Automatica 142 (2022) 110349

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Technical communique

A note on the equivalence of a strongly convex function and its induced contractive differential equation☆

Maxwell Fitzsimmons *, Jun Liu

Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 29 March 2021
Received in revised form 25 February 2022
Accepted 7 March 2022
Available online 6 May 2022

Keywords:
Contraction theory
Lyapunov methods
Convex optimization

ABSTRACT

A strongly convex function naturally induces a gradient flow that is contractive. This paper is a short investigation on when the converse to the previous statement holds. That is, given a differential equation that is contractive, does there exist a strongly convex function that induces the differential equation? We show that, if sufficient smoothness of the vector field is assumed, then the contractivity of such a differential equation with a symmetric Jacobian is equivalent to the existence of a strongly convex function which induces the differential equation as its gradient flow.

1. Introduction

Contraction theory (Lohmiller & Slotine, 1998) characterizes stability of nonlinear systems in terms of how any two trajectories converge to each other. It is shown in Lohmiller and Slotine (1998) that contraction analysis leads to a necessary and sufficient characterization of exponential convergence of any pairs of trajectories. Contraction analysis, as well as closely related notions of incremental stability (Angeli, 2002) and convergent dynamics (Pavlov, Van De Wouw, & Nijmeijer, 2007), has found applications in systems and control theory for observer design (Le Ny, 2020) and controller design (Lohmiller & Slotine, 2000), as well as in the analysis of biological systems (Coogan, 2019; Wang & Slotine, 2005), networked systems (Wang & Slotine, 2006), and the design of robotic systems (chaandar Ravichandar & Dani, 2019).

Recent works (Singh, Majumdar, Slotine, & Pavone, 2017; Wensing & Slotine, 2020) have shown some interesting connections between contraction theory and the analysis of continuous-time gradient-based algorithms for convex optimization. Indeed, a μ-strongly convex function $\mathrm{V} : \mathbb{R}^d \to \mathbb{R}$ naturally induces a differential equation by

$$\dot{\phi} = -\nabla \mathrm{V}(\phi), \quad \phi(0, x) = x, \tag{1}$$

which can be shown to be μ-contractive, i.e., satisfying $\|\phi(t, x) - \phi(t, y)\| \le e^{-\mu t}\|x - y\|$, for all $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$ (cf. Nesterov, 2003, Theorem 2.1.9). This paper is a short investigation on when the converse to the previous sentence holds. That is, given a differential equation that is μ-contractive, does there always exist a μ-strongly convex function that induces the differential equation? We show that, if sufficient smoothness of the vector field is assumed outright, then μ-contractivity of the differential equation with a symmetric Jacobian is equivalent to the existence of a μ-strongly convex function which induces the differential equation.

2. Contractive differential equations and strongly convex functions

Consider the differential equation (DE):

$$\dot{\phi} = \mathrm{f}(\phi), \quad \phi(0, x) = x. \tag{2}$$

Henceforth assume that $\mathrm{f} : \mathbb{R}^d \to \mathbb{R}^d$ is L-Lipschitz on $\mathbb{R}^d$ and f has continuous partial derivatives. We say (2) is contractive if there is a $\mu > 0$ such that

$$(x - y) \cdot (\mathrm{f}(x) - \mathrm{f}(y)) \le -\mu\|x - y\|^2,$$

for all $x, y \in \mathbb{R}^d$ and $\|\cdot\|$ is the Euclidean norm. We say (2) is μ-contractive if the above holds for some specific $\mu > 0$. For convenience we present a number of useful known facts concerning contractive systems.

Proposition 1. *If the DE (2) is μ-contractive, then the following hold:*

☆ This work was partially supported by the NSERC of Canada, the CRC program, and an ERA award. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Iman Shames under the direction of Editor André L. Tits.

* Corresponding author.
E-mail addresses: mfitzsimmons@uwaterloo.ca (M. Fitzsimmons), j.liu@uwaterloo.ca (J. Liu).

https://doi.org/10.1016/j.automatica.2022.110349
0005-1098/

(1) There is a unique equilibrium point x^*.
(2) $\|\phi(t,x)-\phi(t,y)\| \leq e^{-\mu t}\|x-y\|$, for all $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$.
(3) If the unique equilibrium point $x^* = 0$, then $\|\phi(t,x)\| \leq e^{-\mu t}\|x\|$, for all $x \in \mathbb{R}^d$ and $t \in \mathbb{R}$.

Furthermore, item (2) of the above is equivalent to the DE (2) being μ-contractive.

