CONSTRUCTIVE SAFETY USING CONTROL BARRIER FUNCTIONS

Peter Wieland * and **Frank Allgöwer** *

* *Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany*
{wieland,allgower}@ist.uni-stuttgart.de

Abstract: This paper presents a new safety feedback design for nonlinear systems based on barrier certificates and the idea of control Lyapunov functions. In contrast to existing methods, this approach ensures safety independently of abstract high-level tasks that might be unknown or change over time. Leaving as much freedom as possible to the safe system, the authors believe that the flexibility of this approach is very promising. The design is validated using an illustrative example.

Keywords: Safetycritical, Safety Analysis, Decentralized Control

1. INTRODUCTION

Ensuring safety of dynamical systems is crucial in many applications. Safety verification addresses the question whether a given unsafe region of the state space is guaranteed to be avoided by a given system. The subject of this paper is the associated design task, namely finding an appropriate control strategy that can be used to ensure safety for a given control system. Different approaches to solve the analysis problem exist in the literature including the recently proposed use of barrier certificates (Prajna, 2005; Prajna and Jadbabaie, 2004) and the idea of avoidance control (Leitmann and Skowronski, 1977; Stipanovic, Shankaran and Tomlin, 2005) which dates back to the late seventies. In the context of collision avoidance of multiple agents, different solutions for analysis and design have been proposed (Vanualailai, ichi Nakagiri and Ha, 1995; Masoud and Masoud, 2000; Chang, Shadden, Marsden and Olfati-Saber, 2003; Dimarogonas, Loizou, Kyriakopoulos and Zavlanos, 2006). Common to most of the existing solutions for the design task is, that the low-level safety mechanism (e.g. avoid obstacles) is designed together with the abstract high-level tasks (e.g. move from A to B in the plane) of the systems.

In this paper, a different approach is followed by only considering the low-level safety in a first step. The objective is to design a feedback for a given system, that ensures safety under any circumstances while leaving as much freedom to the system as possible. This results in a modular approach where the safety mechanism can be designed independently and without knowledge of the abstract high-level tasks (however, achievement of high-level tasks may depend on the safety mechanism). Hence, safety is not jeopardized by changes in the high-level tasks of the system under consideration. In contrast to Wieland, Ebenbauer and Allgöwer (2007), the approach followed in this paper does not need the assumption of a second order safety metric.

The main ingredients of the proposed approach are barrier certificates as proposed by Prajna and Jadbabaie (2004) and the idea of control Lyapunov functions and Sontag's formula (Artstein, 1983; Sontag, 1989).

Notations: The usual notation $L_\xi \eta(x)$ is used for the Lie-Derivative $\frac{\partial \eta}{\partial x}(x)\xi(x)$ of $\eta(x)$ along the vector field $\xi(x)$. The space of m-times continuously differentiable functions mapping $\mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is denoted by $C^m(\mathcal{X}, \mathbb{R}^n)$. If $n = 1$, the sec-

ond argument is omitted ($C^m(\mathcal{X})$). The space of continuous functions mapping $\mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is denoted $C(\mathcal{X}, \mathbb{R}^n)$ or simply $C(\mathcal{X})$ if $n = 1$.

The remainder of this paper is organized as follows: In Section 2, barrier certificates are briefly introduced and a precise problem statement is given. Section 3 presents the main result, that is subsequently illustrated on an example in Section 4. Section 5 concludes the paper.

2. PRELIMINARIES

2.1 Barrier Certificates

Before proceeding to the problem statement and the results, barrier certificates are very briefly reviewed.

Definition 1. (Barrier Certificate). Given a system $\dot{x} = f(x)$, $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $f \in C(\mathcal{X}, \mathbb{R}^n)$, a set of initial states $\mathcal{X}_0 \subseteq \mathcal{X}$, and a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$. A function $B \in C^1(\mathcal{X})$ satisfying

$$B(x) \leq 0 \qquad \forall x \in \mathcal{X}_0, \tag{1}$$
$$B(x) > 0 \qquad \forall x \in \mathcal{X}_u, \tag{2}$$
$$L_f B(x) \leq 0 \qquad \forall x \in \mathcal{X} \tag{3}$$

is termed *barrier certificate* for the given system.

