Density Functions for Navigation-Function-Based Systems

Savvas G. Loizou, *Member, IEEE*,
and Ali Jadbabaie, *Senior Member, IEEE*

***Abstract*—In this paper, we present a scheme for constructing density functions for systems that are almost globally asymptotically stable (i.e., systems for which all trajectories converge to an equilibrium except for a set of measure zero) using navigation functions (NFs). Although recently-proven converse theorems guarantee the existence of density functions for such systems, such results are only existential and the construction of a density function for almost globally asymptotically stable systems remains a challenging task. We show that for a specific class of dynamical systems that are defined based on an NF, a density function can be easily derived from the system's underlying NF.**

***Index Terms*—Almost gas systems, density functions, dual Lyapunov techniques, navigation functions (NFs).**

I. Introduction

For more than a century, Lyapunov's method has been the major tool used in stability analysis of dynamical systems. Recently, however, a new scheme was proposed by Rantzer [1], which can be thought of as a "dual" to Lyapunov's method. Instead of checking for a positive definite "energy-like" function whose directional derivative along the trajectories of the dynamical system is negative definite, in Rantzer's approach, one searches for a positive "density function" such that the divergence of the vector field $\times$ the density function is positive, almost everywhere. This scalar function has a physical interpretation as the stationary density of a substance that is generated in all points of the state space and flows along the system trajectories. With the stationary density bounded everywhere except at a singularity at the equilibrium, the attractivity of this equilibrium is guaranteed for almost all initial conditions. This is, of course, a weaker result than global asymptotic stability. However, it is a powerful tool for controller synthesis as well as controller composition. This is due to the fact that the synthesis condition for the almost global stability criterion is convex [2]. As a result (at least in the case of polynomial vector fields), convex optimization can be used to search for density functions *and* the controller simultaneously. Furthermore, the convexity argument allows us to compose different controllers and be able to find a density for the composed system. This implies that once we have constructed controllers achieving certain behaviors for our system and we have density functions for those controllers, we can do a parallel composition of behaviors by using a convex combination of the controllers, appropriately weighted by their density function rates [3]. This property can be very useful, e.g., in tasks like multiagent navigation with connectivity constraints, modular composition of complex navigation tasks from simple primitives, and generally navigation tasks that require secondary motion tasks to be run in parallel.

Since the pioneering work of Rantzer, several authors have been able to prove different results analogous to the ones available for asymptotic stability. For example, Rantzer has shown in [2] that given a Lyapunov function which proves global asymptotic stability, one can construct a density function by using the powers of the reciprocal of the Lyapunov function. Also, Monzón [4] and Rantzer [2] have been able to prove converse theorems for almost global stability, similar to converse theorems that guarantee existence of a Lyapunov function for asymptotically stable systems. In [2], Rantzer has proven that existence of density functions is a necessary and sufficient condition for systems that are almost globally stable. Unfortunately, similar to the converse Lyapunov theorems, such results are only existential and cannot be used to construct density functions. Some remarks on the structure of density function candidates are discussed in [5], where it is pointed out that the $\mathcal{C}^1$ continuity requirements on the density functions by converse theorems pose strong constraints in the case of systems with negative divergence in the vicinity of their saddle points.

The purpose of this paper is to show that in certain special cases, such construction is indeed possible. Specifically, we show that for navigation vector fields (NVFs) derived by appropriately transforming the vector field generated by a Rimon–Koditschek navigation function (NF) [6], one can readily construct a density function using the NF.

NFs have been proven extremely useful for rigorously constructing paths that navigate a kinematic robot in a spherical workspace while avoiding spherical obstacles. The construction procedure utilizes Morse theory [7] to construct an artificial potential function, which is zero at the goal state and uniformly maximal at the boundary of the workspace and obstacles. Furthermore, all the critical points of this potential are designed to be saddle points except for the goal state, where the critical point is stable. By constructing a gradient flow based on this potential, it is possible to guarantee that, for almost all initial conditions, the trajectories converge to the goal state while avoiding obstacles.

One can immediately notice parallels between the density function and an NF. This similarity leads us to ask whether it is possible to construct a density function from an NF. We will show that the answer to this question is indeed positive.

The rest of the paper is organized as follows: In Section II, we present some preliminary definitions. Section III presents a review of NFs while Section IV reviews some results on dual Lyapunov Techniques. Our main result is presented in Section V. A simple example is presented in Section VI and the paper concludes with Section VII.

II. Preliminaries

A. Definitions

Let $V : M \to \mathbb{R}$ be a smooth function and $M \subset \mathbb{R}^n$ a smooth manifold with boundary. A point $p \in M$ is called a critical point of V if $\nabla V(p) = 0$, where

$$\nabla V \triangleq \left[\frac{\partial V}{\partial x_1} \cdots \frac{\partial V}{\partial x_n} \right]^T$$

is the gradient of V. The divergence of V is defined as

$$\mathrm{div}(V) \equiv \nabla \cdot V \triangleq \frac{\partial V}{\partial x_1} + \cdots + \frac{\partial V}{\partial x_n}.$$

A critical point p is called nondegenerate iff the matrix $H_V(p) \triangleq [\partial^2 V / \partial x^i \partial x^j]$ is nonsingular. The matrix $H_V(\cdot)$ is called the Hessian of V where $(x^1, \ldots, x^n)$ is a coordinate system. $H_V(\cdot)$ is well known to be symmetric and the nondegeneracy of p does not depend on the chosen coordinate system [7]. A smooth function V is called a *Morse* function if all its critical points are nondegenerate. Function V is called polar if it has a unique minimum in $\mathcal{M}$ and admissible if it attains the unit value uniformly across the boundary of M, that is $\partial\mathcal{M} = \varphi^{-1}(1)$. The boundary of M is denoted by ∂M. The interior of $\mathcal{M}$ is denoted by $\mathring{\mathcal{M}}$. Let the function $f(x) = [f_1(x), \ldots f_n(x)]$ denote a vector field.