Proof. The proof of this proposition can be largely found throughout (Lohmiller & Slotine, 1998) (see also Singh et al., 2017) in a more general setting. For completeness, we provide direct proofs below.

To prove item (1), notice that, for $L_2 \geq \max\{L, \mu\}$, the function $\mathrm{g}(x) = x + \frac{\mu}{L_2^2}\mathrm{f}(x)$ is a contraction map from $\mathbb{R}^d$ to itself. This follows from a direct algebraic manipulation and the μ-contractivity condition. Further, observe that some $\bar{x}$ is a fixed point of g if and only if $\bar{x}$ is an equilibrium point of Eq. (2).

To prove item (2), let $\Delta\phi = \phi(t,x) - \phi(t,y)$ and consider

$$\frac{d}{dt}\|\phi(t,x)-\phi(t,y)\|^2 = 2(\Delta\phi)\cdot(\mathrm{f}(\phi(t,x)) - \mathrm{f}(\phi(t,y))) \leq -2\mu\|\phi(t,x)-\phi(t,y)\|^2.$$

Applying Grönwall's inequality and taking square roots yields the result.

Item (3) is an application of item (2) with $y = x^* = 0$.

To prove the "furthermore" part, notice that $\|\phi(t,x)-\phi(t,y)\|^2 = \|x-y\|^2 + \int_0^t \frac{d}{dt}\|\phi(t,x)-\phi(t,y)\|^2 ds$ and similarly $e^{-2\mu t} = 1 + \int_0^t -2\mu e^{-2\mu s} ds$. Thus,

$$\|\phi(t,x)-\phi(t,y)\|^2 \leq \|x-y\|^2 e^{-2\mu t},$$

$$\|x-y\|^2 + \int_0^t \frac{d}{dt}\|\Delta\phi\|^2 ds \leq \|x-y\|^2 + \int_0^t -2\mu e^{-2\mu s} ds \|x-y\|^2,$$

and we get for all $t \geq 0$

$$\int_0^t 2(\Delta\phi)\cdot(\Delta\,\mathrm{f}\,\phi) + 2\mu e^{-2\mu s}\|x-y\|^2 ds \leq 0,$$

where $\Delta\,\mathrm{f}\,\phi = \mathrm{f}(\phi(s,x)) - \mathrm{f}(\phi(s,y))$. Thus the integrand of the above cannot be strictly positive for all small $s > 0$. Since the integrand is also continuous, it must be non-positive at $s = 0$. Evaluating the integrand at $s = 0$ and dividing by two yields $(x-y)\cdot(\mathrm{f}(x)-\mathrm{f}(y)) \leq -\mu\|x-y\|^2$. □

A function, $\mathrm{V} : \mathbb{R}^d \to \mathbb{R}$ is said to be strongly convex if there is a parameter $\mu > 0$ with

$$(x-y)\cdot(\nabla\mathrm{V}(x) - \nabla\mathrm{V}(y)) \geq \mu\|x-y\|^2.$$

If the previous equation holds for a particular parameter μ, then we say V is μ-strongly convex.

Strongly convex functions are desirable to work with for minimization problems (i.e., $\min_{x\in\mathbb{R}^d}\mathrm{V}(x)$), since they have a global minimum. Gradient-based algorithms for minimizing strongly convex functions often lead to a linear (geometric) convergence rate (Nesterov, 2003). If we cast the gradient descent algorithm

$$x^{(\ell+1)} = x^{(\ell)} - h\nabla\mathrm{V}(x^{(\ell)})$$

as a sequence of Euler steps of a continuous-time DE, we obtain (1). It is easy to show that, if V is μ-strongly convex, then the DE (1) is μ-contractive (see, e.g., Nesterov, 2003, Theorem 2.1.9). The remainder of this paper is devoted to proving that the converse to this statement holds. That is, if the DE (2) is μ-contractive, then there exists some μ-strongly convex function V such that $\nabla\mathrm{V} = -\mathrm{f}$.

3. Finding the potential V

We begin by making some observations on the functions V and f provided $\nabla\mathrm{V} = -\mathrm{f}$.

Proposition 2 (*Necessary Conditions on* V *and* f)**.** *If a function* V *such that* $\nabla\mathrm{V} = -\mathrm{f}$ *exists, then the following hold:*

(a) V is μ-strongly convex if and only if (2) is μ-contractive.
(b) The Jacobian of f is symmetric.
(c) If $\mathrm{V}(0) = 0$, (2) is μ-contractive and $\mathrm{f}(0) = 0$ then, $\mathrm{V} = \mathrm{U}$ where

$$\mathrm{U}(x) := \int_0^\infty \|\,\mathrm{f}(\phi(s,x))\|^2 ds$$

for all $x \in \mathbb{R}^d$.

