Journal of Global Optimization (2005) 31: 85–92

On Absorbing Cycles in Min–Max Digraphs

HARALD GÜNZEL and HUBERTUS TH. JONGEN
Department of Mathematics (C), RWTH Aachen University, D-52056 Aachen, Germany

(Received 14 May 2002; accepted in revised form 15 April 2004)

Abstract. We consider smooth finite dimensional optimization problems with a compact, connected feasible set M and objective function f. The basic problem, on which we focus, is: how to get from one local minimum to all the other ones. To this aim we introduce a bipartite digraph Γ as follows. Its nodes are formed by the set of local minima and maxima of $f|_M$, respectively. Given a smooth Riemannian (i.e. variable) metric, there is an arc from a local minimum x to a local maximum y if the ascent (semi-)flow induced by the projected gradients of f connects points from a neighborhood of x with points from a neighborhood of y. The existence of an arc from y to x is defined with the aid of the descent (semi-)flow. Strong connectedness of Γ ensures that, starting from one local minimum, we may reach any other one using ascent and descent trajectories in an alternating way. In case that no inequality constraints are present or active, it is well known that for a generic Riemannian metric the resulting min-max digraph Γ is indeed strongly connected. However, if inequality constraints are active, then there might appear obstructions. In fact, we show that Γ may contain absorbing two-cycles. If one enters such a cycle, one cannot leave it anymore via ascent and descent trajectories. Moreover, the cycles being constructed are stable with respect to small perturbations (in the C^1-topology) of the Riemannian metric.

Key words: Absorbing cycle, Ascent flow, Descent flow, Global optimization, Min–max digraph, Projected gradient

1. Introduction and Motivation

In this paper we consider smooth finite dimensional optimization problems of the type
$(\mathcal{P})$ Minimize f on the feasible set M, where

$$M := \{x \in \mathbb{R}^n \mid h_i(x) = 0,\ i \in I,\ g_j(x) \geqslant 0,\ j \in J\},$$

and where $f, h_i, g_j : \mathbb{R}^n \to \mathbb{R}$ are smooth (i.e. of class C^∞), $|I| < n$, $|J| < \infty$.

We are concerned with the following basic problem from global optimization: how to get from one local minimum to all the other ones using ascent and descent methods? As main references we use [3, 4]. Throughout the paper we make the following assumption:

ASSUMPTION 1. The feasible set M is non-empty, compact and connected, and the linear independence constraint qualification (LICQ) is satisfied at all points of M. Moreover, all Karush–Kuhn–Tucker points (KKT-points) for $f|_M$ and for $(-f)|_M$ are nondegenerate and they have pairwise different f-values.

LICQ is said to hold at $\bar{x} \in M$ if the vectors $Dh_i(\bar{x})$, $i \in I$, $Dg_j(\bar{x})$, $j \in J_0(\bar{x})$, are linearly independent. Here Dh stands for the row vector of partial derivatives of h and $J_0(\bar{x})$ denotes the set of active inequality constraints, i.e. $J_0(\bar{x}) = \{j \in J \,|\, g_j(\bar{x}) = 0\}$. Under LICQ, a KKT-point $\bar{x}$ for $f|_M$ is said to be nondegenerate if strict complementarity holds and the restriction of the Hessian of the Lagrangian on the tangent space $T_{\bar{x}}M$ is nonsingular, where $T_{\bar{x}}M := \{\xi \in \mathbb{R}^n \,|\, Dh_i(\bar{x})\xi = 0, i \in I, Dg_j(\bar{x}) = 0, j \in J_0(\bar{x})\}$. The number of negative eigenvalues of the latter restriction is called the (quadratic) index of the KKT-point $\bar{x}$.

The following idea from Morse theory plays a key role in global optimization. Let M^a denote the lower level set $\{x \in M \,|\, f(x) \leqslant a\}$. As the level increases, the number of connected components of M^a, say $\sharp$, can only change in the following two cases. Either we pass a level corresponding to a local minimum. In that case, the number $\sharp$ increases by one. Or we pass a level corresponding to a *special* type of KKT-point of index 1, called *decomposition point*. In the latter case, the number $\sharp$ decreases by one. Not all KKT-points of index 1 are decomposition points; only a part of them. Since the feasible set M is *connected*, the number of decomposition points is one less than the number of local minima.

