Exponentially Stable Nonlinear Systems Have Polynomial Lyapunov Functions on Bounded Regions

Matthew M. Peet, *Member, IEEE*

***Abstract*—This paper presents a proof that existence of a polynomial Lyapunov function is necessary and sufficient for exponential stability of a sufficiently smooth nonlinear vector field on a bounded set. The main result states that if there exists an n-times continuously differentiable Lyapunov function which proves exponential stability on a bounded subset of $\mathbb{R}^n$, then there exists a polynomial Lyapunov function which proves exponential stability on the same region. Such a continuous Lyapunov function will exist if, for example, the vector field is at least n-times continuously differentiable. The proof is based on a generalization of the Weierstrass approximation theorem to differentiable functions in several variables. Specifically, polynomials can be used to approximate a differentiable function, using the Sobolev norm $W^{1,\infty}$ to any desired accuracy. This approximation result is combined with the second-order Taylor series expansion to show that polynomial Lyapunov functions can approximate continuous Lyapunov functions arbitrarily well on bounded sets. The investigation is motivated by the use of polynomial optimization algorithms to construct polynomial Lyapunov functions.**

***Index Terms*—Exponential stability, Lyapunov methods, polynomial approximation, polynomials, sum of squares.**

I. Introduction

LYAPUNOV theory strongly influences our understanding of the motion of dynamic systems. This is hardly surprising, given that Lyapunov functions are little more than a generalization of the concept of energy. Construction of a Lyapunov function, then, is simply identification of the essential energy of a system. Control of the motion of a system may be achieved by direct manipulation of this energy. These ideas were first introduced by the works of AM Lyapunov. Lyapunov's thesis appeared originally in Russian in 1892, in French in 1908, and finally, in English, in 1992. In the years since the original publication, a number of extensions of Lyapunov's direct method have been proposed, including the use of multiple Lyapunov functions, application to infinite-dimensional systems (delayed or partial-differential equations), and quadratic Lyapunov functions for linear systems.

There are many results on the properties of Lyapunov functions given certain aspects of the vector field. A number of early results related smoothness and differentiability of the Lyapunov function to that of the vector field. Representative examples include Barbasin [2], Malkin [19] and Kurzweil [15]. Additionally, a summary of many of these results can be found in Massera [21]. The use of smoothing operators on open sets to create infinitely-differentiable functions was explored in the work Wilson [37], and more recently in Lin, Sontag, and Wang [18]. Other innovative results on converse Lyapunov functions can be found in Lakshmikantam and Martynyuk [16] and Teel and Praly [33]. The reader is also referred to Hahn [9] and Krasovskii [13] for a treatment of converse theorems of Lyapunov. As will be seen in Section V, Massera-type converse Lyapunov results are a necessary basis for the results of this paper in that they ensure the existence of a Lyapunov functions with the smoothness properties necessary to apply the approximation theory result proven in this paper.

Recall that the topic of this paper is the existence of polynomial Lyapunov functions on compact sets. At this point it may be tempting to conclude that since previous work has proven the existence of smooth Lyapunov functions, one can simply apply the Weierstrass Approximation Theorem to get a polynomial approximation of the Lyapunov function on the compact set. Unfortunately, there are two major technical difficulties. The first is that the resulting polynomial may not be a Lyapunov function (i.e., it may not be positive semidefinite, etc.) for any level of accuracy. The second is that the derivative of this polynomial is not guaranteed to approximate the derivatives of the original function. This is because Weierstrass has not been extended to the Sobolev norm in n-dimensions. The major technical contribution of this paper is the resolution of these problems through certain extensions of the theory of approximation of functions.

The classical result in approximation theory is the Weierstrass approximation theorem, which was proven in 1885 [36]. This result demonstrated that real-valued polynomial functions can approximate real-valued continuous functions arbitrarily well with respect to the supremum norm on a compact interval. Various structural generalizations of the Weierstrass approximation theorem have focused on generalized mappings, as in Stone [31], and on alternate topologies, as in Krein [14]. Polynomial approximation of differentiable functions has been studied in numerical analysis of differential and partial differential equations. Results of relevance include the Bramble-Hilbert Lemma [3], its generalization in Dupont and Scott [6], and Jackson's theorem [12], generalizations of which can be found in the work of Timan [34].

The approximation of of functions in Sobolev spaces by polynomials has been studied in a number of contexts. In Everitt and Littlejohn [8], the density of polynomials in certain Sobolev spaces was discussed for the single-variable case. Other work

Manuscript received July 02, 2007; revised February 28, 2008. Current version published May 13, 2009. Recommended by Guest Editors G. Chesi and D. Henrion.

The author is with the Department of Mechanical, Materials, and Aerospace Engineering, Illinois Institute of Technology, Chicago, IL 60616 USA (e-mail: mpeet@iit.edu).

Digital Object Identifier 10.1109/TAC.2009.2017116

considers weighted Sobolev spaces, as in Portilla, Quintana, Rodriguez, and Touris [26].

The density of infinitely continuously differentiable test functions in Sobolev spaces has been studied in the context of partial differential equations. Important results include the Meyers-Serrin Theorem [22], extensions of which can be found in Adams [1] or Evans [7].

The existence of quadratic Lyapunov functions for linear systems has been particularly important in recent times. This is because all quadratic Lyapunov functions on a finite-dimensional state-space can be parameterized by the positive matrices. As algorithms for optimization over the positive matrices have become fast and reliable, the analysis and control of many aspects of the behavior of linear finite-dimensional systems has become automatic.

In an effort to replicate the success of linear systems analysis for nonlinear dynamics, significant progress has been made toward the development of efficient algorithms for optimization over the cone of positive polynomials. See Lasserre [17], Nesterov [23], and Parrilo [25] for some early work in this area or Henrion and Garulli [10] and Chesi [4] for more recent advances. Today, there exist a number of software packages for optimization over the positive polynomials, with prominent examples being given by SOSTOOLS [27] and GloptiPoly [11]. As was the case for linear systems and quadratic Lyapunov functions, the development of algorithms for polynomial optimization allows for the efficient analysis of nonlinear systems though the use of polynomial Lyapunov functions.

Consider a nonlinear ordinary differential equation of the form

$$\dot{x}(t) = f(x(t))$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$. A number of authors have discussed the use of polynomial optimization algorithms for constructing polynomial Lyapunov functions to prove stability of this type of system. Substantial contributions to this area include Parrilo [25], Wang [35], Tan [32], and Chesi, *et al.* [5].

The goal of this paper is to resolve the question of whether polynomial Lyapunov functions are necessary and sufficient for stability of nonlinear systems. The main result of this paper is summarized in Theorem 11, where it is shown that if f is n-times continuously differentiable, then exponentially stability of f on a compact region is equivalent to the existence of a polynomial Lyapunov function which proves exponential stability of level sets contained in that region. The smoothness condition is automatically satisfied if f is polynomial.

To establish this converse Lyapunov result, two extensions to the Weierstrass approximation theorem are presented. In Theorem 5, it is shown that for bounded regions, polynomials can be used to approximate continuously differentiable multivariate functions arbitrarily well in a variety of norms, including the Sobolev norm $W^{1,\infty}$. This means that for any continuously differentiable function and any $\gamma > 0$, one can find a polynomial which approximates the function with error γ and whose partial derivatives approximate those of the function with error γ. The proof is based on a construction using approximations to the partial derivatives.