Proof. To show (a) we can see that

$$(x-y)\cdot(\nabla\mathrm{V}(x) - \nabla\mathrm{V}(y)) = -(x-y)\cdot(\mathrm{f}(x)-\mathrm{f}(y)) \geq \mu\|x-y\|^2$$

holds for all $x, y \in \mathbb{R}^d$, in the case of V being μ-strongly convex or (2) being μ-contractive.

For (b), consider the Hessian of V, which is equal to the Jacobian of the gradient of V and also equal to the negative Jacobian of f(i.e., $H(\mathrm{V}) = -\frac{\partial \mathrm{f}}{\partial x}$). Because we assumed that f has continuous partial derivatives, this Hessian matrix must be symmetric by Clairaut's Theorem. Hence, the Jacobian of f must be symmetric if the function V exists with $\nabla\mathrm{V} = -\mathrm{f}$.

Lastly, for (c) notice

$$\int_0^t -\|\,\mathrm{f}(\phi(s,x))\|^2 ds = \int_0^t \nabla\mathrm{V}(\phi(s,x))\cdot\dot{\phi}(s,x) ds = \mathrm{V}(\phi(t,x)) - \mathrm{V}(x).$$

So

$$\int_0^t \|\,\mathrm{f}(\phi(s,x))\|^2 ds = -\mathrm{V}(\phi(t,x)) + \mathrm{V}(x).$$

Now assume that the DE (2) is μ-contractive then, the DE has a unique asymptotically stable equilibrium point in $\mathbb{R}^d$. By assumption, $\mathrm{f}(0) = 0$ and so $\lim_{t\to\infty}\phi(t,x) = 0$ for all $x \in \mathbb{R}^d$. Thus we can see that

$$\mathrm{V}(x) = \int_0^\infty \|\,\mathrm{f}(\phi(s,x))\|^2 ds + \mathrm{V}(0).$$

where we have already assumed that the additive constant $\mathrm{V}(0) = 0$. So $\mathrm{V} = \mathrm{U}$. □

Given that (2) is μ-contractive then, the DE has a unique asymptotically stable equilibrium point in $\mathbb{R}^d$. Without loss of generality, we assume for the reminder of the paper that $\mathrm{f}(0) = 0$.

We see that Proposition 2 item (c) gives us a candidate function for the potential, V. However, there is an even better known construction for the potential—given that f has symmetric Jacobian.

Proposition 3. *Suppose that,* f *has continuous partial derivatives and has symmetric Jacobian then, the function*

$$\mathrm{W}(x) := -\int_0^1 \mathrm{f}(tx)\cdot x\,dt$$

is a potential for f*, ie.* $\nabla\mathrm{W} = -\mathrm{f}$.

Proof. The Poincaré Lemma (Lang, 1999, Theorem 4.1) from differential geometry gives us $\nabla\mathrm{W} = -\mathrm{f}$. Differential geometry can seem cryptic to the uninitiated, so we provide an elementary direct proof of the desired result.

Given a function $g(x)$, let $g_{x_k} = \frac{\partial g}{\partial x_k}$, where x_k is the kth component of x. Assuming that we can take partial derivatives inside the integral (we can here, see Bartle, 2001, Theorem 12.14), we end up with

$$\frac{\partial \mathrm{W}}{\partial x_k} = -\int_0^1 \frac{\partial}{\partial x_k} \mathrm{f}(tx)\cdot x dt = \int_0^1 \frac{\partial \mathrm{f}(tx)}{\partial x_k}\cdot tx + \mathrm{f}_k(tx) dt,$$

where f_k is the kth component of f and $\frac{\partial \mathrm{f}}{\partial x_k} = (\mathrm{f}_{1x_k}, \mathrm{f}_{2x_k}, \ldots, \mathrm{f}_{dx_k})^T$ is the component wise partial derivative with respect to x_k. Note that $\frac{\partial \mathrm{f}}{\partial x_k}$ is the kth column of the Jacobian of f and that the above holds for all $k = 1, 2, \ldots, d$. By definition of matrix multiplication, $(\frac{\partial \mathrm{f}}{\partial x_1}\cdot tx, \frac{\partial \mathrm{f}}{\partial x_2}\cdot tx, \ldots, \frac{\partial \mathrm{f}}{\partial x_d}\cdot tx)^T = \frac{\partial \mathrm{f}}{\partial x}\Big|_{tx}^T tx$. But the Jacobian of f is assumed to be symmetric, this allows us to write

$$\begin{aligned}\nabla \mathrm{W}(x) &= -\int_0^1 \frac{\partial \mathrm{f}}{\partial x}\Big|_{tx} tx + \mathrm{f}(tx) dt \\ &= -\int_0^1 \frac{d}{dt}(t\,\mathrm{f}(tx))dt \\ &= -(t\,\mathrm{f}(tx)|_0^1) \\ &= -\mathrm{f}(x). \quad \square\end{aligned}$$

