

On the stable equilibrium points of gradient systems

P.-A. Absil^{*†} K. Kurdyka[‡]

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.

Key words. Gradient flow, Lyapunov stability, cost function, local minimum.

AMS subject classifications. 93D05 (Systems theory - Stability - Lyapunov), 37N40 (Dynamical systems and ergodic theory - Applications - Dynamical systems in optimization and economics), 37B25 (- Topological dynamics - Lyapunov functions and stability; attractors, repellers), 34D20 (ODE - Stability theory - Lyapunov stability).

1 Introduction

Gradient flows are useful in solving various optimization-related problems. Recent examples deal with principal component analysis [YHS01], optimal control [YTM94, JM96], balanced realizations [HM94], ocean sampling [BL02],

^{*}School of Computational Science, Florida State University, Tallahassee, FL 32306-4120, U.S.A. (www.csit.fsu.edu/~absil).

[‡]Laboratoire de Mathématiques (LAMA), Université de Savoie et CNRS UMR 5127, 73-376 Le Bourget-du-Lac cedex, France (Krzysztof.Kurdyka@univ-savoie.fr).

[†]Part of this work was done 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.

noise reduction [RJ02], pose estimation [BHM94] or the Procrustes problem [TL02]. 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

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

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 $\epsilon > 0$ such that $f(x) \geq f(z)$ for all x such that $\|x - z\| < \epsilon$. If $f(x) > f(z)$ for all x such that $0 < \|x - z\| < \epsilon$, then z is a strict local minimum of f . An equilibrium point z of (1) is (Lyapunov) stable if, for each $\epsilon > 0$, there is $\delta = \delta(\epsilon) > 0$ such that*

$$\|x(0) - z\| < \delta \Rightarrow \|x(t) - z\| < \epsilon, \quad \forall t \geq 0.$$

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

Then we have:

Proposition 2 *(i) There exists 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 exists a function $f \in C^\infty$ 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 neighborhood 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 $f : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

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

where

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

and

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

This function is qualitatively illustrated on Figure 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

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

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

where $m(y) = 1 + \sin \frac{1}{y^2} - 2 \cos \frac{1}{y^2} + y^2 + 2y^4$. 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 $y_0 > 0$ 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), $\dot{x} > \frac{2x}{(1+x^2)^2} g(y_1) y_1^2$ whence

