctan(\frac{3\pi}{2}))\frac{-1}{x_2} \\ -2x_2 \end{pmatrix}. \quad (13)$$

Therefore, (11) is valid also for $x_1 = 0$ by defining $\arctan(\frac{x_2}{x_1}) = \frac{\pi}{2}$ if $x_1 = 0$ and $x_2 > 0$, and $\arctan(\frac{x_2}{x_1}) = \frac{3\pi}{2}$ if $x_1 = 0$ and $x_2 < 0$. From (6), the undesired equilibria lie in $\partial\mathcal{C}$, and therefore are of the form $(r(\theta^*) \cos(\theta^*), r(\theta^*) \sin(\theta^*))$. Furthermore, from (6) and (11), we have

$$\begin{pmatrix} a_{11} + b_1 k_1 & a_{12} + b_1 k_2 \\ a_{21} + b_2 k_1 & a_{22} + b_2 k_2 \end{pmatrix} \begin{pmatrix} r(\theta^*) \cos(\theta^*) \\ r(\theta^*) \sin(\theta^*) \end{pmatrix} = \delta \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}^T \nabla h \begin{pmatrix} r(\theta^*) \cos(\theta^*) \\ r(\theta^*) \sin(\theta^*) \end{pmatrix}.$$

It follows that θ^* satisfies

$$b_2 (a_{11} + b_1 k_1 \ a_{12} + b_1 k_2) \begin{pmatrix} r(\theta^*) \cos(\theta^*) \\ r(\theta^*) \sin(\theta^*) \end{pmatrix} = b_1 (a_{21} + b_2 k_1 \ a_{22} + b_2 k_2) \begin{pmatrix} r(\theta^*) \cos(\theta^*) \\ r(\theta^*) \sin(\theta^*) \end{pmatrix},$$

which implies that θ^* satisfies $\gamma \sin(\theta^*) = \beta \cos(\theta^*)$. Note that there exist exactly two values θ^* in $[0, 2\pi]$ solving $\gamma \sin(\theta^*) = \beta \cos(\theta^*)$. Hence, there exist exactly two potential undesired equilibria (i.e., \mathcal{E} has cardinality 2). Let $\mathbf{x}_*^{(1)}, \mathbf{x}_*^{(2)}$ be such undesired equilibria, with θ_1^* and θ_2^* being their associated θ^* values. From Lemma 4.1, $\mathbf{x}_*^{(1)}$ is an undesired equilibria if and only if $\eta(\mathbf{x}_*^{(1)}) < 0$, or equivalently,

$$\frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) \left((a_{11} - b_1 k_1) \cos(\theta_1^*) + (a_{12} - b_1 k_2) \sin(\theta_1^*) \right) +$$

$$\frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) \left((a_{21} - b_2 k_1) \cos(\theta_1^*) + (a_{22} - b_2 k_2) \sin(\theta_1^*) \right) < 0. \quad (14)$$

If $\gamma \neq 0$, $\sin(\theta_1^*) = \frac{\beta}{\gamma} \cos(\theta_1^*)$, and (14) is equivalent to

$$\begin{aligned} & \frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) \frac{\cos(\theta_1^*)}{\gamma} \left(\gamma(a_{11} - b_1 k_1) + \beta(a_{12} - b_1 k_2) \right) + \\ & \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) \frac{\cos(\theta_1^*)}{\gamma} \left(\gamma(a_{21} - b_2 k_1) + \beta(a_{22} - b_2 k_2) \right) < 0. \end{aligned}$$

On the other hand, if $\gamma = 0$, $\cos(\theta_1^*) = 0$, and (14) is equivalent to

$$\sin(\theta_1^*) \left(\frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) (a_{12} - b_1 k_2) + \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) (a_{22} - b_2 k_2) \right) < 0.$$

Note also that $\gamma = \beta = 0$ leads to no undesired equilibria. Indeed, using the definitions of β and γ ,

$$\gamma(a_{11} - b_1 k_1) + \beta(a_{12} - b_1 k_2) = b_1 (\det(A) - \gamma k_1 - \beta k_2), \quad (15a)$$

$$\gamma(a_{21} - b_2 k_1) + \beta(a_{22} - b_2 k_2) = b_2 (\det(A) - \gamma k_1 - \beta k_2), \quad (15b)$$

which implies that $b_1 \det(A) = b_2 \det(A) = 0$ if $\gamma = \beta = 0$. If $\det(A) = 0$, then $\det(A + BK) = \det(A) - \gamma k_1 - \beta k_2 = 0$ for all K , which contradicts Assumption 2. Hence, $b_1 = b_2 = 0$, in which case (6) implies that $(A - BK)\mathbf{x}_*^{(1)} = (A - BK)\mathbf{x}_*^{(2)} = \mathbf{0}_2$, which can only hold if $\mathbf{x}_*^{(1)} = \mathbf{x}_*^{(2)} = \mathbf{0}_2$ and they are not undesired equilibrium. Hence the rest of the proof focuses on the case $\beta^2 + \gamma^2 > 0$. By the Routh-Hurwitz criterion [26], since $A - BK$ is Hurwitz and is a 2×2 matrix, $\det(A - BK) = \det(A) - \gamma k_1 - \beta k_2 > 0$. Hence, if $\gamma \neq 0$, by using (15), (14) is equivalent to

$$\frac{\cos(\theta_1^*)}{\gamma} \left(\frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) b_1 + \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) b_2 \right) < 0.$$

On the other hand, if $\gamma = 0$, by using (15), (14) is equivalent to

$$\sin(\theta_1^*) \left(\frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) \frac{b_1}{\beta} + \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) \frac{b_2}{\beta} \right) < 0.$$

Using (11), we get that if $\gamma \neq 0$, (14) is equivalent to

$$\left(\frac{\cos(\theta_1^*)}{\gamma} \right)^2 \left(-r(\theta_1^*)(b_1 \gamma + b_2 \beta) + r'(\theta_1^*)(-b_1 \beta + b_2 \gamma) \right) < 0, \quad (16)$$

whereas if $\gamma = 0$, $\cos(\theta_1^*) = 0$, which means that $\sin^2(\theta_1^*) = 1$, and again using (11), (14) is equivalent to

$$\frac{1}{\beta} \left(-r'(\theta_1^*)b_1 - r(\theta_1^*)b_2 \sin^2(\theta_1^*) \right) = \frac{1}{\beta} \left(-r'(\theta_1^*)b_1 - r(\theta_1^*)b_2 \right) < 0, \quad (17)$$

Note that similar expressions to (16) and (17) hold for θ_2^* . In particular, this means that if $\gamma \neq 0$, the sign of $-r(\theta_1^*)(b_1\gamma + b_2\beta) + r'(\theta_1^*)(-b_1\beta + b_2\gamma)$ and $-r(\theta_2^*)(b_1\gamma + b_2\beta) + r'(\theta_2^*)(-b_1\beta + b_2\gamma)$ is the same, and if $\gamma = 0$, the sign of $-r'(\theta_1^*)b_1 - r(\theta_1^*)b_2$ and $-r'(\theta_2^*)b_1 - r(\theta_2^*)b_2$ is the same.

Next, we compute the eigenvalues of the Jacobian of (3) at $\mathbf{x}_*^{(1)}$. By [17, Proposition 11], one of its eigenvalues is $-\alpha'(0)$. By using the expression of the Jacobian of (3) at $\mathbf{x}_*^{(1)}$ also provided in [17, Proposition 11], and using the fact that the trace is the sum of the eigenvalues, we obtain that the other eigenvalue is

$$\lambda_{\mathbf{x}_*^{(1)}} := \frac{\beta \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(1)}) + \gamma \frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)})}{b_1 \frac{\partial h}{\partial x_1}(\mathbf{x}_*^{(1)}) + b_2 \frac{\partial h}{\partial x_2}(\mathbf{x}_*^{(2)})}.$$

Using (11), we get that if $\gamma \neq 0$,

$$\lambda_{\mathbf{x}_*^{(1)}} = \frac{r(\theta_1^*)(\beta^2 + \gamma^2)}{r(\theta_1^*)(b_1\gamma + b_2\beta) - r'(\theta_1^*)(-b_1\beta + b_2\gamma)}, \quad (18)$$

whereas if $\gamma = 0$,

$$\lambda_{\mathbf{x}_*^{(1)}} = \frac{\beta r(\theta_1^*)}{r'(\theta_1^*)b_1 + r(\theta_1^*)b_2}. \quad (19)$$

Now, since if $\gamma \neq 0$, the sign of $-r(\theta_1^*)(b_1\gamma + b_2\beta) + r'(\theta_1^*)(-b_1\beta + b_2\gamma)$ and $-r(\theta_2^*)(b_1\gamma + b_2\beta) + r'(\theta_2^*)(-b_1\beta + b_2\gamma)$ is the same, and if $\gamma = 0$ the sign of $r'(\theta_1^*)b_1 + r(\theta_1^*)b_2$ and $r'(\theta_2^*)b_1 + r(\theta_2^*)b_2$ is the same, $\mathbf{x}_*^{(1)}$ and $\mathbf{x}_*^{(2)}$ have the same stability properties. Note also that since r is strictly positive by definition, $r(\theta_1^*) > 0$, and since we are discussing the case $\beta^2 + \gamma^2 > 0$, $\mathbf{x}_*^{(1)}$ and $\mathbf{x}_*^{(2)}$ cannot be degenerate undesired equilibria. However, according to Proposition 4.10, since \mathcal{C} is compact, connected, and contains the origin, if there exist two undesired equilibria, one must be a saddle point and the other one must be asymptotically stable, which is a contradiction with the fact that $\mathbf{x}_*^{(1)}$ and $\mathbf{x}_*^{(2)}$ have the same stability properties. This implies that $\mathbf{x}_*^{(1)}$ and $\mathbf{x}_*^{(2)}$ cannot be undesired equilibria and therefore (8) does not have any, i.e., $\hat{\mathcal{E}} = \emptyset$. \square

Figure 7 illustrates different examples of safe sets satisfying the assumptions in Proposition 5.1. The following example shows that the conclusions of Proposition 5.1 do not hold if (7) is fully actuated.

Example 5.2. (Undesired equilibrium in convex and bounded safe set, fully actuated system): Consider the planar single integrator, i.e., (8) with $A = \mathbb{0}_2$, $B = \mathbf{I}_2$, and let

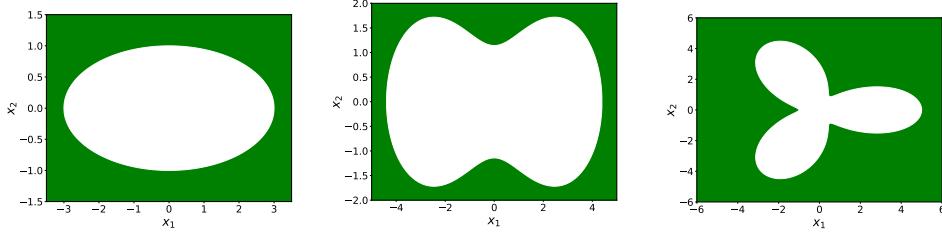


Fig. 7 Examples of safe sets satisfying the assumptions of Proposition 5.1. The unsafe region is colored in green and the safe set in white. For the left figure, $r(\theta) = \frac{1}{\sqrt{3 \cos^2(\theta) + \sin^2(\theta)}}$, for the center figure, $r(\theta) = 3\sqrt{\cos(2\theta) + \sqrt{(1.03)^4 - \sin^2(\theta)}}$, and for the right figure, $r(\theta) = 3 + 2 \cos(3\theta)$.

$K = \begin{bmatrix} 5 & -8 \\ 2 & -3 \end{bmatrix}$, $G(\mathbf{x}) = \mathbf{I}_2$, $h(\mathbf{x}) = 10 - x_1^2 - (x_2 - 2)^2$, and $\alpha(s) = 50s$. Note that $-K$ is Hurwitz. By numerically solving the conditions in (6), it follows that $(3, 3)$ and $(3.161, 2.123)$ are undesired equilibria of (8). Moreover, using the expression of the Jacobian in [17, Proposition 11], we deduce that $(3, 3)$ is asymptotically stable and $(3.161, 2.123)$ is a saddle point. This is illustrated in Figure 8. \triangle

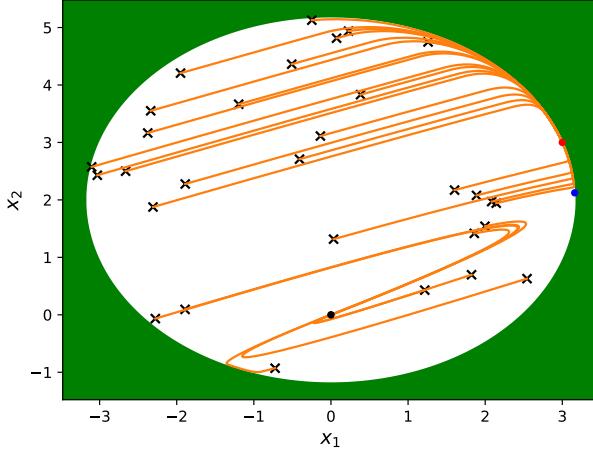


Fig. 8 Plot of different trajectories (in orange) for Example 5.2. The unsafe set is colored in green. Black crosses denote initial conditions, the black dot denotes the origin, the red dot denotes an asymptotically stable undesired equilibrium, and the blue dot denotes an undesired equilibrium that is a saddle point.

We also note that a critical assumption in Proposition 5.1 is that the boundary of the safe set can be parametrized as in (10). In the following example, we show that if such a parametrization does not exist, the closed-loop system (8) can have undesired equilibria even if the system is underactuated and h is a strict CBF.

Example 5.3. (Undesired equilibrium in sets not parametrizable in polar coordinates): Consider (8) with

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}^T,$$

$$h(\mathbf{x}) = b^4 - \left\| \mathbf{x} - \begin{bmatrix} a - c_1 \\ c_2 \end{bmatrix} \right\|^2 \left\| \mathbf{x} + \begin{bmatrix} a + c_1 \\ -c_2 \end{bmatrix} \right\|^2, \quad \alpha(s) = 50s,$$

where $a_{11} = 1.878$, $a_{12} = -6.247$, $a_{21} = -3.189$, $a_{22} = 6.731$, $b_1 = 4.166$, $b_2 = -8.172$, $k_1 = 1.495$, $k_2 = -1.515$, $a = 6.587$, $b = 6.591$, $c_1 = -5$, $c_2 = 0$. The boundary of \mathcal{C} cannot be described as in (10). Note that since B is a column vector, (8) is independent of G . By numerically solving the conditions in (6), it follows that $(5.431, 0.487)$ and $(4.651, 0.417)$ are undesired equilibria of (8). Moreover, using the expression of the Jacobian in [17, Proposition 10], it follows that $(5.431, 0.487)$ is asymptotically stable and $(4.651, 0.417)$ is a saddle point. This is illustrated in Figure 9. \triangle

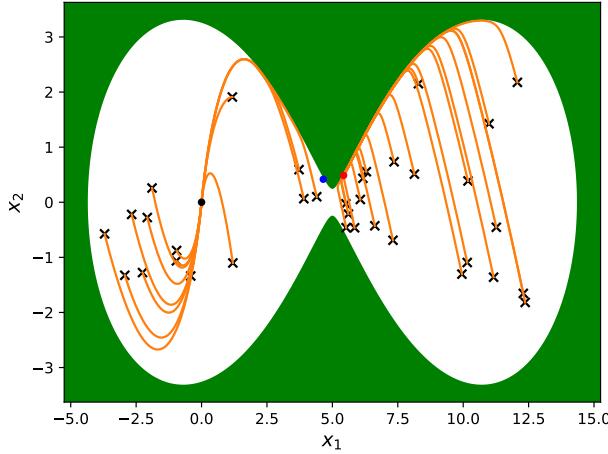


Fig. 9 Plot of different trajectories (in orange) for Example 5.3. The unsafe set is colored in green. Black crosses denote initial conditions, the black dot denotes the origin, the red dot denotes an asymptotically stable undesired equilibrium, and the blue dot denotes an undesired equilibrium that is a saddle point.