If we were able to detect all of those decomposition points or, more generally, all KKT-points of index 1, then the problem of global optimization would be settled. In fact, descending from those points in the two canonical opposite directions of quadratic descent would lead us to all local minima. However, finding KKT-points of index 1 is extremely difficult (cf. [2, 4]). On the other hand, raising the complete lower level set will be computationally intractable from several points of view. For example, KKT-points of higher index may cause a complicated topology in the lower level sets (cf. [4]).

Therefore, it might be more effective trying to jump between local minima via local maxima. This could be done with the aid of certain trajectories of ascent and descent flows, respectively. In fact, given a smooth variable (i.e. Riemannian) metric $\mathcal{R}$ on $\mathbb{R}^n$, we can form the gradient field $\mathrm{grad}_{\mathcal{R}}f$ of f with respect to $\mathcal{R}$. Then at each point $\bar{x} \in M$ we project $\mathrm{grad}_{\mathcal{R}}f$ to the tangent cone $C_{\bar{x}}M$ (with respect to the metric $\mathcal{R}$), where $C_{\bar{x}}M = \{\xi \in \mathbb{R}^n \,|\, Dh_i(\bar{x}) = 0, i \in I, Dg_j(\bar{x})\xi \geqslant 0, j \in J_0(\bar{x})\}$. The resulting projected gradient vector field on M generates a *semi-flow* on M, called the *ascent* (*semi-*)*flow*. Replacing f by $-f$ we obtain the corresponding *descent* (*semi-*)*flow*. The possible transitions from local minima (maxima) to local maxima (minima) via the latter semi-flows are *coded* by means of a *directed* *bipartite graph* $\Gamma(f, \mathcal{R})$, depending on the objective function f and the Riemannian metric $\mathcal{R}$.

The nodes of $\Gamma(f, \mathcal{R})$ are formed by the set of local minima of $f|_M$, say $\{\bar{x}_1, \ldots, \bar{x}_p\}$ and the set the local maxima of $f|_M$, say $\{\bar{y}_1, \ldots, \bar{y}_q\}$. Choose arbitrarily small neighborhoods (germs) $U_{\bar{x}_1}, \ldots, U_{\bar{y}_q}$ of $\bar{x}_1, \ldots, \bar{y}_q$ in M. Then there exists an arc from $\bar{x}_i$ to $\bar{y}_j$ (from $\bar{y}_j$ to $\bar{x}_i$) if the ascent (descent) flow connects some point of $U_{\bar{x}_i}$ ($U_{\bar{y}_j}$) with a point from $U_{\bar{y}_j}$ ($U_{\bar{x}_i}$).

If $\Gamma(f, \mathcal{R})$ is *strongly connected*, then we can reach from any local minimum all the other ones by choosing *alternatingly* certain trajectories of the ascent flow (going upwards form a local minimum to a local maximum) and the descent flow (going downwards form a local maximum to a local minimum). In case that no inequality constraints may become active, strict connectedness can generically be expected, as it is stated by the following theorem.

THEOREM 2 [1, 4]. *Let f and M be given and suppose that all inequality constraints are redundant. Then, for generic Riemannian metric* $\mathcal{R}$, *the min-max digraph* $\Gamma(f, \mathcal{R})$ *is connected.*

Note that, in case that all inequality constraints are redundant, the ascent and descent flows are just opposite to each other. However, the appearance of active inequality constraints may change the situation drastically. In particular, *absorbing two-cycles* may appear. If one enters such an absorbing two-cycle, one cannot leave it anymore via ascent and descent flows.

DEFINITION 3. Let Γ be a directed graph. An *absorbing two-cycle* is a set of nodes x, y such that in Γ there are both an arc from x to y and from y to x (two-cycle), and such that all the other arcs incident with x or y are *incoming* arcs (*absorbing*).

In Figure 1 two examples of absorbing two-cycles are depicted.