Theorem 2. Given a system $\dot{x} = f(x)$, $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $f \in C(\mathcal{X}, \mathbb{R}^n)$, a set of initial states $\mathcal{X}_0 \subseteq \mathcal{X}$, and a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$. If a barrier certificate can be found for that system, safety is guaranteed, i.e., there is no trajectory $x(t)$ of the system such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

Theorem 2 is actually quite intuitive. For a proof and more details on barrier certificates, the reader is referred to Prajna (2005) and the references therein.

2.2 Problem statement

Note that in Definition 1 and Theorem 2, the sets $\mathcal{X}$, $\mathcal{X}_0$ and $\mathcal{X}_u$ were assumed to be known. While this is a natural assumption for $\mathcal{X}$ and the set of unsafe states $\mathcal{X}_u$, the authors argue it is not necessarily a valid assumption for the set of initial states $\mathcal{X}_0$. Usually, the set of unsafe states is defined by some physical constraints or safety requirements. The set of possible initial states can be chosen as some subset of $\mathcal{X}_0^* = \mathcal{X} \setminus \mathcal{X}_u$. It is advantageous to choose $\mathcal{X}_0$ as big as possible in order to put as few restrictions as possible on the system, while safety is guaranteed. But choosing $\mathcal{X}_0 = \mathcal{X}_0^*$ does not always result in a safe system (Wieland et al., 2007). The choice of $\mathcal{X}_0 \subseteq \mathcal{X}_0^*$ will thus be one of the design tasks. Furthermore,

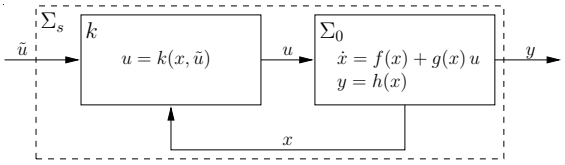

Fig. 1. Block diagram of safe system Σ_s with original system Σ_0 and safety feedback k.

notice that the class of systems considered in Definition 1 and Theorem 2 are closed nonlinear systems. In order to be able to perform a safety feedback design, in this contribution nonlinear input affine systems

$$\dot{x} = f(x) + g(x)u, \quad x \in \mathcal{X} \subseteq \mathbb{R}^n,\ u \in \mathbb{R}^p \tag{4}$$

are considered, where $f \in C(\mathcal{X}, \mathbb{R}^n)$ and $g \in C(\mathcal{X}, \mathbb{R}^{n \times p})$. The problem considered in this paper can thus be stated as follows:

Problem 3. Given a set $\mathcal{X} \subseteq \mathbb{R}^n$, a system (4), and a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$, find a feedback $u = k(x, \tilde{u})$, $\tilde{u} \in \mathbb{R}^p$, satisfying $k(x, \tilde{u}) = \tilde{u}$ "whenever possible", and a set of admissible initial states $\mathcal{X}_0 \subseteq \mathcal{X}_0^*$ such that the system $\dot{x} = f(x) + g(x)k(x, \tilde{u})$ is safe independently of $\tilde{u}$, i.e., there exists no time instant $T \geq 0$ and input trajectory $\tilde{u} : [0, T] \to \mathbb{R}^p$ giving rise to a system trajectory $x(t)$ such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

Looking at the block structure resulting from Problem 3 (see Fig. 1), the feedback $u = k(x, \tilde{u})$ looks just like a feedback stabilizer. Though, the new input $\tilde{u}$ is neither a reference nor a disturbance signal in the classical sense.

Problem 3 can be simplified by fixing the structure of the feedback $u = k(x, \tilde{u})$. In this paper, feedbacks of the form

$$k(x, \tilde{u}) = \sigma(x)k_0(x) + (1 - \sigma(x))\tilde{u} \tag{5}$$

are assumed, where $\sigma(x)$ is a sufficiently smooth function $\sigma : \mathbb{R}^n \to [0, 1]$ satisfying

- $\sigma(x) = 1$ if x is contained in some closed set containing $\partial\mathcal{X}_0 \cap \text{int}(\mathcal{X})$. That is, when trajectories risk to leave the set $\mathcal{X}_0$ while staying in $\mathcal{X}$, $k(x, \tilde{u}) = k_0(x)$.
- $\sigma(x) = 0$ if x is contained in some closed set contained in $\text{int}(\mathcal{X}_0)$. That is, when trajectories are bounded away from $\partial\mathcal{X}_0$ in $\text{int}(\mathcal{X}_0)$, $k(x, \tilde{u}) = \tilde{u}$.