Manuscript received May 29, 2006; revised December 22, 2006 and May 31, 2007. Recommended by Associate Editor D. Nesic. This work was supported by the following Grants ARO-MURI W911NF-05-1-0381, ONR/YIP N00014-04-1-0467, ONR N00014-06-1-436, and NSF-ECS-0347285.

G. Loizou is with the Naval Architecture and Marine Engineering Department of the National Technical University of Athens, Athens 15704, Greece (e-mail: sloizou@ieee.org).

A. Jadbabaie is with the GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: jadbabaie@grasp.upenn.edu).

Digital Object Identifier 10.1109/TAC.2008.917745

The matrix $J_f(x)$ whose ijth element is

$$[J_f(x)]_{ij} = \frac{\partial f_i}{\partial x^j}(x)$$

is called the Jacobian of the vector field f at x. Given any $x_0 \in \mathbb{R}^n$, we denote by $\phi_t(x_0)$, for $t \geq 0$, the solution of $\dot{x}(t) = f(x(t))$ with $x(0) = x_0$. We denote with $\mathrm{C}^n(A, B)$ the set of C^n continuous functions mapping elements of A to elements of B.

III. Navigation Functions

NFs [6] are a special category of potential functions. Their negated gradient vector field is attractive toward the goal configuration and repulsive with respect to obstacles. Consider a system described by a kinematic model as $\dot{q} = u$. The basic idea behind NFs is to use a control law of the form $u = -\nabla\varphi(q)$, where $\varphi(q)$ is an NF, to drive the system to its desired configuration.

It has been shown (Koditschek and Rimon [6]) that strict global navigation (i.e., with a globally attracting equilibrium state) is not possible and a smooth vector field on any sphere world, which has a unique attractor, must have at least as many saddles as obstacles. A sphere world is a compact connected subset of E^n whose boundary is formed from the disjoint union of a finite number of $(n-1)$ spheres. Furthermore, the same authors [6] show that navigation properties are invariant under diffeomorphisms; hence, any world that can be diffeomorphically transformed to a sphere world can accept an NF [8]–[10]. Recent extensions of NFs to the multiple-disk-shaped robots case have been independently proposed by the first author [11] and by the authors in [12].

Formally, an NF is defined as follows:

Definition 1: [6] Let $\mathcal{F} \subset E^n$ be a compact connected analytic manifold with boundary. A map $\varphi : \mathcal{F} \to [0, 1]$, is an NF if it is:

1) analytic on $\mathcal{F}$;
2) polar on $\mathcal{F}$, with minimum at $q_d \in \mathring{\mathcal{F}}$;
3) Morse on $\mathcal{F}$;
4) admissible on $\mathcal{F}$.

The intuition behind property 1 of Definition 1 is that it is preferable to have an analytic form of the gradient of the vector field to encode actuator commands instead of "patching together" closed-form expressions on different portions of space, in order to avoid branching and looping in the control algorithm.

By using smooth vector fields, one cannot do better than almost global navigation [6]. As a result of using a polar function on a compact connected manifold with boundary, all initial conditions will either be brought to a saddle point or to the unique minimum: $\boldsymbol{q}_d$.

The requirement in Definition 1 that an NF be a Morse function, establishes that the initial conditions that bring the system to saddle points are sets of measure zero [7]. In view of this property, all initial conditions other than sets of measure zero are brought to q_d.

The last property of Definition 1 guarantees that the resulting vector field is transverse to the boundary of $\mathcal{F}$, which establishes that the generated trajectories are collision-free.

IV. Dual Lyapunov Techniques

The dual Lyapunov criterion for convergence introduced by Rantzer, can be summarized in the following theorem:

Theorem 1: [1] Given the equation $\dot{x} = f(x(t))$, where $f \in \mathrm{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ and $f(0) = 0$, suppose there exists a nonnegative $\rho \in \mathrm{C}^1(\mathbb{R}^n \setminus \{0\}, \mathbb{R})$ such that $\rho(x)f(x)/|x|$ is integrable on $\{x \in \mathbb{R}^n : |x| \geq 1\}$ and $[\nabla \cdot (f\rho)](x) > 0$ for almost all x. Then, for almost all initial states $x(0)$, the trajectory $x(t)$ exists for $t \geq 0$ and tends to zero as $t \to \infty$. Moreover, if the equilibrium $x = 0$ is stable, then the conclusion remains valid even if ρ takes negative values.

The converse result regarding the necessary and sufficient conditions for almost global stability of nonlinear systems, is stated next.

Theorem 2: [2] Given $f \in \mathrm{C}^2(\mathbb{R}^n, \mathbb{R}^n))$, suppose that the system $\dot{x} = f(x)$ has a stable equilibrium in $x = 0$ and no solutions with finite escape time. Then, the following two conditions are equivalent:

1) For almost all initial states $x(0)$, the solution $x(t)$ tends to zero as $t \to \infty$.
2) There exists a nonnegative $\rho \in \mathrm{C}^1(\mathbb{R}^n \setminus \{0\}, \mathbb{R})$, which is integrable outside a neighborhood of zero and such that $[\nabla \cdot (f\rho)](x) > 0$ for almost all x.

V. Navigation Vector Fields

We define an NVF as a vector field that has navigation-like properties. These properties are captured in the following definition:

Definition 2: Let $\mathcal{F} \subset E^n$ be a compact connected analytic manifold with boundary. The smooth manifold map $f : \mathcal{F} \to T\mathcal{F}$ is an *NVF* if:

1) the origin of system $\dot{x} = f$ is almost GAS;
2) f is transverse across $\partial\mathcal{F}$.

