Available online at www.sciencedirect.com
SCIENCE DIRECT®
Systems & Control Letters 55 (2006) 573–577
SYSTEMS & CONTROL LETTERS
www.elsevier.com/locate/sysconle

On the stable equilibrium points of gradient systems☆

P.-A. Absil[a,b], K. Kurdyka[c,*]

[a]*Département d'ingénierie mathématique, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*
[b]*Peterhouse, University of Cambridge, Cambridge CB2 1RD, UK*
[c]*Laboratoire de Mathematiques (LAMA), Université de Savoie et CNRS UMR 5127, 73-376 Le Bourget-du-Lac, Cedex, France*

Received 7 May 2005; received in revised form 6 January 2006; accepted 14 January 2006
Available online 2 March 2006

Abstract

This paper studies the relations between the local minima of a cost function f and the stable equilibria of the gradient descent flow of f. In particular, it is shown that, under the assumption that f is real analytic, local minimality is necessary and sufficient for stability. Under the weaker assumption that f is indefinitely continuously differentiable, local minimality is neither necessary nor sufficient for stability.

Keywords: Gradient flow; Lyapunov stability; Cost function; Local minimum

1. Introduction

Gradient flows are useful in solving various optimization-related problems. Recent examples deal with principal component analysis [21,15], optimal control [20,9], balanced realizations [7], ocean sampling [3], noise reduction [16], pose estimation [4] or the Procrustes problem [18]. The underlying idea is that the gradient-descent flow will converge to a local minimum of the cost function. It is, however, well known that this property does not hold in general: the initial condition can, e.g. belong to the stable manifold of a saddle point. Not as well known is the fact that, even assuming that the cost function is a C^∞ function, the local minima of the cost function are not necessarily stable equilibria of the gradient-descent system, and vice versa. The main purpose of this paper is to shed some light on this issue.

Specifically, let f be a real, continuously differentiable function on $\mathbb{R}^n$ and consider the continuous-time gradient-descent system

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

where $\nabla f(x)$ denotes the Euclidean gradient of f at x. Define stability and minimality in the standard way:

Definition 1. A point $z \in \mathbb{R}^n$ is a *local minimum* of f if there exists $\varepsilon > 0$ such that $f(x) \geqslant f(z)$ for all x such that $\|x-z\| < \varepsilon$. If $f(x) > f(z)$ for all x such that $0 < \|x-z\| < \varepsilon$, then z is a *strict local minimum* of f. An equilibrium point z of (1) is (*Lyapunov*) *stable* if, for each $\varepsilon > 0$, there is $\delta = \delta(\varepsilon) > 0$ such that

$$\|x(0) - z\| < \delta \quad \Rightarrow \quad \|x(t) - z\| < \varepsilon \quad \forall t \geqslant 0.$$

It is *asymptotically stable* if it is stable, and δ can be chosen such that $\|x(0)\| < \delta \Rightarrow \lim_{t\to\infty} x(t) = z$.

Then we have:

Proposition 2. (i) *There exist a function* $f \in C^\infty$ *and a point* $z \in \mathbb{R}^n$ *such that* z *is a local minimum of* f *and* z *is not a stable equilibrium point of* (1). (ii) *There exist a function* $f \in C^\infty$

☆ This work was initiated while the first author was a Research Fellow with the Belgian National Fund for Scientific Research (Aspirant du F.N.R.S.) at the University of Liège. This work was supported in part by the School of Computational Science of Florida State University through a postdoctoral fellowship, and by Microsoft Research through a Research Fellowship at Peterhouse, Cambridge.
* Corresponding author.
E-mail address: Krzysztof.Kurdyka@univ-savoie.fr (K. Kurdyka)
URL: http://www.inma.ucl.ac.be/∼absil (P.-A. Absil).

0167-6911/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.sysconle.2006.01.002

and a point $z \in \mathbb{R}^n$ such that z is not a local minimum of f and z is a stable equilibrium point of (1).

The proof given in Section 2 consists in producing functions f that satisfy the required properties.

After smoothness, the next stronger condition one may think of imposing on the cost function f is real analyticity (a real function is *analytic* if it possesses derivatives of all orders and agrees with its Taylor series in the neighbourhood of every point). The main result of this paper is that under the analyticity assumption, local minimality becomes a necessary and sufficient condition for stability.

Theorem 3 (*Main result*). *Let f be real analytic in a neighbourhood of $z \in \mathbb{R}^n$. Then, z is a stable equilibrium point of* (1) *if and only if it is a local minimum of f.*

The proof of this theorem, given in Section 3, relies on an inequality by Łojasiewicz that yields bounds on the length of solution curves of the gradient system (1).