The stability properties of undesired equilibria remain the same for a large class of CBFs, cf. [17, Proposition 6.2], and therefore the asymptotically stable undesired equilibria defined in Examples 5.2 and 5.3 exist for a large class of CBFs, not just the ones defined therein.

5.2 Bounded unsafe set

Here we study the case where the unsafe set is bounded. In this case, recall (cf. our discussion after Proposition 4.9) that [25, Proposition 3] implies that there does not exist a safe globally asymptotically stabilizing controller. If no limit cycles exist (for example, under the conditions of Proposition 4.5), by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], this implies that there must exist at least one undesired equilibrium. We next provide two examples that illustrate how system (8) for linear planar plants can give rise in this case to asymptotically stable undesired equilibria, similarly to Examples 5.2 and 5.3.

Example 5.4. (Nonconvex obstacle with asymptotically stable undesired equilibria): Consider the single integrator in the plane, i.e., (8) with $A = \mathbb{0}_2$, $B = \mathbf{I}_2$, and

$$K = \begin{bmatrix} -k_1 & 0 \\ 0 & -k_2 \end{bmatrix}, \quad h(\mathbf{x}) = -b^4 + \left\| \mathbf{x} - \begin{bmatrix} a \\ c_2 \end{bmatrix} \right\|^2 \left\| \mathbf{x} + \begin{bmatrix} a \\ -c_2 \end{bmatrix} \right\|^2, \quad \alpha(s) = 10s,$$

with $k_1 > 0$, $k_2 > 0$, $a > 0$, $b > 0$, $c_2 > 0$ and $\frac{b}{a} \in (1, \sqrt{2})$. It follows that $\mathbf{x}_* = (0, \sqrt{-a^2 + b^2} + c_2)$ and $\delta_{\mathbf{x}_*} = -\frac{c_2 + \sqrt{-a^2 + b^2}}{2\sqrt{-a^2 + b^2}b^2} < 0$ solve (6) and therefore $\mathbf{x}_* \in \hat{\mathcal{E}}$ is an undesired equilibrium. Moreover, by leveraging [17, Proposition 11], the Jacobian of (8) evaluated at \mathbf{x}_* is equal to

$$J(\mathbf{x}_*) = \begin{bmatrix} k_2(b^2 - 2a^2)\frac{c_2 + \sqrt{-a^2 + b^2}}{\sqrt{-a^2 + b^2}b^2} & 0 \\ 0 & -\alpha'(0) \end{bmatrix}.$$

Note that $J(\mathbf{x}_*)$ is a diagonal matrix and since $\alpha'(0) > 0$, $k_2 > 0$, $c_2 > 0$, and $\frac{b}{a} < \sqrt{2}$, all eigenvalues of $J(\mathbf{x}_*)$ are negative and hence \mathbf{x}_* is asymptotically stable. Figure 10 shows different trajectories of the closed-loop system obtained with $k_1 = 1$, $k_2 = 1$, $a = 3$, $b = 1.05a$, $c_2 = 4$.

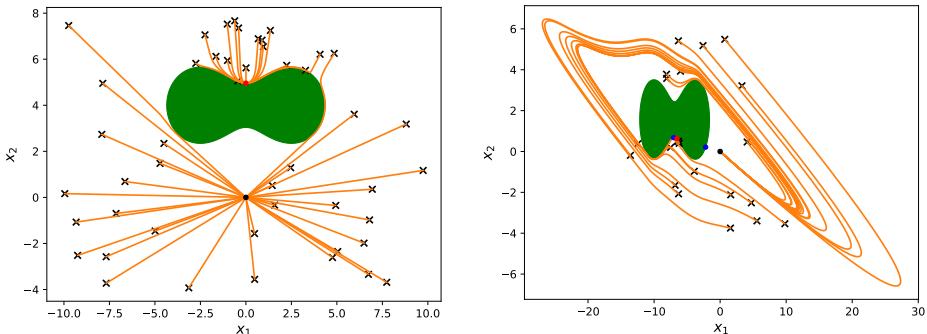


Fig. 10 Plot of different trajectories of (8) for the single integrator system (left) and the underactuated system in (right) in Example 5.4. The unsafe set is colored in green. Trajectories are depicted in orange. Black crosses denote initial conditions, the black dot denotes the origin, and the red dot denotes the asymptotically stable undesired equilibrium.

Next we present a similar example for a linear underactuated system. Consider (8) with

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix},$$

$$h(\mathbf{x}) = -b^4 + \left\| \mathbf{x} - \begin{bmatrix} a - c_1 \\ c_2 \end{bmatrix} \right\|^2 \left\| \mathbf{x} + \begin{bmatrix} a + c_1 \\ -c_2 \end{bmatrix} \right\|^2, \quad \alpha(s) = 10s,$$

where $a_{11} = 0.268$, $a_{12} = 2.866$, $a_{21} = 0.151$, $a_{22} = 1.526$, $b_1 = 0.350$, $b_2 = -0.151$, $k_1 = -41.72$, $k_2 = -174.8$, $a = 3.684$, $b = 3.785$, $c_1 = 6.891$, $c_2 = 1.565$. Note that since B is a column vector, (8) is independent of G . By numerically solving the conditions in (6), we obtain three undesired equilibria, two of which, located at $(-7.062, 0.682)$ and $(-2.197, 0.212)$, are saddle points and another, located at $(-6.519, 0.629)$, which is asymptotically stable. Figure 10 shows that some of the trajectories of (8) converge to $(-6.519, 0.629)$, while some others converge to the origin. \triangle

Example 5.4 shows that the closed-loop system (8) can have asymptotically stable undesired equilibria. A natural question is to ask whether changing the nominal controller would make these equilibria disappear. The following example answers this question in the negative, showing that asymptotically stable undesired equilibria can exist for all nominal stabilizing controllers.

Example 5.5. (Asymptotically stable equilibria can exist for all linear nominal controllers): Consider a linear underactuated system:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}.$$

Given $c_1 > 0$, $c_2 > 0$, $r_1 > 0$, $r = c_1 - r_1$ and $R = c_1 + r_1$, suppose that $R < c_2$ and $r < c_1$. Consider also the sets

$$\mathcal{O}_1 := \{(x_1, x_2) \in \mathbb{R}^2 : h_1(x) := -(x_1 + c_1)^2 - (x_2 - c_2)^2 + r_1^2 > 0, \quad h_2(x) = x_2 - c_2 > 0\},$$

$$\mathcal{O}_2 := \{(x_1, x_2) \in \mathbb{R}^2 : h_3(x) := -x_1^2 - (x_2 - c_2)^2 + (c_1 + r_1)^2 > 0,$$

$$h_4(x) = -x_1^2 - (x_2 - c_2)^2 + (c_1 - r_1)^2 < 0, \quad h_5(x) = c_2 - x_2 > 0\},$$

$$\mathcal{O}_3 := \{(x_1, x_2) \in \mathbb{R}^2 : h_6(x) = r_1^2 - (x_1 - c_1)^2 - x_2^2 > 0, \quad h_7(x) = x_2 - c_2 > 0\},$$

and define the safe set $\mathcal{C} = \mathbb{R}^2 \setminus (\mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{O}_3)$, and $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ be such that $\mathcal{C} = \{x \in \mathbb{R}^2 : h(x) \geq 0\}$, $\partial\mathcal{C} = \{x \in \mathbb{R}^2 : h(x) = 0\}$. By letting $a_{11} = -0.169$, $a_{21} = -1.989$, $a_{12} = -3$, $a_{22} = -1.4$, $b_1 = -2.355$, $b_2 = -1.707$, $c_1 = 6.1$, $c_2 = 10.027$, $r_1 = 2.129$, it follows that $\mathbf{x}_* = (3.157, 7.619)$ is an asymptotically stable undesired equilibrium for any nominal stabilizing controller $\mathbf{u} = -K\mathbf{x}$. We provide a more detailed justification about this in Example 8.6 of the Appendix. We note also that our argument does not preclude the existence of nonlinear nominal stabilizing controllers for which there does not exist any asymptotically stable undesired equilibrium. \triangle

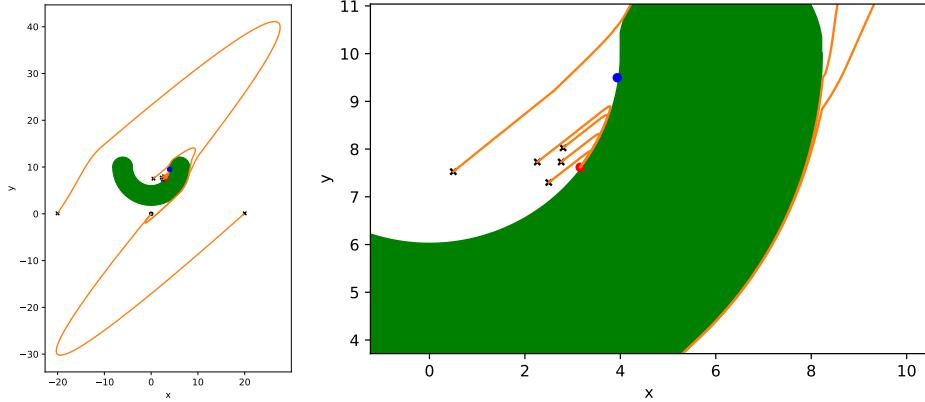


Fig. 11 Plot of different trajectories of (8) for the underactuated system in Example 5.4. The unsafe set is colored in green. Trajectories are depicted in orange. Black crosses denote initial conditions, the black dot denotes the origin, the red dot denotes an asymptotically stable undesired equilibrium, and the blue dot denotes an undesired equilibrium that is a saddle point. The plot on the left is a zoomed-in version of the plot on the right.

5.3 Ellipsoidal unsafe sets

Despite the title of this section, in the following we focus on studying the dynamical properties of safety filters for LTI systems and *circular* obstacles. Accordingly, we consider the circular unsafe set:

$$\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^2 : h(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}_c\|^2 - r^2 \geq 0\},$$

with $\mathbf{x}_c = (x_{c,1}, x_{c,2}) \in \mathbb{R}^2$ the center. This is justified by the following result, which shows that the undesired equilibria of (8) and their stability properties for general ellipsoidal obstacles are equivalent to those of a system with circular obstacles.

Proposition 5.6. (Safety filters with ellipsoidal and circular obstacles have the same dynamical properties): *Let $\mathbf{x}_c \in \mathbb{R}^2$, $P \in \mathbb{R}^{2 \times 2}$ positive definite, $h(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_c)^T P (\mathbf{x} - \mathbf{x}_c) - 1$, and $\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) \geq 0\}$. Suppose that h is a strict CBF and Assumption 1 holds. Further suppose that $P = E^T E$, with $E \in \mathbb{R}^{2 \times 2}$ also positive definite, and define $\hat{\mathbf{x}}_c = E\mathbf{x}_c$, $\hat{h}(\hat{\mathbf{x}}) = (\hat{\mathbf{x}} - \hat{\mathbf{x}}_c)^T (\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) - 1$ and $\hat{\mathcal{C}} = \{\mathbf{x} \in \mathbb{R}^n : \hat{h}(\mathbf{x}) \geq 0\}$. Moreover, let $\hat{A} = EAE^{-1}$, $\hat{B} = EB$, $\hat{G}(\hat{\mathbf{x}}) = G(E^{-1}\hat{\mathbf{x}})$ and $\hat{\eta}(\hat{\mathbf{x}}) = \nabla \hat{h}(\hat{\mathbf{x}})^T (\hat{A} - \hat{B}KE^{-1})\hat{\mathbf{x}} + \alpha(\hat{h}(\hat{\mathbf{x}}))$. Consider the system*

$$\dot{\hat{\mathbf{x}}} = \hat{F}(\hat{\mathbf{x}}) := (\hat{A} - \hat{B}KE^{-1})\hat{\mathbf{x}} + \hat{B}\hat{v}(\hat{\mathbf{x}}), \quad (20)$$

where

$$\hat{v}(\hat{\mathbf{x}}) = \begin{cases} 0, & \text{if } \hat{\eta}(\hat{\mathbf{x}}) \geq 0, \\ -\frac{\hat{\eta}(\hat{\mathbf{x}})\hat{G}(\hat{\mathbf{x}})^{-1}(\hat{\mathbf{x}})\hat{B}^T \nabla \hat{h}(\hat{\mathbf{x}})}{\|\hat{B}^T \nabla \hat{h}(\hat{\mathbf{x}})\|_{\hat{G}(\hat{\mathbf{x}})^{-1}}^2}, & \text{if } \hat{\eta}(\hat{\mathbf{x}}) < 0 \end{cases} \quad (21)$$

Then,

- (i) $\hat{\mathcal{C}}$ is forward invariant under system (20) and \mathcal{C} is forward invariant under system (8);
- (ii) system (20) is locally Lipschitz and system (8) is locally Lipschitz;
- (iii) (A, B) is stabilizable if and only if (\hat{A}, \hat{B}) is stabilizable;
- (iv) $\hat{\mathbf{x}}_* \in \mathbb{R}^2$ is an undesired equilibrium of (20) if and only if $\mathbf{x}_* := E^{-1}\hat{\mathbf{x}}_*$ is an undesired equilibrium of (8);
- (v) the Jacobian of \hat{F} at $\hat{\mathbf{x}}_*$ and the Jacobian of F at \mathbf{x}_* are similar.

Proof. Re: (i), by construction, system (20) satisfies $\nabla \hat{h}(\hat{\mathbf{x}})^T \hat{F}(\hat{\mathbf{x}}) + \alpha(\hat{h}(\hat{\mathbf{x}})) \geq 0$ for all $\hat{\mathbf{x}} \in \hat{\mathcal{C}}$; similarly, system (8) satisfies $\nabla h(\mathbf{x})^T F(\mathbf{x}) + \alpha(h(\mathbf{x})) \geq 0$ for all $\mathbf{x} \in \mathcal{C}$. These two conditions imply, respectively, that $\hat{\mathcal{C}}$ is forward invariant under system (20) and \mathcal{C} is forward invariant under system (8) [3, Theorem 2].

