A Generalized Chain Rule and a Bound on the Continuity of Solutions and Converse Lyapunov Functions

Matthew M. Peet

***Abstract*— This paper gives a bound on the continuity of solutions to nonlinear ordinary differential equations. Continuity is measured with respect to an arbitrary Sobolev norm. This result is used to give a bound on the continuity of a common converse Lyapunov function. A major technical contribution of this paper is to give an explicit formula for n^{th}-degree derivatives of the composition of differentiable mappings from $\mathbb{R}^n$ to $\mathbb{R}^n$. This is a generalization of the formula of Faa di Bruno which dealt with differentiable mappings from $\mathbb{R}$ to $\mathbb{R}$. It is expected that continuity bounds of the type given in this paper can be used to prove the existence of bounded-degree polynomial Lyapunov functions or give bounds on the Lyapunov exponent.**

I. Introduction

The question of the continuity of solutions of ordinary differential equations has been studied for some time. Indeed, it is well-known that for an ordinary differential equation of the form

$$\dot{x}(t) = f(x(t)), \qquad x(0) = y$$

The continuity properties of the solution are inherited from the continuity of the vector field. Specifically, assume existence and uniqueness and denote the solution by $g(t, y)$, where we have made clear the dependence of the solution on the initial condition, y. Setting aside the question of continuity of the solution with respect to t, we examine continuity of the solution with respect to the initial condition, y. This problem arises often in areas such as sensitivity analysis and finding Lyapunov exponents. There are many excellent references on the the fundamental properties of nonlinear ordinary differential equations, e.g. [6], [3] and [16]. Concisely stated, the solution, g, lies in the same class of continuously differentiable functions as that of f. This result can be found in [1]. However, although the continuity of g is equivalent to that of f, a bound on the continuity properties of f will not, in general, translate to a bound on the continuity of g. Precisely: if L is a global bound on the Lipschitz continuity of f, then L will not generally be a bound on the Lipschitz continuity of g. Moreover, there seems to be no result in the literature yielding any kind of bound on the m^{th} derivative of g in the general n-dimensional case.

The contribution of this paper is to use an extension of the generalized chain rule and a bound on the Lipschitz continuity of f to give a bound on the Lipschitz continuity of the m^{th} derivative of g in the general n-dimensional case. This result is used to give a bound on the m^{th} derivative of a common form of converse Lyapunov functional. This result is obtained by developing a generalized chain-rule to bound the derivatives of the solution map.

Although there are many reasons that one might be interested in obtaining a bound on the continuity of the derivatives of g, or of converse Lyapunov functionals, there is one reason that is of particular interest. Recently, researchers have investigated the use of semidefinite programming to construct polynomial Lyapunov functions for nonlinear ordinary differential equations. Among the proposed methods is the use of "sum-of-squares" of polynomials. Given a degree bound, the set of such polynomials can be parameterized using positive semidefinite matrices. Although not all positive polynomials are sums-of-squares (SOS), the gap between sufficient and necessary can often be reduced through the use of "Positivstellensatz"-type results, at the expense of additional complexity. In [14], we showed that if a locally exponentially stable system has a vector field which is three times continuously differentiable, then there will always exist a polynomial Lyapunov function which proves exponential stability. While this result is intriguing, the practical effect is limited by the fact that computation always occurs over a set of polynomials of bounded degree. Therefore, conservatism will generally arise through the choice of degree bound. There is strong evidence to suggest, however, that a bound on the continuity of the derivatives of the converse Lyapunov function will allow one to find a bound on the degree of the associated polynomial Lyapunov function. Note, however, that application of the results of this paper to obtain this bound is significantly beyond the scope of the paper.

Naturally, the general question of the properties of converse Lyapunov functions has been well studied, and in particular the differentiability of the Lyapunov function has been explored. Some early examples include [2], [11] and [8]. A summary and extension of many of these results can be found in [12]. In a different vein, the use of smoothing operators on open sets to create infinitely-differentiable functions was explored in the work [5], and more recently in [10]. Other innovative results on converse Lyapunov functions can be found in [9] and [15]. The reader is also referred to [3] and [7] for a treatment of converse theorems of Lyapunov.