The aforesaid definition is motivated by the properties of NFs. Obviously $-\nabla\varphi$ is an NVF since it satisfies both requirements. Clearly, the first requirement establishes the almost everywhere convergence of the system $\dot{x} = f$ while the second property (transversality) establishes that the vector field is perpendicular on $\partial\mathcal{F}$ guaranteeing that any trajectory will be safely brought to the origin without collisions. Our next step is to propose a construction of such a vector field that we will call a "canonical" NVF.

A. Construction

Let $\lambda_{\min,i}(x_{s,i})$ be the minimum eigenvalue of $H_\varphi(x_{s,i})$ at the saddle point $x_{s,i}$. The corresponding unit eigenvector is u_i. Let $d_i(x) = \|x - x_{s,i}\|^2$ be the squared metric distance of point x from the saddle point i for $i \in \{1 \ldots n_s\}$ where n_s is the number of saddle points. Let I denote the $n \times n$ identity matrix and n the workspace dimension. We can now define the matrix $U_i = u_i u_i^T + \varepsilon I$ for $i \in \{1 \ldots n_s\}$, where $0 < \varepsilon \leq 1$. Since the matrix $u_i u_i^T$ is positive semidefinite, the matrix U_i will be positive definite for any positive ε. Define $U_{n_s+1} = U_{n_s+2} = I$. A metric of the distance from the destination configuration can be encoded by using the NF, so we can define $d_{n_s+1} = \varphi$, and since the NF $\varphi(\partial\mathcal{F}) = 1$, we can encode a metric of the distance from the workspace boundary by denoting $d_{n_s+2} = 1 - \varphi$. Define $\bar{d}_j = \prod_{i=1 i\neq j}^{n_s+2} d_i$. Then D_φ is defined as

$$D_\varphi = \mu \sum_{i=1}^{n_s+2} \frac{\bar{d}_i}{\bar{d}_i + d_i} U_i \tag{1}$$

where μ is a positive constant. The function $\bar{d}_i / \bar{d}_i + d_i$ is an analytic switch that takes values between zero and 1. The properties of the matrix D_φ are provided in the following:

Lemma 1: The matrix $D_\varphi(x)$ defined in (1) has the following properties

1) a) $D_\varphi(x_{s,i}) = \mu U_{s,i}$, b) $D_\varphi(\partial\mathcal{F}) = \mu I$, c) $D_\varphi(0) = \mu I$;
2) a) $(\partial/\partial x) D_\varphi(x_{s,i}) = 0$, b) $(\partial/\partial x) D_\varphi(0) = 0$;
3) $D_\varphi > 0$;
4) $q^T D_\varphi q \leq 2(n_s + 2)\mu \|q\|^2, \forall q \in \mathbb{R}^n$.

Proof: Property 1: a) By direct computation, we have that, at the saddle point i, $d_i(x_{s,i}) = 0$, $\bar{d}_i \neq 0$, $\bar{d}_j = 0$ for $j \neq i$. Hence, $D_\varphi(x_{s,i}) = \mu U_{s,i}$;

b) At the workspace boundary, it holds that $\varphi(\partial\mathcal{F}) = 1$. Hence, $d_{n_s+2} = 0$, $\bar{d}_{n_s+2} \neq 0$ and $\bar{d}_j = 0$, $j \neq n_s + 2$ and $D_\varphi(\partial\mathcal{F}) = \mu I$;

c) At the origin $\varphi = 0$. Hence, $d_{n_s+1} = 0$, $\bar{d}_{n_s+1} \neq 0$ and $\bar{d}_j = 0$, $j \neq n_s + 1$; therefore, $D_\varphi(0) = \mu I$.

Property 2: a, b) For this property, observe that $d'_i(x_{s,i}) = 0$ for $i \in \{1 \ldots n_s\}$ where $f'(x) = \frac{\partial f(x)}{\partial x}$ and $d'_{n_s+1}(0) = 0$ since $\nabla\varphi(0) = 0$. Also, note that

$$\frac{\partial}{\partial x}\frac{\bar{d}_i}{\bar{d}_i + d_i} = \frac{\bar{d}'_i d_i - \bar{d}_i d'_i}{(\bar{d}_i + d_i)^2}.$$

Hence, at $x_{s,i}$ and at 0, it will hold that $d_i = d'_i = 0$ and $\bar{d}_j = \bar{d}'_j = 0$ for $j \neq i$ since they will either contain d_i or d'_i; therefore, $D_\varphi = 0$ at those locations.

Property 3: Since the matrix $u_i u_i^T$ is a matrix with one eigenvalue equal to unity and the rest eigenvalues zero, it follows that the matrix $U_i = u_i u_i^T + \varepsilon I$ is positive definite for $\varepsilon > 0$. Since the matrix D_φ is the sum of positive definite matrices multiplied by positive scalars, it will still be positive definite.

Property 4: First, observe that

$$0 \leq \frac{\bar{d}_i}{\bar{d}_i + d_i} \leq 1.$$

Multiplying D_φ left and right with the unit vectors $\hat{q}$, we get

$$\begin{aligned}\hat{q}^T D_\varphi \hat{q} &= \mu \sum_{i=1}^{n_s+2} \frac{\bar{d}_i}{\bar{d}_i + d_i}\hat{q}^T U_i \hat{q} \leq \mu \sum_{i=1}^{n_s+2} \hat{q}^T U_i \hat{q} \\ &\leq \mu \sum_{i=1}^{n_s+2} (1+\varepsilon) \leq 2(n_s+2)\mu.\end{aligned}$$

Multiplying both sides by $\|q\|^2$, we get the result

$$q^T D_\varphi q \leq 2(n_s+2)\mu \|q\|^2, \forall q \in \mathbb{R}^n.$$

The main feature of the matrix D_φ is that it allows for local modifications of the vector field in the vicinity of the saddle points. Without loss of generality, we assume in the following analysis that the destination configuration of the NF is the origin.