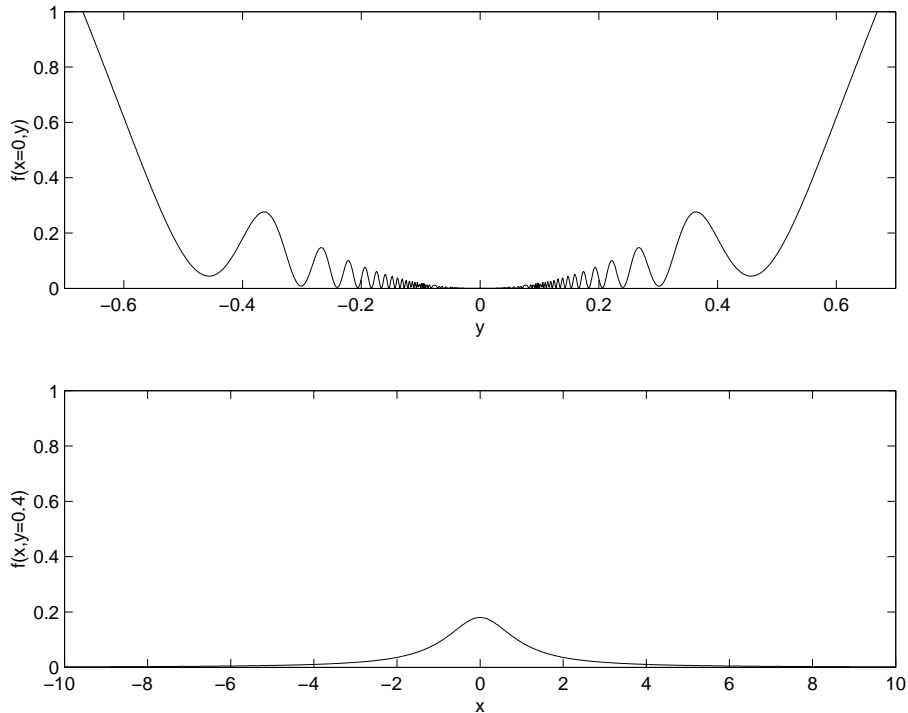


Figure 1: Plots of $f(x, y)$ along the line $x = 0$ (above) and $y = 0.4$ (below). The function f is the one in (2), where $g(y)$ has been replaced by y^2 for clarity of the illustration.

$\lim_{t \rightarrow +\infty} x(t) = +\infty$. 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 $f : \mathbb{R} \rightarrow \mathbb{R}$ given by

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

where the function g is given by (3). This function f is illustrated on Figure 2 and straightforwardly verifies the properties in point (ii) of Proposition 2.

Notice that both functions f defined in (2) and (5) are nonanalytic at $(0, 0)$. This is not coincidental in view of Theorem 3 which we prove in the next section.

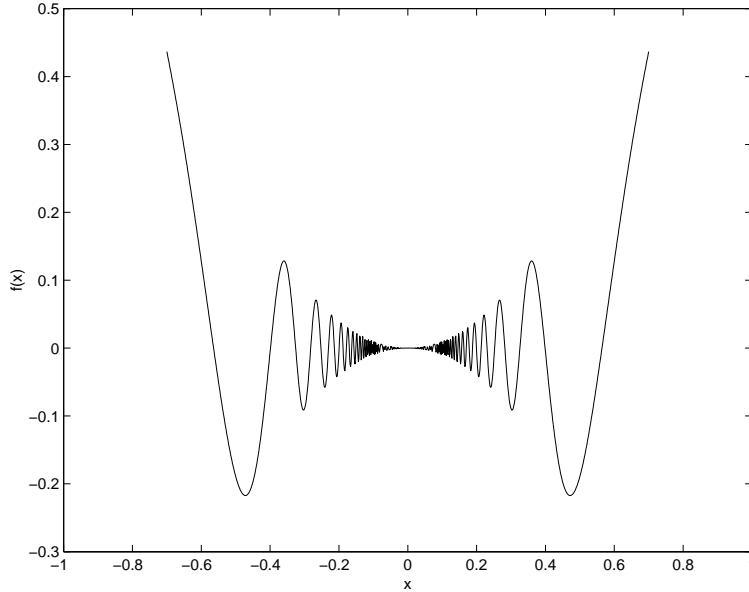


Figure 2: Plot of $f(x)$ defined in (5), where $g(y)$ has been replaced by y^2 for clarity of the illustration.

3 Analytic cost function

This section is dedicated to proving Theorem 3.

We assume throughout, without loss of generality, that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is analytic on an open set U containing the origin, that $f(0) = 0$ and that $\nabla f(0) = 0$, and we will 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 (Łojasiewicz’s inequality) *Let f be a real analytic function on a neighbourhood of z in \mathbb{R}^n . Then there are constants c, ρ such that $0 \leq \rho < 1$ and*

$$\|\nabla f(x)\| \geq c|f(x) - f(z)|^\rho$$

in some neighbourhood of z .

Proof. See [Loj65], [BM88], or the short proof in [KP94]. \square

We first prove the “if” part of Theorem 3, i.e., we assume that the origin is a local minimum of f and we show that the origin is a stable equilibrium

point (Definition 1) of the gradient system (1). The rationale is based on Łojasiewicz's argument [Loj84] which provides a bound on the length of the trajectories of (1).