The difference between previous results and the current work is that while previous results proved differentiability of the converse Lyapunov function, no bound was given on the size of this derivative, whereas in the present work, we give an explicit bound for the size of the derivatives by using properties of the vector field. See Section V. The first step in bounding the continuity of converse Lyapunov functions is to bound the continuity of solutions to ordinary

Matthew M. Peet is an Assistant Professor in the department of Mechanical, Materials, and Aerospace Engineering, the Illinois Insitute of Technology, 10 W. 32nd st., Chicago, IL, 60521, USA

differential equations. This work is given in Section IV. Finally, a major technical contribution of this paper is to construct a generalization of Faa di Bruno's famous formula which gave an explicit formula for the k^{th} derivatives of a composition of differentiable functions. Whereas Faa di Bruno was interested in derivatives of $f(g(x))$ where $f, g : \mathbb{R} \to \mathbb{R}$, we are interested in the case where $f, g : \mathbb{R}^n \to \mathbb{R}^n$. This work also generalizes and unifies several previous results on the generalization of the Faa di Bruno's formula, such as [13] and [4]. See Section III.

II. 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, let $\prod_i h_i$ 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 denote partial derivatives by

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

where naturally, $\partial^0 f / \partial x_i^0 = f$. When it is not clear which variable is being differentiated, we will make this explicit theough the notation D_x^α where x is the vector of variables to differentiate. For $\Omega \subset \mathbb{R}^n$, we define the following sets of differentiable functions.

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

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

As is customary, we refer to a function f as r-times continuously differentiable on Ω if $f \in \mathcal{C}_1^r(\Omega)$. $\mathcal{C}^\infty(\Omega)$ is the extension to infinitely continuously differentiable functions. 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 $|\alpha|_1 \leq k$ and norm

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

The n^{th} Bell number, denoted B_n, is the number of partitions of a set with n members. The Stirling number of the second kind, $S(n, k)$ is defined as the number of ways to partition a set of n elements into k nonempty subsets. More explicitly,

$$S(n, k) = \sum_{i=1}^k (-1)^{k-i} \frac{j^{n-1}}{(i-1)!(k-i)!}$$

and

$$B_n = \sum_{k=1}^n S(n, k).$$

III. Previous Work and a Generalized Chain Rule

The results in this paper are based on the application of a generalized chain rule in combination with the Gronwall-Bellman Inequality. The derivation of a generalized chain rule is the major technical contribution of this paper. However, before presenting this result, we give a statement of the well-known Gronwall-Bellman Lemma. This result is a classic tool for investigating properties of the solution.

Lemma 1 (Gronwall-Bellman): Let λ be continuous and μ be continuous and nonnegative. Let y be continuous and satisfy for $t \leq b$,

$$y(t) \leq \lambda(t) + \int_a^t \mu(s) y(s) ds.$$

Then

$$y(t) \leq \lambda(t) + \int_a^t \lambda(s)\mu(s) \exp\left[\int_s^t \mu(\tau) d\tau\right] ds$$

If λ and μ are constants, then

$$y(t) \leq \lambda e^{\mu t}.$$

The second topic discussed in this section is a generalization of the chain rule to higher derivatives and to the composition of maps from $\mathbb{R}^n \to \mathbb{R}^n$. Alternatively, this is also a generalization of the well-known Faà di Bruno's formula

$$\frac{d^n}{dx^n} f(g(x)) = \sum_{\pi \in \Pi} f^{(|\pi|)}(g(x)) \cdot \prod_{B \in \pi} g^{(|B|)}(x).$$

where $f, g : \mathbb{R} \to \mathbb{R}$, Π is the set of partitions of $\{1, \ldots, n\}$, and $|\cdot|$ denotes cardinality. There are two existing generalization of Faà di Bruno's formula. The first extension was given in [13] and considers the case when $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}^n$. We will not list the formula given in [13] because the presentation is not given in combinatoric form and is therefore quite lengthy. The second extension of Faà di Bruno's formula considers the case when $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ can be found in [4].

$$\frac{\partial^n}{\partial x_1 \cdots \partial x_n} f(g(x)) = \sum_{\pi \in \Pi} f^{(|\pi|)}(g(x)) \cdot \prod_{B \in \pi} \frac{\partial^{|B|} g}{\prod_{j \in B} \partial x_j}$$

In this paper, we give a generalization of Faà di Bruno's formula to the case when $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^n$. We rely on the following notation.

Definition 2: Let Ω_r^i denote the set of partitions of $(1,\ldots,r)$ into i non-empty subsets.

In the following, we decompose the multi-index $\alpha \in \mathbb{N}^n$, into a sequence of r-dimensional unit vectors γ_i. Although the set of such vectors is unique, the ordering, naturally, is not. It is possible that this aspect of the formula may admit some future simplification.