Re: (ii), let us prove that for all $\hat{\mathbf{x}} \in \hat{\mathcal{C}}$, there exists $\hat{\mathbf{u}} \in \mathbb{R}^m$ such that $\nabla \hat{h}(\hat{\mathbf{x}})^T (\hat{A}\hat{\mathbf{x}} + \hat{B}\hat{\mathbf{u}}) + \alpha(\hat{h}(\hat{\mathbf{x}})) > 0$, which implies by [19, Lemma III.2] that (20) is locally Lipschitz. Indeed, suppose that $(EB)^\top (\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) = 0$. Since for any $\hat{\mathbf{x}} \in \mathbb{R}^n$, there exists $\mathbf{x} \in \mathbb{R}^n$ such that $\hat{\mathbf{x}} = E\mathbf{x}$ and h is a strict CBF, we have $0 = (EB)^\top (\hat{\mathbf{x}} - \hat{\mathbf{x}}_c) = B^\top E^\top E(\mathbf{x} - \mathbf{x}_c) = BP(\mathbf{x} - \mathbf{x}_c) = 0$, which means that $(\hat{\mathbf{x}} - \hat{\mathbf{x}}_c)^\top EAE^{-1}\hat{\mathbf{x}} = (\mathbf{x} - \mathbf{x}_c)^\top E^\top EA\mathbf{x} = (\mathbf{x} - \mathbf{x}_c)^\top P^\top A\mathbf{x} > 0$. Hence for all $\hat{\mathbf{x}} \in \hat{\mathcal{C}}$, there exists $\hat{\mathbf{u}} \in \mathbb{R}^m$ such that $\nabla \hat{h}(\hat{\mathbf{x}})^T (\hat{A}\hat{\mathbf{x}} + \hat{B}\hat{\mathbf{u}}) + \alpha(\hat{h}(\hat{\mathbf{x}})) > 0$, from which it follows that (20) is locally Lipschitz.

Re: (iii), this follows from the observation that if $A - BK$ is Hurwitz then $\hat{A} - \hat{B}KE^{-1} = E(A - BK)E^{-1}$ is also Hurwitz. This is because E is nonsingular and similar matrices have the same eigenvalues [27, Corollary 1.3.4].

Re: (iv), note that $\hat{\mathbf{x}}_*$ satisfies the conditions in Lemma 4.1 for (20) if and only if \mathbf{x}_* satisfies the conditions in Lemma 4.1 for (8).

Re: (v), we note that $F(E^{-1}\hat{\mathbf{x}}) = E^{-1}\hat{F}(\hat{\mathbf{x}})$ for any $\hat{\mathbf{x}} \in \mathbb{R}^2$. Since the safety filter is active at undesired equilibria, $\eta(\mathbf{x}_*) < 0$. Now, let J be the Jacobian of (8) at \mathbf{x}_* , and let \hat{J} be the Jacobian of (20) at $\hat{\mathbf{x}}_*$. By the chain rule, $\hat{J} = EJE^{-1}$, which implies that J and \hat{J} are similar. \square

5.3.1 Underactuated LTI planar systems

In the under-actuated case, we write

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad K = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}^T. \quad (22)$$

Throughout the section, we let $\beta = a_{11}b_2 - b_1a_{21}$, $\gamma = a_{22}b_1 - b_2a_{12}$, and $T_3 = -\gamma x_{c,2} + \beta x_{c,1}$. We note also that since in this case G is a scalar, (8) is independent of G . The following results give conditions on h and system (22) that ensure that Assumption 1 holds and h is a strict CBF.

Lemma 5.7 (Conditions for Assumption 1). *Assumption 1 holds if and only if $\|\mathbf{x}_c\|^2 > r^2$.*

The proof of this result is straightforward.

Proposition 5.8 (Conditions for h to be a strict CBF). *Let $\alpha_0 > 0$, $T_1 := b_2\beta + b_1\gamma + \frac{1}{2}\alpha_0(b_2^2 + b_1^2)$, and $T_2 := (\beta x_{c,1} - \gamma x_{c,2})^2 + 2\alpha_0 r^2 T_1$. Suppose that $r > 0$, $b_1^2 + b_2^2 > 0$, $T_1 > 0$, and*

$$\frac{r}{\sqrt{b_2^2 + b_1^2}} > \frac{|T_3| + \sqrt{T_2}}{2T_1}.$$

Then, h is a strict CBF with the linear extended class \mathcal{K}_∞ function $\alpha(s) = \alpha_0 s$.

Proof. We need to ensure that all $\mathbf{x} \in \mathbb{R}^2$ such that $h(\mathbf{x}) \geq 0$ and $B^T(\mathbf{x} - \mathbf{x}_c) = 0$, satisfy $2(\mathbf{x} - \mathbf{x}_c)^T A \mathbf{x} + \alpha(h(\mathbf{x})) > 0$. First suppose $b_1 \neq 0$. Equivalently, we need to ensure that

$$(x_2 - x_{c,2})^2 \left((a_{11} + \frac{\alpha_0}{2}) \frac{b_2^2}{b_1^2} - \frac{b_2}{b_1}(a_{12} + a_{21}) + a_{22} + \frac{\alpha_0}{2} \right) + (x_2 - x_{c,2}) \left(a_{22}x_{c,2} - \frac{b_2}{b_1}a_{11}x_{c,1} - \frac{b_2}{b_1}a_{12}x_{c,2} + a_{21}x_{c,1} \right) - \frac{1}{2}\alpha_0 r^2 > 0 \quad (23)$$

whenever $(x_2 - x_{c,2})^2 \geq r^2 / ((b_2^2/b_1^2) + 1)$. This follows by assumption. The condition $T_1 > 0$ ensures that the coefficient of $x_2 - x_{c,2}$ of (23) is positive, and the condition $T_2 > 0$ ensures that the discriminant of (23) is positive. Now, by calculating the roots of the quadratic equation in $x_2 - x_{c,2}$ we observe that the rest of the conditions in the statement ensure that (23) holds whenever $(x_2 - x_{c,2})^2 \geq r^2 / ((b_2^2/b_1^2) + 1)$. The case $b_1 = 0$ follows by a similar argument. \square

The following result shows that for circular obstacles and linear planar underactuated systems, (6) can be solved explicitly, characterizing the undesired equilibria of the closed-loop system.

Proposition 5.9. (Equilibria in Underactuated Systems with Circular Obstacles): *Suppose that Assumptions 1, 2 and the conditions in Proposition 5.8 hold. Define $\mathbf{x}_{*,+} := (\gamma z_+, \beta z_+)$ and $\mathbf{x}_{*,-} := (\gamma z_-, \beta z_-)$, where*

$$z_+ = \frac{\gamma x_{c,1} + \beta x_{c,2} + \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2}}{\gamma^2 + \beta^2},$$

$$z_- = \frac{\gamma x_{c,1} + \beta x_{c,2} - \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2}}{\gamma^2 + \beta^2}.$$

Then, $\gamma x_{c,1} + \beta x_{c,2} \neq 0$, and

- (i) if $\gamma x_{c,1} + \beta x_{c,2} < 0$, $\hat{\mathcal{E}} = \{\mathbf{x}_{*,+}\}$ is the only undesired equilibrium of the closed-loop system (8). Moreover, $\mathbf{x}_{*,+}$ is a saddle point;
- (ii) if $\gamma x_{c,1} + \beta x_{c,2} > 0$, $\hat{\mathcal{E}} = \{\mathbf{x}_{*,-}\}$ is the only undesired equilibrium of the closed-loop system (8). Moreover, $\mathbf{x}_{*,-}$ is a saddle point.

Proof. The fact that $\gamma x_{c,1} + \beta x_{c,2} \neq 0$ is shown in Lemma 8.5. The same result also implies that the expressions for $\mathbf{x}_{*,+}$ and $\mathbf{x}_{*,-}$ are well defined (note that if

$\gamma^2 + \beta^2 = 0$, then Assumption 2 would not hold). Moreover, it follows from Lemma 4.1 that $\mathcal{E} = \{\mathbf{x}_{*,+}, \mathbf{x}_{*,-}\}$ is the set of potential undesired equilibria for system (8) when the system data is of the form (22). In order to ensure that $\mathbf{x}_{*,+}$ is an undesired equilibrium of the closed-loop system, the condition $(\mathbf{x} - \mathbf{x}_c)^T(A - BK)\mathbf{x}|_{\mathbf{x}=\mathbf{x}_{*,+}} < 0$ should hold. By using the expression of $\mathbf{x}_{*,+}$, the condition is equivalent to

$$z_+ T_4 \left(b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2}) \right) < 0, \quad (24)$$

where $T_4 = a_{11}a_{22} - a_{12}a_{21} - k_1\gamma - k_2\beta$. Since $A - BK$ is Hurwitz, $a_{11}a_{22} - a_{12}a_{21} - \gamma k_1 - \beta k_2 > 0$. This implies that $T_4 > 0$ and therefore (24) is equivalent to

$$z_+ (b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2})) < 0. \quad (25)$$

Now, let us show that $b_1(\gamma z_+ - x_{c,1}) + b_2(\beta z_+ - x_{c,2}) > 0$. Indeed, this is equivalent to

$$T_3(\gamma b_2 - \beta b_1) + (\gamma b_1 + \beta b_2) \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2} > 0,$$

and since $\gamma b_1 + \beta b_2 > 0$ as argued in the proof of Lemma 8.5, this could only not hold if $T_3(\gamma b_2 - \beta b_1) < 0$ and $(\gamma b_1 + \beta b_2)^2(r^2(\gamma^2 + \beta^2) - T_3^2) \leq T_3^2(\gamma b_2 - \beta b_1)^2$. However, this last condition is in contradiction with (28) in the proof of Lemma 8.5. Therefore, (25) holds if and only if $z_+ < 0$, which is equivalent to: $\gamma x_{c,1} + \beta x_{c,2} < 0$ and

$$|\gamma x_{c,1} + \beta x_{c,2}| > \sqrt{r^2(\gamma^2 + \beta^2) - T_3^2}. \quad (26)$$

Note that since $r^2(\gamma^2 + \beta^2) - T_3^2 = (x_{c,1}\gamma + x_{c,2}\beta)^2 - (\gamma^2 + \beta^2)(x_{c,1}^2 + x_{c,2}^2 - r^2) < (x_{c,1}\gamma + x_{c,2}\beta)^2$ (where in the last inequality we have used the fact that $x_{c,1}^2 + x_{c,2}^2 > r^2$), it follows that the last of the inequalities in (26) always holds. Therefore, $\mathbf{x}_{*,+}$ is an undesired equilibrium of the closed-loop system if and only if $\gamma x_{c,1} + \beta x_{c,2} < 0$. The fact that $\mathbf{x}_{*,+}$ is the unique undesired equilibrium and is a saddle point follows from Proposition 4.9. An analogous argument shows that $\mathbf{x}_{*,-}$ is the unique undesired equilibrium if and only if $\gamma x_{c,1} + \beta x_{c,2} > 0$, in which case it is a saddle point. \square

Note that the statement of Proposition 5.9 is consistent with Proposition 4.9, since it also states that, provided that all the undesired equilibria are not degenerate, their number is odd.

Remark 5.10. (Almost global asymptotic stability): The Stable Manifold Theorem [18, Ch. 2.7] ensures that if \mathbf{x}_* is a saddle point in \mathbb{R}^2 , the local stable manifold is 1-dimensional. Therefore, it has measure zero. Moreover, the global stable manifold must also have measure zero. If this were not the case, solutions would have to intersect. However this is not possible due to the uniqueness of solutions. Hence the global stable manifold is exactly equal to $\{\mathbf{x}_0 \in \mathbb{R}^n : \lim_{t \rightarrow \infty} \mathbf{x}(t; \mathbf{x}_0) = \mathbf{0}_n\}$. It follows that the set of initial conditions whose associated trajectory converges to \mathbf{x}_* has measure zero. Hence, by appropriately tuning the class \mathcal{K}_∞ function to rule out limit cycles (cf. Proposition 4.5), the Poincaré-Bendixson Theorem [22, Chapter 7, Thm.

\tilde{A} diagonalizable,	SP	DE	ASE	\tilde{A} non-diagonalizable	SP	DE	ASE
$(\mathbf{v}_i^T \mathbf{v}_j)^2 < 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \ \mathbf{x}_c\ ^2}$	1	0	0	$(\mathbf{v}_1^T \mathbf{v}_2)^2 < 1 - \frac{r^2}{\lambda^2 \ \mathbf{x}_c\ ^2}$	1	0	0
$(\mathbf{v}_i^T \mathbf{v}_j)^2 = 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \ \mathbf{x}_c\ ^2}$	1	1	0	$(\mathbf{v}_1^T \mathbf{v}_2)^2 = 1 - \frac{r^2}{\lambda^2 \ \mathbf{x}_c\ ^2}$	1	1	0
$(\mathbf{v}_i^T \mathbf{v}_j)^2 > 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \ \mathbf{x}_c\ ^2}$	2	0	1	$(\mathbf{v}_1^T \mathbf{v}_2)^2 > 1 - \frac{r^2}{\lambda^2 \ \mathbf{x}_c\ ^2}$	2	0	1

(a)

(b)

Table 1 Characterization of undesired equilibria (SP: saddle point, DE: degenerate equilibrium, ASE: asymptotically stable equilibrium) when \mathbf{x}_c is an eigenvector of \tilde{A} . In (a), \tilde{A} is diagonalizable, i.e., $\tilde{A}\mathbf{x}_c = \lambda_i \mathbf{x}_c$, $\mathbf{v}_i = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\tilde{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$, $\|\mathbf{v}_j\| = 1$, $i, j = \{1, 2\}$, and $\{\mathbf{v}_i, \mathbf{v}_j\}$ linearly independent. We assume that $\frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2} \neq 1$. In (b), \tilde{A} is not diagonalizable, i.e., $\tilde{A}\mathbf{v}_2 = \lambda \mathbf{v}_2 + \mathbf{v}_1$, $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\tilde{A}\mathbf{x}_c = \lambda \mathbf{x}_c$, and $\|\mathbf{v}_2\| = 1$.

4.1] implies that the origin is almost globally asymptotically stable (i.e., asymptotically stable with a region of attraction equal to \mathbb{R}^2 minus a set of measure zero). \square

5.3.2 Fully actuated LTI planar systems

Here we consider the system (8) and assume that $n = 2$, $m = 2$, and $B \in \mathbb{R}^{2 \times 2}$ is invertible; in this case h is a strict CBF and Assumption 2 is satisfied. Throughout the section, $\tilde{A} = A - BK$. The following result summarizes the different possible undesired equilibria of (8) in the special case where \mathbf{x}_c , the center of the circular unsafe set, is an eigenvector of \tilde{A} .

Proposition 5.11. (Characterization of undesired equilibria): *Suppose Assumption 1 is satisfied, B is invertible, and \tilde{A} is Hurwitz. Suppose also that \mathbf{x}_c is an eigenvector of \tilde{A} . Then one of the following is true:*

- (i) $|\mathcal{E}| = 2$, $|\hat{\mathcal{E}}| = 1$, and $\hat{\mathcal{E}}$ consists of a degenerate equilibrium;
- (ii) $|\mathcal{E}| = 2$, $|\hat{\mathcal{E}}| = 1$, and $\hat{\mathcal{E}}$ consists of a saddle point;
- (iii) $|\mathcal{E}| = 3$, $|\hat{\mathcal{E}}| = 2$, and $\hat{\mathcal{E}}$ consists of a saddle point and a degenerate equilibrium;
- (iv) $|\mathcal{E}| = 4$, $|\hat{\mathcal{E}}| = 3$, and $\hat{\mathcal{E}}$ consists of an asymptotically stable equilibrium and two saddle points.

The proof of Proposition 5.11 is provided in the Appendix. Table 1 outlines the majority of the cases discussed in the proof of Proposition 5.11. For different numerical examples illustrating the different cases outlined in Table 1, we refer the reader to [1, Figure 1]. We build on this result to show that the eigenvalues of \tilde{A} do not determine the stability properties of undesired equilibria.