B. Main Result

The following is the main result of this paper

Proposition 1: Consider the system

$$\dot{x} = -D_\varphi \nabla\varphi \tag{2}$$

with $D_\varphi(x)$ constructed according to (1). Then, there exists an $a_0 > 0$ and an $\varepsilon_0 > 0$ such that the function $\rho = \varphi^{-a}$ is a density function for system (2) for any $a \geq a_0$ and $0 < \varepsilon \leq \varepsilon_0$.

Proof: Our analysis will be performed for the 2-D case, but the results can be readily extended to higher dimensions. The first observation is that the proposed density function is integrable outside a neighborhood of zero. This can be inferred by the fact that $\varphi(q)$ is analytic and bounded away from zero for $q \in \mathcal{F} - \mathcal{B}(0)$. By construction, ρ is positive definite. Setting $f = -D_\varphi \nabla\varphi$ from the divergence criterion, we get

$$\nabla \cdot (\rho f) = \nabla\rho f + \rho\nabla \cdot (f).$$

We have that

$$\nabla\rho = -\frac{a}{\varphi^{a+1}}\nabla\varphi.$$

Hence

$$\nabla \cdot (\rho f) = \frac{1}{\varphi^{a+1}}\left(a\nabla^T\varphi D_\varphi \nabla\varphi - \varphi\nabla \cdot (D_\varphi \nabla\varphi)\right). \tag{3}$$

Expanding the term $\nabla \cdot (D_\varphi \nabla\varphi)$ we get

$$\nabla \cdot (D_\varphi \nabla\varphi) = \nabla \cdot \left(\begin{bmatrix} d_{11}\varphi_x + d_{12}\varphi_y \\ d_{21}\varphi_x + d_{22}\varphi_y \end{bmatrix}\right)$$

where the notations f_x and f_{xx} denote the first and second derivatives of f with respect to x and d_{ij} is the ijth element of D_φ. For an NF, all critical points, except the origin, are saddle points [6]. At a saddle point $x_{s,i}$, we have that $\nabla\varphi(x_{s,i}) = 0$; therefore, the terms that contain first-order derivatives of φ are canceled. Also

$$\varphi_{xx} + \varphi_{yy} = \lambda_{\min} + \lambda_{\max}$$

since the trace of the Hessian is invariant. Thus, we have the following

$$\nabla \cdot \left(\begin{bmatrix} d_{11}\varphi_x + d_{12}\varphi_y \\ d_{21}\varphi_x + d_{22}\varphi_y \end{bmatrix}\right) = d_{11}\varphi_{xx} + d_{12}\varphi_{yx} + d_{21}\varphi_{xy} + d_{22}\varphi_{yy}.$$

Note that by Lemma 1, Property 1a), we have that, at the saddle point

$$D_\varphi(x_{s,i}) = \mu u_i u_i^T + \varepsilon\mu I.$$

Hence, $d_{11} = \mu(u_i^x)^2 + \varepsilon\mu$, $d_{21} = d_{12} = \mu(u_i^x)(u_i^y)$, and $d_{22} = \mu(u_i^y)^2 + \varepsilon\mu$ where u_i^x and u_i^y are x and y components of the vector u_i. Hence, $\nabla \cdot (D_\varphi \nabla\varphi)(x_{s,i})$ can be written as

$$\nabla \cdot (D_\varphi \nabla\varphi)(x_{s,i}) = \mu u_i^T H_\varphi(x_{s,i}) u_i + \varepsilon\mu(\lambda_{\min,i} + \lambda_{\max,i}).$$

Since by construction u_i is the eigenvector corresponding to the minimum eigenvalue of $H_\varphi(x_{s,i})$, the quadratic form $u_i^T H_\varphi(x_{s,i}) u_i = \lambda_{\min,i}$; hence, we get

$$\nabla \cdot (D_\varphi \nabla\varphi)(x_{s,i}) = \mu\lambda_{\min,i} + \varepsilon\mu(\lambda_{\min,i} + \lambda_{\max,i}).$$

Since $x_{s,i}$ is a saddle point, the minimum eigenvalue $\lambda_{\min,i}$ of the Hessian is necessarily negative (existence of unstable submanifold). By setting

$$\varepsilon < \left|\frac{\lambda_{\min,i}}{\lambda_{\min,i} + \lambda_{\max,i}}\right| \triangleq \varepsilon_{0,i}$$

and in combination with (3), exactly on the saddle points

$$\nabla \cdot (\rho f)(x_{s,i}) = -\mu\lambda_{min,i}\left(1 \pm \frac{\varepsilon}{\varepsilon_{0,i}}\right) > 0.$$

Close to the destination configuration, both $\nabla\varphi(0) = 0$ and $\varphi(0) = 0$; therefore, we need to analyze both terms of (3) to understand its behavior. Noting that (see [6], Proof of Proposition 3.2) $H_\varphi(0) = 2\beta^{-1/k}(0)I$ and from Lemma 1, property 1, we have that $D_\varphi(0) = \mu I$, and from property 2, that

$$\frac{\partial}{\partial x}D_\varphi(0) = 0$$

the Taylor expansions of φ and D_φ around the origin are as follows

$$\varphi(x) = \beta^{-1/k}(0)\|x\|^2 + O\left(\|x\|^3\right) D_\varphi(x) = \mu \mathrm{I} + O\left(\|x\|^2\right).$$