The second extension combines a second order Taylor series expansion with the Weierstrass approximation theorem to find polynomials which approximate functions with a pointwise weight on the error given by

$$w(x) = \frac{1}{x^T x}.$$

These two extensions are combined into the main polynomial approximation result, which is Theorem 8. The application to Lyapunov functions is given in Proposition 9. Proposition 9 states that if there exists a sufficiently smooth continuous Lyapunov function which proves exponential stability on a bounded set, then there exists a polynomial Lyapunov function which proves exponential stability on the same set. In Section V, a converse Lyapunov theorem given along with a brief discussion of the implications for the use of polynomial optimization to prove stability of nonlinear ordinary differential equations.

Notation and Background

Let $\mathbb{N}^n$ denote the set of length n vectors of non-negative natural numbers. Denote the unit cube in $\mathbb{N}^n$ by $Z^n := \{\alpha \in \mathbb{N}^n : \alpha_i \in \{0,1\}\}$. For $x \in \mathbb{R}^n$, $\|x\|_\infty = \max_i |x_i|$ and $\|x\|_2 = \sqrt{x^T x}$. Define the unit cube in $\mathbb{R}^n$ by $B := \{x \in \mathbb{R}^n : \|x\|_\infty \le 1\}$. Let $\mathcal{C}(\Omega)$ be the Banach space of scalar continuous functions defined on $\Omega \subset \mathbb{R}^n$ with norm

$$\|f\|_\infty := \sup_{x\in\Omega} \|f(x)\|_\infty.$$

For operators $h_i : X \to X$, let $\prod_i h_i : X \to X$ denote the sequential composition of the h_i. i.e.,

$$\prod_i h_i := h_1 \circ h_2 \circ \cdots \circ h_{n-1} \circ h_n.$$

For a sufficiently regular function $f : \mathbb{R}^n \to \mathbb{R}$ and $\alpha \in \mathbb{N}^n$, we will often use the following shorthand to denote the partial derivative:

$$D^\alpha f(x) := \frac{\partial^\alpha}{\partial x^\alpha} f(x) = \prod_{i=1}^n \frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}} f(x)$$

where naturally, $\partial^0 f/\partial x_i^0 = f$. The gradient operator will be denoted as ∇ and is defined as

$$\nabla := \begin{bmatrix} D^{(1,0,\ldots,0)} \\ \vdots \\ D^{(0,\ldots,0,1)} \end{bmatrix}.$$

For $\Omega \subset \mathbb{R}^n$, we define the following sets of differentiable functions:

$$\mathcal{C}_1^i(\Omega) := \Big\{ f : D^\alpha f \in \mathcal{C}(\Omega) \quad \text{for any } \alpha \in \mathbb{N}^n \text{ such that } \|\alpha\|_1 = \sum_{j=1}^n \alpha_j \le i. \Big\}$$

$$\mathcal{C}_\infty^i(\Omega) := \Big\{ f : D^\alpha f \in \mathcal{C}(\Omega) \text{ for any } \alpha \in \mathbb{N}^n \text{ such that } \|\alpha\|_\infty = \max_j \alpha_j \le i. \Big\}$$

$\mathcal{C}^\infty(\Omega) = \mathcal{C}_1^\infty(\Omega) = \mathcal{C}_\infty^\infty(\Omega)$ is the logical extension to infinitely continuously differentiable functions. Note that in n-dimensions, $\mathcal{C}_1^i(\Omega) \subset \mathcal{C}_\infty^i(\Omega) \subset \mathcal{C}_1^{i \cdot n}(\Omega)$. We will occasionally refer to the Banach spaces $W^{k,p}(\Omega)$, which denote the standard Sobolev spaces of locally summable functions $u : \Omega \to \mathbb{R}$ with *weak* derivatives $D^\alpha u \in L_p(\Omega)$ for $|\alpha|_1 \le k$ and norm

$$\|u\|_{W^{k,p}} := \sum_{|\alpha|_1 \le k} \|D^\alpha u\|_{L_p}.$$

Here, $L_p(\Omega)$ refers to the standard space of Lebesgue-measurable functions with norm, for $p < \infty$

$$\|f\|_p := \left(\int_\Omega |f(s)|^p ds\right)^{(1/p)}$$

and $\|f\|_\infty := \sup_{s\in\Omega} |f(s)|$. The following version of the Weierstrass approximation theorem in multiple variables comes from Timan [34].

Theorem 1: Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and $G \subset \mathbb{R}^n$ is compact. Then there exists a sequence of polynomials which converges to f uniformly in G.

II. Approximation of Differentiable Functions

Most of the technical results of this paper concern a constructive method of approximating a differentiable function of several variables using approximations to the partial derivatives of that function. Specifically, the result states that if one can find polynomials which approximate the partial derivatives of a given function to accuracy $\gamma/2^n$, then one can construct a polynomial which approximates the given function to accuracy γ, and all of whose partial derivatives approximate those of the given function to accuracy γ.

The difficulty of this problem is a consequence of the fact that, for a given function, not all the partial derivatives of that function are independent. For example, we have the identities

$$\begin{aligned}&D^{(1,1,0)} f(x,y,z) = \\ &D^{(1,1,0)}\left(f(x,y,0) + \int_0^z D^{(0,0,1)} f(x,y,s) ds\right) \\ &= D^{(1,0,0)}\left(D^{(0,1,0)} f(x,y,z)\right)\end{aligned}$$

among others. Therefore, given approximations to the partial derivatives of a function, these approximations will, in general, not be the partial derivatives of any function. The problem becomes, for each partial derivative approximation, how to extract the information which is unique to that partial derivative in order to form an approximation to the original function. The following construction shows how this can be done.

Definition 2: An element of X is defined to be a set of 2^n continuous functions, indexed as $f_\alpha \in \mathcal{C}(\Omega)$ for $\alpha \in Z^n$, and denoted $\{f_\alpha\}_{\alpha\in Z^n} \in X$. Define the linear map $K : X \to \mathcal{C}_\infty^1(\Omega)$ as

$$K\left(\{f_\alpha\}_{\alpha\in Z^n}\right) = \sum_{\alpha\in Z^n} \left(\prod_{i=1}^n g_{i,\alpha_i}\right) f_\alpha = \sum_{\alpha\in Z^n} G_\alpha f_\alpha$$

where $G_\alpha : \mathcal{C}(\Omega) \to \mathcal{C}_\infty^1(\Omega)$ is given by

$$G_\alpha h = \left(\prod_{i=1}^n g_{i,\alpha_i}\right) h$$

and where

$$\begin{aligned}&(g_{i,j}h)(x_1,\ldots,x_n) \\ &\quad = \begin{cases} h(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n) & j=0 \\ \int_0^{x_i} h(x_1,\ldots,x_{i-1},s,x_{i+1},\ldots,x_n) ds & j=1. \end{cases}\end{aligned}$$

It is useful to think of the functions f_α as representing either the partial derivatives of a function or approximations to those partial derivatives. The following examples illustrate the construction.