Lemma 3 (Generalized Chain Rule): Suppose $f : \mathbb{R}^n \to \mathbb{R}$ and $z : \mathbb{R}^n \to \mathbb{R}^n$ are r-times continuously differentiable. Let $\alpha \in \mathbb{N}^n$ with $|\alpha|_1 = r$. Let $\{\gamma_i\}_{i=1}^r \in \mathbb{N}^n$ be any sequence of bases such that $\alpha = \sum_{i=1}^r \gamma_i$ and $\|\gamma_i\|_1 = 1$.

$$D_x^\alpha f(z(x)) = \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \times \sum_{\beta\in\Omega_r^i} \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x)$$

Proof: Assume that the statement is true for some r. Then for $r+1$, we apply the chain rule to obtain the following two terms.

$$\begin{aligned} D^\alpha f(z(x)) &= D^{\gamma_{r+1}} D^{\alpha-\gamma_{r+1}} f(z(x)) \\ &= \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \left(\sum_{j_{i+1}=1}^n \frac{\partial^{i+1}}{\partial x_{j_1}\cdots\partial x_{j_{i+1}}} f(z(x)) \times D^{\gamma_{r+1}} z_{j_{i+1}}(x) \right) \sum_{\beta\in\Omega_r^i} \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &\quad + \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \times D^{\gamma_{r+1}} \sum_{\beta\in\Omega_r^i} \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &= f_1(x) + f_2(x) \end{aligned}$$

Now, we examine the first term. We first make the substitution $i = i'-1$. We then use the identity $\Omega_r^i \oplus \{r+1\} = \{\beta \in \Omega_{r+1}^{i+1} : \beta_{i+1} = \{r+1\}\}$. This is the set of partitions of $(1,\cdots,r+1)$ into $i+1$ distinct subsets, such that for each partition, one of the elements of the partition is simply $\{r+1\}$. We have the following.

$$\begin{aligned} f_1(x) &= \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \sum_{j_{i+1}=1}^n \frac{\partial^{i+1}}{\partial x_{j_1}\cdots\partial x_{j_{i+1}}} f(z(x)) \times \sum_{\beta\in\Omega_r^i} D^{\gamma_{r+1}} z_{j_{i+1}}(x) \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &= \sum_{i'=2}^{r+1} \sum_{j_1=1}^n \cdots \sum_{j_{i'}=1}^n \frac{\partial^{i'}}{\partial x_{j_1}\cdots\partial x_{j_{i'}}} f(z(x)) \times \sum_{\beta\in\Omega_r^{i'-1}} D^{\gamma_{r+1}} z_{j_{i'}}(x) \prod_{k=1}^{i'-1} D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &= \sum_{i'=2}^{r+1} \sum_{j_1=1}^n \cdots \sum_{j_{i'}=1}^n \frac{\partial^{i'}}{\partial x_{j_1}\cdots\partial x_{j_{i'}}} f(z(x)) \times \sum_{\substack{\beta\in\Omega_{r+1}^{i'} \\ \beta_{i'}=\{r+1\}}} \prod_{k=1}^{i'} D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \end{aligned}$$

For the second term, we take the set of partitions of $(1,\cdots,r)$ into i distinct subsets. Then for each partition β, create i new partitions of $(1,\cdots,r+1)$ by adding $r+1$ to each of the i subsets. This new set of partitions can be denoted using the cartesian product as $\Omega_r^i \times \{r+1\}$.

$$\begin{aligned} f_2(x) &= \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \times \sum_{\beta\in\Omega_r^i} D^{\gamma_{r+1}} \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &= \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \times \sum_{\beta\in\Omega_r^i} \sum_{h=1}^i D^{\gamma_r + \sum_{l\in\beta_h}\gamma_l} z_{j_h}(x) \prod_{\substack{k=1 \\ k\neq h}}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\ &= \sum_{i=1}^r \sum_{j_1=1}^n \cdots \sum_{j_i=1}^n \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \times \sum_{\beta\in\Omega_r^i\times\{r+1\}} \prod_{k=1}^i D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \end{aligned}$$