The feedback (5) ensures, that trajectories can not leave $\mathcal{X}_0$ while staying in $\mathcal{X}$ without crossing parts of the state space where $\sigma(x) = 1$ and thus $k(x, \tilde{u}) = k_0(x)$. This leads to the following obvious fact:

Fact 4. The feedback $u = k(x, \tilde{u})$ as defined by (5) ensures safety for system (4) independently of $\tilde{u}$ if and only if the feedback $u = k_0(x)$ ensures safety for system (4).

Using Fact 4, Problem 3 can be replaced by the following simpler problem, for which a solution is proposed in this paper.

Problem 5. Given a set $\mathcal{X} \subseteq \mathbb{R}^n$, a system (4), and a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$, find a feedback $u = k_0(x)$ and a set of admissible initial states $\mathcal{X}_0 \subseteq \mathcal{X} \setminus \mathcal{X}_u$ such that the system $\dot{x} = f(x) + g(x)k_0(x)$ is safe, i.e., there exists no system trajectory $x(t)$ such that $x(0) \in \mathcal{X}_0$, $x(T) \in \mathcal{X}_u$ for some $T \geq 0$, and $x(t) \in \mathcal{X}$ for all $t \in [0, T]$.

3. CONTROL BARRIER FUNCTIONS

Inspired by control Lyapunov functions (CLF) (Artstein, 1983), the idea of this paper is to use what we term control barrier functions to construct a safety feedback $k_0(x)$ and determine the set of admissible initial states $\mathcal{X}_0$ that solve Problem 5. We start by defining control barrier functions.

Definition 6. (Control Barrier Function). Given a system (4) and a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$. A function $B \in C^1(\mathcal{X})$ satisfying

$$x \in \mathcal{X}_u \;\Rightarrow\; B(x) > 0, \tag{6}$$

$$L_g B(x) = 0 \;\Rightarrow\; L_f B(x) < 0, \tag{7}$$

$$\{x \in \mathcal{X} \,|\, B(x) \leq 0\} \neq \emptyset \tag{8}$$

is termed *control barrier function* (CBF).

Using CBFs, the main contribution of this paper is stated in the following theorem.

Theorem 7. Given a system (4), a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$, and a CBF $B(x)$ for that system. Define

$$k_0(x) = \begin{cases} -\dfrac{a + \sqrt{a^2 + \kappa^2 b^T b}}{b^T b} b & \text{if } b \neq 0, \\ 0 & \text{if } b = 0, \end{cases} \tag{9}$$

with $a(x) := L_f B(x)$, $b^T(x) := L_g B(x)$, and $\kappa > 0$ a free design parameter. Then the set of initial states can be taken as

$$\mathcal{X}_0 = \{x \in \mathcal{X} \,|B(x) \leq 0\} \tag{10}$$

and the control $u = k_0(x)$ is continuous in x and ensures safety for the system $\dot{x} = f(x)+g(x)k_0(x)$.

PROOF. As in the case of CLFs, the CBF $B(x)$ serves as a barrier certificate for the closed loop system. Conditions (1) and (2) are satisfied by the definition of CBFs and the construction of $\mathcal{X}_0$. It remains to verify, that condition (3) holds for the closed loop system $\dot{x} = f(x)+g(x)k_0(x) =: f_{\text{cl}}(x)$. One obtains

$$\begin{aligned} L_{f_{\text{cl}}} B(x) &= L_f B(x) + L_g B(x) k_0(x) \\ &= \begin{cases} -\sqrt{a^2 + \kappa^2 b^T b} & \text{if } b \neq 0, \\ a & \text{if } b = 0. \end{cases} \end{aligned}$$

If $b = 0$, it follows $a < 0$ by (7), if $b \neq 0$, clearly $-\sqrt{a^2 + \kappa^2 b^T b} \leq -\kappa\sqrt{b^T b} < 0$, hence, in both cases $L_{f_{\text{cl}}} B(x) < 0$. This proves in fact, that $\mathcal{X}_0$ is invariant for the considered system.