Proposition 5.12. (Spectrum of \tilde{A} does not determine stability properties of undesired equilibria): *Suppose Assumption 1 is satisfied and B is invertible. Then for any given negative λ_1 and λ_2 , there exists $K_1 \in \mathbb{R}^{2 \times 2}$ and $K_2 \in \mathbb{R}^{2 \times 2}$ in the set $\{K \in \mathbb{R}^{2 \times 2} : \lambda_1, \lambda_2 = \text{spec}(A - BK)\}$, such that (8) with $K = K_1$ has an undesired asymptotically stable equilibrium and (8) with $K = K_2$, has a single undesired equilibrium, which is a saddle point.*

Proof. First let us construct K_1 . Start by supposing that $\lambda_1 = \lambda_2$. Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$ and $\theta = \arccos \sqrt{Q_1}$, $Q_1 := \max\{0, 1 - \frac{r^2}{2\lambda_1^2 \|\mathbf{x}_c\|}\}$. Then, $\mathbf{v}_1^T \mathbf{v}_2 = \cos(\theta) \|\mathbf{v}_1\|^2 = \cos(\theta)$. We let

$$K_1 := B^{-1} \left(A - [\mathbf{v}_1 \ \mathbf{v}_2] \begin{bmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{bmatrix} [\mathbf{v}_1 \ \mathbf{v}_2]^{-1} \right).$$

By the third row of Table 1(b), (8) has an undesired asymptotically stable equilibrium. Next, suppose that $\lambda_1 \neq \lambda_2$. Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\mathbf{v}_2 = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \mathbf{v}_1$, where $\theta = \arccos \sqrt{Q_2}$, $Q_2 := \max\{0, 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{2\lambda_1^2 \|\mathbf{x}_c\|^2}\}$. Then $\mathbf{v}_1^T \mathbf{v}_2 = \cos \theta \|\mathbf{v}_1\|_2^2 = \cos \theta$ and

$$K_1 := B^{-1} \left(A - [\mathbf{v}_1 \ \mathbf{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} [\mathbf{v}_1 \ \mathbf{v}_2]^{-1} \right).$$

By the third row of Table 1(a), (8) has an undesired asymptotically stable equilibrium.

To construct K_2 , we assume without loss of generality that $\lambda_1 \leq \lambda_2$. Note that since both λ_1 and λ_2 are negative, $1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2} > 0$. Let $\mathbf{v}_i = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, $\mathbf{v}_j = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \mathbf{v}_i$, so that $\mathbf{v}_i^T \mathbf{v}_j = 0$ and define

$$K_2 := B^{-1} \left(A - [\mathbf{v}_i \ \mathbf{v}_j] \begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_j \end{bmatrix} [\mathbf{v}_i \ \mathbf{v}_j]^{-1} \right).$$

By the first row of Table 1(a), (8) has a single undesired equilibrium and it is a saddle point. \square

Interestingly, even though one can characterize the global stability properties of the origin based on the eigenvalues of \tilde{A} for the system without a safety filter, this is no longer the case for the system with a safety filter. On the other hand, as a consequence of Proposition 5.12, we deduce that it is always possible to choose a nominal controller $\mathbf{u} = -K\mathbf{x}$ such that \tilde{A} has negative eigenvalues and the set of trajectories of (8) that do not converge to the origin has measure zero (cf. Remark 5.10). Note that, as shown in [17, Proposition 10] and Table 1, the extended class \mathcal{K}_∞ function only affects the rate of decay in the stable manifold of the undesired equilibria and it does not affect the existence and stability of undesired equilibria. Therefore, the choice of nominal controller $\mathbf{u} = -K\mathbf{x}$ determines in which of the cases we fall into. Ideally, such nominal controller should be designed so that there exists only one undesired equilibrium and it is a saddle point.

We conclude this section studying the case when \mathbf{x}_c is not an eigenvector of \tilde{A} , which requires a more involved technical analysis. The following result characterizes the number of undesired equilibria under appropriate sufficient conditions.

Proposition 5.13. (Number of undesired equilibria when \mathbf{x}_c is not an eigenvector): Suppose Assumption 1 is satisfied, B is invertible, $G(\mathbf{x}) = B^\top B$, \tilde{A} is Hurwitz and \mathbf{x}_c

is not an eigenvector of \tilde{A} . Then $1 \leq |\hat{\mathcal{E}}| \leq 3$ and $|\mathcal{E} \setminus \hat{\mathcal{E}}| \geq 1$. In addition, if $\lambda_1 \leq \lambda_2$, there exists $\mathbf{x}_* \in \hat{\mathcal{E}}$ with indicator $\delta < \frac{\lambda_1}{2}$.

The proof of Proposition 5.13 is given in the Appendix.

The following result establishes the stability properties of undesired equilibria in the case where \mathbf{x}_c is not an eigenvector under some additional assumptions.

Proposition 5.14. (Stability properties of undesired equilibria when \mathbf{x}_c is not an eigenvector): *Suppose Assumption 1 is satisfied, B is invertible, $G(\mathbf{x}) = B^\top B$, \tilde{A} is Hurwitz with two real eigenvalues $\lambda_1 < \lambda_2$ and \mathbf{x}_c is not an eigenvector of \tilde{A} . Then there is no undesired equilibrium with indicator $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. Moreover, if \mathbf{v}_1 and \mathbf{v}_2 are the eigenvectors associated with λ_1 and λ_2 , respectively, and $\mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$, and $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$, the following holds.*

- (i) *If $\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, then for any undesired equilibrium \mathbf{x}_* with indicator $\delta_{\mathbf{x}_*}$ such that $\delta < \frac{\lambda_1}{2}$, \mathbf{x}_* is a saddle point.*
- (ii) *If $\beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 + \beta_2^2 \geq 0$, then for any undesired equilibrium \mathbf{x}_* with indicator $\delta_{\mathbf{x}_*}$ such that $\frac{\lambda_2}{2} < \delta < 0$, \mathbf{x}_* is asymptotically stable.*
- (iii) *Define $F_1 : \mathbb{R} \rightarrow \mathbb{R}$ as:*

$$F_1(\delta) := -|\lambda_1 - 2\delta|^2 |\lambda_2 - 2\delta|^2 r^2 + |\lambda_1 \beta_1|^2 |\lambda_2 - 2\delta|^2 + |\lambda_2 \beta_2|^2 |\lambda_1 - 2\delta|^2 \quad (27)$$

$$+ \lambda_1^* \beta_1^* \lambda_2 \beta_2 (\lambda_2 - 2\delta)^* (\lambda_1 - 2\delta) \mathbf{v}_1^* \mathbf{v}_2 + \lambda_1 \beta_1 \lambda_2^* \beta_2^* (\lambda_2 - 2\delta) (\lambda_1 - 2\delta)^* \mathbf{v}_2^* \mathbf{v}_1.$$

If the third-order polynomial $\frac{dF_1(\delta)}{d\delta}$ has only one real root¹ and $\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \geq 0$, then there exists only one undesired equilibrium and it is a saddle point.

The proof of Proposition 5.14 is given in the Appendix.

6 Conclusions

This work has characterized different dynamical properties of CBF-based safety filters. We have provided a characterization of the undesired equilibria in the corresponding closed-loop system through the solution of an algebraic equation. Next, we have shown that by appropriately designing the parameters of the safety filter, the closed-loop trajectories are bounded and in the planar case, no limit cycles exist. We have shown through a counterexample that in dimension greater than 2, limit cycles can exist and may not be removed by changing the parameters of the safety filter. Moreover, for general planar systems under general assumptions, we have characterized the parity of the number of undesired equilibria, as well as the number of such undesired equilibria that are saddle points. Finally, for linear planar systems, we have shown that if the system is underactuated and the safe set is parametrizable in polar coordinates, undesired equilibria do not exist, but if any of these conditions fail, undesired equilibria may exist and might even be asymptotically stable. In the special case of ellipsoidal obstacles, we have provided explicit expressions for the undesired equilibria, and characterized their stability properties. Future work will focus on characterizing

¹For third-order polynomial $ax^3 + bx^2 + cx + d$, its discriminant is defined as $18abcd - 4b^3d + b^2c^2 - 4ac^3 - 27a^2d^2$. If $a \neq 0$ and the discriminant is negative, the third-order polynomial only has one real root.

more explicitly the regions of attraction of the origin and the different undesired equilibria by using other techniques like normal forms, utilizing Euler indices and other methods from algebraic topology to improve our understanding of undesired equilibria, and performing a similar analysis for other CBF-based control designs from the literature, such as ones leveraging control Lyapunov functions.

7 Acknowledgements

This work was supported by AFOSR Award FA9550-23-1-0740.

References

- [1] Chen, Y., Mestres, P., Dall’Anese, E., Cortés, J.: Characterization of the dynamical properties of safety filters for linear planar systems. In: IEEE Conf. on Decision and Control, Milan, Italy, pp. 2397–2402 (2024)
- [2] Wieland, P., Allgöwer, F.: Constructive safety using control barrier functions. IFAC Proceedings Volumes **40**(12), 462–467 (2007)
- [3] Ames, A.D., Coogan, S., Egerstedt, M., Notomista, G., Sreenath, K., Tabuada, P.: Control barrier functions: theory and applications. In: European Control Conference, Naples, Italy, pp. 3420–3431 (2019)
- [4] Xiao, W., Cassandras, C.G., Belta, C.: Safe Autonomy with Control Barrier Functions: Theory and Applications. Synthesis Lectures on Computer Science. Springer, New York (2023)
- [5] Reis, M.F., Aguilar, A.P., Tabuada, P.: Control barrier function-based quadratic programs introduce undesirable asymptotically stable equilibria. IEEE Control Systems Letters **5**(2), 731–736 (2021)
- [6] Tan, X., Dimarogonas, D.V.: On the undesired equilibria induced by control barrier function based quadratic programs. Automatica **159**, 111359 (2024)
- [7] Mestres, P., Cortés, J.: Optimization-based safe stabilizing feedback with guaranteed region of attraction. IEEE Control Systems Letters **7**, 367–372 (2023)
- [8] Yi, Y., Koga, S., Gavrea, B., Atanasov, N.: Control synthesis for stability and safety by differential complementarity problem. IEEE Control Systems Letters **7**, 895–900 (2023)
- [9] Notomista, G., Saveriano, M.: Safety of dynamical systems with multiple non-convex unsafe sets using control barrier functions. IEEE Control Systems Letters **6**, 1136–1141 (2022)
- [10] Sontag, E.D.: Mathematical Control Theory: Deterministic Finite Dimensional Systems, 2nd edn. Texts in Applied Mathematics, vol. 6. Springer, New York

(1998)

- [11] Ames, A.D., Xu, X., Grizzle, J.W., Tabuada, P.: Control barrier function based quadratic programs for safety critical systems. *IEEE Transactions on Automatic Control* **62**(8), 3861–3876 (2017)
- [12] Jankovic, M.: Robust control barrier functions for constrained stabilization of nonlinear systems. *Automatica* **96**, 359–367 (2018)
- [13] Castañeda, F., Choi, J.J., Zhang, B., Tomlin, C.J., Sreenath, K.: Pointwise feasibility of Gaussian process-based safety-critical control under model uncertainty. In: *IEEE Conf. on Decision and Control*, Austin, Texas, USA, pp. 6762–6769 (2021)
- [14] Agrawal, D.R., Panagou, D.: Safe control synthesis via input constrained control barrier functions. In: *IEEE Conf. on Decision and Control*, Austin, TX, USA, pp. 6113–6118 (2021)
- [15] Wang, L., Ames, A., Egerstedt, M.: Safety barrier certificates for collisions-free multirobot systems. *IEEE Transactions on Robotics* **33**(3), 661–674 (2017)
- [16] Cortez, W.S., Dimarogonas, D.V.: On compatibility and region of attraction for safe, stabilizing control laws. *IEEE Transactions on Automatic Control* **67**(9), 7706–7712 (2022)
- [17] Chen, Y., Mestres, P., Cortés, J., Dall’Anese, E.: Equilibria and their stability do not depend on the control barrier function in safe optimization-based control. *Automatica* (2024). submitted
- [18] Perko, L.: *Differential Equations and Dynamical Systems*, 3rd edn. Texts in Applied Mathematics, vol. 7. Springer, New York (2000)
- [19] Alyaseen, M., Atanasov, N., Cortés, J.: Continuity and boundedness of minimum-norm CBF-safe controllers. *IEEE Transactions on Automatic Control* **70**(6) (2025). To appear.
- [20] Khalil, H.: *Nonlinear Systems*, 3rd Ed. Prentice Hall, Englewood Cliffs, NJ (2002)
- [21] Meiss, J.D.: *Differential Dynamical Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA (2007)
- [22] Hartman, P.: *Ordinary Differential Equations*, 2nd edn. Classics in Applied Mathematics, vol. 38. SIAM, Philadelphia, PA (2002)
- [23] Bhat, S.P., Bernstein, D.S.: Nontangency-based Lyapunov tests for convergence and stability in systems having a continuum of equilibria. *SIAM Journal on Control and Optimization* **42**(5), 1745–1775 (2003)

- [24] Chen, Y., Mestres, P., Cortés, J., Dall’Anese, E.: Equilibria and their stability do not depend on the control barrier function in safe optimization-based control. *Automatica* (2024). Submitted. <https://arxiv.org/abs/2409.06808>
- [25] Koditschek, D.E.: Exact robot navigation by means of potential functions: some topological considerations. In: IEEE Int. Conf. on Robotics and Automation, Raleigh, NC, USA, pp. 1–6 (1987)
- [26] Hurwitz, A.: Ueber die bedingungen, unter welchen eine gleichung nur wurzeln mit negativen reellen theilen besitzt. *Mathematische Annalen* **46**(1), 273–284 (1895)
- [27] Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, New York, USA (2012)

8 Appendix

8.1 Auxiliary results for Section 4

This section provides a number of supporting results for the technical treatment of Section 4.

Lemma 8.1. (Trajectories of globally asymptotically stable system): *Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be locally Lipschitz, and suppose that the origin is globally asymptotically stable for the differential equation $\dot{\mathbf{x}} = F(\mathbf{x})$. Then, for any $\mathbf{x}_0 \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$, it holds that $\lim_{t \rightarrow -\infty} \|\mathbf{x}(t; \mathbf{x}_0)\| = \infty$.*

Proof. Since F is locally Lipschitz and the origin is globally asymptotically stable, by [20, Theorem 4.17] there exists a smooth, positive definite, and radially unbounded function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ and a continuous positive definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\nabla V(\mathbf{x})^T f(\mathbf{x}) \leq -W(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Now, let $t \geq 0$ and note that $\frac{d}{dt} V(\mathbf{x}(-t; \mathbf{x}_0)) = -\nabla V(\mathbf{x}(-t; \mathbf{x}_0))^T f(\mathbf{x}(-t; \mathbf{x}_0)) \geq W(\mathbf{x}(-t; \mathbf{x}_0)) \geq 0$. Hence, $V(\mathbf{x}(-t; \mathbf{x}_0)) \geq V(\mathbf{x}_0)$. By letting $\bar{w} = \min_{\{\mathbf{y} \in \mathbb{R}^n : V(\mathbf{y}) \geq V(\mathbf{x}_0)\}} W(\mathbf{y})$, it follows that $\frac{d}{dt} V(\mathbf{x}(-t; \mathbf{x}_0)) \geq \bar{w} > 0$ for all $t \geq 0$. This implies that $\lim_{t \rightarrow \infty} V(\mathbf{x}(-t; \mathbf{x}_0)) = \infty$ and since V is radially unbounded, the result follows. \square

Lemma 8.2. (Convergence and tangency properties of stable manifold): *Let h be a strict CBF and suppose Assumption 1 holds. Let $\mathbb{R}^2 \setminus \mathcal{C}$ be a bounded connected set. Consider (3) with $n = 2$ and let \mathbf{x}_* be an undesired equilibrium of (3) which is a saddle point. Let γ be a subset of the one-dimensional local stable manifold of \mathbf{x}_* such that γ corresponds to a maximal trajectory of (3) and there exists $T > 0$ with $\gamma(t) \in \text{Int}(\mathcal{C})$ for all $t > T$. Then,*

- (i) γ is not tangent to $\partial\mathcal{C}$ at any point;
- (ii) if (ω_-, ∞) is the interval of definition of γ , then
 - (a) if $\omega_- = -\infty$, then either $\lim_{t \rightarrow -\infty} \|\gamma(t)\| = \infty$, $\lim_{t \rightarrow -\infty} \|\gamma(t)\| = \mathbf{x}'_*$, with \mathbf{x}'_* another equilibrium of (3), or $\lim_{t \rightarrow -\infty} \|\gamma(t)\|$ converges to a limit cycle;

(b) if $\omega_- > -\infty$, then $\lim_{t \rightarrow \omega_-} \gamma(t) \notin \mathcal{C}$.