For the term $\nabla^T\varphi D_\varphi \nabla\varphi$, we obtain

$$\nabla^T\varphi D_\varphi \nabla\varphi = 4\mu\beta^{-2/k}\|x\|^2 + O(\|x\|^3)$$

and for the term $\varphi\nabla \cdot (D_\varphi \nabla\varphi)$, we obtain

$$\varphi\nabla \cdot (D_\varphi \nabla\varphi) = 4\mu\beta^{-2/k}\|x\|^2 + O(\|x\|^3).$$

From (3), we get that

$$\varphi^{a+1}\nabla \cdot (\rho f) = (a-1)\,4\mu\beta^{-2/k}\|x\|^2 + O(\|x\|^3).$$

So, choosing $a > 1$ will render $\nabla \cdot (\rho f) > 0$ in a neighborhood of zero.

We have until now established the positivity of (3) in the vicinity of critical points. To establish the global positivity of (3), since D_φ is positive definite (property 3 in Lemma 1), we require that

$$a > \frac{\max_{x \in \mathcal{F}} \varphi \nabla \cdot (D_\varphi \nabla \varphi)}{\min_{x \in \{\mathcal{F} - \mathcal{B}_\varepsilon(\mathcal{C})\}} \{\nabla^T \varphi D_\varphi \nabla \varphi\|} \triangleq a_1.$$

Since the workspace is bounded and the functions D_φ, φ are smooth, existence of a finite a_1 is guaranteed. Let $\varepsilon_1 = \min_{i \in \{1,\dots n_s\}} \varepsilon_{0,i}$. The positivity of the divergence criterion of Theorem 1 is satisfied by choosing $a_0 = \max\{1, a_1\}$ and $\varepsilon_0 = \min\{1, \varepsilon_1\}$, and the result is proven. ■

We can now state some properties of the proposed vector field.

Proposition 2: The vector field $f = -D_\varphi \nabla \varphi$ defined in Proposition 1 with $0 < \varepsilon < \varepsilon_0$, where ε_0 is defined in the proof of Proposition 1, is an *NVF*.

Proof: By Proposition 1, choosing an $a \geq a_0$, the function $\rho = \varphi^{-a}$ is a density function for (2).

Applying the dual criterion (Theorem 1) establishes the almost GAS property of $\dot{x} = f$.

For the transversality property we have that by property 1 of Lemma 1, it holds that $D_\varphi(\partial \mathcal{F}) = \mu I$. Hence, (2) becomes $\dot{x} = -\mu \nabla \varphi$ since $\mu > 0$ and by property 4 of Definition 1 we have that the vector field on the workspace boundary is transverse. ■

Some additional properties of the vector field $-D_\varphi \nabla \varphi$ are provided by the following.

Corollary 1: The NVF established in Proposition 2 assuming appropriate choice of parameters, vanishes only at the critical points of φ while its Jacobian is nondegenerate over the critical set of φ

Proof: Since by Property 3 of Lemma 1, $D_\varphi > 0$, the vector field vanishes only when $\nabla \varphi = 0$, which is true only at the set of critical points of φ. We have that

$$D_\varphi \nabla \varphi = \begin{bmatrix} d_{11}\varphi_x + d_{12}\varphi_y \\ d_{21}\varphi_x + d_{22}\varphi_y \end{bmatrix}.$$

Taking the Jacobian at a critical point, since $\varphi_x = \varphi_y = 0$, we have that

$$\begin{aligned} \frac{\partial}{\partial x}(D_\varphi \nabla \varphi) &= \begin{bmatrix} d_{11}\varphi_{xx} + d_{12}\varphi_{xy} & d_{11}\varphi_{xy} + d_{12}\varphi_{yy} \\ d_{21}\varphi_{xx} + d_{22}\varphi_{yx} & d_{21}\varphi_{xy} + d_{22}\varphi_{yy} \end{bmatrix} \\ &= D_\varphi H_\varphi. \end{aligned}$$

We know by the Morse property of φ that $\det(H_\varphi) \neq 0$ at every critical point. By using the relation $\det(AB) = \det(A)\det(B)$ we only need to prove that $\det(D_\varphi) \neq 0$ at the critical points. From Property 3 of Lemma 1, $D_\varphi > 0$, so the determinant is always positive and the Jacobian is nondegenerate at the critical points. ■

Due to the similarities of $-D_\varphi \nabla \varphi$ with $\nabla \varphi$ and the capability of $-D_\varphi \nabla \varphi$ to maintain the (Morse) index of the initial vector field while enforcing a positive definite Jacobian at the vicinity of the saddle points, we will call the vector field $-D_\varphi \nabla \varphi$ a "canonical" NVF and the system that this vector field is applied to a "canonical" navigation system.

A comparison of the convergence properties of canonical navigation systems with NF-based systems is provided by the following result that will allow us to reason about the NF-based system by examining the canonical system.