To see that $k_0(x)$ is continuous in x on $\mathcal{X}$, note that it is continuous in a, b since (9) is analytic for $b \neq 0 \vee a < 0$. Continuity of a, b in x follows, since $B(x)$ is continuously differentiable and the vector fields f, g are continuous by assumption. □

Note, that (9) is nothing but a version of Sontag's formula (Sontag, 1989), slightly differing from the commonly used version $-(a + \sqrt{a^2 + (b^T b)^2})/b$, where the square of $b^T b$ is taken to guarantee continuity of the feedback. As will be shown, continuity of $k_0(x)$ is not needed everywhere to ensure safety, hence, for our purpose formula (9) is adequate. In fact, continuity of solutions $x(t)$ of (4) leads to the following corollary:

Corollary 8. Given a system (4), a set of unsafe states $\mathcal{X}_u \subseteq \mathcal{X}$, and $\bar{\mathcal{X}}_u$ a closed subset of the interior of $\mathcal{X}_u$. If a barrier function that satisfies conditions (6) – (8) for the reduced sets $\mathcal{X} \setminus \bar{\mathcal{X}}_u$ and $\mathcal{X}_u \setminus \bar{\mathcal{X}}_u$ can be found, the system is safe. Thus, conditions (6) – (8) and continuous differentiability of $B(x)$ in Theorem 7 are not required to hold for $x \in \bar{\mathcal{X}}_u$. Moreover discontinuities of $k_0(x)$ for $x \in \bar{\mathcal{X}}_u$ do not pose any problems.

Equivalently, consider $\bar{\mathcal{X}}_0$ a closed subset of the interior of $\mathcal{X}_0$ as defined by (10). Then, condition (7), continuous differentiability of $B(x)$, and continuity of $k_0(x)$ are not required to hold for $x \in \bar{\mathcal{X}}_0$.

Remark 9. In the proof of Theorem 7, it was shown that a CBF $B(x)$ for a system (4) serves as a barrier certificate for the closed-loop system with feedback (9). Actually, $B(x)$ satisfies stronger conditions then those of Definition 1 because condition (3) is strictly satisfied, i.e. $L_{f_{\text{cl}}} B(x) < 0$. This is a result of the strict inequality in condition (7), which is necessary in order to guarantee continuity of the feedback $k_0(x)$. In the single input case, this can be seen by noting that

$$\frac{a + \sqrt{a^2 + \kappa^2 b^T b}}{b^T b} b = \frac{a + \sqrt{a^2 + (\kappa b)^2}}{b} = \kappa \cot(\varphi)$$

where the angle φ is defined by the relation

$$\kappa b = a \tan(2\varphi)$$

with an obvious discontinuity at $a = b = 0$.

So far, CBFs where introduced, the set $\mathcal{X}_0$ was characterized depending on the CBF, and an explicit formula for a safety feedback, which is due to Sontag (1989), was given. Hence, if a CBF for a system (4) can be found, a solution

to Problem 5 is given by (10), (9). Nothing was said so far on how to find a CBF for a given system. Furthermore, from the approach followed in this paper, it is obvious that no guarantees can be given on optimality of a specific CBF in the sense that it leads to the largest possible set $\mathcal{X}_0$ of admissible initial states. Actually, if different CBFs $B_i(x)$, $i = 1, \ldots, l$ are known for system (4) and $\mathcal{X}_{0,i}$, $i = 1, \ldots, l$ are the corresponding sets of admissible initial states, it is possible to ensure safety for system (4) with the set of initial states chosen as $\mathcal{X}_0 = \bigcup_{i=1}^{l} \mathcal{X}_{0,i}$. However, it is not clear which feedback function $k_0(x)$ can be used to ensure safety for this set of initial conditions. Additionally, input constraints may limit the possible choice of $B(x)$ and $\mathcal{X}_0$. Yet, these constraints cannot be taken into account explicitely in the CBF-approach so far. From the above, it turns out that finding a CBF is a difficult task in general. Fortunately, some problem insight can often be used to construct CBFs and input constraints can be taken care of by shaping the CBF $B(x)$ iteratively as illustrated in Section 4.

This section is concluded by two additonal remarks concerning safety for multiple systems with multiple unsafe parts of the state space and necessity of CBFs for ensuring safety by feedback.