1) Example: If $p = K(\{q_\alpha\}_{\alpha\in Z^2})$, then

$$\begin{aligned}p(x_1,x_2) = &\int_0^{x_1}\int_0^{x_2} q_{(1,1)}(s_1,s_2) ds_2 ds_1 \\ &+ \int_0^{x_1} q_{(1,0)}(s_1,0) ds_1 \\ &+ \int_0^{x_2} q_{(0,1)}(0,s_2) ds_2 \\ &+ q_{(0,0)}(0,0).\end{aligned}$$

If $n = 3$, then

$$\begin{aligned}p(x_1,x_2,x_3) = &\int_0^{x_1}\int_0^{x_2}\int_0^{x_3} q_{(1,1,1)}(s_1,s_2,s_3) ds_3 ds_2 ds_1 \\ &+ \int_0^{x_1}\int_0^{x_2} q_{(1,1,0)}(s_1,s_2,0) ds_2 ds_1 \\ &+ \int_0^{x_1}\int_0^{x_3} q_{(1,0,1)}(s_1,0,s_3) ds_3 ds_1 \\ &+ \int_0^{x_2}\int_0^{x_3} q_{(0,1,1)}(0,s_2,s_3) ds_3 ds_2 \\ &+ \int_0^{x_1} q_{(1,0,0)}(s_1,0,0) ds_1 \\ &+ \int_0^{x_2} q_{(0,1,0)}(0,s_2,0) ds_2 \\ &+ \int_0^{x_3} q_{(0,0,1)}(0,0,s_3) ds_3 \\ &+ q_{(0,0,0)}(0,0,0).\end{aligned}$$

Notice that this structure gives a natural way of approximating the function p and all of its partial derivatives. For example, if p is defined as above for $n = 2$, then

$$\frac{\partial}{\partial x_1} p(x_1,x_2) = \int_0^{x_2} q_{(1,1)}(x_1,s_2) ds_2 + q_{(1,0)}(x_1,0).$$

Now, if one approximates q_α by $\tilde{q}_\alpha$ and uses the construction to create $\tilde{p}$, then

$$\frac{\partial}{\partial x_1} \tilde{p}(x_1,x_2) = \int_0^{x_2} \tilde{q}_{(1,1)}(x_1,s_2) ds_2 + \tilde{q}_{(1,0)}(x_1,0).$$

Since the integral is a bounded operator, and since $q_\alpha \approx \tilde{q}_\alpha$, then $(\partial/\partial x_1)\tilde{p} \approx (\partial/\partial x_1)p$. A formal argument is as follows.

The first lemma shows that when the f_α are the partial derivatives of a function, we recover the original function. The proof is simple and works by using a single integral identity to repeatedly expand the function.

Lemma 3: For $v \in \mathcal{C}^1_\infty(\Omega)$, $K(\{D^\alpha v\}_{\alpha\in Z^n}) = v$.

Proof: Let e_i denote the ith unit vector in $\mathbb{N}^n$. Observe the following identity which holds for any differentiable function h:

$$h = g_{i,0}h + g_{i,1}D^{e_i}h.$$

By assumption, $v \in \mathcal{C}^1_\infty(\Omega)$ and so $D^\alpha v \in \mathcal{C}(\Omega)$ for all $\alpha \in Z^n$. We proceed by induction. Suppose the following identity holds for some $d < n$:

$$v = \sum_{\alpha\in Z^d}\left(\prod_{i=1}^d g_{i,\alpha_i}\right)\left(\prod_{i=1}^d D^{\alpha_i e_i}\right)v.$$

Then

$$\begin{aligned} v &= \sum_{\alpha\in Z^d}\left(\prod_{i=1}^d g_{i,\alpha_i}\right)g_{d+1,0}\left(\prod_{i=1}^d D^{\alpha_i e_i}\right)v \\ &\quad + \sum_{\alpha\in Z^d}\left(\prod_{i=1}^d g_{i,\alpha_i}\right)g_{d+1,1}D^{e_{d+1}}\left(\prod_{i=1}^d D^{\alpha_i e_i}\right)v \\ &= \sum_{\substack{\alpha\in Z^{d+1}\\ \alpha_{d+1}=0}}\left(\prod_{i=1}^{d+1} g_{i,\alpha_i}\right)\left(\prod_{i=1}^{d+1} D^{\alpha_i e_i}\right)v \\ &\quad + \sum_{\substack{\alpha\in Z^{d+1}\\ \alpha_{d+1}=1}}\left(\prod_{i=1}^{d+1} g_{i,\alpha_i}\right)\left(\prod_{i=1}^{d+1} D^{\alpha_i e_i}\right)v \\ &= \sum_{\alpha\in Z^{d+1}}\left(\prod_{i=1}^{d+1} g_{i,\alpha_i}\right)\left(\prod_{i=1}^{d+1} D^{\alpha_i e_i}\right)v. \end{aligned}$$

Thus, if the identity holds for d, the identity holds for $d+1$. The identity holds for $d = 1$ since v is differentiable. Therefore

$$v = \sum_{\alpha\in Z^n}\left(\prod_{i=1}^n g_{i,\alpha_i}\right)\left(\prod_{i=1}^n D^{\alpha_i e_i}\right)v = \sum_{\alpha\in Z^n}\left(\prod_{i=1}^n g_{i,\alpha_i}\right)D^\alpha v$$

as desired. ■

The following lemma states that the linear map K is Lipschitz continuous in an appropriate sense. This means that a small error in the partial derivatives, f_α, results in a small error of the construction $K(\{f_\alpha\}_{\alpha\in Z^n})$ and all of its partial derivatives.

Lemma 4: Suppose $p = \{p_\alpha\}_{\alpha\in Z^n}$ and $q = \{q_\alpha\}_{\alpha\in Z^n}$, where $p_\alpha, q_\alpha \in \mathcal{C}(B)$, then

$$\max_{\beta\in Z^n}\left\|D^\beta Kp - D^\beta Kq\right\|_\infty \le 2^n \max_{\alpha\in Z^n}\|p_\alpha - q_\alpha\|_\infty.$$

Proof: This proof works by noticing that the map from any function f_α to any of the partial derivatives of $K(\{f_\alpha\}_{\alpha\in Z^n})$ is defined by the composition of the operators $g_{i,j}$, all of which have small gain. To see this, first note the following:

$$\frac{\partial}{\partial x_i}g_{j,k}f = \begin{cases} g_{j,k}\frac{\partial}{\partial x_i}f & i \ne j \\ f & i = j, k = 1 \\ 0 & i = j, k = 0. \end{cases}$$

Then for all $\alpha, \beta \in Z^n$

$$\frac{\partial^\beta}{\partial x^\beta}G_\alpha f = \begin{cases} 0 & \alpha_i < \beta_i \text{ for some } i \\ \left(\prod_{\substack{i=1\\ \beta_i\ne 1}}^n g_{i,\alpha_i}\right)f & \text{otherwise.}\end{cases}$$

Now, for any $f \in \mathcal{C}(B)$, it follows from the mean value theorem that for any $x \in B$:

$$\begin{aligned} |(g_{i,1}f)(x)| &= \left|\int_0^{x_i} f(x_1,\ldots,x_{i-1},\nu,x_{i+1},\ldots,x_n)d\nu\right| \\ &\le \sup_{s\in[-1,1]}|f(x_1,\ldots,x_{i-1},s,x_{i+1},\ldots,x_n)| \\ &\le \|f\|_\infty. \end{aligned}$$

