Exponentially Stable Nonlinear Systems have Polynomial Lyapunov Functions on Bounded Regions

Matthew M. Peet [a],

[a] INRIA-Rocquencourt, Domaine de Voluceau, Rocquencourt BP105, 78153 Le Chesnay Cedex, France

Abstract

This paper presents a proof that existence of a polynomial Lyapunov function is necessary and sufficient for exponential stability of sufficiently smooth nonlinear ordinary differential equations on bounded sets. 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 right-hand side of the differential equation is polynomial or 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, we show how to use polynomials to approximate a differentiable function in the Sobolev norm $W^{1,\infty}$ to any desired accuracy. We combine this approximation result with the second-order Taylor series expansion to find that polynomial Lyapunov functions can approximate continuous Lyapunov functions arbitrarily well on bounded sets. Our investigation is motivated by the use of polynomial optimization algorithms to construct polynomial Lyapunov functions.

Key words: Nonlinear Systems, Stability, Lyapunov functions, Semidefinite Programming, Polynomial Optimization, Polynomials, Polynomial Approximation, Sobolev Spaces.

1 Introduction

The Weierstrass approximation theorem was proven in 1885 [22]. 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 [20], and on alternate topologies, as in Krein [10]. 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 [2] and its generalization in Dupont and Scott [3], and Jackson's theorem [8], generalizations of which can be found in the work of Timan [21].

The approximation of Sobolev spaces by smooth functions has been studied in a number of contexts. In Everitt and Littlejohn [5], the density of polynomials of a single variable in certain Sobolev spaces was discussed. Other work considers weighted Sobolev spaces, as in Portilla, Quintana, Rodriguez, and Touris [15].

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 [12], extensions of which can be found in Adams [1] or Evans [4].

Recently, the density of polynomials in the space of continuous functions has been used to motivate research on optimization over parameterized sets of positive polynomial functions. Relevant results include an extension of the

⋆ Corresponding author M. M. Peet.
Email address: matthew.peet@inria.fr (Matthew M. Peet).

Weierstrass approximation theorem to polynomials subject to affine constraints [14]. Examples of algorithms for optimization over cones of positive polynomials include SOSTOOLS [16], which optimizes over the cone of sums of squares of polynomials, and Gloptipoly [7], which optimizes over the dual space to obtain bounds on the primal objective.

One application of polynomial optimization which has received particular attention has been construction of polynomial Lyapunov functions for ordinary nonlinear differential equations of the form

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

where $f : \mathbb{R}^n \to \mathbb{R}^n$. Several converse theorems have shown that local exponential stability of this system implies the existence of a Lyapunov function with certain properties of continuity. The reader is referred to Hahn [6] and Krasovskii [9] for an extensive treatment of converse theorems of Lyapunov. It is not generally known, however, under what conditions an exponentially stable system has a polynomial Lyapunov function.

The main conclusion of this paper is summarized in Theorem 11, where we show that if f is n-times continuously differentiable, then exponentially stability of f on a ball is equivalent to the existence of a polynomial Lyapunov function which proves exponential stability on that ball. The smoothness condition is automatically satisfied if f is polynomial.

To establish this converse Lyapunov result, we prove two extensions to the Weierstrass approximation theorem. In Theorem 5, we show 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 $\gamma > 0$, we 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.

Our 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 5, we interpret our work as a converse Lyapunov theorem and briefly discuss implications for the use of polynomial optimization to prove stability of nonlinear ordinary differential equations.

2 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 $B := \{x \in \mathbb{R}^n : \|x\|_\infty \leq 1\}$. 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

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

where naturally, $\partial^0 f/\partial x_i^0 = f$. For $\Omega \subset \mathbb{R}^n$, we define the following sets of differentiable functions.

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

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

$\mathcal{C}^\infty(\Omega) = \mathcal{C}_1^\infty(\Omega) = \mathcal{C}_\infty^\infty(\Omega)$ is the logical extension to infinitely continuously differentiable functions. Note that in n-dimensions, $\mathcal{C}_1^i(\Omega) \subset \mathcal{C}_\infty^i(\Omega) \subset \mathcal{C}_1^{in}(\Omega)$. 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

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

The following version of the Weierstrass approximation theorem in multiple variables comes from Timan [21].

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.*

3 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, the partial derivatives of that function are not independent. For example, we have the identities

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

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. Then 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. Let X be the space of 2^n-tuples of continuous functions indexed using the 2^n elements $\alpha \in Z^n$. Thus if functions $f_\alpha \in \mathcal{C}(\Omega)$ for all $\alpha \in Z^n$, then these functions define an element of X, 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

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

In practice, the functions f_α represent either the partial derivatives of a function or approximations to those partial derivatives. The following examples illustrate the construction.

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

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

Notice that this structure automatically gives a way of approximating the partial derivatives of p. e.g.