Now, since

$$\Omega_r^i \times \{r+1\} \bigcap \Big\{ \beta \in \Omega_{r+1}^i : \ \beta_i = \{r+1\} \Big\} = \Omega_{r+1}^i$$

for $i = 2,\cdots,r$, and

$$\Omega_r^i \times \{r+1\} = \Omega_{r+1}^i$$

for $i=1$ and

$$\{\beta \in \Omega_{r+1}^i : \ \beta_i = \{r+1\}\} = \Omega_{r+1}^i$$

for $i = r+1$, when we combine the terms, we can also combine β summations for $i = 2,\cdots,r$ (for $i = 1, r+1$, only a single appears) to get.

$$
\begin{aligned}
&f(x) = f_1(x) + f_2(x) \\
&= \Bigg(\sum_{i=1}^{1} \sum_{\beta\in\Omega_r^i\times\{r+1\}} \\
&\qquad + \sum_{i=2}^{r} \left(\sum_{\substack{\beta\in\Omega_{r+1}^i \\ \beta_i=\{r+1\}}} + \sum_{\beta\in\Omega_r^i\times\{r+1\}} \right) \\
&\qquad\qquad + \sum_{i=r+1}^{r+1} \sum_{\substack{\beta\in\Omega_{r+1}^i \\ \beta_i=\{r+1\}}} \Bigg)\times \\
&\sum_{j_1=1}^{n}\cdots\sum_{j_i=1}^{n} \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \prod_{k=1}^{i} D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \\
&= \sum_{i=1}^{r+1}\sum_{j_1=1}^{n}\cdots\sum_{j_i=1}^{n} \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x))\times \\
&\qquad \sum_{\beta\in\Omega_{r+1}^i} \prod_{k=1}^{i} D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x)
\end{aligned}
$$

Thus we have completed the inductive step. Since the statement is true for $r = 1$ by the chain rule, we have completed the proof. ∎

We can use Lemma 3 to obtain bounds on the derivatives of a function. This is done in the following two corollaries.

Corollary 4: Suppose $z : \mathbb{R}^n \to \mathbb{R}^n$ is r-times continuously differentiable. Let $\alpha \in \mathbb{N}^n$ with $|\alpha|_1 = r$. Then

$$
\|D^\alpha z(x)^T z(x)\| \le 2^r \max_{\beta\le\alpha} \left\|D^\beta z(x)\right\|^2 .
$$

Proof: Let $\{\gamma_i\} \in Z^r$ be any decomposition of α so that $\alpha = \sum_{i=1}^r \gamma_i$. If $f(z) = z^T z$, then by Lemma 3, we have the following.

$$
\begin{aligned}
&D_x^\alpha z(x)^T z(x) \\
&= \sum_{j=1}^{n}\sum_{i=1}^{r} \frac{d^i}{dz_j} z_j(x)^2 \sum_{\beta\in\Omega_r^i}\prod_{k=1}^{i} D^{\sum_{l\in\beta_k}\gamma_l} z_j(x) \\
&= 2z(x)^T D^\alpha z(x) \\
&\quad + 2\sum_{\beta\in\Omega_r^2} (D^{\sum_{l\in\beta_1}\gamma_l} z(x))^T (D^{\sum_{l\in\beta_2}\gamma_l} z(x))
\end{aligned}
$$

The second line is the expansion of the first for $i = 1, 2$. In the second equality, we have furthermore noted that $\Omega_r^1 = \{1, \cdots, r\}$. Naturally, since z_l is scalar, we could have used the formula in [4]. Finally, we have the following.

$$
\begin{aligned}
&\|D_x^\alpha z(x)^T z(x)\| \le 2\,\|z(x)\|\,\|D^\alpha z(x)\| \\
&+ 2\sum_{\beta\in\Omega_r^2} \left\|D^{\sum_{l\in\beta_1}\gamma_l} z(x)\right\| \left\|D^{\sum_{l\in\beta_2}\gamma_l} z(x)\right\| \\
&\le (2S(r,2)+2) \max_{\beta\le\alpha} \left\|D^\beta z(x)\right\|^2 \\
&= 2^r \max_{\beta\le\alpha} \left\|D^\beta z(x)\right\|^2
\end{aligned}
$$

Recall here that $S(r, n)$ denotes the Stirling number of the second kind. In the final equation, we have used the identity $S(r, 2) = 2^{r-1} - 1$. ∎