We note, that in the context of the above proposition, W can be thought of as the potential of the conservative force field f. Under this viewpoint, the defining integral can be written as $-\int_0^x \mathrm{f}(r)\cdot dr$. Because the force field is conservative, we can compute the aforementioned integral by integrating over the straight line from 0 to x; doing this will yield the definition for W.

We can now state and prove the main result.

Theorem 4. *Suppose that* f *has continuous partial derivatives. Then the following are equivalent:*

(1) The DE (2) is μ-contractive and $\frac{\partial \mathrm{f}}{\partial x}\big|_x = \left(\frac{\partial \mathrm{f}}{\partial x}\big|_x\right)^T$ for all $x \in \mathbb{R}^d$.
(2) The function $\mathrm{W}(x) = -\int_0^1 \mathrm{f}(tx)\cdot x dt$ is μ-strongly convex and $\nabla \mathrm{W} = -\mathrm{f}$.

Furthermore, any μ-strongly convex function, $\mathrm{V} : \mathbb{R}^d \to \mathbb{R}$, with $\nabla \mathrm{V} = -\mathrm{f}$ is equal to both W *and* U *up to an additive constant. Lastly, if (1) and therefore (2) holds then, $\mathrm{W}(x)$ is a Lyapunov function for the DE (2) which shows global exponential stability of the origin.*

Proof. (2) $\Longrightarrow$ (1): By items (a) and (b) of Proposition 2.

Assume now that (1) holds. By Proposition 3 we have that $\nabla \mathrm{W} = -\mathrm{f}$. By item (a) of Proposition 2 we have that W is μ-strongly convex.

To prove the "furthermore", if we are given a μ-strongly convex function V with $\nabla \mathrm{V} = -\mathrm{f}$ then, (1) holds by the same argument as in the (2) $\Longrightarrow$ (1) proof above. Since (1) and (2) are equivalent we know that $\nabla \mathrm{W} = -\mathrm{f}$. Thus, $\nabla \mathrm{W} = \nabla \mathrm{V}$ and, by standard calculus results, differ by a constant. Lastly, $\mathrm{W}(x) = \mathrm{U}(x)$ by item (c) of Proposition 2.

To show that W is a Lyapunov function for the DE (2) which proves global exponential stability of the origin, we must show that (see Khalil, 1992, Theorem 4.10) $k_1\|x\|^2 \le \mathrm{W}(x) \le k_2\|x\|^2$ and $\frac{d}{dt}\mathrm{W}\circ\phi(t,x) \le -k_3\|\phi(t,x)\|^2$ for some constants $k_1, k_2, k_3 > 0$. Since f is μ-contractive we have,

$$(x-y)\cdot(\mathrm{f}(x)-\mathrm{f}(y)) \le -\mu\|x-y\|^2$$

set $y = 0$ (assuming that $\mathrm{f}(0) = 0$ as always) and multiply both sides by -1,

$$-x\cdot \mathrm{f}(x) \ge \mu\|x\|^2. \tag{3}$$

From here we can tell

$$\begin{aligned}\mathrm{W}(x) &= \int_0^1 -\mathrm{f}(tx)\cdot x dt \\ &= \int_0^1 \frac{-\mathrm{f}(tx)\cdot tx}{t} dt \\ &\ge \int_0^1 \mu\|tx\|^2 t^{-1} dt = \frac{\mu\|x\|^2}{2}.\end{aligned}$$

Since f is L Lipschitz and $\mathrm{f}(0) = 0$,

$-\mathrm{f}(tx)\cdot x \le |\mathrm{f}(tx)\cdot x| \le Lt\|x\|^2$.