Thus

$$\|g_{i,1}f\|_\infty \le \|f\|_\infty$$

for any i. Also, it is clear that for any i

$$|(g_{i,0}f)(x)| = |f(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_n)| \le \|f\|_\infty.$$

Therefore the $g_{i,j}$ have small gain, since $\|g_{i,j}f\|_\infty \le \|f\|_\infty$ for any i, j. Now since for any $\beta \in Z^n$

$$\frac{\partial^\beta}{\partial x^\beta}G_\alpha$$

is the composition of $g_{i,j}$, induction can be used to prove the the following for all $\alpha, \beta \in Z^n$

$$\left\|\frac{\partial^\beta}{\partial x^\beta}G_\alpha f\right\| \le \|f\|.$$

Therefore

$$\begin{aligned} \left\|D^\beta Kp - D^\beta Kq\right\|_\infty &= \left\|\frac{\partial^\beta}{\partial x^\beta}K(p-q)\right\|_\infty \\ &\le \sum_{\alpha\in Z^n}\left\|\frac{\partial^\beta}{\partial x^\beta}G_\alpha(p_\alpha - q_\alpha)\right\| \\ &\le \sum_{\alpha\in Z^n}\|p_\alpha - q_\alpha\| \\ &\le 2^n\max_{\alpha\in Z^n}\|p_\alpha - q_\alpha\| \end{aligned}$$

for any $\beta \in Z^n$, as desired. ■

The following theorem combines Lemmas 3 and 4 with the Weierstrass approximation theorem. It says that the polynomials are dense in $\mathcal{C}^1_\infty$ with respect to the Sobolev norm for $W^{1,\infty}$, among others.

Theorem 5: Suppose $v \in \mathcal{C}^1_\infty(B)$. Then for any $\epsilon > 0$, there exists a polynomial p, such that

$$\max_{\alpha\in Z^n}\|D^\alpha p - D^\alpha v\|_\infty \le \epsilon.$$

Proof: Since $v \in \mathcal{C}^1_\infty(B)$, $D^\alpha v \in \mathcal{C}(B)$ for all $\alpha \in Z^n$. By the Weierstrass approximation theorem, there exist polynomials

q_α such that

$$\max_{\alpha\in Z^n} \|q_\alpha - D^\alpha v\| \le \frac{\epsilon}{2^n}.$$

Let $q = \{q_\alpha\}_{\alpha\in Z^n}$ and $p = Kq$. Since the q_α are polynomial, p is polynomial. Let $f = \{D^\alpha v\}_{\alpha\in Z^n}$. By Lemma 3, $v = Kf$. Thus by Lemma 4, we have that

$$\begin{aligned}\max_{\alpha\in Z^n} \|D^\alpha p - D^\alpha v\| &= \max_{\alpha\in Z^n} \|D^\alpha Kq - D^\alpha Kf\| \\ &\le 2^n \max_{\alpha\in Z^n} \|q_\alpha - D^\alpha v\| \le \epsilon.\end{aligned}$$

■

Theorem 5 shows that for any continuously differentiable function, f, there exists an arbitrarily good polynomial approximation to the function, with error defined using the norm $\max_{\alpha\in Z^n}\|D^\alpha f\|_\infty$. The proof can be made constructive by using the Bernstein polynomials to approximate the partial derivatives. If the partial derivatives are Lipschitz continuous, then this method also gives explicit bounds on the error. In practice, numerical experiments indicate that our constructions, at least in 2 dimensions, tend to have error roughly equivalent to the standard Bernstein polynomial approximations.

III. Polynomial Lyapunov Functions

In this section, we demonstrate that polynomial Lyapunov functions can be used to approximate continuous Lyapunov functions. To be a Lyapunov function, a polynomial approximation must satisfy certain constraints. In particular, suppose that v is a Lyapunov function and p is a polynomial approximation to v. In order to also be a Lyapunov function, p must satisfy an error bound of the form

$$\left\|\frac{v(x)-p(x)}{x^Tx}\right\|_\infty \le \epsilon.$$

For p to prove exponential stability, the derivatives of p and v must satisfy a similar error bound. Conceptually, this error bound requires that the error should be everywhere bounded on a compact set, but in addition must decay to zero near the origin.

The idea behind the proof of the existence of such a polynomial, p, is to combine a Taylor series approximation with the Weierstrass approximation theorem. Specifically, a second order Taylor series expansion about a point, x_0, has the property that the error, or residue, R, satisfies

$$\frac{R(x,x_0)}{x^Tx} \to 0$$

as $x \to x_0$. However, for the error in the Taylor series to converge uniformly requires infinite differentiability of the function with uniform bounds on smoothness, an assumption too strict for the purposes of this paper. The Weierstrass approximation theorem, on the other hand, gives approximations which will always converge uniformly on a compact set, but in general no Weierstrass approximation will have the residual convergence property mentioned above for any point, x_0. The approach, then, is to use a second order Taylor series expansion to guarantee accuracy near the origin. A Weierstrass polynomial approximation to the error between the Taylor series and the function away from the origin is then used to cancel out this error and guarantee a uniform bound. An approach similar to that taken in Lemma 4 and Theorem 5 is then used to show that the map K can be used to construct polynomial approximations to differentiable functions in the norm

$$\max_{\alpha\in Z^n}\left\|\frac{D^\alpha v(x)}{x^Tx}\right\|_\infty.$$

The proof begins by combining the second order Taylor series expansion and the Weierstrass approximation theorem.

Lemma 6: Suppose $v \in \mathcal{C}_1^2(B)$. Then for any $\epsilon > 0$, there exists a polynomial p such that

$$\left\|\frac{p(x)-v(x)}{x^Tx}\right\|_\infty \le \epsilon.$$

Proof: Let the polynomial m be defined using the second order Taylor series expansion for v about $x = 0$ as

$$m(x) = v(0) + \sum_{i=1}^n x_i \frac{\partial v}{\partial x_i}(0) + \frac{1}{2}\sum_{i,j=1}^n x_i x_j \frac{\partial^2 v}{\partial x_i \partial x_j}(0).$$

Then m approximates v near the origin and specifically

$$(v-m)(0) = \frac{\partial(v-m)}{\partial x_i}(0) = \frac{\partial^2(v-m)}{\partial x_i\partial x_j}(0) = 0$$

for $i,j = 1,\ldots,n$.

Now define

$$h(x) = \begin{cases} 0 & x = 0 \\ \frac{v(x)-m(x)}{x^Tx} & \text{otherwise.}\end{cases}$$

Then from Taylor's theorem (See, e.g., [20]), we have that

$$v(x) = v(0) + \sum_{i=1}^n x_i \frac{\partial v}{\partial x_i}(0) + \frac{1}{2}\sum_{i,j=1}^n x_i x_j \frac{\partial^2 v}{\partial x_i \partial x_j}(0) + R_2(x)$$

where $\left(R_2(x)/x^Tx\right) \to 0$ as $x \to 0$. Therefore

$$h(x) = \frac{v(x)-m(x)}{x^Tx} = \frac{R_2(x)}{x^Tx} \to 0$$

as $x \to 0$ and so $h(x)$ is continuous at 0. Since $v(x) - m(x)$ and x^Tx are continuous and $x^Tx \ne 0$ on every domain not containing $x = 0$ and every point $x \ne 0$ has a neighborhood not containing $x = 0$, we conclude that $h(x)$ is continuous at every point $x \in \mathbb{R}^n$.