Remark 10. It is a common feature of safety critical systems that several (independent) subsystems need to be considered and several unsafe parts of the state space exist. An example for such a scenario is collsion avoidance of mobile agents. CBFs as proposed in this paper are applicable for such scenarios. Multiple systems can be handled by considering the overall system. Choosing the sets involved in the design carefully, the safety feedback for each subsystem will only depend on the states of very few "neighboring" subsystems. Thus, good scalability can be achieved. Multiple unsafe parts of the state space can be approached in two ways. A first (direct) way consistes in considering the unsafe set to be the union of all the parts of the statespace that are unsafe. This is possible, because connectedness of $\mathcal{X}_u$ is not assumed in the proposed approach. A second way is due to the structure of the safety feedback defined by (5). Assume the unsafe parts of the state space are such that one safety feedback $k_i(x, \tilde{u}) = \sigma_i(x)k_{0,i}(x) + (1 - \sigma_i(x))\tilde{u}$ for each of them can be designed. If the functions $\sigma_i(x), i = 1, \ldots, l$ satisfy $\sigma_i(x) \neq 0 \Rightarrow \sigma_j(x) = 0$ for all $j \neq i$, i.e. only one safety feedback is active at each time instant, the feedback $k(x, \tilde{u}) = \sum_{i=1}^{l} k_i(x, \tilde{u})$ ensures safety for all subsystems.

Remark 11. Theorem 7 gives a sufficient condition for the existence of a feedback that ensures safety for system (4). A natural question to ask is whether existence of a CBF is also necessary for the possibility to ensure safety by feedback. Note that given a closed system $\dot{x} = f(x) + g(x)k(x)$ and a barrier certificate $B(x)$ that proves safety for that system with strict inequality in (3), $B(x)$ is a CBF for the open system $\dot{x} = f(x) + g(x)u$. Hence, necessity of CBFs for ensuring safety by feedback is equivalent to necessity of barrier certificates for safety that satisfy (3) strictly. If discontinuities in the safety feedback are acceptable, the inequality constraint in (3) may also be non-strict (see Remark 9) resulting in a weak CBF. It was shown by Prajna and Rantzer (2005) that barrier certificates are necessary and sufficient for safety of $\dot{x} = f(x)$, if $\mathcal{X}, \mathcal{X}_0, \mathcal{X}_u$ are compact and there exists a function $\tilde{B} \in C^1(\mathbb{R}^n)$ such that $L_f\tilde{B}(x) < 0$ for all $x \in \mathcal{X}$. This yields the following theorem which is stated without proof:

Theorem 12. Assuming compactness of $\mathcal{X}$ and $\mathcal{X}_u$ and existance of a function $\tilde{B} \in C^1(\mathbb{R}^n)$: $L_f\tilde{B}(x) < 0 \, \forall x \in \mathcal{X}$, existance of a weak CBF is a necessary and sufficient condition for the existance of a possibly discontinuous safety feedback.

4. ILLUSTRATING EXAMPLE

The example chosen to illustrate the proposed approach is fairly small and simple in order to easily demonstrate the ideas. In particular, a linear model is considered for the seek of easy understanding, while the method itself was derived for general nonlinear input affine models.

Consider the system governed by

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & -\mu \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u, \quad x \in \mathbb{R}^2, u \in \mathbb{R}, \tag{11}$$

which describes e.g. force-controlled one-dimensional motion under viscous damping with position x_1, velocity x_2 and damping coefficient $\mu > 0$. If the control is constrained to $|u| \leq u_{\max}$, the state is naturally constrained by $|x_2| \leq \frac{u_{\max}}{\mu}$. Thus, we restrict ourselves to the part of the state space given by

$$\mathcal{X} = \{x \in \mathbb{R}^2 | -\frac{u_{\max}}{\mu} \leq x_2 \leq \frac{u_{\max}}{\mu}\}.$$

The unsafe part of the state-space is assumed as

$$\mathcal{X}_u = \{x \in \mathcal{X} | -d_u < x_1 < d_u\}, \tag{12}$$

i.e., the position x_1 must stay at least d_u units away from the (geometric) origin.