So $\mathrm{W}(x) \le \int_0^1 Lt\|x\|^2 dt = \frac{1}{2}L\|x\|^2$. Lastly, consider

$$\frac{d}{dt}\mathrm{W}\circ\phi(t,x) = \nabla \mathrm{W}(\phi(t,x))\cdot\dot{\phi}(t,x) = -\|\mathrm{f}(\phi(t,x))\|^2$$

Notice that by Cauchy–Schwartz inequality and Eq. (3), $\mu\|x\|^2 \le -x\cdot \mathrm{f}(x) \le \|x\|\|\mathrm{f}(x)\|$, dividing this by $\|x\|$ and squaring both sides gives us $\mu^2\|x\|^2 \le \|\mathrm{f}(x)\|^2$. Using this,

$$\frac{d}{dt}\mathrm{W}\circ\phi(t,x) = -\|\mathrm{f}(\phi(t,x))\|^2 \le -\mu^2\|\phi(t,x)\|^2.$$

This completes the proof. $\square$

It is possible to directly show that the integral function $\mathrm{U}(x) = \int_0^\infty \|\mathrm{f}(\phi(s,x))\|^2 ds$ is a potential for f.

Remark 5. As stated, the main result is clearly related to Lyapunov theory. It is somewhat rare for a general Lyapunov function, like $\mathrm{W}(x)$, to not explicitly depend on solutions to the DE. In fact, it is more typical for Lyapunov functions to involve integrating solutions, like in $\mathrm{U}(x)$. Such Lyapunov functions are called non-constructive and are primarily of theoretical significance in the context of converse Lyapunov theorems. However, sometimes a non-constructive Lyapunov function can be approximated constructively via linear programming in a way which produces a practically usable Lyapunov function, see Hafstein (2004).

In the very specific scenario of this paper, we have shown that the non-constructive $\mathrm{U}(x)$ is equal to the constructive $\mathrm{W}(x)$. This suggests modifications to $\mathrm{W}(x)$ may provide inspiration for another algorithm which provides constructive Lyapunov functions.

Acknowledgments

This work was partially supported by the NSERC of Canada, the CRC program, and an ERA award.

References

Angeli, David (2002). A Lyapunov approach to incremental stability properties. *IEEE Transactions on Automatic Control*, *47*(3), 410–421.
Bartle, Robert Gardner (2001). *A modern theory of integration, Vol. 32*. American Mathematical Soc..
chaandar Ravichandar, Harish, & Dani, Ashwin (2019). Learning position and orientation dynamics from demonstrations via contraction analysis. *Autonomous Robots*, *43*(4), 897–912.
Coogan, Samuel (2019). A contractive approach to separable Lyapunov functions for monotone systems. *Automatica*, *106*, 349–357.
Hafstein, Sigurdur Freyr (2004). A constructive converse Lyapunov theorem on exponential stability. *Discrete & Continuous Dynamical Systems*, *10*(3), 657–678.
Khalil, H. K. (1992). *Nonlinear systems* (3rd ed.). Macmillan Publishing Company.
Lang, Serge (1999). *Graduate texts in mathematics*: *vol. 191*, *Fundamentals of differential geometry* (1st ed. 1999). (p. 137). New York, NY: Springer New York.
Le Ny, Jerome (2020). Differentially private nonlinear observer design using contraction analysis. *International Journal of Robust and Nonlinear Control*, *30*(11), 4225–4243.

Lohmiller, Winfried, & Slotine, Jean-Jacques E. (1998). On contraction analysis for non-linear systems. *Automatica*, *34*(6), 683–696.
Lohmiller, Winfried, & Slotine, Jean-Jacques E. (2000). Control system design for mechanical systems using contraction theory. *IEEE Transactions on Automatic Control*, *45*(5), 984–989.
Nesterov, Yurii (2003). *Introductory lectures on convex optimization: A basic course, Vol. 87*. Springer Science & Business Media.
Pavlov, A., Van De Wouw, Nathan, & Nijmeijer, Henk (2007). Global nonlinear output regulation: convergence-based controller design. *Automatica*, *43*(3), 456–463.
Singh, Sumeet, Majumdar, Anirudha, Slotine, Jean-Jacques E., & Pavone, Marco (2017). Robust online motion planning via contraction theory and convex optimization. In *Proc. of ICRA* (pp. 5883–5890). IEEE.
Wang, Wei, & Slotine, Jean-Jacques E. (2005). On partial contraction analysis for coupled nonlinear oscillators. *Biological Cybernetics*, *92*(1), 38–53.
Wang, Wei, & Slotine, Jean-Jacques E. (2006). Contraction analysis of time-delayed communications and group cooperation. *IEEE Transactions on Automatic Control*, *51*(4), 712–717.
Wensing, Patrick M., & Slotine, Jean-Jacques E. (2020). Beyond convexity—Contraction and global convergence of gradient descent. *Plos One*, *15*(8), Article e0236661.