Proposition 3: Consider the system

$$\dot{x} = -K \nabla \varphi \tag{4}$$

where K a positive gain. Then, there exists a $0 < \mu \leq \mu_0$ such that for almost all the same initial conditions $x_{(4)}(0) = x_{(2)}(0)$, the trajectories of (4) are bounded by the trajectories of (2) as follows: $\varphi(x_{(4)}(t)) \leq \varphi(x_{(2)}(t)),\ \forall t \geq 0$. Moreover, there exists a spherical neighborhood $\mathcal{B}(0)$ around the origin for which, for all $x_{(4)}(0) = x_{(2)}(0) \in \mathcal{B}(0)$, it holds that

$$\left\| x_{(4)}(t) \right\| \leq \left\| x_{(2)}(t) \right\| \ \forall t \geq 0.$$

Proof: Taking the time derivative of φ across the trajectories of system (4), we get:

$$\dot{\varphi}_{(4)} = -K \left\| \nabla \varphi_{(4)} \right\|^2. \tag{5}$$

The time derivative of φ across the trajectories of (2) is $\dot{\varphi}_{(2)} = -\nabla \varphi^T D_\varphi \nabla \varphi \geq -2(n_s + 2)\mu \left\| \nabla \varphi \right\|^2$ by use of Property 4 of Lemma 1. Setting $\mu = \mu_1 \frac{K}{2(n_s+2)}$ with $0 < \mu_1 < 1$, we get

$$\dot{\varphi}_{(2)} > -\mu_1 K \left\| \nabla \varphi_{(2)} \right\|^2. \tag{6}$$

To prove the first part of the Proposition, we need to establish that $\dot{\varphi}_{(4)}(x_{(4)}(t)) \leq \dot{\varphi}_{(2)}(x_{(2)}(t))$ for all $t \geq 0$ given that $x_{(4)}(0) = x_{(2)}(0)$. By (5) and (6), we have for $t = 0$ that

$$\dot{\varphi}_{(2)}(x_{(2)}(0)) > -\mu_1 \dot{\varphi}_{(4)}(x_{(4)}(0)). \tag{7}$$

By smoothness arguments, there exists a neighborhood of $\mathcal{B}_\varepsilon(x_{(2)}(0))$ around $x_{(2)}(0)$ such that the inequality (7) still holds as long as the initial conditions are not exactly on the saddle point. So, in this neighborhood, we have that $\varphi(x_{(4)}(t)) \leq \varphi(x_{(2)}(t)),\ t \in [0, \delta(\varepsilon)]$ for some increasing function $\delta(\cdot)$. By the selection of μ_1, we have that

$$\left\| D_\varphi \nabla \varphi \right\| \leq K \left\| \nabla \varphi \right\|.$$

Hence, we can assert that $x_{(4)}$ will exit $\mathcal{B}_\varepsilon(x_{(2)}(0))$ first. Let

$$g_{\max}(a) = \max_{x \in \varphi^{-1}(a)} \left\| \nabla \varphi(x) \right\|$$

and

$$g_{\min}(a) = \min_{x \in \left\{ \varphi^{-1}(a) - \mathcal{B}_\varepsilon(\mathcal{S}) \right\}} \left\| \nabla \varphi(x) \right\|$$

where $\mathcal{S}$ is the set of saddle points. Since the reachable set of initial conditions, excluding the set $\mathcal{B}_\varepsilon(x_{(2)}(0))$ is bounded away from saddle points, $g_{\min}$ is nonzero. Since the workspace is bounded, and φ is smooth, the maximum value of $\nabla \varphi$ is finite. Hence, function $r(a) = (g_{\min}(a)/g_{\max}(a))$ is well defined everywhere, except at $a = 0$, where the limit exists and is

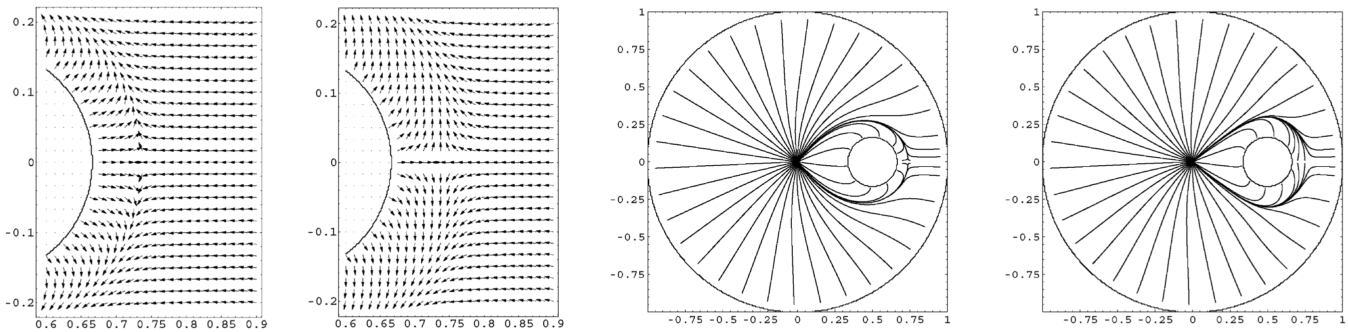
$$\lim_{x \to 0} r(x) = \frac{\lambda_{\min}(0)}{\lambda_{\max}(0)}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalues of the Hessian of φ. This can be verified by considering that the origin is a nondegenerate critical point, hence a quadratic one for appropriate coordinates near the origin $\varphi(x) = \lambda_{\min} x_1^2 + \lambda_{\max} x_2^2$. By setting

$$\mu_1 \leq \min_{a \in (0,1]} r(a) \triangleq \mu_2$$

we have that whenever $\varphi(x_{(4)}(t)) = \varphi(x_{(2)}(t))$, system (4) will have a higher velocity than system (2). Hence, $\dot{\varphi}(x_{(4)}(t)) < \dot{\varphi}(x_{(2)}(t))$. This means that as long as at some t it is true that $\varphi(x_{(4)}(t)) < \varphi(x_{(2)}(t))$, then it be true for all $t' \geq t$. However, since $x_{(4)}$ will exit first $\mathcal{B}_\varepsilon(x_{(4)}(0))$, we have that

$$\varphi\left(x_{(4)}(t)\right) \leq \varphi\left(x_{(2)}(t)\right) \ \forall t \geq 0.$$



(a) (b) (c) (d)

Fig. 1. (a) Vector field (normalized) near the saddle point for an original Rimon–Koditschek. (b) A canonical vector field. (c) Trajectories corresponding to the same initial conditions for a system under the influence of an original Rimon–Koditschek vector field [system (8)]. (d) A canonical vector field (system (9)).