As a first guess for a CBF we take

$$B_0(x) = d_u - |x_1|,$$

which is chosen to satisfy (6) and (8) while $\mathcal{X}_0 = \mathcal{X} \setminus \mathcal{X}_u$. By Corollary 8, non-differentiability of $B_0(x)$ at $x_1 = 0$ does not pose any problem. With

$$a_0(x) := L_f B_0(x) = -\text{sign}(x_1)x_2,$$
$$b_0(x) := L_g B_0(x) = 0,$$

condition (7) is satisfied only if $x_1 x_2 > 0$. If $x_1 x_2 < 0$, the velocity is directed into the unsafe

set $\mathcal{X}_u$. Not surprisingly, $B_0(x)$ is not a CBF for (11). Thus, the CBF candidate has to be changed such that $\mathcal{X}_0$ becomes smaller. Physically, some braking distance is needed. This motivates the CBF candidate

$$B(x) = \begin{cases} B_0(x), & \text{if } a_0(x) \leq -\beta, \\ B_0(x) + \alpha(a_0(x)+\beta)^2, & \text{if } a_0(x) > -\beta, \end{cases}$$

where $\alpha > 0$ and $\beta > 0$ are parameters to be determined. Clearly, $B(x)$ is continuously differentiable if $x_1 \neq 0$. Again, conditions (6) and (8) are satisfied and now

$$a(x) = \begin{cases} a_0(x), & \text{if } a_0(x) \leq -\beta, \\ -x_2 - 2\alpha\mu(x_2+\beta)x_2, & \text{if } a_0(x) > -\beta, \\ & x_1 > 0, \\ x_2 - 2\alpha\mu(x_2-\beta)x_2, & \text{if } a_0(x) > -\beta, \\ & x_1 < 0, \end{cases}$$

and

$$b(x) = \begin{cases} 0, & \text{if } a_0(x) \leq -\beta, \\ 2\alpha(x_2 + \text{sign}(x_1)\beta), & \text{if } a_0(x) > -\beta. \end{cases}$$

Observe that $B(x)$ is a CBF for (11). The feedback $k_0(x)$ is obtained from $a(x)$ and $b(x)$ using (9). The lengthy expression is omitted, due to limited space. It can be observed that $k_0(x) = -k_0(-x)$ and $k_0(x)$ depends on $\text{sign}(x_1)$ and x_2 only, but not on $|x_1|$. The parameters $\alpha > 0$, $\beta > 0$, and $\kappa > 0$ can be used to shape $B(x)$ such that input constraints $|u| \leq u_{\max}$ are satisfied. It turns out that

$$\sup_{x \in \mathcal{X}} |k_0(x)| \geq \kappa, \quad \forall\, \alpha > 0,\ \beta > 0.$$

Equality holds if and only if $2\alpha\beta\mu = 1$. In this case, the supremum is attained at $x_2 = 0$. Hence, the choice $\beta = 1/(2\alpha\mu)$ and $\kappa \leq u_{\max}$ ensures that input constraints are not violated, i.e. there exists a safety feedback for arbitrarily small $u_{\max}$. However, it is not always best to choose α and β such that $2\alpha\beta\mu = 1$ as this may result in unnecessarily small sets $\mathcal{X}_0$.

In the sequel, the problem parameters are assumed as $\mu = 1$, $u_{\max} = 10$, and $d_u = 1$. The feedback parameters are chosen as $\alpha = \frac{1}{20}$, $\beta = 1$ (note that $2\alpha\beta\mu \neq 1$), and $\kappa = 1$. The resulting values of $k_0(x)$ (for $x_1 > 0$) are depicted in Fig. 2. Decreasing values of $k_0(x)$ for increasing $|x_2|$ can be explained by increasing influence of the damping term $-\mu x$. Clearly $|k_0(x)| < u_{\max}$ for all $x \in \mathcal{X}$. This fact indicates some conservatism (or robustness) in the proposed solution: If the largest possible set of initial states was chosen, one would expect $|k_0(x)| = u_{max}$ on the boundary of this set. Varying κ, $\max_{x\in\mathcal{X}} |k_0(x)|$ can only be reduced very little for the considered set of parameters. Hence, the choice of κ does not introduce significant conservatism. The sets $\mathcal{X}_0$ and $\mathcal{X}_u$ obtained with the above parameters are depicted in Fig. 3.