We can now use the Weierstrass approximation theorem, which states that there exists some polynomial q such that

$$\|q - h\|_\infty \le \epsilon.$$

The Taylor and Weierstrass approximations are now combined as $p(x) = m(x) + q(x)x^Tx$. Then p is polynomial and

$$\begin{aligned}\left\|\frac{p(x)-v(x)}{x^Tx}\right\|_\infty &= \left\|\frac{m(x)+q(x)x^Tx - v(x)}{x^Tx}\right\|_\infty \\ &= \left\|\frac{m(x)-v(x)}{x^Tx} + h(x) + (q(x)-h(x))\right\|_\infty \\ &= \|q(x)-h(x)\|_\infty \le \epsilon.\end{aligned}$$

■

The proof of the following lemma closely follows that of Lemma 4. However, the presence of the $1/x^Tx$ term poses significant technical challenges. In particular, small gain of the operators $g_{i,j}$ is no longer sufficient. Instead we use an inductive reasoning, similar to small gain, which is described in the proof.

Lemma 7: Let $p = \{p_\alpha\}_{\alpha \in Z^n}$ and $q = \{q_\alpha\}_{\alpha \in Z^n}$ with $p_\alpha, q_\alpha \in \mathcal{C}(B)$. Then

$$\max_{\beta \in Z^n} \left\| \frac{D^\beta K p(x) - D^\beta K q(x)}{x^Tx} \right\|_\infty \le 2^n \max_{\alpha \in Z^n} \left\| \frac{p_\alpha(x) - q_\alpha(x)}{x^Tx} \right\|_\infty.$$

Proof: Recall that from the definition of $g_{j,k}$, we have that

$$\frac{\partial}{\partial x_i} g_{j,k} f = \begin{cases} g_{j,k} \frac{\partial}{\partial x_i} f & i \ne j \\ f & i = j, k = 1 \\ 0 & i = j, k = 0 \end{cases}$$

which implies

$$\frac{\partial^\beta}{\partial x^\beta} G_\alpha f = \begin{cases} 0 & \alpha_i < \beta_i \text{ for some } i \\ \left(\prod_{\substack{i=1 \\ \beta_i \ne 1}}^n g_{i,\alpha_i} \right) f & \text{otherwise.} \end{cases}$$

Now consider the term

$$\frac{1}{x^Tx} (g_{i,j} f)(x).$$

We would like to obtain bounds on this function. For $j = 1$, and for any $x \in B$

$$\begin{aligned} &\left| \frac{1}{x^Tx} (g_{i,1} f)(x) \right| \\ &= \left| \int_0^{x_i} \frac{f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)}{\sum\limits_{k=1}^n x_k^2} dt \right| \\ &\le \sup_{\nu \in [-|x_i|, |x_i|]} \frac{|f(x_1, \ldots, x_{i-1}, \nu, x_{i+1}, \ldots, x_n)|}{\sum\limits_{k=1}^n x_k^2} \\ &\le \sup_{\nu \in [-|x_i|, |x_i|]} \frac{|f(\ldots, x_{i-1}, \nu, x_{i+1}, \ldots)|}{\nu^2 + \sum\limits_{\substack{k=1 \\ k \ne i}}^n x_k^2} \\ &\le \left\| \frac{f(s)}{s^Ts} \right\|_\infty. \end{aligned}$$

Here the first inequality is due to the mean value theorem and that $|x_i| \le 1$ and the second inequality follows since $x_i^2 \ge \nu^2$ for $\nu \in [-|x_i|, |x_i|]$. Therefore

$$\left\| \frac{1}{x^Tx} (g_{i,1} f)(x) \right\|_\infty \le \left\| \frac{1}{x^Tx} f(x) \right\|_\infty.$$

Similarly, if $j = 0$, then for any $x \in B$

$$\begin{aligned} &\left| \frac{1}{x^Tx} (g_{i,0} f)(x) \right| \\ &= \left| \frac{f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)}{x^Tx} \right| \\ &\le \left| \frac{f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)}{\sum\limits_{\substack{k=1 \\ k \ne i}}^n x_k^2} \right| \\ &\le \left\| \frac{f(s)}{s^Ts} \right\|_\infty \end{aligned}$$

where the first inequality follows since $x_i^2 \ge 0$. Therefore, for $j \in \{0, 1\}$ and $i = 1, \ldots n$

$$\left\| \frac{1}{x^Tx} (g_{i,j} f)(x) \right\|_\infty \le \left\| \frac{1}{x^Tx} f(x) \right\|_\infty.$$

Since the terms G_α are compositions of the $g_{i,j}$, the bounds can be applied inductively. Specifically, for any $\beta \in Z^n$

$$\begin{aligned} &\left\| \frac{1}{x^Tx} \left(\frac{\partial^\beta}{\partial x^\beta} G_\alpha f \right)(x) \right\|_\infty \\ &= \left\| \frac{1}{x^Tx} \left(\left(\prod_{\substack{i=1 \\ \beta_i \ne 1}}^n g_{i,\alpha_i} \right) f \right)(x) \right\|_\infty \\ &\le \left\| \frac{1}{x^Tx} \left(\left(\prod_{\substack{i=2 \\ \beta_i \ne 1}}^n g_{i,\alpha_i} \right) f \right)(x) \right\|_\infty \\ &\cdots \le \left\| \frac{f(x)}{x^Tx} \right\|_\infty. \end{aligned}$$

Now, given the bounds on the G_α, the triangle inequality can be used to deduce that for any $\beta \in Z^n$

$$\begin{aligned} &\left\| \frac{D^\beta K p(x) - D^\beta K q(x)}{x^Tx} \right\|_\infty \\ &= \left\| \frac{1}{x^Tx} \frac{\partial^\beta}{\partial x^\beta} K(p - q)(x) \right\|_\infty \\ &\le \sum_{\alpha \in Z^n} \left\| \frac{1}{x^Tx} \left(\frac{\partial^\beta}{\partial x^\beta} G_\alpha (p_\alpha - q_\alpha) \right)(x) \right\|_\infty \\ &\le \sum_{\alpha \in Z^n} \left\| \frac{p_\alpha(x) - q_\alpha(x)}{x^Tx} \right\|_\infty \\ &\le 2^n \max_{\alpha \in Z^n} \left\| \frac{p_\alpha(x) - q_\alpha(x)}{x^Tx} \right\|_\infty. \end{aligned}$$

■

The following theorem gives the main approximation result of the paper. It combines Lemmas 6 and 7 to show that polynomials are dense in the space $\mathcal{C}_1^{n+2}$ with respect to the $W^{1,\infty}$ norm with weight $1/x^T x$.