Now let $\rho_{\max}$ be the maximum radius of a disk centered at the origin that has no intersections with obstacles. Then, this circle contains no saddle points, since saddles occur between workspace boundary and obstacles. Alternatively, the radius $\rho_{\max}$ can be fixed so that the circle is bounded away from saddle points and obstacles. Moreover, we constrain $\rho_{\max}$ even more such that the Hessian of φ in the disk defined by $\rho_{\max}$ is everywhere positive definite and its minimum eigenvalue is greater than $\lambda_0 > 0$. Now since the Hessian is positive definite, the level sets of φ inside the circle are convex. Moreover, the nonzero minimum eigenvalue establishes that intersections of level sets of φ with circles centered at the origin will be performed at obtuse angles. Hence, the unit vector of the gradient $-\widehat{\nabla\varphi}$ will have a positive projection on the inside pointing unit vector that is perpendicular to the circle's circumference. Denote the value of this projection by $p(x)$. For $\rho \leq \rho_{\max}$, define

$$g'_{\max}(\rho) = \max_{\|x\|=\rho} \|\nabla\varphi(x)\| \quad \text{and } g'_{\min}(\rho) = \min_{\|x\|=\rho} \|p(x)\nabla\varphi(x)\|.$$

Obviously $g'_{\max}(\rho)$ and $g'_{\min}(\rho)$ are nonzero except at the origin and are bounded due to smoothness and compactness arguments. Hence, the function

$$r'(\rho) = \frac{g'_{\min}(\rho)}{g'_{\max}(\rho)}$$

is well defined, finite, and nonzero everywhere, except at $a = 0$, where the limit exists and is

$$\lim_{x\to 0} r'(x) = \frac{\lambda_{\min}(0)}{\lambda_{\max}(0)}.$$

By setting

$$\mu_1 \leq \min_{\rho\in(0,\rho_{\max}]} r'(\rho) \triangleq \mu_3$$

we have that whenever $\|x_{(4)}(t)\| = \|x_{(2)}(t)\|$, system (4) will have a velocity whose projection on the perpendicular of the circle's circumference will be higher than the velocity of system (2). This means that as long as at some t it is true that $\|x_{(4)}(t))\| \leq \|(x_{(2)}(t))\|$, then it will be true for all $t' \geq t$. However, since the initial conditions are the same, we have that

$$\|x_{(4)}(t)\| \leq \|x_{(2)}(t)\| \;\forall t \geq 0.$$

Choosing

$$\mu_0 = \frac{K}{2(n_s+2)} \min\{1, \mu_2, \mu_3\}$$

completes the proof. ■

VI. Example

To demonstrate the navigation properties of the canonical navigation field, we present in this section a simple example for navigation in a spherical world with one obstacle.

The workspace is centered at the origin which is the destination configuration and its radius $r_w = 1$, while the obstacle is centered at $x_o = (\frac{1}{2}, 0)$ and its radius is $r_o = \frac{1}{6}$ units [see Fig. 1(c)]. A Rimon–Koditschek NF φ is constructed on this workspace with tuning parameter $k = 5$. The NF is given by

$$\varphi(x) = \frac{\|x\|^2}{\left(\|x\|^{2k} + \left(r_w^2 - \|x\|^2\right)\left(\|x - x_o^T\|^2 - r_o^2\right)\right)^{1/k}}.$$

The unique saddle point is located at $x_s \approx (0.74, 0.0)$. The Hessian at x_s is evaluated as

$$H_\varphi(x_s) \approx \begin{bmatrix} 8.51 & 0.0 \\ 0.0 & -1.98 \end{bmatrix}.$$

Its eigenvalues are $\lambda_{\min,1}(x_s) \approx -1.98$ with corresponding eigenvector $u_1 = (0, 1)$ and $\lambda_2(x_s) \approx 8.51$ with corresponding eigenvector $v_2 = (1, 0)$. Following the construction process in Section V-A, we create the matrix D_φ from (1) with $\varepsilon = 0.1$ and $\mu = 1$.

The vector field corresponding to the Rimon–Koditschek NF for the system

$$\dot{x} = -\nabla\varphi(x) \tag{8}$$

in the saddle point neighborhood, is depicted in Fig. 1(a). The vector field corresponding to the canonical navigation field for the system

$$\dot{x} = -D_\varphi(x)\nabla\varphi(x) \tag{9}$$

is depicted in Fig. 1(b). As can be seen in the modified vector field in Fig. 1(b), the vectors belonging to the subspace corresponding to the negative eigenvalue at the saddle point have an increased magnitude. This is due to the effect of the D_φ operator that tends to exaggerate the vectors lying in that subspace, forcing the divergence of $-D_\varphi\nabla\varphi$ to attain positive values in the neighborhood of the saddle point, as expected by our analysis in the previous section.

Fig. 1(c) depicts the trajectories of the system (8) based on the Rimon–Koditschek NVF, while in Fig. 1(d), we can see the trajectories of the system (9) which is under the influence of the canonical NVF. The initial conditions are the same for both trajectory sets. Observe that the canonical NVF enjoys the same navigation properties as the original (i.e., safety and convergence), while the net effect of the D_φ operator is to locally enhance (at the saddle point neighborhood) the vector field in the submanifold defined by the eigendirection of the negative eigenvalue at the saddle point, making the system depart faster from the stable submanifold.