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

If $n = 3$, then

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

The following 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 $v \in \mathcal{C}^1_\infty(\Omega)$, if $f_\alpha = D^\alpha v$ for all $\alpha \in Z^n$, then $K(\{f_\alpha\}_{\alpha \in Z^n}) = v$.*

Proof. Let e_i denote the i^{th} unit vector. Observe the following identity which holds for any function $h \in \mathcal{C}(\Omega)$

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

By assumption, $v \in \mathcal{C}^1_\infty(\Omega)$ and so $D^\alpha v \in \mathcal{C}(\Omega)$ exist 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

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

Thus the identity holds for $d+1$. The identity holds for $d = 1$ since $v \in \mathcal{C}(\Omega)$. 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*

$$\max_{\beta \in Z^n} \|D^\beta K p - D^\beta K q\|_\infty \leq 2^n \max_{\alpha \in Z^n} \|p_\alpha - q_\alpha\|_\infty.$$

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

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

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

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

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

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

Thus

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

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

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

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

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

Therefore

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

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*

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

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

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

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

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

□

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.

4 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, if v is a Lyapunov function and p is a polynomial approximation to v, then p is also a Lyapunov function if it satisfies an error bound of the form

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

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

The idea behind our 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

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

as $x \to x_0$. However, the error in the Taylor series is not guaranteed to converge uniformly over an arbitrary compact set as the order of the expansion increases. The Weierstrass approximation theorem, on the other hand, gives approximations which 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. Our approach, then, is to use a second order Taylor series expansion to guarantee accuracy near the origin. We then use a Weierstrass polynomial approximation to the error between the Taylor series and the function away from the origin to cancel out this error and guarantee a uniform bound. We then use an approach similar to that taken in Lemmas 4 and 5 to show that the map K can be used to construct polynomial approximations to differentiable functions in the norm

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

We begin by combining the second order Taylor series expansion and the Weierstrass approximation.

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

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

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

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

Then from Taylor's theorem(See, e.g. [11]), 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 $\frac{R_2(x)}{x^T x} \to 0$ as $x \to 0$. Therefore

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

as $x \to 0$ and so $h(x)$ is continuous at 0. Since $v(x) - m(x)$ and $x^T x$ are continuous and $x^T x \neq 0$ on every domain not containing $x = 0$ and every point $x \neq 0$ 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

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

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

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

□

The proof of the following lemma closely follows that of Lemma 4. However, the presence of the $1/x^T x$ term poses significant technical challenges. In particular, small gain of the operators $g_{i,j}$ is no longer sufficient. We instead 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*

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

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

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

which implies

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

Now consider the term

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

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

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

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

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

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

$$\begin{aligned}
&\left| \frac{1}{x^T x}(g_{i,0} f)(x) \right| \\
&= \left| \frac{f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)}{x^T x} \right| \\
&\leq \left| \frac{f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)}{\sum_{k \neq i}^n x_k^2} \right| \\
&\leq \left\| \frac{f(s)}{s^T s} \right\|_\infty ,
\end{aligned}$$

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

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

Since the terms G_α are compositions of the $g_{i,j}$, we can apply the above bounds inductively. Specifically, we see that for any $\beta \in Z^n$,

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

Now that we have bounds on the G_α, we can use the triangle inequality to deduce that for any $\beta \in Z^n$,

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

□

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 weighted $W^{1,\infty}$ norm with weight $1/x^T x$, among others.

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, we have

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

as desired. ☐

We now conclude the section by using Theorem 8 to show that the existence of a sufficiently smooth Lyapunov function which proves exponential stability on a bounded set implies the existence of a polynomial Lyapunov function which proves exponential stability on the set.

Proposition 9. *Let $\Omega \subset \mathbb{R}^n$ be bounded with radius r in norm $\|\cdot\|_\infty$ and $f(x)$ be uniformly bounded on $B_r := \{x \in \mathbb{R}^n : \|x\|_\infty \leq r\}$. 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*

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

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

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

for all $x \in \Omega$.

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

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

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

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

and

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

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

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

Likewise,

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

Finally,

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

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.

5 Lyapunov Stability

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 $\|x_0\|_\infty \le r$, Equation (1) has a unique solution for all $t \ge 0$. We 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 Equation (1) with initial condition y. We can prove the following Lyapunov stability theorem.

Theorem 10. *Consider the system defined by Equation* (1) *and suppose that* $f \in \mathcal{C}_1^k(\mathbb{R}^n)$ *for some integer* $k \ge 1$*. Suppose that there exist constants* $\mu, \delta, r > 0$ *such that*

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

for all $t \ge 0$ *and* $\|x_0\|_2 \le r$*. Then there exists a* $\mathcal{C}_1^k(\mathbb{R}^n)$ *function* $V : \mathbb{R}^n \to \mathbb{R}$ *and constants* $\alpha, \beta, \gamma, \mu > 0$ *such that*

$$\begin{aligned} \alpha \|x\|_2^2 \le& V(x) \le \beta \|x\|_2^2 \\ \frac{\partial}{\partial t} V((Ax)(t)) \le& -\gamma \|x\|_2^2 \end{aligned}$$

for all $\|x\|_2 \le r$*.*

The following gives a converse Lyapunov result which may be taken as the main conclusion of the paper.