From $k_0(x)$, a solution to Problem 3 is obtained constructing a safety feedback $k(x, \tilde{u})$ with structure as defined in (5) using the barrier function $B(x)$. To do so, the function $\sigma(x)$ is defined as

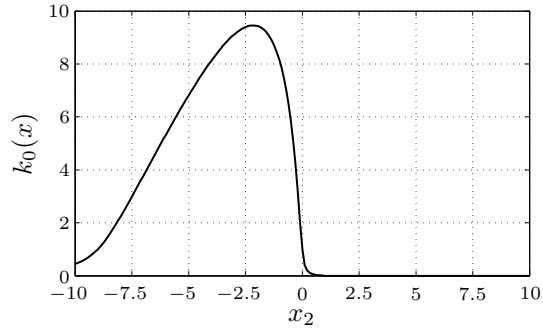

Fig. 2. Safety feedback $k_0(x)$ for $x_1 > 0$.

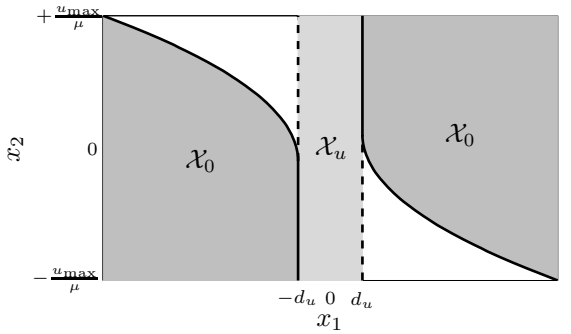

Fig. 3. Sets of initial and unsafe states.

$$\sigma(x) = \begin{cases} 0 & \text{if } B(x) \leq -\epsilon, \\ \tilde{\sigma}(x) & \text{if } B(x) \in (-\epsilon, 0), \\ 1 & \text{if } B(x) \geq 0, \end{cases} \tag{13}$$

$$\tilde{\sigma}(x) = -2\left(\frac{B(x)}{\epsilon}\right)^3 - 3\left(\frac{B(x)}{\epsilon}\right)^2 + 1$$

for some (small) $\epsilon > 0$. Observe that $\sigma(x)$ is a continuous approximation of the unit step of $B(x)$.

To illustrate the safety feature, the new input is taken to be $\tilde{u} = k_P(x_1^d - x_1)$ with $k_P = 20$ and $x_1^d = 1.5$. Fig. 4 shows simulation results for the original system and the safe system. It is observed that trajectories of the original system penetrate the set of unsafe states $\mathcal{X}_u$ while trajectories of the safe system do not, as desired. Both systems asymptotically reach $x_1 = x_1^d$.

Safety of (11) is considered for a similar scenario in Wieland et al. (2007) where the fact that (12) can be characterized with a second order safety pseudo-metric is used to construct the set $\mathcal{X}_0$ explicitely. The safety feedback is given as $k_0(x) = \text{sign}(x_1)u_{\max}$. Using CBFs, a safety feedback satisfying $|k_0(x)| < u_{\max}$ can be applied to ensure safety. A major difference compared to Wieland et al. (2007) is that the CBF-approach implicitely considers the damping, which is helpful in ensuring safety, while damping was neglected in the prior paper.

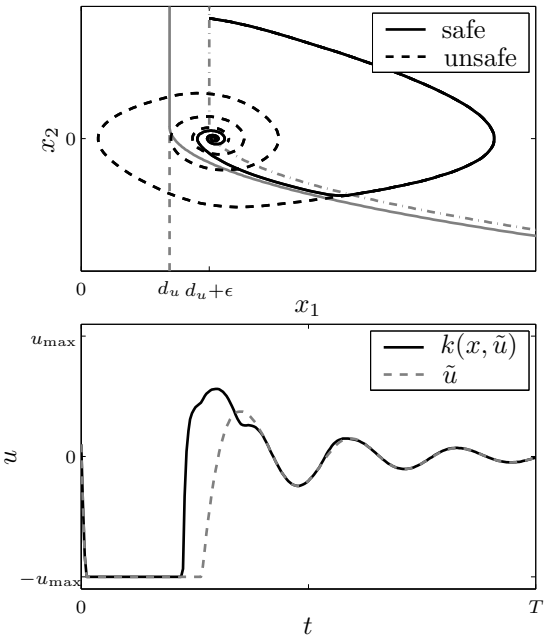

Fig. 4. Simulation results. Top: phase plot for safe and unsafe system with level curve $B(x) = -\epsilon$ (grey dash-dotted), level curve $B(x) = 0$ (grey solid), and line $d(x) = d_u$ (grey dashed); Bottom: input trajectories for safe system.