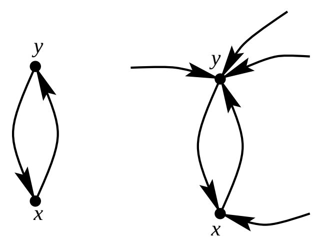

Figure 1. Absorbing two-cycles.

The next theorem is the main result of the paper. Its proof will be given in Section 2. Note that the dimension $\dim(M)$ of M equals $n - |I|$.

THEOREM 4. *Suppose that* $\dim(M) \geqslant 2$ *and* $J_0(x) \neq \emptyset$ *at some* $x \in M$. *Let* $m \in \mathcal{N}$. *Then there exists a pair* $(f, \mathcal{R})$ *and a* (C^2, C^1)*-neighborhood* $\mathcal{O}$ of $(f, \mathcal{R})$ *such that for all* $(\tilde{f}, \tilde{\mathcal{R}}) \in \mathcal{O}$ *the min–max digraph* $\Gamma(\tilde{f}, \tilde{\mathcal{R}})$ *contains at least m absorbing two-cycles.*

Although Theorem 4 is quite discouraging, we also have the following positive result.

THEOREM 5. ([3]). *For given f there exists a Riemannian metric* $\mathcal{R}$ *such that* $\Gamma(f, \mathcal{R})$ *is strongly connected.*

Concrete metrics resulting in a strongly connected min–max digraph are presented in [5]. In fact, the idea in [5] is based on automatic adaptation (via constraints) of a given metric $\mathcal{R}$ in the spirit of interior point methods.

2. Proof of Theorem 4

The possible disconnectedness of the min–max digraph has been shown for the first time by Horst Zank [6]. His example was two-dimensional with a minimal number of KKT-points for $f|_M$ and $(-f)|_M$. In fact, let M be a two-dimensional disc and suppose that $f|_M$ has two local minima and two local maxima (all of them in the boundary ∂M) and in addition, one saddle point (in the interior $\overset{\circ}{M}$). The Riemannian metric is chosen such that the separatrices (stable and unstable manifolds) of the saddle point intersect ∂M outside the chosen neighborhoods of the local minima and maxima. See Figure 2, in which also some level lines of f have been sketched. The resulting min–max digraph is not connected. In particular it consist of two absorbing two-cycles of the type sketched in Figure 1 on the left. Moreover, the latter situation is stable with respect to small perturbations of

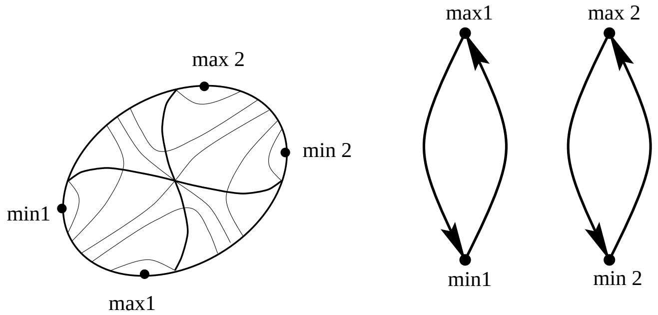

Figure 2. Zank's example.

both the Riemannian metric $\mathcal{R}$ (in the C^1-topology) and the objective function f (in the C^2-topology). The latter stability follows from a continuity argument using the theory of dependence of solutions of ordinary differential equations on perturbations of the right handside as well as the differentiable dependence of the KKT-points of $f|_M$ and $(-f)|_M$ on f itself (cf. [4]).

Note that the saddle point in Figure 2 serves as a decomposition point for both $f|_M$ and $(-f)|_M$. This phenomenon can only occur in dimension 2. In higher dimensions there would be more KKT-points for $f|_M$ and $(-f)|_M$ if the number of absorbing two-cycles would be greater than one.

The idea of our construction for proving Theorem 4 in all dimensions greater than 1 is the following. First we construct a basic (stable) model of a two-cycle. Then, by means of some surgery on a given $f|_M$, this model is locally implanted at m places in M. In this way we obtain a new f and a new Riemannian metric $\mathcal{R}$ which, as a pair $(f, \mathcal{R})$, fulfil the requirements as stated in Theorem 4. In the following we focus on the main geometric-topological ideas. For missing technical details we refer to the book [4].