Theorem 8: Suppose v is a function with partial derivatives

$$D^\alpha v \in \mathcal{C}_1^2(B)$$

for all $\alpha \in Z^n$. Then for any $\epsilon > 0$, there exists a polynomial p, such that

$$\max_{\alpha \in Z^n} \left\| \frac{D^\alpha p(x) - D^\alpha v(x)}{x^T x} \right\|_\infty \le \epsilon.$$

Proof: The proof is similar to that for Theorem 5. By assumption, $D^\alpha v \in \mathcal{C}_1^2(B)$ for all $\alpha \in Z^n$. By Lemma 6, there exist polynomial functions r_α such that

$$\max_{\alpha \in Z^n} \left\| \frac{r_\alpha(x) - D^\alpha v(x)}{x^T x} \right\|_\infty \le \frac{\epsilon}{2^n}.$$

Let $r = \{r_\alpha\}_{\alpha \in Z^n}$ and $p = Kr$. Then p is polynomial since the r_α are polynomial. Let $h = \{D^\alpha v\}_{\alpha \in Z^n}$. Then by Lemma 3, $v = Kh$. Therefore by Lemma 7

$$\begin{aligned}
&\max_{\alpha \in Z^n} \left\| \frac{D^\alpha p(x) - D^\alpha v(x)}{x^T x} \right\|_\infty \\
&= \max_{\alpha \in Z^n} \left\| \frac{D^\alpha K r(x) - D^\alpha K h(x)}{x^T x} \right\|_\infty \\
&\le 2^n \max_{\alpha \in Z} \left\| \frac{r_\alpha(x) - D^\alpha v(x)}{x^T x} \right\|_\infty \le \epsilon
\end{aligned}$$

as desired. ∎

IV. Lyapunov Stability

This section begins by showing that the existence of a sufficiently smooth Lyapunov function implies the existence of a polynomial Lyapunov function.

Theorem 9: Let $\Omega \subset \mathbb{R}^n$ be bounded with radius r in norm $\|\cdot\|_\infty$ and $f(x)$ be uniformly bounded on $B_r := \{x \in \mathbb{R}^n : \|x\|_\infty \le r\}$. Suppose there exists a $v : B_r \to \mathbb{R}$ with $D^\alpha v \in \mathcal{C}_1^2(B_r)$ for all $\alpha \in Z^n$ and such that

$$\begin{aligned}
\beta_0 \|x\|^2 &\le v(x) \le \gamma_0 \|x\|^2 \\
\nabla v(x)^T f(x) &\le -\delta_0 \|x\|^2
\end{aligned}$$

for some $\beta_0 > 0$, $\gamma_0 > 0$ and $\delta_0 > 0$ and all $x \in \Omega$. Then for any $\beta < \beta_0$, $\gamma > \gamma_0$ and $\delta < \delta_0$ there exists a polynomial p such that we have the following for $x \in \Omega$:

$$\begin{aligned}
\beta \|x\|^2 &\le p(x) \le \gamma \|x\|^2 \\
\nabla p(x)^T f(x) &\le -\delta \|x\|^2 .
\end{aligned}$$

Proof: Let $\hat{v}(x) = v(rx)$ and

$$b = \|f\|_\infty = \sup_{\|x\|_\infty \le r} \|f(x)\|_\infty .$$

Choose $0 < d < \min\{\beta_0 - \beta, \gamma - \gamma_0, (\delta_0 - \delta/nb)\}$. By Theorem 8, there exists a polynomial, $\hat{p}$, such that for $\|x\|_\infty \le 1$

$$\left| \frac{\hat{p}(x) - \hat{v}(x)}{x^T x} \right| \le \frac{d}{r^2}$$

and

$$\left| \frac{\frac{\partial \hat{p}}{\partial x^i}(x) - \frac{\partial \hat{v}}{\partial x^i}(x)}{x^T x} \right| \le \frac{d}{r^2}$$

for $i = 1, \ldots, n$. Now let $p(x) = \hat{p}(x/r)$. Then for $x \in \Omega$, $\|x\|_\infty \le r$ and so $\|x/r\|_\infty \le 1$. Therefore we have the following for all $x \in \Omega$:

$$\begin{aligned}
p(x) &= v(x) + \hat{p}\left(\frac{x}{r}\right) - \hat{v}\left(\frac{x}{r}\right) \\
&= v(x) + \frac{\hat{p}\left(\frac{x}{r}\right) - \hat{v}\left(\frac{x}{r}\right)}{\left(\frac{x}{r}\right)^T \left(\frac{x}{r}\right)} r^2 x^T x \\
&\ge (\beta_0 - d) x^T x \ge \beta x^T x.
\end{aligned}$$

Likewise

$$\begin{aligned}
p(x) &= v(x) + \frac{\hat{p}\left(\frac{x}{r}\right) - \hat{v}\left(\frac{x}{r}\right)}{\left(\frac{x}{r}\right)^T \left(\frac{x}{r}\right)} r^2 x^T x \\
&\le (\gamma_0 + d) x^T x \le \gamma x^T x.
\end{aligned}$$

Finally

$$\begin{aligned}
&\nabla p(x)^T f(x) \\
&\quad = \frac{\nabla(\hat{p}\left(\frac{x}{r}\right) - \hat{v}\left(\frac{x}{r}\right))^T f}{x^T x} x^T x + \nabla v(x)^T f(x) \\
&\quad \le \sum_{i=1}^n \left(r^2 \frac{\frac{\partial \hat{p}}{\partial x_i}\left(\frac{x}{r}\right) - \frac{\partial \hat{v}}{\partial x_i}\left(\frac{x}{r}\right)}{\left(\frac{x}{r}\right)^T \left(\frac{x}{r}\right)} f_i(x) \right) x^T x - \delta_0 x^T x \\
&\quad \le n d b x^T x - \delta_0 x^T x \\
&\quad \le -\delta x^T x.
\end{aligned}$$

Thus the proposition holds for $x \in \Omega$. ∎

A consequence of Proposition 9 is that when estimating exponential rates of decay, using polynomial Lyapunov functions does not result in a reduction of accuracy. i.e., if there exists a continuous Lyapunov function proving an exponential rate of decay with bound α_0, then for any $0 < \alpha < \alpha_0$, there exists a polynomial Lyapunov function which proves an exponential rate of decay with bound α. Consider the system

$$\dot{x}(t) = f(x(t)) \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(0) = 0$ and $x(0) = x_0$. We assume that there exists an $r \ge 0$ such that for any $\|x_0\|_\infty \le r$, (1) has a unique solution for all $t \ge 0$. Define the solution map $A : \mathbb{R}^n \to \mathcal{C}([0, \infty))$ by

$$(Ay)(t) = x(t)$$

for $t \ge 0$, where x is the unique solution of (1) with initial condition y. The following Massera-type [21] converse Lyapunov theorem can be proven through application of the Gronwall-Bellman inequality.

Theorem 10: Consider the system defined by (1) where $D^\alpha f \in \mathcal{C}(\mathbb{R}^n)$ for any $\|\alpha\|_1 \leq s$. Suppose that there exist constants $\mu, \delta, r > 0$ such that

$$\|(Ax_0)(t)\|_2 \leq \mu \|x_0\|_2 e^{-\delta t}$$

for all $t \geq 0$ and $\|x_0\|_2 \leq r$. Then there exists a function $V : \mathbb{R}^n \to \mathbb{R}$ and constants $\alpha, \beta, \gamma > 0$ such that

$$\begin{aligned} \alpha \|x\|_2^2 &\leq V(x) \leq \beta \|x\|_2^2 \\ \nabla V(x)^T f(x) &\leq -\gamma \|x\|_2^2 \end{aligned}$$

for all $\|x\|_2 \leq r$. Furthermore, $D^\alpha V \in \mathcal{C}(\mathbb{R}^n)$ for any α with $\|\alpha\|_1 \leq s$.