Since the origin is a local minimum of f , it follows that there exists a neighbourhood U_m of 0 contained in U such that $f(x) \geq 0$ for all $x \in U_m$. Let x_0 be in U_m and let $x(t)$ be the solution trajectory of the gradient system (1) with initial condition $x(0) = x_0$. We shall parameterize $x(t)$ by its arc-length s starting from x_0 . By Łojasiewicz's inequality around the origin, in a neighbourhood U_L of the origin

$$\|\nabla f\| \geq c|f|^\rho$$

for some $\rho < 1$ and $c > 0$. Thus in U_L we have on the trajectory $x(s)$

$$\frac{df}{ds} = \langle \nabla f, \frac{dx}{ds} \rangle = \langle \nabla f, -\frac{\nabla f}{\|\nabla f\|} \rangle = -\|\nabla f\| \leq -c|f|^\rho.$$

In particular $f(x(s))$ is decreasing and in $U_m \cap U_L$

$$\frac{df^{1-\rho}}{ds} \leq -c(1-\rho) < 0.$$

By integration of this inequality, if $x(s)$ lies in $U_m \cap U_L$ for $s \in [s_1, s_2]$ then the length of the segment of curve between s_1 and s_2 is bounded by

$$c_1(f(x(s_1)))^{1-\rho} - f(x(s_2))^{1-\rho} \leq c_1(f(x(s_1)))^{1-\rho}$$

where $c_1 = (c(1-\rho))^{-1}$. Now let B_ϵ be a ball of radius $\epsilon > 0$ centered on the origin such that B_ϵ is in $U_m \cap U_L$. By continuity of f , there exists $\delta < \epsilon/2$ such that $f(x) < (\epsilon/2c_1)^{1/(1-\rho)}$ for all $x \in B_\delta$. If x_0 belongs to B_δ then the length of the trajectory $x(s)$ inside B_ϵ is smaller than $c_1(f(x_0))^{1-\rho}$, which is smaller than $\epsilon/2$. But since $\delta < \epsilon/2$ the distance between x_0 and the boundary of B_ϵ is greater than $\epsilon/2$. Thus $x(t)$ remains in B_ϵ for all t , and this is Lyapunov stability. We have proven that minimality is sufficient for stability.

The proof of the “only if” part of Theorem 3 uses the following classical convergence result of gradient systems.

Lemma 5 *Assume that f is a C^2 function and let $x(t)$ be a solution trajectory of the gradient system (1) contained in a compact set $K \subset \mathbb{R}^n$. Then $x(t)$ approaches the critical set $C_K = \{y \in K : \nabla f(y) = 0\}$ as t approaches infinity, that is, $\lim_{t \rightarrow +\infty} \inf_{y \in C_K} \|x(t) - y\| = 0$.*

Proof. The limit points of the trajectories of gradient flows are stationary points [HS74, Thm 9.4.4]. If the solution is bounded then it approaches its positive limit set [Kha96, Lemma 3.1]. \square

We suppose that the origin is not a local minimum of f and we show that the origin is not a stable equilibrium point of the gradient system (1). Since we have assumed that $f(0) = 0$ and $\nabla f(0) = 0$, it directly follows from Łojasiewicz's inequality that there exist an $\epsilon > 0$ such that f is zero on the set $\{y \in \bar{B}_\epsilon : \nabla f(y) = 0\}$, where \bar{B}_ϵ denotes the closed ball of radius ϵ centered on the origin. Since the origin is not a local minimum of f , it follows that for all $\delta > 0$, there exists x_0 in B_δ with $f(x_0) < f(0) = 0$. Then we show that the solution trajectory starting from x_0 leaves \bar{B}_ϵ and the proof that minimality is necessary for stability will be complete. Suppose for contradiction that the solution trajectory of (1) starting from x_0 stays in \bar{B}_ϵ for all t . Then, by Lemma 5, $x(t)$ approaches the set $\{y \in \bar{B}_\epsilon : \nabla f(y) = 0\}$ as t approaches infinity. By continuity of f , it follows that $f(x(t))$ approaches 0. But $f(x(t)) \leq f(x(0)) < 0$ for all $t > 0$, a contradiction.

Note that the “only if” portion of the proof did not make use of Łojasiewicz's inequality in its full strength. Rather, it only used the following consequence: in some neighbourhood of z , if $f(x) \neq f(z)$ then $\nabla f(x) \neq 0$. This result also follows from the fact that analytic functions admit a finite number of critical values on any compact set. This weaker result will again be sufficient for proving the results discussed in Section 4.