Proof. To show (i), we reason by contradiction. Figure 12 serves as visual aid for the different elements defined in the proof. Suppose that γ is tangent to $\partial\mathcal{C}$ at a point $\bar{\mathbf{q}}$. Then, $\bar{\mathbf{q}} = \gamma(t_{\bar{\mathbf{q}}})$ for some $t_{\bar{\mathbf{q}}} \in \mathbb{R}$, and there exists $T_{\bar{\mathbf{q}}} > 0$ small enough with $\gamma(t) \in \text{Int}(\mathcal{C})$ for all $t \in (t_{\bar{\mathbf{q}}} - T_{\bar{\mathbf{q}}}, t_{\bar{\mathbf{q}}})$. Now, note that there exists a sufficiently small neighborhood $N_{\bar{\mathbf{q}}}$ of $\bar{\mathbf{q}}$ such that one of the arches ζ between $\partial\mathcal{C}$ and γ defined by $N_{\bar{\mathbf{q}}}$ is such that all trajectories of (3) with initial condition in ζ stay between $\partial\mathcal{C}$ (because \mathcal{C} is forward invariant) and γ (because of uniqueness of solutions), and stay inside $N_{\bar{\mathbf{q}}}$ at all times (by continuity of the vector field (3), $\gamma'(t_{\bar{\mathbf{q}}}) \neq \mathbf{0}_n$, and because $N_{\bar{\mathbf{q}}}$ is sufficiently small). However, it is not possible that such trajectories stay inside the region defined by ζ , $\partial\mathcal{C}$ and γ by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], because $\bar{\mathbf{q}}$ is not an equilibrium point and the region defined by ζ , $\partial\mathcal{C}$ and γ does not contain limit cycles, because such limit cycle would encircle only points in the interior of \mathcal{C} [21, Corollary 6.26]. Hence, we have reached a contradiction, which means that γ is not tangent to $\partial\mathcal{C}$ at any point. Let us now show (ii). Part (a) follows by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], whereas part (b) follows from the fact that if $\lim_{t \rightarrow \omega_-} \gamma(t) \in \mathcal{C}$, since (3) is well-defined at all points in the safe set, then γ is well-defined as $t \rightarrow \omega_-$, and the interval of definition of γ can be increased, contradicting the assumption that γ is a maximal solution. \square

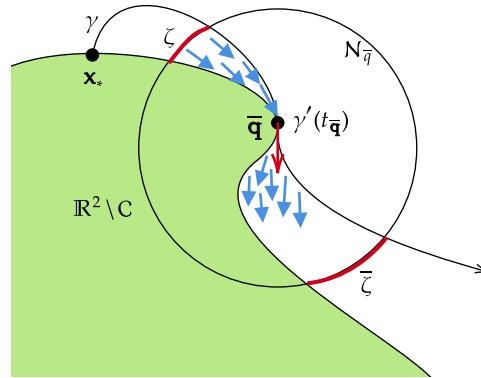


Fig. 12 Sketch of the setting considered in the proof of Lemma 8.2. The unsafe set is depicted in green, whereas the closed-loop vector field at the point of tangency is the red arrow, the blue arrows denote the vector field elsewhere, and the arches ζ and $\bar{\zeta}$ are depicted in red.

Lemma 8.3. (Existence of stable manifold exiting set with no limit cycles): *Let h be a strict CBF and suppose Assumption 1 holds. Let $\mathbb{R}^2 \setminus \mathcal{C}$ be a bounded connected set. Consider (3) with $n = 2$ and let $\{\mathbf{x}_*^{(i)}\}_{i=1}^k \subset \partial\mathcal{C}$ be its set of undesired equilibria. Denote by $\mathcal{L} \subset [k]$ the index set of undesired equilibria that are saddle points and let $\tilde{\Phi}$ be a compact forward invariant set containing the origin and $\mathbb{R}^2 \setminus \mathcal{C}$ and such that $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycles. For each $j \in \mathcal{L}$, let γ_j be a subset of the one-dimensional local stable manifold of (3) at $\mathbf{x}_*^{(j)}$ such that γ_j corresponds to a maximal*

trajectory and there exists $T > 0$ with $\gamma_j(t) \in \text{Int}(\mathcal{C})$ for all $t > T$. Then, there exists at least one $j \in \mathcal{L}$ and $\tilde{t}_j \in \mathbb{R}$ such that $\gamma_j(\tilde{t}_j) \in \partial\tilde{\Phi}$ and $\gamma_j(t) \notin \tilde{\Phi}$ for all $t < t_j$.

Proof. Let us reason by contradiction. If the statement does not hold, then using Lemma 8.2(ii), we deduce that for all $j \in \mathcal{L}$, one of the following holds:

- (i) there exists $t_j > 0$ such that $\lim_{t \rightarrow -t_j} \gamma_j(t) \in \mathbb{R}^2 \setminus \mathcal{C}$;
- (ii) $\gamma_j(t)$ converges to a limit cycle in $\tilde{\Phi} \cap \mathcal{C}$;
- (iii) $\lim_{t \rightarrow -\infty} \gamma_j(t) = \mathbf{x}_*^{(j)}$ for some $\bar{j} \in [k]$.

Since $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycles, case (ii) is not possible. If for all $j \in \mathcal{L}$, either (i) or (iii) hold, there exists a compact set $\tilde{\mathcal{O}}$, whose boundary is comprised of $\partial(\mathbb{R}^2 \setminus \mathcal{C})$ and the union of the trajectories γ_j for all $j \in \mathcal{L}$. If all trajectories with initial condition in $\tilde{\Phi} \setminus \tilde{\mathcal{O}}$ converge to the origin, that contradicts [25, Proposition 3], which shows that there cannot exist a continuous dynamical system, forward invariant in a set whose complement is compact, and with such set being the region of attraction of an asymptotically stable equilibrium. Note also that if $\tilde{\Phi} \setminus \tilde{\mathcal{O}}$ contains trajectories belonging to the regions of attraction of different asymptotically stable equilibria, by [20, Theorem 8.1], there exists a trajectory of (3) in the boundary of these regions of attraction. By definition, this trajectory does not belong to any region of attraction and by the Poincaré-Bendixson Theorem [22, Chapter 7, Thm. 4.1], this trajectory can only converge to a limit cycle. However, $\tilde{\Phi}$ is forward invariant and $\tilde{\Phi} \cap \mathcal{C}$ does not contain any limit cycle, hence reaching a contradiction. \square

8.2 Auxiliary results for Section 5

This section provides a number of supporting results for the technical treatment of Section 5. We start with an auxiliary result used in the proof of Proposition 5.1.

Lemma 8.4. (If a strict CBF exists, all CBFs are strict): *Let \mathcal{C} be a compact set and assume h is a strict CBF of \mathcal{C} . Then any other CBF \tilde{h} of \mathcal{C} is also strict.*

Proof. By [17, Lemma 2.2], there exists a function $\zeta : \partial\mathcal{C} \rightarrow \mathbb{R}_{>0}$ such that $\nabla\tilde{h}(\mathbf{x}) = \zeta(\mathbf{x})\nabla h(\mathbf{x})$ for all \mathbf{x} in \mathcal{C} . Since h is strict, for all $\mathbf{x} \in \partial\mathcal{C}$, there exist $\mathbf{u}_x \in \mathbb{R}^m$ such that $\nabla h(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}_x) > 0$. This implies that $\nabla\tilde{h}(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}_x) = \zeta(\mathbf{x})\nabla h(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}_x) > 0$ for all $\mathbf{x} \in \partial\mathcal{C}$. Now, since $\nabla\tilde{h}$, f , and g are continuous, there exists a neighborhood \mathcal{N}_x of each $\mathbf{x} \in \partial\mathcal{C}$ such that $\nabla\tilde{h}(\mathbf{y})^\top(f(\mathbf{y}) + g(\mathbf{y})\mathbf{u}_x) > 0$ for all $\mathbf{y} \in \mathcal{N}_x$. Therefore, there exists a neighborhood \mathcal{N} of $\partial\mathcal{C}$ where the CBF condition for \tilde{h} is strictly feasible. Now, since \mathcal{C} is compact, we can choose α as a linear function with a sufficiently large slope to ensure that $\nabla\tilde{h}(\mathbf{x})^\top(f(\mathbf{x}) + g(\mathbf{x})\mathbf{u}_x) + \alpha(\tilde{h}(\mathbf{x})) > 0$ holds for all $\mathbf{x} \in \mathcal{C} \setminus \mathcal{N}$ and hence for all $\mathbf{x} \in \mathcal{C}$, making \tilde{h} a strict CBF. \square

We next give a technical result used in the proof of Proposition 5.9. We use the same notation.

Lemma 8.5 (Conditions for β and γ): *Let Assumption 2 hold. Furthermore, suppose that the conditions in Proposition 5.8 hold. Then, $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$, and in particular $\gamma^2 + \beta^2 > 0$. Moreover, if Assumption 1 holds, then $\gamma x_{c,1} + \beta x_{c,2} \neq 0$.*

Proof. Let us show that $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$, which implies that $\gamma^2 + \beta^2 > 0$. By noting that $|T_3| + \sqrt{T_2} > 0$, and squaring both sides of the condition $\frac{r}{\sqrt{b_1^2 + b_2^2}} > \frac{|T_3| + \sqrt{T_2}}{2T_1}$ in Proposition 5.8, we get: $\left(\frac{2T_1 r}{\sqrt{b_1^2 + b_2^2}} - |T_3|\right)^2 > T_2$, which is equivalent to $\frac{4T_1^2 r^2}{b_1^2 + b_2^2} + T_3^2 - \frac{4T_1 r |T_3|}{\sqrt{b_1^2 + b_2^2}} > T_2^2 + 2\alpha_0 r^2 T_1$. Rearranging terms, this yields

$$|T_3| < \frac{(b_2 \beta + b_1 \gamma)r}{\sqrt{b_1^2 + b_2^2}}. \quad (28)$$

Note that (28) requires $b_2 \beta + b_1 \gamma > 0$ since otherwise the conditions in (28) would not be feasible for any T_3 . Now, by using condition (28) and applying the Cauchy-Schwartz inequality, we get $T_3 > -\sqrt{b_1^2 + b_2^2}r$, $T_3 < \sqrt{b_1^2 + b_2^2}r$, from which it follows that $r^2(\gamma^2 + \beta^2) - T_3^2 > 0$. Finally suppose that $\|\mathbf{x}_c\|^2 > r^2$ and $\gamma x_{c,1} + \beta x_{c,2} = 0$. Note that $T_3^2 = (-\gamma x_{c,2} + \beta x_{c,1})^2 = (-\gamma x_{c,2} + \beta x_{c,1})^2 + (\gamma x_{c,1} + \beta x_{c,2})^2 = (\gamma^2 + \beta^2) \|\mathbf{x}_c\|^2$. Since $\|\mathbf{x}_c\|^2 > r^2$, this implies that $r^2(\gamma^2 + \beta^2) - T_3^2 < 0$, which is a contradiction. \square

Next we add details to the Example 5.5. In particular, we elaborate further on the stability properties of undesired equilibria.

Example 8.6. (Example 5.5 continued): First, since the boundary of \mathcal{C} is given by a union of semicircles, by following an argument similar to that of the proof of Proposition 5.8, one has that the following are sufficient conditions for h to be a strict CBF:

$$T_1 > 0, b_1^2 + b_2^2, \quad (29a)$$

$$\frac{r_1}{\sqrt{b_2^2 + b_1^2}} > \frac{\frac{b_1}{|b_1|}(-\gamma c_2 - \beta c_1) + \sqrt{(\gamma c_2 + \beta c_1)^2 + 2\alpha_0 r_1^2 T_1}}{2T_1}, \quad (29b)$$

$$-\frac{r}{\sqrt{b_2^2 + b_1^2}} > \frac{\frac{b_1}{|b_1|}(-\gamma c_2) - \sqrt{\gamma^2 c_2^2 + 2\alpha_0 r^2 T_1}}{2T_1}, \quad (29c)$$

$$\frac{r_1}{\sqrt{b_2^2 + b_1^2}} > \frac{\frac{b_1}{|b_1|}(-\gamma c_2 + \beta c_1) + \sqrt{(-\gamma c_2 + \beta c_1)^2 + 2\alpha_0 r_1^2 T_1}}{2T_1}, \quad (29d)$$

$$-\frac{R}{\sqrt{b_2^2 + b_1^2}} < \frac{\frac{b_1}{|b_1|}(-\gamma c_2) - \sqrt{\gamma^2 c_2^2 + 2\alpha_0 R^2 T_1}}{2T_1} \quad (29e)$$

where $\beta = a_{11}b_2 - b_1a_{21}$, $\gamma = a_{22}b_1 - b_2a_{12}$, and $T_1 := b_2\beta + b_1\gamma + \frac{1}{2}\alpha_0(b_2^2 + b_1^2)$. Further, suppose that $r^2(\gamma^2 + \beta^2) - \gamma^2 c_2^2 \geq 0$, and let

$$z_{+,2} = \frac{\beta c_2 + \sqrt{r^2(\gamma^2 + \beta^2) - \gamma^2 c_2^2}}{\gamma^2 + \beta^2}.$$

It follows that the point $\mathbf{x}_{*,+,2} = (\gamma z_{+,2}, \beta z_{+,2})$ is in $\partial\mathcal{C}$ and satisfies (6) for some $\delta \in \mathbb{R}$; this, in turn, means that $\mathbf{x}_{*,+,2} \in \mathcal{E}$. To show whether $\mathbf{x}_{*,+,2}$ is an undesired

equilibrium, we need to check if that $-(\mathbf{x}_{*,+,2} - \mathbf{x}_c)^\top (A - BK)\mathbf{x}_{*,+,2} < 0$. By using the expression of $\mathbf{x}_{*,+,2}$, this condition is equivalent to

$$z_{+,2} T_4 (b_1 \gamma z_{+,2} + b_2 (\beta z_{+,2} - c_2)) > 0, \quad (30)$$

where $T_4 = a_{11}a_{22} - a_{12}a_{21} - k_1\gamma - k_2\beta$. Since $A - BK$ is Hurwitz, $T_4 > 0$, and therefore (30) is independent of K and equivalent to

$$z_{+,2} (b_1 \gamma z_{+,2} + b_2 (\beta z_{+,2} - c_2)) > 0, \quad (31)$$

Now, note that Example 5.5 satisfies (31). Therefore, $\mathbf{x}_{*,+,2} = (3.157, 7.619)$ is an undesired equilibrium for any K . To show that it is asymptotically stable, note that Example 5.5 satisfies

$$(-\gamma c_2)(\gamma b_2 - \beta b_1) + (\gamma b_1 + \beta b_2) \sqrt{r^2(\gamma^2 + \beta^2) - \gamma^2 c_2^2} < 0. \quad (32)$$

Then, by following the same argument as in the proof of [1, Proposition 4], the Jacobian of (8) at $\mathbf{x}_{*,+,2}$ is

$$J(\mathbf{x}_{*,+,2}) = A - \frac{B}{(\mathbf{x}_c - \mathbf{x}_{*,+,2})^\top B} ((\mathbf{x}_{*,+,2} - \mathbf{x}_c)^\top (A + \alpha_0 \mathbf{I}_n)),$$

and (32) implies that J has two negative eigenvalues. Moreover, since J is independent of K , this implies that $\mathbf{x}_{*,+,2} = (3.157, 7.619)$ in Example 5.5 is asymptotically stable for any choice of linear stabilizing nominal controller. \triangle

The following results concern Section 5.3.2, i.e., the case when the LTI system is fully actuated. We employ the same notation. We start by stating two auxiliary results that determine the eigenvalue other than $-\alpha'(0)$ of the Jacobian.