The following gives a converse Lyapunov result. This can be taken as the main contribution of the paper.

Theorem 11: Consider the system defined by (1) where $D^\alpha f \in C_1^2(\mathbb{R}^n)$ for all $\alpha \in Z^n$. Suppose there exist constants $\mu, \delta, r > 0$ such that

$$\|Ax_0(t)\|_2 \leq \mu \|x_0\|_2 e^{-\delta t}$$

for all $t \geq 0$ and $\|x_0\|_2 \leq r$.

Then there exists a **polynomial** $v : \mathbb{R}^n \to \mathbb{R}$ and constants $\alpha, \beta, \gamma, \mu > 0$ such that

$$\begin{aligned} \alpha \|x\|_2^2 &\leq v(x) \leq \beta \|x\|_2^2 \\ \nabla v(x)^T f(x) &\leq -\gamma \|x\|_2^2 \end{aligned}$$

for all $\|x\|_2 \leq r$.

Proof: By Theorem 10, there exists a Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}$ with $D^\alpha V \in \mathcal{C}_1^2(\mathbb{R}^n)$ for all $\alpha \in Z^n$ and $\alpha_0, \beta_0, \gamma_0 > 0$ such that

$$\begin{aligned} \alpha_0 \|x\|_2^2 &\leq V(x) \leq \beta_0 \|x\|_2^2 \\ \nabla V(x)^T f(x) &\leq -\gamma_0 \|x\|_2^2 \end{aligned}$$

for $\|x\|_2 \leq r$. Then by Theorem 9, there exists a polynomial v and constants $\alpha, \beta, \gamma > 0$ such that

$$\begin{aligned} \alpha \|x\|_2^2 &\leq v(x) \leq \beta \|x\|_2^2 \\ \nabla v(x)^T f(x) &\leq -\gamma \|x\|_2^2 \end{aligned}$$

for all $\|x\|_2 \leq r$. ■

A sufficient condition for a function, f to satisfy the conditions of Theorem 11 is that f is $(n+2)$-times continuously differentiable. This is conservative, however, as f only need be 3-times continuously differentiable in any given variable. An important corollary of Theorem 11 is that ordinary differential equations defined by polynomials have polynomial Lyapunov functions. Since polynomial optimization is typically applied to systems defined by polynomials, this means that the assumption of a polynomial Lyapunov function is not conservative.

In polynomial optimization, it is common to use Positivstellensatz results to find locally positive polynomial Lyapunov functions in a manner similar to the S-procedure. When the polynomial v can be assumed to be positive, i.e., $v(x) > 0$ for all x, these conditions are necessary and sufficient. See Stengle [30], Schmüdgen [29], and Putinar [28] for strong theoretical contributions. Unfortunately, polynomial Lyapunov functions are not strictly positive since $v(0) = 0$, and so these conditions are only sufficient. However, Positivstellensatz results still allow us to search over polynomial Lyapunov functions in a manner which has proven very effective in practice.

Definition 12: A polynomial, p, is **sum-of-squares**, if there exists a $K > 0$ and polynomials g_i for $i = 1, \ldots, K$ such that

$$p(x) = \sum_{i=1}^{K} g_i(x)^2.$$

Proposition 13: Consider the system defined by (1) where f is polynomial. Suppose there exists a polynomial $v : \mathbb{R}^n \to \mathbb{R}$, a constant $\epsilon > 0$, and sum-of-squares polynomials $s_1, s_2, t_1, t_2 : \mathbb{R}^n \to \mathbb{R}$ such that

$$\begin{aligned} v(x) - s_1(x)(r - x^T x) - s_2(x) - \epsilon x^T x &= 0 \\ -\nabla v(x)^T f(x) - t_1(x)(r - x^T x) - t_2(x) - \epsilon x^T x &= 0. \end{aligned}$$

Then there exist constants $\mu, \delta, r > 0$ such that

$$\|(Ax_0)(t)\|_2 \leq \mu \|x_0\|_2 e^{-\delta t}$$

for all $t \geq 0$ and $x_0 \in Y(v, r)$ where $Y(v, r)$ is the largest sublevel set of v contained in the ball $\|x\|^2 \leq r$.

See Papachristodoulou and Prajna [24] for a proof and more details on using semidefinite programming to construct solutions to this polynomial optimization problem.

Example 1: To illustrate the "sum-of-squares" approach, we give a brief numerical example. Consider the vector field

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} -x_1^3 - x_1 x_3^2 \\ -x_2 - x_1^2 x_2 + x_3^2 \\ -x_3 + 3x_1 x_3 - 3x_3 \end{bmatrix}.$$

Using the software package SOSTOOLS [27], the following polynomial Lyapunov function can be found to prove global stability of the system:

$$\begin{aligned} V(x) = 2.164x_3^2 + 4.0451x_1^2 + 4.4676x_2^2 + 4.1204x_1^2 x_3^2 \\ +.53267x_3^4 + 2.3763x_2^2 x_1^2 + 4.1617x_1^4. \end{aligned}$$

Example 2: To illustrate the use of Proposition 13 to prove local stability, consider the following rather trivial example:

$$\dot{x}(t) = -x + x^3.$$

Using the software package SOSTOOLS [27], the following polynomial Lyapunov function can be found to prove local exponential stability of the system for $r = .9$ and $\epsilon = .01$:

$$v(x) = .939x^2.$$

The sum-of-squares variables used are

$$\begin{aligned} s_1(x) &= 0 \\ s_2 &= .929x^2 \\ t_1 &= 1.878x^2 \\ t_2 &= .1778x^2. \end{aligned}$$

The solution can be readily verified using the equations in Proposition 11. For more sophisticated examples and a thorough review of numerical methods related to the construction of Lyapunov functions for polynomial systems, the reader is referred to Wang [35] or Tan [32].

V. Conclusion

This paper gives a proof that any locally exponentially stable system with a thrice differentiable vector field will have a polynomial Lyapunov function which decreases exponentially on that region. Although we know that the polynomial degree will be finite, we can as yet give no a-priori upper bound on the degree of the polynomial. Work on the problem of generating degree bounds is currently underway. A corollary of this result is that ordinary differential equations defined by polynomials have polynomial Lyapunov functions. An important application of optimization of polynomials is the search for a polynomial Lyapunov function which proves local exponential stability. Our results, therefore, tend to support continued research into improving polynomial optimization algorithms.

In addition, as a byproduct of the proof, we gave a method for constructing polynomial approximations to differentiable functions. The interesting feature of this construction is the guaranteed convergence of the derivatives of the approximation. Another consequence of this paper is that Theorem 5 states that the polynomials are dense in $\mathcal{C}_{\infty}^{1}(B)$ with respect to the Sobolev norm $W^{1,\infty}(B)$.