Corollary 5: Suppose $f : \mathbb{R}^n \to \mathbb{R}$ and $z : \mathbb{R}^n \to \mathbb{R}^n$ are r-times continuously differentiable. Let $\alpha \in \mathbb{N}^n$ with $|\alpha|_1 = r$. Let B_r denote the r^{th} Bell number. Then

$$
\begin{aligned}
&\|D^\alpha f(z(x))\| \le \\
&\quad n^r B_r \max_{|\beta|\le r}\|D^\beta f\|_\infty \max_{\beta<\alpha} \max_k \|D^\beta z_k(x)\|_\infty^r \\
&\qquad + n \max_j \|\frac{\partial}{\partial x_j} f\|_\infty \max_k \|D^\alpha z_k(x)\|.
\end{aligned}
$$

Proof: Let $\{\gamma_i\}_{i=1}^r \in \mathbb{N}^n$ be any sequence of bases such that $\alpha = \sum_{i=1}^r \gamma_i$ and $\|\gamma_i\|_1 = 1$. By Lemma 3, we have the following.

$$
\begin{aligned}
&\|D_x^\alpha f(z(x))\| \\
&= \Bigg\| \sum_{i=1}^{r}\sum_{j_1=1}^{n}\cdots\sum_{j_i=1}^{n}\sum_{\beta\in\Omega_r^i} \\
&\frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \prod_{k=1}^{i} D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \Bigg\| \\
&\le \sum_{i=2}^{r}\sum_{\beta\in\Omega_r^i}\sum_{j_1=1}^{n}\cdots\sum_{j_i=1}^{n} \\
&\quad \left\| \frac{\partial^i}{\partial x_{j_1}\cdots\partial x_{j_i}} f(z(x)) \right\| \prod_{k=1}^{i} \left\| D^{\sum_{l\in\beta_k}\gamma_l} z_{j_k}(x) \right\| \\
&\quad + \sum_{j=1}^{n} \left\| \frac{\partial}{\partial x_j} f(z(x)) D^\alpha z_j(x) \right\| \\
&\le n^r B_r \max_{|\beta|\le r}\|D^\beta f\|_\infty \max_{\substack{\beta<\alpha \\ k=1,\cdots,n}} \|D^\beta z_k\|_\infty^r \\
&\quad + n \max_j \|\frac{\partial}{\partial x_j} f\|_\infty \max_k \|D^\alpha z_k(x)\|
\end{aligned}
$$

∎

These corollaries give bounds on the derivatives of a composite function given bounds on the lower-order derivatives of the component functions. Corollary 5 is used as an inductive step several times in the following proofs. The bound given clearly increase quickly with the number of states and the order of the derivative. In Corollary 4, there does not appear to be any conservatism. In Corollary 5, however, we have used the inequality

$$
\sum_{i=1}^{r} S(r,i) n^i \le \sum_{i=1}^{r} S(r,i) n^r = B_r n^r .
$$

Although this inequality is conservative, it is clearly not conservative by more than a factor of r.

IV. Bounds on the Derivatives of the Solution

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 $r \geq 0$ such that for any $\|x_0\|_\infty \leq r$, Equation (1) has a unique solution for all $t \geq 0$. We define the solution map $A : \mathbb{R}^n \to \mathcal{C}([0,\infty))$ by

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

for $t \geq 0$, where x is the unique solution of Equation (1) with initial condition y.

A. Existence of solutions

We briefly mention the question of existence of sufficiently continuously differentiable solutions. Proof of the existence of these solutions can be found in a number of sources. In particular, we quote the following work of Arnol'd [1].

Lemma 6 (Arnol'd): Suppose that $f \in \mathcal{C}_1^r(X)$ for $X \subset \mathbb{R}^n$. Then any solution $g : \mathbb{R}^+ \times X \to Y$, where $Y \subset \mathbb{R}^n$, and such that

$$\frac{\partial}{\partial t} g(t,y) = f(g(t,y))$$
$$g(0,y) = y$$

satisfies, for any fixed T, $g(T,\cdot) \in \mathcal{C}_1^r(Y)$.

In this paper, we make use of the assumption that the vector field is sufficiently differentiable. This ensures that the solutions are likewise differentiable.

B. Bounded Solutions

In the following lemma, we give a bound on the derivatives of the solution of a nonlinear ordinary differential equation. For reasons which will become clear when consider the converse Lyapunov function, we are obliged to keep the form general at the expense of some additional complexity. In particular, the bound is a function of

- A prior bound, w, on the solution $z(t,y)$ for $y \in Y$, $t \in [0,T]$. Such bounds are readily obtained even for unstable systems. Obviously, the tighter the bound on z, the better the overall bound. Hence we keep this general as opposed to giving a bound a-priori.
- A time frame, $[0,T]$.
- The maximum order of the derivative, r.
- A Lipschitz continuity factor, L.