Lemma 8.7. (Eigenvalue of the Jacobian when \tilde{A} is not diagonalizable): *Assume that $\lambda_1 = \lambda_2$ and \tilde{A} is not diagonalizable. Then, there exists $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$ such that $\|\mathbf{v}_1\|_2 = \|\mathbf{v}_2\|_2 = 1$, $\tilde{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $\tilde{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2 + \mathbf{v}_1$. For any $\mathbf{x}_* \in \hat{\mathcal{E}}$, if the associated indicator $\delta_{\mathbf{x}_*} \neq \frac{\lambda_1}{2}$, $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$ and $\mathbf{x}_* = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$, then it holds that $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} + \frac{2\delta_{\mathbf{x}_*}\beta_2}{(\lambda_1 - 2\delta_{\mathbf{x}_*})^2}$, $\beta_4 - \beta_2 = \frac{-\lambda_1\beta_2}{\lambda_1 - 2\delta_{\mathbf{x}_*}}$ and the eigenvalue other than $-\alpha'(0)$ of the Jacobian of (3) at \mathbf{x}_* is*

$$\lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_4 - \beta_2)}{r^2} ((\beta_3 - \beta_1) + (\beta_4 - \beta_2) \mathbf{v}_2^\top \mathbf{v}_1).$$

Proof. Let $J(\mathbf{x})$ be the Jacobian of (3) evaluated at \mathbf{x} . If we write $J(\mathbf{x}_*)\mathbf{v}_1 = d_{11}\mathbf{v}_1 + d_{21}\mathbf{v}_2$ and $J(\mathbf{x}_*)\mathbf{v}_2 = d_{12}\mathbf{v}_1 + d_{22}\mathbf{v}_2$, then the other eigenvalue of $J(\mathbf{x}_*)$ is equal to $d_{11} + d_{22} + \alpha'(0)$.

Using the expression for the Jacobian in [17, Proposition 11],

$$d_{11} = -\frac{(\beta_3 - \beta_1)\alpha'(0)}{r^2} ((\beta_3 - \beta_1) + (\beta_4 - \beta_2) \mathbf{v}_2^\top \mathbf{v}_1) +$$

$$\begin{aligned} \lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_3 - \beta_1)(\lambda_1 - 2\delta_{\mathbf{x}_*})}{r^2} & ((\beta_3 - \beta_1) + (\beta_4 - \beta_2)\mathbf{v}_2^\top \mathbf{v}_1), \\ d_{22} = -\frac{(\beta_4 - \beta_2)\alpha'(0)}{r^2} & ((\beta_3 - \beta_1)\mathbf{v}_1^\top \mathbf{v}_2 + (\beta_4 - \beta_2)) \\ - \frac{(\beta_4 - \beta_2)(\lambda_1 - 2\delta_{\mathbf{x}_*})}{r^2} & ((\beta_3 - \beta_1)\mathbf{v}_1^\top \mathbf{v}_2 + (\beta_4 - \beta_2)) \\ + \lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_4 - \beta_2)}{r^2} & ((\beta_3 - \beta_1) + (\beta_4 - \beta_2)\mathbf{v}_2^\top \mathbf{v}_1) \end{aligned}$$

Since $(\mathbf{x}_*, \delta_{\mathbf{x}_*})$ is an solution of (6), it follows that $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} + \frac{2\delta_{\mathbf{x}_*}\beta_2}{(\lambda_1 - 2\delta_{\mathbf{x}_*})^2}$ and $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}}$, by (6a); additionally, by (6b), one has that

$$\left\| \left(\frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} + \frac{2\delta_{\mathbf{x}_*}\beta_2}{(\lambda_1 - 2\delta_{\mathbf{x}_*})^2} \right) \mathbf{v}_1 + \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} \mathbf{v}_2 \right\|^2 - r^2 = 0.$$

Thus, it follows that:

$$d_{11} + d_{22} + \alpha'(0) = \lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_4 - \beta_2)}{r^2} ((\beta_3 - \beta_1) + (\beta_4 - \beta_2)\mathbf{v}_2^\top \mathbf{v}_1).$$

□

Lemma 8.8. (The other eigenvalue of Jacobian when \tilde{A} is diagonalizable): *Assume that $\lambda_1 \neq \lambda_2$. Then, there exists $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^2$ such that $\|\mathbf{v}_1\|_2 = \|\mathbf{v}_2\|_2 = 1$, $\tilde{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $\tilde{A}\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. Additionally, for any $\mathbf{x}_* \in \hat{\mathcal{E}}$, if the associated indicator $\delta_{\mathbf{x}_*} \notin \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$, $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$ and $\mathbf{x}_* = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$, it holds that $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}}$, $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}}$ and the eigenvalues other than $-\alpha'(0)$ of the Jacobian of (3) at \mathbf{x}_* is*

$$\lambda_1 + \lambda_2 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_3 - \beta_1)\lambda_1}{r^2} \Delta_1 - \frac{(\beta_4 - \beta_2)\lambda_2}{r^2} \Delta_2,$$

where $\Delta_1 := (\beta_3 - \beta_1)^* + (\beta_4 - \beta_2)^* \mathbf{v}_2^* \mathbf{v}_1$, $\Delta_2 := (\beta_3 - \beta_1)^* \mathbf{v}_1^* \mathbf{v}_2 + (\beta_4 - \beta_2)^*$. Equivalently, the eigenvalue other than $\alpha'(0)$ can be expressed as $\lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_3 - \beta_1)(\lambda_1 - \lambda_2)}{r^2} \Delta_1$ and $\lambda_2 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_4 - \beta_2)(\lambda_2 - \lambda_1)}{r^2} \Delta_2$.

Proof. If we write $J(\mathbf{x}_*)\mathbf{v}_1 = d_{11}\mathbf{v}_1 + d_{21}\mathbf{v}_2$ and $J(\mathbf{x}_*)\mathbf{v}_2 = d_{12}\mathbf{v}_1 + d_{22}\mathbf{v}_2$, then the other eigenvalue of $J(\mathbf{x}_*)$ is equal to $d_{11} + d_{22} + \alpha'(0)$. Using the expression for the Jacobian in [17, Proposition 11],

$$\begin{aligned} d_{11} = \lambda_1 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_3 - \beta_1)\alpha'(0)}{r^2} & ((\beta_3 - \beta_1)^* + (\beta_4 - \beta_2)^* \mathbf{v}_2^* \mathbf{v}_1) \\ - \frac{(\beta_3 - \beta_1)(\lambda_1 - 2\delta_{\mathbf{x}_*})}{r^2} & ((\beta_3 - \beta_1)^* + (\beta_4 - \beta_2)^* \mathbf{v}_2^* \mathbf{v}_1), \end{aligned}$$

and

$$d_{22} = \lambda_2 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_4 - \beta_2)\alpha'(0)}{r^2} ((\beta_3 - \beta_1)^* \mathbf{v}_1^* \mathbf{v}_2 + (\beta_4 - \beta_2)^*)$$

$$-\frac{(\beta_4 - \beta_2)(\lambda_2 - 2\delta_{\mathbf{x}_*})}{r^2}((\beta_3 - \beta_1)^* \mathbf{v}_1^* \mathbf{v}_2 + (\beta_4 - \beta_2)^*).$$

Note that since $(\mathbf{x}_*, \delta_{\mathbf{x}_*})$ is a solution of (6), it follows that $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}}$, $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}}$, and

$$\left\| \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} \mathbf{v}_1 + \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} \mathbf{v}_2 \right\|^2 - r^2 = 0.$$

Then,

$$\begin{aligned} d_{11} + d_{22} + \alpha'(0) &= \lambda_1 + \lambda_2 - 2\delta_{\mathbf{x}_*} - \frac{(\beta_3 - \beta_1)\lambda_1}{r^2}((\beta_3 - \beta_1)^* + (\beta_4 - \beta_2)^* \mathbf{v}_2^* \mathbf{v}_1) \\ &\quad - \frac{(\beta_4 - \beta_2)\lambda_2}{r^2}((\beta_3 - \beta_1)^* \mathbf{v}_1^* \mathbf{v}_2 + (\beta_4 - \beta_2)^*). \end{aligned}$$

By leveraging $\beta_3 - \beta_1 = \frac{-\lambda_1\beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}}$ and $\beta_4 - \beta_2 = \frac{-\lambda_2\beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}}$, we get the two remaining expressions. \square

To prove Proposition 5.11, we need to determine the Jacobian evaluated at the undesired equilibrium and analyze its spectrum. Applying [17, Proposition 6.2] and Lemma 4.1 to system (8), we have the following result.

Lemma 8.9. (The Jacobian evaluated at undesired equilibria): *Suppose Assumption 1 is satisfied, B is invertible, and $G = B^\top B$. The Jacobian of (8) at any undesired equilibrium \mathbf{x}_* is*

$$J(\mathbf{x}_*) = \tilde{A} - 2\delta_{\mathbf{x}_*} \mathbf{I} - \frac{(\mathbf{x}_* - \mathbf{x}_c)(\mathbf{x}_* - \mathbf{x}_c)^\top}{\|\mathbf{x}_* - \mathbf{x}_c\|^2} (\tilde{A} - (2\delta_{\mathbf{x}_*} - \alpha'(0))\mathbf{I}).$$

Proof of Proposition 5.11:

Denote λ the eigenvalue associated with \mathbf{x}_c . Then $\lambda = \lambda_i$, $i = 1$ or 2 ; and both λ_1 and λ_2 are real. We first determine the solution for (6) with $\delta \notin \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. Since $\delta \notin \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$, by the first equation in (6), it follows that $\mathbf{x}_* = \frac{2\delta}{2\delta - \lambda} \mathbf{x}_c$. Plugging this in the second equation in (6), we can solve for δ . This leads to the potential undesired equilibria $\mathbf{x}_{*,-} := (1 + \frac{r}{\|\mathbf{x}_c\|})\mathbf{x}_c$, with associated value of δ equal to $\frac{\lambda}{2} + \frac{\lambda\|\mathbf{x}_c\|}{2r}$, and $\mathbf{x}_{*,+} := (1 - \frac{r}{\|\mathbf{x}_c\|})\mathbf{x}_c$, with associated value of δ equal to $\frac{\lambda}{2} - \frac{\lambda\|\mathbf{x}_c\|}{2r}$. We note that the value of δ associated with $\mathbf{x}_{*,-}$ is negative and the value of δ associated with $\mathbf{x}_{*,+}$ is positive, so $\mathbf{x}_{*,-}$ is an undesired equilibrium while $\mathbf{x}_{*,+}$ is not an undesired equilibrium. By Lemma 8.9, the Jacobian at $\mathbf{x}_{*,-}$ is

$$J(\mathbf{x}_{*,-}) = \tilde{A} - 2\delta \mathbf{I} - \frac{\mathbf{x}_c \mathbf{x}_c^\top}{\|\mathbf{x}_c\|^2} (\tilde{A} - (2\delta - \alpha'(0))\mathbf{I}).$$

where $\delta = \frac{\lambda_1}{2} + \frac{\lambda_1\|\mathbf{x}_c\|}{2r}$.

In the following, we determine if there exist solutions of (6) with $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$ and discuss the stability of the corresponding undesired equilibria case by case.

Case 1: \tilde{A} is not diagonalizable.

In this case, we first show that $\mathbf{x}_{*,-}$ is always a saddle point. We note that we must have $\lambda_1 = \lambda_2$. Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, \mathbf{v}_2 be a vector such that $\|\mathbf{v}_2\| = 1$, $\tilde{A}\mathbf{v}_2 = \lambda_1\mathbf{v}_2 + \mathbf{v}_1$. If we write $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$, then $\beta_1 = \|\mathbf{x}_c\|$ and $\beta_2 = 0$. By Lemma 8.7, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_2 - 2\delta_{\mathbf{x}_{*,-}} = \lambda_1 - \lambda_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{r} > 0$, implying that $\mathbf{x}_{*,-}$ is a saddle point.

Next, we determine if there exists a solution with $\delta = \frac{\lambda_1}{2}$. We write $\mathbf{x}_* = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$. Hence the first equation of (6) with $\delta = \frac{\lambda_1}{2}$ can be rewritten as $\beta_4 = -\lambda_1\|\mathbf{x}_c\|$. Plugging the value of β_4 into the second equation of (6), and defining $\hat{\beta}_3 := \beta_3 - \|\mathbf{x}_c\|$ and $\tau_1 := \lambda_1\|\mathbf{x}_c\|$, it follows that

$$\hat{\beta}_3^2 - 2\tau_1\mathbf{v}_1^\top\mathbf{v}_2\hat{\beta}_3 + \tau_1^2 - r^2 = 0. \quad (33)$$

Note that the discriminant of the quadratic equation (33) is

$$\Delta := 4(\tau_1^2(\mathbf{v}_1^\top\mathbf{v}_2)^2 - \tau_1^2 + r^2) = 4(\tau_1^2((\mathbf{v}_1^\top\mathbf{v}_2)^2 - 1) + r^2)$$

This leads to the following three subcases.

Case 1.1 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 < 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, there does not exist a solution associated with $\delta = \frac{\lambda_1}{2}$.

Case 1.2 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 = 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, then there exists one solution $\mathbf{x}_{*,1} = (\tau_1\mathbf{v}_1^\top\mathbf{v}_2 + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2$ with $\delta_{\mathbf{x}_{*,1}} = \frac{\lambda_1}{2}$, equal to

$$\mathbf{x}_{*,1} = (\tau_1\mathbf{v}_1^\top\mathbf{v}_2 + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2$$

Since $(\mathbf{x}_{*,1} - \mathbf{x}_c)^\top\mathbf{v}_1 = 0$ and $(\tilde{A} - 2\delta_{\mathbf{x}_{*,1}}\mathbf{I})\mathbf{v}_1 = 0$, it follows that $J(\mathbf{x}_{*,1})^\top\mathbf{v}_1 = 0$, by Lemma 8.9. Therefore in this case, there is another undesired equilibrium $\mathbf{x}_{*,1}$, at which the Jacobian has a negative eigenvalue and a zero eigenvalue.