VII. Conclusion

We have derived a density function for an NF-based system. The function is derived for a transformed, smooth vector field that enjoys the navigation properties of the original NF vector field. Under some assumptions, the convergence results derived on the transformed vector field are propagated to the original. This result will enable exploitation of several features of dual Lyapunov techniques to robotic navigation. Initial results from applying this approach to robotic navigation are reported in [13]. Further research includes finding density functions that are directly applicable to the primary navigation system.

References

[1] A. Rantzer, "A dual to Lyapunov's stability theorem," *Syst. Control Lett.*, vol. 42, no. 3, pp. 161–168, 2001.
[2] A. Rantzer, "A converse theorem for density functions," in *Proc. IEEE Conf. Decis. Control*, 2002, pp. 1890–1891.
[3] A. Rantzer and F. Ceragioli, "Smooth blending of nonlinear controllers using density functions," presented at the Eur., Porto, Portugal, 2001.
[4] P. Monzon, "On necessary conditions for almost global stability," *IEEE Trans. Autom. Control*, vol. 48, no. 4, pp. 631–634, Apr. 2003.
[5] D. Angeli, "Some remarks on density functions for dual Lyapunov methods," in *Proc. 42nd IEEE Conf. Decis. Control*, 2003, pp. 5080–5082.
[6] D. E. Koditschek and E. Rimon, "Robot navigation functions on manifolds with boundary," *Adv. Appl. Math.*, vol. 11, pp. 412–442, 1990.
[7] J. Milnor, *Morse Theory* (Annals of Mathematics Studies Series). Princeton, NJ: Princeton University Press, 1963.
[8] E. Rimon and D. E. Koditschek, "Exact robot navigation using artificial potential functions," *IEEE Trans. Robot. Autom.*, vol. 8, no. 5, pp. 501–518, Oct. 1992.
[9] E. Rimon and D. E. Koditschek, "The construction of analytic diffeomorphisms for exact robot navigation on star worlds," *Trans. Amer. Math. Soc.*, vol. 327, no. 1, pp. 71–115, 1991.
[10] H. G. Tanner, S. G. Loizou, and K. J. Kyriakopoulos, "Nonholonomic navigation and control of cooperating mobile manipulators," *IEEE Trans. Robot. Autom.*, vol. 19, no. 1, pp. 53–64, Feb. 2003.
[11] S. Loizou and K. Kyriakopoulos, "A feedback-based multiagent navigation framework," *Int. J. Syst. Sci.*, vol. 37, no. 6, pp. 377–384, 2006.
[12] C. Karagöz, H. Bozma, and D. Koditschek, "Coordinated navigation of multiple independent disk-shaped robots," EECS Dept., Univ. Michigan, Ann Arbor, MI, Tech. Rep. CSE-TR-486-04, 2004.
[13] S. Loizou and V. Kumar, "Weak input-to-state stability properties for navigation function based controllers," in *Proc. IEEE Int. Conf. Decis. Control*, 2006, pp. 1800–1805.

Output Feedback Control of Bilinear Systems via a Bilinear LTR Observer

Min-Shin Chen
and Chi-Che Chen

***Abstract*—In the literature, most observer-based output feedback controls for bilinear systems are only applicable to open-loop (neutrally) stable systems. This paper proposes a new observer-based output feedback control that can be applied to open-loop unstable systems. The key component of the new control is an exponentially stable *bilinear* loop transfer recovery (LTR) observer that derives from the linear LTR observer.**

***Index Terms*—Bilinear observer, bilinear system, dyadic bilinear system, loop transfer recovery (LTR) observer, output feedback control.**

I. Introduction

Bilinear systems exist in many physical phenomena that are of considerable interest to human activities [1], [2]. Recent applications of bilinear system control include heating, air conditioning control [3], power converter control [4], electromagnetic actuator control [5], and quantum system control using finite-dimensional bilinear models [6] or infinite-dimensional bilinear models [7], [8]. Even though a variety of control designs have been developed for bilinear systems, most of them are based on state feedback [9]–[15]. If only part of the state variables are accessible for measurement, one has to resort to output feedback control. Unfortunately, most output feedback controls in the literature require that the open-loop bilinear system be stable [16], neutrally stable [17] or dissipative [18]. The reason for requiring this open-loop stable condition is that they all assume the stabilizing control signal be of small magnitude so that their bilinear observer designs can be successful. There are a few bilinear observer designs proposed in the literature for the open-loop unstable bilinear system without imposing the small control condition. For example, an open-loop dead-beat observer for state estimation of open-loop unstable bilinear systems is suggested in [19], but the system must satisfy the existence condition of a control Lyapunov function [20]. In [21] and [22], bilinear observers can be constructed with the state estimation error converging independent of the control input under a set of system matrix equalities.

This paper proposes a new output feedback control for unstable bilinear systems. The key element is a bilinear loop transfer recovery (LTR) observer that derives from the linear LTR observer [23]. The new bilinear LTR observer is exponentially stable without imposing the small control condition, the existence of a control Lyapunov function, or extra matrix equalities on the system matrices. Hence, it relaxes the stringent conditions imposed by previous bilinear observer designs. Then, by combining this new bilinear LTR observer with the state feedback division control in [15], one obtains a stabilizing output feedback control for bilinear systems that may be open-loop unstable.

The remainder of this paper is arranged as follows. Section II introduces the new bilinear LTR observer. Section III presents the observer-based output feedback control, and its stability analysis. Section IV concludes the paper.

Manuscript received March 7, 2006; revised November 27, 2006, May 9, 2007, and October 12, 2007. Recommended by Associate Editor D. Angeli.
The authors are with the Department of Mechanical Engineering, National Taiwan University, Taipei 106, Taiwan, R.O.C. (e-mail: mschen@ntu.edu.tw; d93522014@ntu.edu.tw).
Digital Object Identifier 10.1109/TAC.2008.917652