The result of the lemma is to give a uniform bound on the derivatives of the solution, z, up to order r.

Lemma 7: Suppose the elements of $f : \mathbb{R}^n \to \mathbb{R}^n$ are r-times continuously differentiable and $\|D^\alpha f(z)\| \leq L$ on $\|z\| \leq w$ for any $|\alpha|_1 \leq r$. If $z : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ satisfies

$$\frac{\partial}{\partial t} z(t,y) = f(z(t,y)) \qquad z(0,y) = y$$

and

$$\|z(\cdot,y)\| \leq w \quad \text{for all } y \in Y,\ t \in [0,T]$$

Then

$$\max_{|\alpha|_1 \leq r} \max_k \|D_y^\alpha z_k(T,y)\|_\infty \leq M(T,w,n,L,r)$$

for all $y \in Y$, where

$$M(T,w,n,L,r) = (\max\{1,w\} + TnLw)e^{nLT} \prod_{i=2}^{r} \left(Tn^i B_i L e^{nLT}\right)^{r!/i!}.$$

Proof: Note the identity

$$z(t,y) = y + \int_0^t f(z(s,y))ds.$$

Then

$$D_y^\alpha z(t,y) = D^\alpha y + \int_0^t D_y^\alpha f(z(s,y))ds.$$

By Corollary 5, we have

$$\|D_x^\alpha f_i(z(x))\| \leq n^r B_r \max_{|\beta|\leq r} \|D^\beta f_i\|_\infty \max_{\substack{\beta<\alpha \\ k=1,\cdots,n}} \|D^\beta z_k\|_\infty^r + n \max_j \|\frac{\partial}{\partial x_j} f_i\|_\infty \max_k \|D^\alpha z_k(x)\|.$$

We proceed by induction. We show the statement holds for $r=0$ and $r=1$, then consider $r>1$. By assumption, the bound is satisfied for $\alpha = 0$. Now consider the case of $|\alpha|_1 = r = 1$. Applying the inequality from Corollary 5, we get

$$\begin{aligned}
&\max_k \|D_y^\alpha z_k(t,y)\| \\
&\leq \max_k \|D^\alpha y_k\| + \int_0^t \max_k \|D_y^\alpha f_k(z(s,y))\| ds \\
&\leq 1 + nL \int_0^t w + \max_j \|D^\alpha z_j(x)\| ds \\
&= 1 + nLwt + \int_0^t nL \max_k \|D^\alpha z_k(x)\| ds,
\end{aligned}$$

and so by Gronwall-Bellman,

$$\max_k \|D_y^\alpha z_k(t,y)\| \leq (1 + tnLw)e^{nLt} \leq M(T,w,n,L,r)$$

Which holds for $r=1$. Now, we consider the inductive argument for $r>1$. Suppose $\max_{\beta<\alpha} \max_k \|D_y^\beta z_k(t,y)\|_\infty \leq M(t,w,n,c,r-1)$ on $y \in Y$. Then we have

$$D_y^\alpha z_k(t,y) = \int_0^t D_y^\alpha f_k(z(s,y))ds$$

and so, using the formula from Corollary 5, we have

$$\max_k \|D_y^\alpha z_k(t,y)\| \leq Tn^r B_r L M(T,w,n,L,r-1)^r + \int_0^t nL \max_k \|D^\alpha z_k(s,y)\| ds.$$

Similarly, by Gronwall-Bellman, this implies

$$\begin{aligned}
&\max_k \|D_y^\alpha z_k(t,y)\| \le Tn^r B_r LM(T,w,n,L,r-1)^r e^{nLt} \\
&\le Tn^r B_r LM(T,w,n,L,r-1)^r e^{nLT} \\
&= (1+TnLw)e^{nLT} Tn^r B_r Le^{nLT} \times \\
&\qquad \prod_{i=2}^{r-1} \left(Tn^i B_i Le^{nLT}\right)^{r(r-1)!/i!} \\
&= (1+TnLw)e^{nLT} \prod_{i=2}^{r} \left(Tn^i B_i Le^{nLT}\right)^{r!/i!} \\
&= M(T,w,n,L,r).
\end{aligned}$$

∎

Although the bounds given in Lemma 7 are complicated, there is no obvious simplification unless we consider a special case. Such a simplification must inevitably arise from the choice of the function f. In the following sections, the bounds given for the solution map will be used to determine bounds for a converse Lyapunov function.