Case 1.3 if $(\mathbf{v}_1^\top\mathbf{v}_2)^2 > 1 - r^2/\tau_1^2 = 1 - \frac{r^2}{\lambda_1^2\|\mathbf{x}_c\|^2}$, there exist two solutions $\hat{\beta}_3 = \hat{\beta}_3^{(1)}$ and $\hat{\beta}_3 = \hat{\beta}_3^{(2)}$ for (33). This implies that there exist two extra undesired equilibria given by $\mathbf{x}_{*,2} = (\hat{\beta}_3^{(1)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2$ and $\mathbf{x}_{*,3} = (\hat{\beta}_3^{(2)} + \|\mathbf{x}_c\|)\mathbf{v}_1 - \tau_1\mathbf{v}_2$.

Notice that in this sub-case, $\hat{\beta}_3^{(1)} + \hat{\beta}_3^{(2)} = 2\tau_1\mathbf{v}_1^\top\mathbf{v}_2$, we can assume that $\hat{\beta}_3^{(1)} < \tau_1\mathbf{v}_1^\top\mathbf{v}_2$ and $\hat{\beta}_3^{(1)} > \tau_1\mathbf{v}_1^\top\mathbf{v}_2$.

Using the same technique in the proof of Lemma 8.8, we can show that $J(\mathbf{x}_{*,2})$, has an eigenvalue $\frac{\tau_1}{r^2}(\hat{\beta}_3^{(1)} - \tau_1\mathbf{v}_1^\top\mathbf{v}_2) > 0$; and $J(\mathbf{x}_{*,3})$ has an eigenvalue $\frac{\tau_1}{r^2}(\hat{\beta}_3^{(2)} - \tau_1\mathbf{v}_1^\top\mathbf{v}_2) < 0$.

Hence in this case, there are another two undesired equilibria, one of which is stable and the other one is saddle point.

Case 2: \tilde{A} is diagonalizable and $\lambda_1 \leq \lambda_2 < 0$, $\tilde{A}\mathbf{x}_c = \lambda_1\mathbf{x}_c$

In this case, we first show that $\mathbf{x}_{*,-}$ is always a saddle point. Let $\mathbf{v}_1 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, \mathbf{v}_2 be an eigenvector associated with λ_2 satisfying $\|\mathbf{v}_2\| = 1$, $\mathbf{v}_1^\top \mathbf{v}_2 \geq 0$. If we write $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$, then $\beta_1 = \|\mathbf{x}_c\|$ and $\beta_2 = 0$. By Lemma 8.8, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_2 - \lambda_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{r} > 0$, implying that $\mathbf{x}_{*,-}$ is a saddle point. Next, we determine if there exists a solution with $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. We write $\mathbf{x}_* = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$ and then the first equation of (6) can be rewritten as

$$\begin{aligned} (\lambda_1 - 2\delta)(\beta_3 - \|\mathbf{x}_c\|) &= -\lambda_1 \|\mathbf{x}_c\| \\ (\lambda_2 - 2\delta)\beta_4 &= 0 \end{aligned} \tag{34}$$

from which it follows that $\delta \neq \frac{\lambda_1}{2}$.

If $\delta = \frac{\lambda_2}{2}$, from the first equation of (34) it follows that $\beta_3 = \frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2}$. Plugging the value of β_3 into the equation $h(\mathbf{x}) = 0$ from (6), and by defining $\tau_2 := \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2}$, it follows that

$$\beta_4^2 - 2\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4 + \tau_2^2 - r^2 = 0. \tag{35}$$

Note that the discriminant of quadratic equation (35) is

$$\Delta := 4(\tau_2^2(\mathbf{v}_1^\top \mathbf{v}_2)^2 - \tau_2^2 + r^2) = 4(\tau_2^2((\mathbf{v}_1^\top \mathbf{v}_2)^2 - 1) + r^2),$$

which leads to the following three subcases.

Case 2.1 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 < 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$, there does not exist a solution associated with $\delta = \frac{\lambda_2}{2}$.

Case 2.2 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$, then there exists an undesired equilibrium

$$\mathbf{x}_{*,4} = \frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1^\top \mathbf{v}_2 \mathbf{v}_2$$

with corresponding δ equal to $\frac{\lambda_2}{2}$. We note that $(\mathbf{x}_{*,4} - \mathbf{x}_c)^\top \mathbf{v}_2 = 0$ and $(\tilde{A} - 2\delta_{\mathbf{x}_{*,4}} \mathbf{I}) \mathbf{v}_2 = 0$. Hence, by Lemma 8.9, we have $J(\mathbf{x}_{*,4}) \mathbf{v}_2 = 0$. Thus in this case, the Jacobian evaluated at $\mathbf{x}_{*,4}$ has a negative eigenvalue and a zero eigenvalue.

Case 2.3 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - r^2/\tau_2^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$, there exist two solutions $\beta_4 = \beta_4^{(1)}$ and $\beta_4 = \beta_4^{(2)}$ for (33). Then there exist two undesired equilibria $\mathbf{x}_{*,5} = \frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(1)} \mathbf{v}_2$ and $\mathbf{x}_{*,6} = \frac{-\lambda_2 \|\mathbf{x}_c\|}{\lambda_1 - \lambda_2} \mathbf{v}_1 + \beta_4^{(2)} \mathbf{v}_2$ with associated value of δ equal to $\frac{\lambda_2}{2}$. Notice that in this sub-case, $\beta_4^{(1)} + \beta_4^{(2)} = 2\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 > 0$ and $\beta_4^{(1)} \beta_4^{(2)} = \tau_2^2 - r^2 > 0$, we can assume that $0 < \beta_4^{(1)} < \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$ and $\beta_4^{(2)} > \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$. It follows that $-\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} + \tau_2^2 - r^2 = -\beta_4^{(1)} \beta_4^{(1)} + \tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} > 0$ and $-\tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} + \tau_2^2 - r^2 = -\beta_4^{(2)} \beta_4^{(2)} + \tau_2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} < 0$. Using the same technique in the proof of Lemma 8.7, we can show that $J(\mathbf{x}_{*,4})$ has an eigenvalue $\frac{\lambda_2 - \lambda_1}{r^2}(\tau_2^2 - \tau_2^2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(1)} - r^2) > 0$, and $J(\mathbf{x}_{*,5})$ has an eigenvalue $\frac{\lambda_2 - \lambda_1}{r^2}(\tau_2^2 - \tau_2^2 \mathbf{v}_1^\top \mathbf{v}_2 \beta_4^{(2)} - r^2) < 0$. Hence in this case, there are two extra undesired equilibria, one of which is stable and the other one is a saddle point.

Case 3: \tilde{A} diagonalizable, $\lambda_1 < \lambda_2 < 0$, $\tilde{A}\mathbf{x}_c = \lambda_2\mathbf{x}_c$.

Let $\mathbf{v}_2 = \frac{\mathbf{x}_c}{\|\mathbf{x}_c\|}$, \mathbf{v}_1 be an eigenvector associated with λ_1 and $\|\mathbf{v}_1\| = 1$, $\mathbf{v}_1^\top \mathbf{v}_2 \geq 0$. If we write $\mathbf{x}_c = \beta_1\mathbf{v}_1 + \beta_2\mathbf{v}_2$, then $\beta_2 = \|\mathbf{x}_c\|$ and $\beta_1 = 0$. By Lemma 8.8, it follows that the Jacobian at $\mathbf{x}_{*,-}$ has an eigenvalue equal to $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r}$. We determine the sign of $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r}$ later. First, let us determine if there exists a solution with $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. We write $\mathbf{x}_c = \|\mathbf{x}_c\|\mathbf{v}_2$, $\mathbf{x} = \beta_3\mathbf{v}_1 + \beta_4\mathbf{v}_2$ and then from (6) it follows that

$$\begin{aligned} (\lambda_1 - 2\delta)\beta_3 &= 0 \\ (\lambda_2 - 2\delta)(\beta_4 - \|\mathbf{x}_c\|) &= -\lambda_2\|\mathbf{x}_c\|, \end{aligned} \tag{36}$$

from which it follows that $\delta \neq \frac{\lambda_2}{2}$. If $\delta = \frac{\lambda_1}{2}$, it follows from (36) that $\beta_4 = \frac{-\lambda_1\|\mathbf{x}_c\|}{\lambda_2 - \lambda_1}$. Plugging the value of β_4 into the equation $h(\mathbf{x}) = 0$ from (6), and by letting $\tau_3 := \frac{\lambda_2\|\mathbf{x}_c\|}{\lambda_2 - \lambda_1}$, it follows that

$$\beta_3^2 - 2\tau_3\mathbf{v}_1^\top \mathbf{v}_2\beta_3 + \tau_3^2 - r^2 = 0. \tag{37}$$

Note that the discriminant of quadratic equation (37) is

$$\Delta := 4(\tau_3^2(\mathbf{v}_1^\top \mathbf{v}_2)^2 - \tau_3^2 + r^2) = 4(\tau_3^2((\mathbf{v}_1^\top \mathbf{v}_2)^2 - 1) + r^2),$$

which leads to the following three subcases.

Case 3.1 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 < 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$, there does not exist a solution associated with $\delta = \frac{\lambda_1}{2}$. Recall also that the eigenvalue (other than $-\alpha'(0)$) of Jacobian at $\mathbf{x}_{*,-}$ is $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r}$. In this subcase, we have $1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} > (\mathbf{v}_1^\top \mathbf{v}_2)^2 \geq 0$, which implies that $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r} > 0$. Hence in this case, we only have one undesired equilibrium $\mathbf{x}_{*,-}$, which is a saddle point.

Case 3.2 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} \neq 0$, there exists an undesired equilibrium equal to

$$\mathbf{x}_{*,7} = \frac{\lambda_2\|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_1^\top \mathbf{v}_2 \mathbf{v}_1 - \frac{\lambda_1\|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2,$$

with associated δ equal to $\delta = \frac{\lambda_1}{2}$. We note that $(\mathbf{x}_{*,7} - \mathbf{x}_c)^\top \mathbf{v}_1 = 0$ and $(\tilde{A} - 2\delta_{\mathbf{x}_{*,7}} \mathbf{I})\mathbf{v}_2 = 0$. Hence, by Lemma 8.9, we get that $J(\mathbf{x}_{*,7})\mathbf{v}_1 = 0$. Therefore, $\mathbf{x}_{*,7}$ is an undesirable equilibrium and a degenerate equilibrium. Recall that the eigenvalue (other than $-\alpha'(0)$) of the Jacobian at $\mathbf{x}_{*,-}$ is $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r}$. In this subcase, we still have $1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} = (\mathbf{v}_1^\top \mathbf{v}_2)^2 > 0$, which implies that $\lambda_1 - \lambda_2 - \frac{\lambda_2\|\mathbf{x}_c\|}{r} > 0$. Hence in this subcase, there are two undesired equilibria $\mathbf{x}_{*,-}$ and $\mathbf{x}_{*,7}$, where $\mathbf{x}_{*,-}$ is an saddle point and $\mathbf{x}_{*,7}$ is a degenerate equilibrium.

Case 3.3 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 = 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} = 0$, there exists one solution associated with $\delta = \frac{\lambda_1}{2}$, which is

$$\mathbf{x}_{*,8} = -\frac{\lambda_1\|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2.$$

Notice that $1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} = 0$ implies that $\frac{-\lambda_2}{\lambda_2 - \lambda_1} = \frac{r}{\|\mathbf{x}_c\|}$, from which it follows that

$$\mathbf{x}_{*,8} = -\frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2 = (1 - \frac{\lambda_2}{\lambda_2 - \lambda_1}) \mathbf{x}_c = (1 + \frac{r}{\|\mathbf{x}_c\|}) \mathbf{x}_c = \mathbf{x}_{*,-},$$

and $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} = 0$. Thus in this case, there is only one undesired equilibrium $\mathbf{x}_{*,-}$, which is a degenerate equilibrium.

Case 3.4 if $(\mathbf{v}_1^\top \mathbf{v}_2)^2 > 1 - r^2/\tau_3^2 = 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$, there exist two solutions $\beta_3 = \beta_3^{(1)}$ and $\beta_3 = \beta_3^{(2)}$ for (37). Then there exist two extra undesired equilibria: $\mathbf{x}_{*,9} = \beta_3^{(1)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2$ and $\mathbf{x}_{*,10} = \beta_3^{(2)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2$. Notice that in this sub-case, we have $\beta_3^{(1)} + \beta_3^{(2)} = 2\tau_3 \mathbf{v}_1^\top \mathbf{v}_2 < 0$. Then we can assume that $\beta_3^{(1)} < \tau_3 \mathbf{v}_1^\top \mathbf{v}_2$ and $\beta_3^{(2)} > \tau_2 \mathbf{v}_1^\top \mathbf{v}_2$. Using the same technique in the proof of Lemma 8.7, we can show that the $J(\mathbf{x}_{*,9})$ has an eigenvalue

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(1)} - r^2);$$

and $J(\mathbf{x}_{*,10})$ has an eigenvalue

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(2)} - r^2).$$

Recall that the Jacobian evaluated at $\mathbf{x}_{*,-}$ has an eigenvalue $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r}$, and then we only need to determine the sign of these three eigenvalues case by case.

Case 3.4.1 If $0 < \tau_3^2 - r^2 = \frac{\lambda_2^2 \|\mathbf{x}_c\|^2}{(\lambda_2 - \lambda_1)^2} - r^2$, it is easy to check that $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} > 0$. In addition, similar to **Case 3.3**, we can show that $\{\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(i)} - r^2) : i = 1, 2\}$ contains one positive number and one negative number. Thus in this case, there are three undesired equilibria in total, two of which are saddle points and one of which is asymptotically stable.

Case 3.4.2 If $0 = \tau_3^2 - r^2 = \frac{\lambda_2^2 \|\mathbf{x}_c\|^2}{(\lambda_2 - \lambda_1)^2} - r^2$, it follows that $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} = 0$. In addition, we have $\beta_3^{(2)} = 0$ and the point $\mathbf{x}_{*,10} = \beta_3^{(2)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2$ is equal to $\mathbf{x}_{*,-}$. The point $\mathbf{x}_{*,9} = \beta_3^{(1)} \mathbf{v}_1 - \frac{\lambda_1 \|\mathbf{x}_c\|}{\lambda_2 - \lambda_1} \mathbf{v}_2$ is a saddle point since the eigenvalue

$$\begin{aligned} \frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(1)} - r^2) &= 2 \frac{\lambda_2 - \lambda_1}{r^2} \tau_3^2 (\mathbf{v}_1^\top \mathbf{v}_2)^2 \\ &> 2 \frac{\lambda_2 - \lambda_1}{r^2} \tau_3^2 \left(1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}\right) > 0. \end{aligned}$$

Thus in this case, there are two undesired equilibria $\mathbf{x}_{*,-}$ and $\mathbf{x}_{*,9}$, where $\mathbf{x}_{*,-}$ is a degenerate equilibrium and $\mathbf{x}_{*,9}$ is an saddle point .