4 Strict minimality and asymptotic stability

The previous results were concerned with (simple) Lyapunov stability and (nonstrict) minimality. In this section, we also consider asymptotic stability and strict minimality. A characterization of the relations between Lyapunov stability, asymptotic stability, and various notions of minimality, is displayed in Figure 3; see the caption of the figure for details.

In the remainder of this section, we briefly review the relations displayed on Figure 3; see the caption of Figure 3 for the meaning of the initials used below. As before, we assume that $f(0) = 0$ and $\nabla f(0) = 0$, and we study the stability of the equilibrium point 0 of the gradient system (1).

We first consider the case $f \in C^\infty$. LMICP \Rightarrow ASE follows from Lyapunov's stability theorem (see [Kha96, Thm. 3.1]). To show ILM \nRightarrow ASE, take $n = 1$ and consider the function $f(x) = \int_0^{|x|} g(\xi)(1 + \sin(1/\xi^2))d\xi$, g as in (3), which exists since the integrand is bounded; any neighbourhood

of the origin contains points where ∇f vanishes, hence the origin is not asymptotically stable¹. $\text{SLM} \Rightarrow \text{SE}$ also follows from Lyapunov's stability theorem (see [Kha96, Thm. 3.1]). $\text{LM} \not\Rightarrow \text{SE}$ is point (i) of Proposition 2. $\text{ASE} \Rightarrow \text{LMICP}$: if the origin is not a LMICP, then either it is not a LM, or it is not an isolated critical point; in both cases, the origin is not an ASE. $\text{SE} \not\Rightarrow \text{LM}$ is point (ii) of Proposition 2.

Now we consider the case where f is real analytic. Then, as can be directly shown from Łojasiewicz's inequality, $\text{SLM} \Rightarrow \text{LMICP}$, and therefore the three properties LMICP, ILM and SLM are equivalent. Since analytic functions are C^∞ , the relations that hold for the C^∞ case remain valid, hence the three properties are also equivalent to ASE. Finally, the relation $\text{LM} \Leftrightarrow \text{SE}$ corresponds to Theorem 3.

5 Final remarks

For a general cost function $f \in C^p$, $p \in \{2, 3, \dots\} \cup \{\infty\}$, the classical way of studying the stability of an equilibrium point (say $x = 0$) of the gradient descent flow (1) is to consider the Hessian of f at $x = 0$. If the Hessian is positive definite, then $x = 0$ is a local minimum and an isolated critical point of f (LCICP in Figure 3); it follows from Figure 3 that the origin is asymptotically stable. But the converse is not true, as the simple example $f(x) = x^4$ shows (the origin is asymptotically stable but the Hessian vanishes). What is more, we have shown that local minima of f are not necessarily stable minima of the gradient system (1), and vice versa. However, the main result of this paper (Theorem 3) ensures that if f is real analytic then the stable points of the gradient descent flow (1) and the local minima of f coincide; the same holds for strict local minima and asymptotically stable equilibria. We refer to Figure 3 for a characterization of the relations between various notions of minimality and stability.

This paper is just one step towards understanding the behaviour of gradient flows. Previous advances include: the result by Łojasiewicz [Loj84] that the trajectories of gradient flows cannot have more than one limit point; the proof by Kurdyka *et al.* [KMP00] of the gradient conjecture of R. Thom stating that the limit of secants exists (see illustration on Figure 4); and the universal bounds for gradient trajectories of polynomial and definable functions given by D'Acunto and Kurdyka [DK04].

¹Note in this respect that the corollary in section 9.4 of [HS74] is subject to possible misinterpretation: the minimum must be isolated *as a critical point* to guarantee asymptotic stability.

These results, including our Figure 3, remain valid when the Euclidean metric is replaced by a (nondegenerate) Riemannian metric, i.e., when (1) becomes

$$\dot{x}_i = - \sum_{j=1}^n Q_{ij}(x) \partial_j f(x)$$

where $Q(x)$ is a smooth symmetric positive-definite matrix function.