Theorem 11. *Consider the system defined by Equation* (1) *where* $f \in C_1^{n+2}(\mathbb{R}^n)$ *. Then the following are equivalent.*

1. *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 $\|x_0\|_2 \leq r$.

2. *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. That 2 implies 1 is a standard result. For 1 implies 2, we use Theorem 10 to prove the existence of a Lyapunov function $V : \mathbb{R}^n \to \mathbb{R}$ with $V \in \mathcal{C}_1^{n+2}(\mathbb{R}^n)$ satisfying the conditions of 2 on $\Omega := \{x : \|x\|_2 \leq r\}$. Since $\mathcal{C}_\infty^1(\mathbb{R}^n) \subset \mathcal{C}_1^n(\mathbb{R}^n)$, Theorem 9 proves the existence of a polynomial function v which satisfies 2. □

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 [19], Schmüdgen [18], and Putinar [17] for strong theoretical contributions. Unfortunately, the polynomial Lyapunov functions are not positive since $v(0) = 0$, and so these conditions are no longer necessary and 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 Equation* (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*

$$v(x) - s_1(x)(r - x^T x) - s_2(s) - \epsilon\, x^T x = 0$$

and

$$-\nabla v(x)^T f(x) - t_1(x)(r - x^T x) - t_2(x) - \epsilon\, x^T x = 0$$

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 $\|x_0\|_2 \leq r$.

See Papachristodoulou and Prajna [13] for a proof and more details on using semidefinite programming to construct solutions to this polynomial optimization problem.

6 Conclusion

The main result of this paper is a proof that exponential stability of a sufficiently smooth nonlinear ordinary differential equation on a bounded region implies the existence of a polynomial Lyapunov function which proves exponential stability on the region. A corollary of this result is that ordinary differential equations defined by polynomials have polynomial Lyapunov functions. An important application of polynomial programming 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 our proof, we were able to give 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 the results of this paper is that the polynomials are dense in $\mathcal{C}^1_\infty(B)$ with respect to the Sobolev norm $W^{1,\infty}(B)$.

References

[1] R. A. Adams. *Sobolev Spaces*. Academic Press, 1975.

[2] J.H. Bramble and S.R. Hilbert. Bounds for a class of linear functionals with application to Hermite interpolation. *Numer. Math.*, 16:362–369, 1971.

[3] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Mathematics of Computation*, 34(150):441–463, April 1980.

[4] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[5] W.N. Everitt and L. L. Littlejohn. Orthogonal polynomials and approximation in Sobolev spaces. *Journal of Computational and Applied Mathematics*, 48:69–90, 1993.

[6] W. Hahn. *Stability of Motion*. Springer-Verlag, 1967.

[7] D. Henrion and J.-B. Lassere. GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. In *IEEE Conference on Decision and Control*, pages 747–752, 2001.

[8] D. Jackson. *Über genauigkeit der annäherung stetiger funktionen durch ganze rationale funktionen gegebenen grades und trigonometrische summengegebener ordnung*. PhD thesis, Göttingen, 1911.

[9] N. N. Krasovskii. *Stability of Motion*. Stanford University Press, 1963.

[10] M. G. Krein. On a problem of extrapolation of A.N. Kolmogorov. *Dokl. Akad. Nauk. SSSR*, 46:306–309, 1945.

[11] J. E. Marsden and A. J. Tromba. *Vector Calculus*. W. H. Feeman and Company, 3rd edition, 1988.

[12] N. G. Meyers and J. Serrin. H=W. *Proceedings of the National Academy of Sciences*, 51(6):1055–1056, June 1964.

[13] A. Papachristodoulou and S. Prajna. On the construction of Lyapunov functions using the sum of squares decomposition. In *Proceedings IEEE Conference on Decision and Control*, 2002.

[14] M. Peet and P.-A. Bliman. An extension of the weierstrass approximation theorem to linear varieties: Application to delay systems. In *Seventh IFAC Workshop on Time-Delay Systems*, 2007.

[15] A. Portilla, Y. Quintana, J.M. Rodriguez, and E. Touris. Weierstrass' theorem in weighted Sobolev spaces with K derivatives. *Rocky Mountain Journal of Mathematics*. to appear.

[16] Stephen Prajna, A. Papachristodoulou, P. Seiler, and P. A. Parrilo. New developments in sum of squares optimization and SOSTOOLS. In *Proceedings of the American Control Conference*, 2004.

[17] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.*, 42(3):969–984, 1993.

[18] C. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(2): 203–206, 1991.

[19] G. Stengle. A nullstellensatz and a positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207: 87–97, 1973.

[20] M. H. Stone. The generalized Weierstrass approximation theorem. *Mathematics Magazine*, pages 167–183 and 237–254, 1948.

[21] A.F. Timan. *Theory of approximation of functions of a real variable*. Pergamon Press Ltd., 1960.

[22] K. Weierstrass. Über die analytische darstellbarkeit sogenannter willkürlicher functionen einer reellen veränderlichen. *Sitzungsberichte der Akademie zu Berlin*, pages 633–639 and 789–805, 1885.