V. Bounds on the Derivative of a Converse Lyapunov Function

In this section, we examine a classic converse Lyapunov function given by

$$V(x) = \int_0^T \|z(s,x)\|^2 ds,$$

where $z(t,x_0)$ is the solution of $\dot{z}(t,x_0) = f(z(t,x_0))$ for $t \ge 0$, with initial condition $z(0,x_0) = x_0$. Clearly, the smoothness properties of $V(x)$ with respect to x are inherited from the smoothness properties of the solution map $z(t,x)$. Therefore, we can give bounds on the derivatives of the solution $z(t,x_0)$ with respect to the initial condition x_0 and use these results to bound the derivative of the converse Lyapunov function.

A. Converse Lyapunov Functions

The classical Lyapunov function given above can be found in a number of sources. Typical is the following analysis given in [6].

Theorem 8 (Khalil): Suppose that $\|\partial f/\partial x\|_2 \le L$, $f(0) = 0$, and there exists a unique function $z : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\begin{aligned}
&\frac{\partial}{\partial t} z(t,y) = f(z(t,y)) \\
&z(0,y) = y
\end{aligned}$$

and

$$\|z(t,y)\| \le k\|y\|e^{-\lambda t} \quad \text{for all } \|y\| \le d$$

for $k, \lambda, d > 0$. Then there exist constants $c_1, c_2, c_3 > 0$ and a function $V : \mathbb{R}^n \to \mathbb{R}$ such that for any $\|x\| \le d$,

$$\begin{aligned}
c_1\|x\|^2 \le V(x) \le c_2\|x\|^2 \\
\nabla V(x)^T f(x) \le -c_3\|x\|^2.
\end{aligned}$$

Furthermore, V can be chosen as

$$V(x) = \int_0^T \|z(s,x)\|^2 ds$$

with $T = \frac{\ln 2k^2}{2\lambda}$, $c_1 = \frac{1-e^{-2LT}}{2L}$, $c_2 = \frac{k^2(1-e^{-2\lambda T})}{2\lambda}$, and $c_3 = 1/2$.

It is important to note that the Lyapunov function used here will not necessarily be optimal for the purposes of estimating the exponential rate of decay. This is partially due to the fact that the integral is over a finite time, $[0,T]$. The finite time horizon is necessary, however, to construct the smoothness bounds given in the following Theorem.

Theorem 9: Suppose that $\|D^\alpha f\|_\infty \le L$ on $\|x\| \le d$ for $|\alpha|_1 \le r$, $f(0) = 0$, and there exists a unique function $z : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\begin{aligned}
&\frac{\partial}{\partial t} z(t,y) = f(z(t,y)) \\
&z(0,y) = y
\end{aligned}$$

and

$$\|z(t,y)\| \le k\|y\|e^{-\lambda t} \quad \text{for all } \|y\| \le d$$

for $k, \lambda, d > 0$. Then there exists constants $c_1, c_2, c_3 > 0$ and a function $V : \mathbb{R}^n \to \mathbb{R}$ such that for any $\|x\| \le d$,

$$c_1\|x\|^2 \le V(x) \le c_2\|x\|^2 \tag{2}$$

$$\nabla V(x)^T f(x) \le -c_3\|x\|^2. \tag{3}$$

Furthermore,

$$\max_{|\alpha|_1 < r} \|D^\alpha V(x)\| \le 2^r \frac{\ln 2k^2}{2\lambda} M\left(\frac{\ln 2k^2}{2\lambda}, kd, n, L, r\right)^2$$

for all $\|x\| \le d$ where

$$M(T,w,n,L,r) = (1+TnLw)e^{nLT} \prod_{i=2}^{r} \left(Tn^i B_i Le^{nLT}\right)^{r!/i!}.$$

Proof: Let

$$V(x) = \int_0^T z(s,x)^T z(s,x) ds$$

with $T = \frac{\ln 2k^2}{2\lambda}$, $c_1 = \frac{1-e^{-2LT}}{2L}$, $c_2 = \frac{k^2(1-e^{-2\lambda T})}{2\lambda}$, and $c_3 = 1/2$. By Theorem 8, V satisfies conditions 2 and 3.