Several questions remain open, for instance concerning the existence of $\lim_{t \rightarrow +\infty} \dot{x}/\|\dot{x}\|$, the eventual monotonicity of $\|x(t) - x_\infty\|$, and the case of degenerate Riemannian metrics which has a particular importance for inequality-constrained optimization problems (see, e.g., [AS04]).

Finally, note that the relation between minimality and stability for the system $\ddot{x}(t) = -\nabla f(x(t))$ —which differs from (1) by the double dot on $x(t)$ —is a classical problem in mechanics. By the Lagrange-Dirichlet theorem, the equilibrium position is stable if the potential f has a strict local minimum at this position. But (nonstrict) minimality is not sufficient for stability [Wie92]. Moreover, the converse of the Lagrange-Dirichlet is not true, but several authors have proposed additional constraints, apart from the absence of a potential minimum, that make the equilibrium unstable; see [RS94] and references therein.

Acknowledgements

The authors thank A. L. Tits and R. Sepulchre for several useful comments. Part of this work was done while the first author was visiting the second author at the Laboratoire de Mathématiques de l'Université de Savoie. The hospitality of the members of the Laboratory is gratefully acknowledged.

References

- [AS04] P.-A. Absil and R. Sepulchre, *Continuous dynamical systems that realize discrete optimization on the hypercube*, Systems Control Lett. **52** (2004), no. 3-4, 297–304. MR MR2064925 (2005b:90075)
- [BHM94] M. Baeg, U. Helmke, and J. B. Moore, *Gradient flow techniques for pose estimation of quadratic surfaces*, Proc. of the World Congress in Computational Methods and Applied Mathematics, 1994.

- [BL02] R. Bachmayer and N. E. Leonard, *Vehicle networks for gradient descent in a sampled environment*, Proc. 41st IEEE Conf. Decision and Control, 2002.
- [BM88] E. Bierstone and P. D. Milman, *Semianalytic and subanalytic sets*, Inst. Hautes études Sci. Publ. Math **67** (1988), 5–42.
- [DK04] Didier D’Acunzio and Krzysztof Kurdyka, *Bounds for gradient trajectories of polynomial and definable functions with applications*, submitted, 2004.
- [HM94] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, Springer, 1994.
- [HS74] M. W. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra*, Academic Press, 1974.
- [JM96] Danchi Jiang and J. B. Moore, *A gradient flow approach to decentralised output feedback optimal control*, Systems Control Lett. **27** (1996), no. 4, 223–231. MR 97a:93006
- [Kha96] H. K. Khalil, *Nonlinear systems, second edition*, Prentice Hall, 1996.
- [KMP00] K. Kurdyka, T. Mostowski, and A. Parusiński, *Proof of the gradient conjecture of R. Thom*, Ann. of Math. (2) **152** (2000), no. 3, 763–792.
- [KP94] Krzysztof Kurdyka and Adam Parusiński, *\mathbf{w}_f -stratification of subanalytic functions and the Lojasiewicz inequality*, C. R. Acad. Sci. Paris Sér. I Math. **318** (1994), no. 2, 129–133. MR MR1260324 (95d:32012)
- [Loj65] S. Lojasiewicz, *Ensembles semi-analytiques*, Inst. Hautes Études Sci., Bures-sur-Yvette, 1965.
- [Loj84] ———, *Sur les trajectoires du gradient d’une fonction analytique*, Seminari di Geometria 1982-1983 (Università di Bologna, Istituto di Geometria, Dipartimento di Matematica), 1984, pp. 115–117.
- [RJ02] D. Ridout and K. Judd, *Convergence properties of gradient descent noise reduction*, Physica D **165** (2002), 26–47.