References

[1] R. A. Adams, *Sobolev Spaces*. New York: Academic Press, 1975.
[2] E. A. Barbasin, "The method of sections in the theory of dynamical systems," *Rec. Math. (Mat. Sbornik) N. S.*, vol. 29, pp. 233–280, 1951.
[3] J. Bramble and S. Hilbert, "Bounds for a class of linear functionals with application to hermite interpolation," *Numer. Math.*, vol. 16, pp. 362–369, 1971.
[4] G. Chesi and S. o. polynomials, "On the gap between positive polynomials and SOS of polynomials," *IEEE Trans. Automat. Control*, vol. 52, no. 6, pp. 1066–1072, Jun. 2007.
[5] G. Chesi, A. Tesi, A. Vicino, and R. Genesio, "On convexification of some minimum distance problems," in *Proc. 5th Eur. Control Conf.*, Karlsruhe, Germany, 1999, [CD ROM].
[6] T. Dupont and R. Scott, "Polynomial approximation of functions in Sobolev spaces," *Math. Computat.*, vol. 34, no. 150, pp. 441–463, Apr. 1980.
[7] L. C. Evans, *Partial Differential Equations*. Providence, RI: American Mathematical Society, 1998.
[8] W. N. Everitt and L. L. Littlejohn, "Orthogonal polynomials & approximation in Sobolev spaces," *J. Computat. Appl. Math.*, vol. 48, pp. 69–90, 1993.
[9] W. Hahn, *Stability of Motion*. New York: Springer-Verlag, 1967.
[10] *Positive Polynomials in Control*, D. Henrion and A. Garulli, Eds. New York: Springer, 2005, vol. 312.
[11] D. Henrion and J.-B. Lassere, "GloptiPoly: Global optimization over polynomials with Matlab & SeDuMi," in *Proc. IEEE Conf. Decision Control*, 2001, pp. 747–752.
[12] D. Jackson, "Über Genauigkeit der Annäherung Stetiger Funktionen Durch Ganze Rationale Funktionen Gegebenen Grades und Trigonometrische Summengegebener Ordnung," Ph.D. dissertation, Univ. Göttingen, Göttingen, Germany, 1911.
[13] N. N. Krasovskii, *Stability of Motion*. Stanford, CA: Stanford Univ. Press, 1963.
[14] M. G. Krein, "On a problem of extrapolation of A.N. Kolmogorov," *Dokl. Akad. Nauk. SSSR*, vol. 46, pp. 306–309, 1945.
[15] J. Kurzweil, "On the inversion of Lyapuonv's second theorem on stability of motion," (in English) *Amer. Math. Soc. Transl.*, vol. 2, no. 24, pp. 19–77, 1963.
[16] V. Lakshmikantam and A. A. Martynyuk, "Lyapunov's direct method in stability theory (review)," *Int. Appl. Mech.*, vol. 28, no. 3, pp. 135–144, Mar. 1992.
[17] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, 2001.
[18] Y. Lin, E. Sontag, and Y. Wang, "A smooth converse Lyapunov theorem for robust stability," *Siam J. Control Optim.*, vol. 34, no. 1, pp. 124–160, 1996.
[19] I. Malkin, "On the question of the reciprocal of Lyapunov's theorem on asymptotic stability," *Prikl. Mat. Meh.*, vol. 18, pp. 129–138, 1954.
[20] J. E. Marsden and A. J. Tromba, *Vector Calculus*, 3rd ed. New York: W. H. Feeman & Company, 1988.
[21] J. L. Massera, "Contributions to stability theory," *Annals Math.*, vol. 64, no. 1, pp. 182–206, Jul. 1956.
[22] N. G. Meyers and J. Serrin, "$H = W$," *Proc. Nat. Acad. Sci.*, vol. 51, no. 6, pp. 1055–1056, Jun. 1964.
[23] Y. Nesterov, *High Performance Optimization*, ser. Applied Optimization. New York: Springer, 2000, vol. 33, ch. Squared Functional Systems and Optimization Problems.
[24] A. Papachristodoulou and S. Prajna, "On the construction of Lyapunov functions using the sum of squares decomposition," in *Proc. IEEE Conf. Decision Control*, Dec. 2002, pp. 3482–3487.
[25] P. A. Parrilo, "Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization," Ph.D. dissertation, California Institute of Technology, Pasadena, 2000.
[26] A. Portilla, Y. Quintana, J. Rodriguez, and E. Touris, "Weierstrass' theorem in weighted Sobolev spaces with K derivatives," *Rocky Mountain J. Math.*, vol. 37, no. 6, pp. 1989–2024, 2007.
[27] S. Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo, "New developments in sum of squares optimization & SOSTOOLS," in *Proc. Amer. Control Conf.*, 2004, pp. 5606–5611.
[28] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana Univ. Math. J.*, vol. 42, no. 3, pp. 969–984, 1993.
[29] C. Schmüdgen, "The K-moment problem for compact semi-algebraic sets," *Mathematische Annalen*, vol. 289, no. 2, pp. 203–206, 1991.
[30] G. Stengle, "A nullstellensatz and a positivstellensatz in semialgebraic geometry," *Mathematische Annalen*, vol. 207, pp. 87–97, 1973.
[31] M. H. Stone, "The generalized Weierstrass approximation theorem," *Math. Mag.*, pp. 167–183, 1948.
[32] W. Tan, "Nonlinear Control Analysis and Synthesis Using Sum-of-Squares Programming," Ph.D. dissertation, University of California, Berkeley, 2006.
[33] A. R. Teel and L. Praly, "Results on Converse Lyapunov Functions From Class-KL Estimates," in *Proc. 38th IEEE Conf. Decision Control.*, 1999, pp. 2545–2550.
[34] A. Timan, *Theory of Approximation of Functions of a Real Variable*. New York: Pergamon Press, 1960.
[35] T.-C. Wang, "Polynomial Level-Set Methods for Nonlinear Dynamics and Control," Ph.D. dissertation, Stanford University, Stanford, CA, 2007.
[36] K. Weierstrass, "Über die Analytische Darstellbarkeit Sogenannter Willkürlicher Functionen Einer Reellen veränderlichen," *Sitzungsberichte der Akademie zu Berlin*, pp. 633–639, 1885.
[37] F. W. Wilson, "Smoothing derivatives of functions and applications," *Trans. Amer. Math. Soc.*, vol. 139, pp. 413–428, May 1969.

Matthew M. Peet (S'02–M'06) received the B.S. degrees in physics and in aerospace engineering from the University of Texas at Austin in 1999 and the M.S. and the Ph.D. degree in aeronautics and astronautics from Stanford University, Stanford, CA, in 2001 and 2006, respectively.

He was a Postdoctoral Fellow at the National Institute for Research in Computer Science and Control (INRIA) near Paris, France, from 2006 to 2008, where he worked in the SISYPHE and BANG Groups. He is currently an Assistant Professor in the Mechanical, Materials, and Aerospace Engineering Department, Illinois Institute of Technology, Chicago, and Director of the Cybernetic Systems and Controls Laboratory. His current research interests are in the role of computation as it is applied to the understanding and control of complex and large-scale systems. He has studied the use of optimization algorithms and sum-of-squares programming for the analysis of nonlinear systems and partial differential equations and has worked with applications in information networks and cancer therapy.