Moreover, we give in Section 4 a complete characterization of the relations between (isolated, strict) local minima and (asymptotically) stable equilibria for gradient flows of both C^∞ and analytic cost functions. Final remarks are presented in Section 5.

2. Smooth cost function

In this section we prove Proposition 2. Consider $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x, y) = \frac{1}{1+x^2} g(y) h(y), \tag{2}$$

where

$$g(y) = \begin{cases} e^{-1/y^2} & \text{if } y \neq 0, \\ 0 & \text{if } y = 0, \end{cases} \tag{3}$$

and

$$h(y) = \begin{cases} y^2 + 1 + \sin \dfrac{1}{y^2} & \text{if } y \neq 0, \\ 1 & \text{if } y = 0. \end{cases}$$

This function is qualitatively illustrated in Fig. 1. We show that this function f satisfies the properties of point (i) of Proposition 2 with $z = (0, 0)$. It is routine to check that $f \in C^\infty$, and it is clear that the origin is a local minimum of f, since f is nonnegative and $f(0) = 0$. The gradient system (1) becomes

$$\dot{x} = \frac{2x}{(1+x^2)^2} g(y) h(y), \tag{4a}$$

$$\dot{y} = -\frac{1}{1+x^2} \frac{g(y)}{y^3} m(y), \tag{4b}$$

where $m(y) = 1 + \sin 1/y^2 - 2\cos 1/y^2 + y^2 + 2y^4$. Let $(x(t), y(t))$ be the solution trajectory of (4) with initial conditions $(x(0), y(0)) = (x_0, y_0)$ where we pick $y_0 > 0$

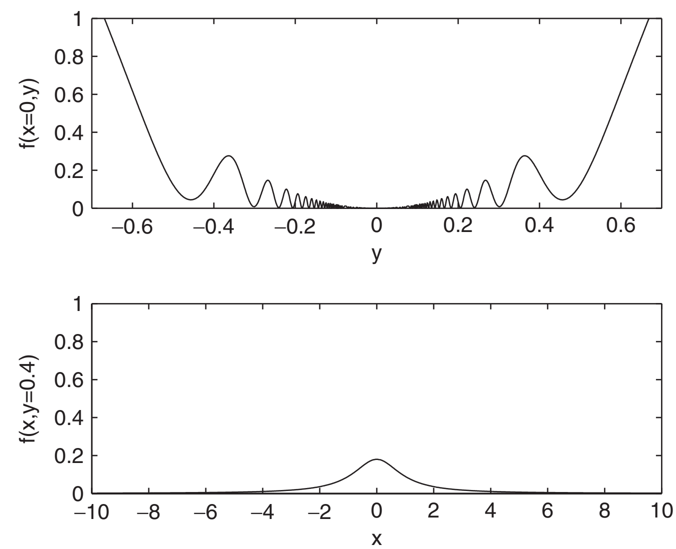

Fig. 1. Plots of $f(x, y)$ along the line $x = 0$ (above) and $y = 0.4$ (below). The function f is the one in (2), where $g(y)$ has been replaced by y^2 for clarity of the illustration.

and $x_0 > 0$. Then there exists y_1 such that $0 < y_1 < y_0$ and $m(y_1) = 0$. Therefore, $y(t) > y_1$ for all t. Then from (4a), $\dot{x} > (2x/(1+x^2)^2) g(y_1) y_1^2$ whence $\lim_{t \to +\infty} x(t) = +\infty$. We have shown that from an initial point arbitrarily close to the origin the solution of (4) escapes to infinity. That is, the origin is not a stable equilibrium point of (4).

Point (ii) of Proposition 2 is easier to show. Take $f : \mathbb{R} \to \mathbb{R}$ given by

$$f(x) = \begin{cases} g(x) \sin \dfrac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases} \tag{5}$$

where the function g is given by (3). This function f has (infinitely many) local minima in any neighbourhood of $x = 0$. Since solution trajectories of (1) are bounded by the local minima, it follows that $x = 0$ is a Lyapunov stable point of (1); yet $x = 0$ is not a local minimum of f.

Notice that both functions f defined in (2) and (5) are nonanalytic at the origin. This is not coincidental in view of Theorem 3 which we prove in the next section.

3. Analytic cost function

This section is dedicated to proving Theorem 3. We assume throughout, without loss of generality, that $f : \mathbb{R}^n \to \mathbb{R}$ is analytic on an open set U containing the origin, that $f(0) = 0$ and that $\nabla f(0) = 0$, and we study the stability of the equilibrium point 0 of the gradient system (1).

The proof relies on the following fundamental property of analytic functions.

Lemma 4 (*Lojasiewicz's inequality*). *Let f be a real analytic function on a neighbourhood of z in $\mathbb{R}^n$. Then there are*