- [RS94] V. V. Rumyantsev and S. P. Sosnitskiĭ, *On the instability of the equilibrium of holonomic conservative systems*, J. Appl. Math. Mech. **57** (1994), no. 6, 1101–1122.
- [TL02] N. T. Trendafilov and R. A. Lippert, *The multimode Procrustes problem*, Linear Algebra Appl. **349** (2002), 245–264.
- [Wie92] Z. Wiener, *Instability of a non-isolated equilibrium*, Arch. Rational Mech. Anal. **116** (1992), no. 4, 301–305.
- [YHS01] S. Yoshizawa, U. Helmke, and K. Starkov, *Convergence analysis for principal component flows*, Int. J. Appl. Math. Comput. Sci. **11** (2001), no. 1, 223–236.
- [YTM94] Wei Yong Yan, Kok L. Teo, and John B. Moore, *A gradient flow approach to computing LQ optimal output feedback gains*, Optimal Control Appl. Methods **15** (1994), no. 1, 67–75. MR 1 263 418

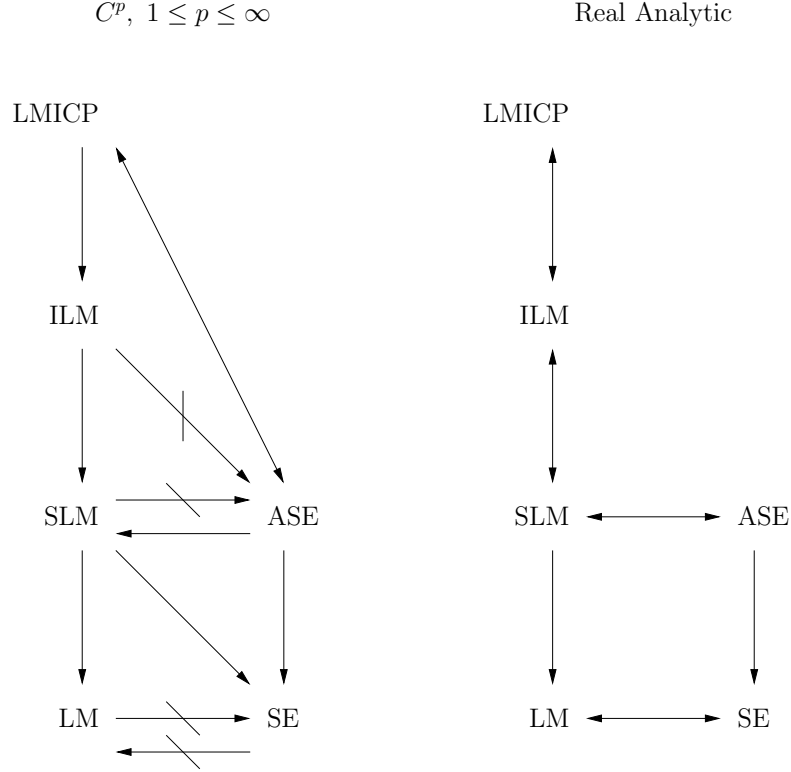


Figure 3: Relations between minimality properties of a critical point z of a cost function f and its stability as an equilibrium point of the gradient descent system (1) for f . LM: local minimum; SLM: strict local minimum; ILM: isolated local minimum; LMICP: local minimum and isolated critical point; SE: stable equilibrium; ASE: asymptotically stable equilibrium (see Definition 1 for details). The left-hand graph holds under the assumption that $f \in C^p$ with $p \in \{1, 2, \dots\} \cup \{\infty\}$ and the right-hand graph holds for f real analytic. A relation $A \rightarrow B$ means that property A implies property B, and $A \nrightarrow B$ means that property A is not sufficient for property B. All the relations that cannot be deduced by transitivity are \nrightarrow .

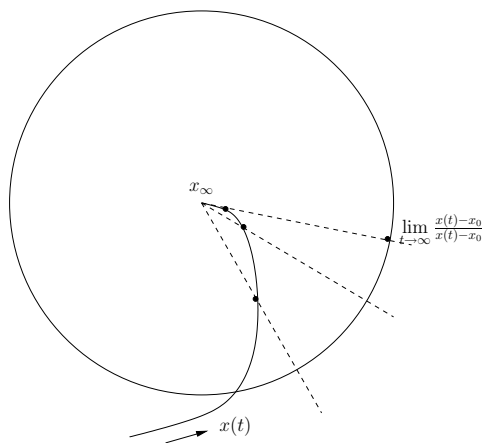


Figure 4: Illustration of the gradient conjecture of R. Thom.