Now, by Corollary 4 we have that

$$\begin{aligned}
\max_{|\alpha|_1 \le r} \|D_x^\alpha z(s,x)^T z(s,x)\| &\le 2^r \max_{|\alpha|_1 \le r} \|D^\alpha z(s,x)\|^2 \\
&\le 2^r n \max_{|\alpha|_1 \le r} \max_k \|D^\alpha z_k(s,x)\|^2
\end{aligned}$$

where z_k is the k-th component of z. Noting that $\|z_k(s,y)\| \le kd$, by Lemma 7, we have that

$$\max_{|\alpha|_1 \le r} \max_k \|D_y^\alpha z_k(s,y)\|_\infty \le M(s,kd,n,L,r).$$

Therefore, since the function $M(s)$ is increasing,

$$\begin{aligned}
&\max_{|\alpha|_1 \le r} \|D^\alpha V(x)\| \\
&\le \int_0^T \max_{|\alpha|_1 \le r} \|D^\alpha z(s,x)^T z(s,x)\| ds \\
&\le \int_0^T 2^r \max_{|\alpha|_1 \le r} \|D^\alpha z(s,x)\|^2 ds \\
&\le T 2^r n M(T, w, n, L, r) \\
&= 2^r \frac{\ln 2k^2}{2\lambda} M n \left(\frac{\ln 2k^2}{2\lambda}, kd, n, L, r \right)^2 .
\end{aligned}$$

■

To conclude this section, we note that the bound given here may be conservative due to a number of sources. First, there is a factor of r conservatism in Corollary 5. In addition, there may be conservatism due to our choice of the particular structure of Lyapunov function. In particular, the bound grows exponentially with the value of T in the integration. Converse functions which reduce this value will yield better bounds.

The exponential growth with respect to T is due to the the fact that exponential stability of the solution does not imply exponential stability of the derivatives of the solution. An interesting question, therefore, is under what condition the derivatives of the solution are also stable. Are there classes of system for which we can prove stability?

VI. Conclusion

In this paper, we have presented significant technical results. The first result is a generalization of the formula of Faa di Bruno to maps from $\mathbb{R}^n$ to $\mathbb{R}^n$. This is a generalization of the chain rule to higher derivatives and multiple dimensions. The second result uses the first result to obtain bounds on the derivatives of solutions of ordinary differential equations and thereby to obtain bounds on the derivatives of a common converse Lyapunov function. These bounds are obtained using a bound on the derivatives of the vector field. Aside from purely theoretical interest, these bounds can possibly be used to prove the existence of a polynomial Lyapunov function with a bound on the degree of the polynomial. Work on this topic is beyond the scope of this paper.

References

[1] V. I. Arnol'd, *Ordinary Differential Equations*. The MIT Press, 1978.

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

[4] M. Hardy, "Combinatorics of partial derivatives," *The Electronic Journal of Combinatorics*, vol. 13, 2006.

[5] F. W. W. Jr., "Smoothing derivatives of functions and applications," *Transactions of the American Mathematical Society*, vol. 139, pp. 413–428, May 1969.

[6] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

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

[8] J. Kurzweil, "On the inversion of Lyapuonv's second theorem on stability of motion," *Amer. Math. Soc. Transl.*, vol. 2, no. 24, pp. 19–77, 1963, english Translation. Originally appeared 1956.

[9] V. Lakshmikantam and A. A. Martynyuk, "Lyapunov's direct method in stability theory (review)," *International Applied Mechanics*, vol. 28, no. 3, pp. 135–144, March 1992.

[10] 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.

[11] I. Malkin, "On the question of the reciprocal of Lyapunov's theorem on asymptotic stability," *Prikl. Mat. Meh.*, vol. 18, pp. 129–138, 1954.

[12] J. L. Massera, "Contributions to stability theory," *Annals of Mathematics*, vol. 64, no. 1, pp. 182–206, July 1956.

[13] R. L. Mishkov, "Generalization of the formula of Faa di Bruno for a composite function with a vector argument," *International Journal of Mathematics and Mathematical Sciences*, vol. 24, no. 7, pp. 481–491, 2000.

[14] M. M. Peet, "Exponentially stable nonlinear systems have polynomial Lyapunov functions on bounded regions," *IEEE Transactions on Automatic Control*, vol. 52, no. 5, 2009.

[15] A. R. Teel and L. Praly, "Results on converse lyapunov functions from class-KL estimates," 1999, pp. 2545–2550.

[16] M. Vidyasagar, *Nonlinear Systems Analysis*. Prentice-Hall, 1993.