Case 3.4.3 If $0 > \tau_3^2 - r^2 = \frac{\lambda_2^2 \|\mathbf{x}_c\|^2}{(\lambda_2 - \lambda_1)^2} - r^2$, it is easy to check that $\lambda_1 - \lambda_2 - \frac{\lambda_2 \|\mathbf{x}_c\|}{r} < 0$, which implies that $\mathbf{x}_{*,-}$ is asymptotically stable. By $\beta_3^{(1)} \beta_3^{(2)} = \tau_3^2 - r^2 < 0$ and

$\beta_3^{(1)} < \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 < 0$, it follows that $\beta_3^{(2)} > 0$. Using the fact that $\beta_3^{(1)} < \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 < 0$, we can show that

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(1)} - r^2) > 0.$$

On the other hand, using the fact that $\beta_3^{(1)} > 0 > \tau_3 \mathbf{v}_1^\top \mathbf{v}_2$, we can show that

$$\frac{\lambda_1 - \lambda_2}{r^2} (\tau_3^2 - \tau_3 \mathbf{v}_1^\top \mathbf{v}_2 \beta_3^{(2)} - r^2) > 0.$$

Thus in this case, there are three undesired equilibria in total, two of which are saddle points and one of which is asymptotically stable.

Table 1 summarizes the cases discussed in the proof, except for **Case 3.3** and **Case 3.4.2**. \square

Proof of Proposition 5.13

Denote the eigenvalues of \tilde{A} as $\lambda_1, \lambda_2 \in \mathbb{C}$. We note that the conditions in (6) can be rewritten as follows:

$$\begin{aligned} (\tilde{A} - 2\delta \mathbf{I}_{2 \times 2})(\mathbf{x} - \mathbf{x}_c) &= -\tilde{A}\mathbf{x}_c \text{ and,} \\ \|\mathbf{x} - \mathbf{x}_c\|^2 - r^2 &= 0. \end{aligned} \tag{38}$$

Next, we consider two cases.

- *Case #1* (\tilde{A} is diagonalizable): Recall that \mathbf{x}_c is not an eigenvector, so it holds that $\lambda_1 \neq \lambda_2$. Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^2$ be eigenvectors such that $\tilde{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\tilde{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$, $\|\mathbf{v}_1\| = \|\mathbf{v}_2\| = 1$. Write \mathbf{x}_c as $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$ and $\mathbf{x} = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$. Hence, the first equation in (38) can be rewritten as:

$$\begin{aligned} (\lambda_1 - 2\delta)(\beta_3 - \beta_1) &= -\lambda_1 \beta_1 \\ (\lambda_2 - 2\delta)(\beta_4 - \beta_2) &= -\lambda_2 \beta_2. \end{aligned} \tag{39}$$

Note that $\beta_1 \neq 0$ and $\beta_2 \neq 0$ as \mathbf{x}_c is not an eigenvector of $A - BK$; it follows that there is no solution with $\delta \in \{\frac{\lambda_1}{2}, \frac{\lambda_2}{2}\}$. For any solution $(\mathbf{x}, \delta_{\mathbf{x}})$ of (38), we have that $\beta_3 - \beta_1 = \frac{-\lambda_1 \beta_1}{\lambda_1 - 2\delta_{\mathbf{x}}}$, $\beta_4 - \beta_2 = \frac{-\lambda_2 \beta_2}{\lambda_2 - 2\delta_{\mathbf{x}}}$ and $\left\| \frac{-\lambda_1 \beta_1}{\lambda_1 - 2\delta_{\mathbf{x}}} \mathbf{v}_1 + \frac{-\lambda_2 \beta_2}{\lambda_2 - 2\delta_{\mathbf{x}}} \mathbf{v}_2 \right\|^2 - r^2 = 0$, which is equivalent to $F_1(\delta) = 0$, where $F_1(\delta)$ is defined in (27).

We first note that $F_1(\delta) = 0$ can have at most 4 solutions. Therefore, there are at most four solutions for (38). In addition, notice that $F_1(-\infty) < 0$, $F_1(+\infty) < 0$ and $F_1(0) = (\|\mathbf{x}_c\|^2 - r^2)\|\lambda_1 \lambda_2\|^2 > 0$, it follows that there exists at least one solution of (38) with positive δ and at least one solution with negative δ . If $\lambda_1 \leq \lambda_2$, we have $F_1(-\infty) < 0$, and $F_1(\frac{\lambda_1}{2}) > 0$ and there exists at least one solution for (38) with $\delta < \frac{\lambda_1}{2}$.

- *Case #2* (\tilde{A} is not diagonalizable): In this case, we have $\lambda_1 = \lambda_2$. Note that both eigenvalues are negative and \mathbf{x}_c is not an eigenvector of \tilde{A} . Let $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$ be vectors of length 1, such that $\tilde{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\tilde{A}\mathbf{v}_2 = \lambda_1 \mathbf{v}_2 + \mathbf{v}_1$. We write $\mathbf{x}_c = \beta_1 \mathbf{v}_1 + \beta_2 \mathbf{v}_2$ and

$\mathbf{x}_* = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$. Hence, the first equation in (38) can be rewritten as

$$\begin{aligned} (\lambda_1 - 2\delta)(\beta_3 - \beta_1) + (\beta_4 - \beta_2) &= -\lambda_1 \beta_1 - \beta_2 \\ (\lambda_2 - 2\delta)(\beta_4 - \beta_2) &= -\lambda_2 \beta_2. \end{aligned} \quad (40)$$

Note that $\beta_2 \neq 0$ as \mathbf{x}_c is not an eigenvector of \tilde{A} ; it follows that there is no solution with $2\delta = \lambda_1$. For any solution $(\mathbf{x}_*, \delta_{\mathbf{x}_*})$ of equation (38), we have $\beta_3 - \beta_1 = \frac{-\lambda_1 \beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} + \frac{2\delta_{\mathbf{x}_*} \beta_2}{(\lambda_1 - 2\delta_{\mathbf{x}_*})^2}$, $\beta_4 - \beta_2 = \frac{-\lambda_2 \beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}}$ and $\left\| \left(\frac{-\lambda_1 \beta_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} + \frac{2\delta_{\mathbf{x}_*} \beta_2}{(\lambda_1 - 2\delta_{\mathbf{x}_*})^2} \right) \mathbf{v}_1 + \frac{-\lambda_2 \beta_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} \mathbf{v}_2 \right\|^2 - r^2 = 0$, which is equivalent to $F_2(\delta_{\mathbf{x}_*}) = 0$, where F_2 is defined as

$$\begin{aligned} F_2(\delta) &:= -(\lambda_1 - 2\delta)^4 r^2 + (\lambda_1 \beta_2)^2 (\lambda_1 - 2\delta)^2 \\ &\quad + 2(\lambda_1 \beta_1 (\lambda_1 - 2\delta)^2 - 2\delta(\lambda_1 - 2\delta)\beta_2) \lambda_1 \beta_2 \mathbf{v}_1^\top \mathbf{v}_2 \\ &\quad + (2\delta \beta_2 - \lambda_1 (\lambda_1 - 2\delta) \beta_1)^2. \end{aligned}$$

We first note that $F_2(\delta) = 0$ can have at most 4 solutions. Therefore, there are four solutions at most for (38). In addition, notice that $F_2(+\infty) < 0$ and $F_2(0) = (\|\mathbf{x}_c\|^2 - r^2) \lambda_1^4 > 0$; it follows that there exists at least a solution for (38) with positive δ . Similarly, we have that $\frac{1}{(\lambda_1 - 2\delta)^4} F_2(\delta) < 0$ as $\delta \rightarrow -\infty$, and $\frac{1}{(\lambda_1 - 2\delta)^4} F_2(\delta) \rightarrow +\infty$ as $\delta \rightarrow \frac{\lambda_1^-}{2}$; then, there exists at least one solution for (38) with negative $\delta < \frac{\lambda_1^-}{2}$. \square

Proof of Proposition 5.14

Let $\mathbf{x}_* \in \hat{\mathcal{E}}$ with indicator $\delta_{\mathbf{x}_*} < \frac{\lambda_1}{2}$, and write $\mathbf{x}_* = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$; then, it follows by Lemma 8.8 that the Jacobian evaluated at \mathbf{x}_* has an eigenvalue greater than $\frac{(\lambda_2 - \lambda_1)}{r^2} \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} ((\beta_3 - \beta_1)\beta_1 + (\beta_4 - \beta_2)\beta_1 \mathbf{v}_2^\top \mathbf{v}_1)$. Notice that $\frac{(\lambda_2 - \lambda_1)}{r^2} \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} > 0$ and

$$\begin{aligned} (\beta_3 - \beta_1)\beta_1 + (\beta_4 - \beta_2)\beta_1 \mathbf{v}_2^\top \mathbf{v}_1 &= \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} \beta_1^2 + \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} \beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 \\ &\geq \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} (\beta_1^2 + \beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1) \geq 0. \end{aligned}$$

Hence, the Jacobian evaluated at \mathbf{x}_* has a positive eigenvalue and, thus, \mathbf{x}_* is a saddle point. On the other hand, for any $\mathbf{x}_* \in \hat{\mathcal{E}}$ with indicator $\frac{\lambda_2}{2} < \delta_{\mathbf{x}_*} < 0$, write $\mathbf{x}_* = \beta_3 \mathbf{v}_1 + \beta_4 \mathbf{v}_2$; then, by Lemma 8.7, it follows that the Jacobian evaluated at \mathbf{x}_* has an eigenvalue less than $\frac{(\lambda_2 - \lambda_1)}{r^2} \frac{2\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} ((\beta_3 - \beta_1)\beta_2 \mathbf{v}_1^\top \mathbf{v}_2 + (\beta_4 - \beta_2)\beta_2)$. Notice that $\frac{(\lambda_2 - \lambda_1)}{r^2} \frac{2\lambda_2}{\lambda_1 - 2\delta_{\mathbf{x}_*}} > 0$ and

$$\begin{aligned} (\beta_3 - \beta_1)\beta_2 \mathbf{v}_1^\top \mathbf{v}_2 + (\beta_4 - \beta_2)\beta_2 &= \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} \beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 + \frac{-\lambda_2}{\lambda_2 - 2\delta_{\mathbf{x}_*}} \beta_2^2 \\ &\leq \frac{-\lambda_1}{\lambda_1 - 2\delta_{\mathbf{x}_*}} (\beta_1 \beta_2 \mathbf{v}_2^\top \mathbf{v}_1 + \beta_2^2) \leq 0. \end{aligned}$$

Besides, by [17, Proposition 10], $-\alpha'(0)$ is another eigenvalue. Hence, all the eigenvalues of the Jacobian evaluated at \mathbf{x}_* are negative, which means that \mathbf{x}_* is an undesired asymptotically stable equilibrium.

To prove the last claim, let δ_0 denote the only real root of the third-order polynomial $\frac{dF_1(\delta)}{d\delta}$. It follows that $F_1(\delta)$ is monotonically increasing on $(-\infty, \delta_0)$ and monotonically decreasing on $(\delta_0, +\infty)$; this implies that $F_1(\delta) = 0$ only has two solutions. By Lemma 8.7, there is only one undesired equilibrium and its indicator satisfies $\delta < \frac{\lambda_1}{2}$. Since $\beta_1^2 + \beta_1\beta_2\mathbf{v}_1^\top\mathbf{v}_2 \geq 0$, there is only one undesired equilibrium and it is a saddle point. \square

Remark 8.10. (Connection between the eigenvector-eigenvalue structure): Propositions 5.11 and 5.14 describe the number and stability properties of undesired equilibria based on the specific eigenvector-eigenvalue structure. Proposition 5.11 (cf. Table 1) consists of the cases where \mathbf{x}_c is an eigenvector of \tilde{A} , while Proposition 5.14 focuses on the cases where \mathbf{x}_c is not. We clarify the relationship between them here. Proposition 5.14(i) and (ii) correspond to the first row and last row of Table 1(a), respectively, but apply to cases where \mathbf{x}_c is not an eigenvector. When the two eigenvalues are “highly distinct”, $\lambda_1 \ll \lambda_2$, Proposition 5.14(i) aligns with the first row of Table 1(a). Specifically, suppose that $\tilde{A}\mathbf{x}_c = \lambda_i\mathbf{x}_c$ and $(\mathbf{v}_i^T\mathbf{v}_j)^2 < 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2}$, then we have $|\beta_1| = 0$ or $|\beta_2| = 0$. However, $|\beta_1| = 0$ cannot hold with $\lambda_1 \ll \lambda_2$, since this implies that $1 - \frac{(\lambda_2 - \lambda_1)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2} < 0$, contradicting $(\mathbf{v}_1^T\mathbf{v}_2)^2 < 1 - \frac{(\lambda_2 - \lambda_1)^2 r^2}{\lambda_2^2 \|\mathbf{x}_c\|^2}$. Hence if $\lambda_1 \ll \lambda_2$, the condition of the first row of Table 1(a) says that $|\beta_2| = 0$ and $\mathbf{v}_1^\top\mathbf{v}_2$ is small enough. If $|\beta_1/\beta_2| \geq 1$ (i.e., \mathbf{x}_c is “essentially” an eigenvector associated with λ_1), then $\beta_1^2 + \beta_1\beta_2\mathbf{v}_1^\top\mathbf{v}_2 \geq |\beta_1|(|\beta_1| - |\beta_2||\mathbf{v}_1^\top\mathbf{v}_2|) \geq |\beta_1|(|\beta_1| - |\beta_2|) \geq 0$. Thus the condition $\beta_1^2 + \beta_1\beta_2\mathbf{v}_1^\top\mathbf{v}_2 \geq 0$ can be viewed as a counterpart to the first row of Table 1(a), when $\lambda_1 \ll \lambda_2$.

When the two eigenvectors \mathbf{v}_1 and \mathbf{v}_2 are “highly distinct”, i.e., $\mathbf{v}_1^\top\mathbf{v}_2$ is small enough, Proposition 5.14(ii) corresponds to the last row of Table 1(a). Similarly, we have $|\beta_1| = 0$ or $|\beta_2| = 0$ if \mathbf{x}_c is an eigenvector. However, if $|\beta_2| = 0$, one has that $1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2} > 0$ together with small enough $\mathbf{v}_1^\top\mathbf{v}_2$ violate the inequality $(\mathbf{v}_1^T\mathbf{v}_2)^2 > 1 - \frac{(\lambda_1 - \lambda_2)^2 r^2}{\lambda_1^2 \|\mathbf{x}_c\|^2}$. Hence if $\mathbf{v}_1^\top\mathbf{v}_2$ is small enough, $\tilde{A}\mathbf{x}_c = \lambda_i\mathbf{x}_c$ and $(\mathbf{v}_i^T\mathbf{v}_j)^2 > 1 - \frac{(\lambda_i - \lambda_j)^2 r^2}{\lambda_i^2 \|\mathbf{x}_c\|^2}$ imply that $|\beta_1| = 0$ and $\frac{|\lambda_1|}{|\lambda_2|} > \frac{\|\mathbf{x}_c\|}{r} \sqrt{1 - (\mathbf{v}_1^T\mathbf{v}_2)^2} + 1$, i.e., $\lambda_1 \ll \lambda_2$. If $|\beta_2/\beta_1| \geq 1$ (i.e., \mathbf{x}_c is “essentially” an eigenvector associated with λ_2), then $\beta_2^2 + \beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1 \geq |\beta_2|(|\beta_2| - |\beta_1||\mathbf{v}_1^\top\mathbf{v}_2|) \geq |\beta_2|(|\beta_2| - |\beta_1|) \geq 0$. Hence, provided that $\mathbf{v}_1^\top\mathbf{v}_2$ is small enough, the condition $\beta_1\beta_2\mathbf{v}_2^\top\mathbf{v}_1 + \beta_2^2 \geq 0$ can be viewed as a counterpart to the last row of Table 1(a). \square