5. CONCLUSIONS

The goal of the present paper was to ensure safety for dynamical systems by appropriate feedback strategies. It is desired to obtain safety feedbacks that are independent of higher level tasks of the system. For that purpose, these feedback strategies must leave as much freedom as possible to the system. The problem of finding such feedback strategies was solved using barrier certificates and the idea of control Lyapunov functions. Merging these two main ingredients, the new concept of CBFs was proposed in Section 3. Once a CBF is found for a given system, it can be used to define the set of admissible initial states and a feedback strategy that ensures safety for that system. In contrast to existing solutions, the CBF approach implicitly takes into account parts of the system dynamics that are beneficial for system safety, e.g. damping as seen in Section 4.

To illustrate the approach, a simple example was examined in Section 4. With the help of this example, some more insight into the approach, concering e.g. input constraints and questions of conservatism, was gained. More complex systems can be treated analogous to what was illustrated in Section 4. Especially two-dimensional motion and multiple independent subsystem with the objective of collision avoidance are easy to handle with the proposed approach similarly to Wieland et al. (2007).

In the example presented in this paper, it was fairly simple to find a CBF. However, as is typical for Lyapunov based approaches, for bigger and more complex systems, finding CBFs may be a difficult task. Thus, future research could be directed towards finding systematic ways to constructing CBFs, similar to those existing for construction of CLFs. Furthermore, it is not completely understood what kind of interdependencies between a safety feedback and a high level task can occur. Problems may range from windup phenomena to deadlocks. Further research is needed to solve these interesting open problems.

REFERENCES

Artstein, Z. (1983). Stabilization with relaxed controls, *Nonlinear Analysis* **7**: 1163–1173.

Chang, D. E., Shadden, S. C., Marsden, J. E. and Olfati-Saber, R. (2003). Collision avoidance for multiple agent systems, *Proc. 42nd IEEE Conf. Dec. and Contr.*, pp. 539–543.

Dimarogonas, D. V., Loizou, S. G., Kyriakopoulos, K. J. and Zavlanos, M. M. (2006). A feedback stabilization and collision avoidance scheme for multiple independent non-point agents, *Automatica* **42**(2): 229–243.

Leitmann, G. and Skowronski, J. (1977). Avoidance control, *J. of Opt. Theory and Appl.* **23**(4): 581–591.

Masoud, S. A. and Masoud, A. A. (2000). Constrained motion control using vector potential fields, *IEEE Trans. Syst., Man, Cybern. A* **30**(3): 251–272.

Prajna, S. (2005). *Optimization-Based Methods for Nonlinear and Hybrid Systems Verification*, PhD thesis, California Institute of Technology.

Prajna, S. and Jadbabaie, A. (2004). Safety verification of hybrid systems using barrier certificates, *in* R. Alur and G. J. Pappas (eds), *Hybrid Systems: Computation and Control*, Vol. 2993 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 477–492.

Prajna, S. and Rantzer, A. (2005). On the necessity of barrier certificates, *Proc. IFAC World Congress*. Topic 2.3 Paper We-M14-T0/1.

Sontag, E. D. (1989). A universal construction of arstein's theorem on nonlinear stabilization, *Sys. and Contr. Letters* **13**: 117–123.

Stipanovic, D. M., Shankaran, S. and Tomlin, C. J. (2005). Multi-agent avoidance control using an m-matrix property, *Electronic Journal of Linear Algebra* **12**: 64–72.

Vanualailai, J., ichi Nakagiri, S. and Ha, J.-H. (1995). Collision avoidance in a two-point system via liapunov's second method, *Mathematics and Computers in Simulation* **39**: 125–141.

Wieland, P., Ebenbauer, C. and Allgöwer, F. (2007). Task-independent safety for multi-agent systems by feedback, *Proc. American Control Conf.*