2.1. THE BASIC MODEL OF A TWO-CYCLE

For $k \geqslant 2$ we decompose $\mathbb{R}^k$ as follows: $\mathbb{R}^k = \mathbb{R}^{k-1} \times \mathbb{R}$. Let $D^n := \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 \leqslant 1\}$ denote the Euclidean unit ball in $\mathbb{R}^n$ and consider the cylinder $C := D^{k-1} \times D^1$ in $\mathbb{R}^k$. In a neighborhood of C the function h is well-defined, where

$$h(x_1, \ldots, x_k) = x_k \left(2 - \sum_{i=1}^{k-1} x_i^2 \right).$$

Note that the point $(0, -1) \in \mathbb{R}^{k-1} \times \mathbb{R}$ is the global minimum of $h|_C$ and moreover, it is the only KKT-point of $h|_C$. The function value at the global minimum might be varied arbitrarily by adding a suitable constant to the function h.

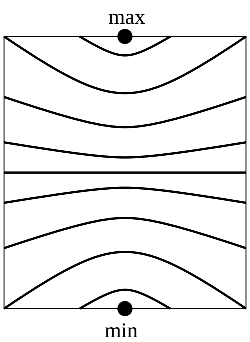

Figure 3. Some level sets.

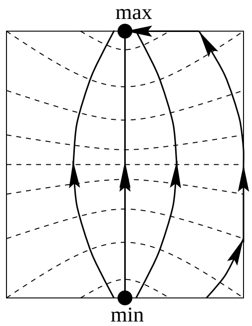

Figure 4. Typical ascent trajectories.

On the other hand, the point $(0, 1)$ is the global maximum of $h\,|\,_C$ and moreover, it is the only KKT-point of $(-h)\,|\,_C$. The points $(0, -1)$ and $(0, 1)$ are nondegenerate KKT-points for $h\,|\,_C$ and $(-h)\,|\,_C$, respectively. See Figure 3 for some level lines in case $k = 2$. As a Riemannian metric we choose the *constant* metric given by the standard scalar product $\langle u, v \rangle := \sum_{i=1}^{k} u_i v_i$.

See Figure 4 for some trajectories of the associated ascent (semi-)flow. Note that the ascent flow converges to the maximum, whereas the descent flow converges to the minimum.

Note, in particular, the boundary behaviour. Of particular interest for our construction are the flows in the interior $\overset{\circ}{C}$ of C. In $\overset{\circ}{C}$ the trajectories of the descent and the ascent (semi-)flow coincide, but have opposite directions. Let us fix in C two small neighborhoods (germs) $U_{\min}$ and $U_{\max}$ of $(0, -1)$ and $(0, 1)$, respectively, such that not any semi-flow starting in one of them will touch the cylindrical part of the boundary $\partial D^{k-1} \times D^1$. Because the trajectory $\gamma := \{0\} \times (-1, 1)$ lies in the interior of C, continuity arguments guaranty the existence of the latter germs.

The latter situation is stable under small perturbations of both the (constant) metric (in the C^1 *– topology*) and the function h (in the C^2-topology). Strictly speaking, the (germ)-neighborhoods also change, since the minimum and the maximum of $h\,|\,_C$ may shift; but this is only a technical detail.

2.2. THE U-SHAPE CONSTRUCTION

In order to perform a surgery on a given $f|_M$ by implanting the basic two-cycle model, it is convenient to make the following intermediate step.

Let ϕ be a smooth diffeomorphism of an open neighborhood of the cylinder C onto an open set in $\mathbb{R}^k$ having the following properties (see Figure 5). The cylinder C is mapped to the upper half space $\mathbb{R}^{k-1} \times \mathbb{H}$, where $\mathbb{H} := \{x \in \mathbb{R} \,|\, x \geqslant 0\}$. The interval $\overline{\gamma} = \{0\} \times [-1, 1]$ is mapped onto a U-shape curve $\phi(\overline{\gamma})$, meeting the subspace $\mathbb{R}^{k-1} \times \{0\}$ transversally. Finally, the balls $D(t) := D^{k-1} \times \{t\}$, $t \in [-1, 1]$, are mapped on balls as

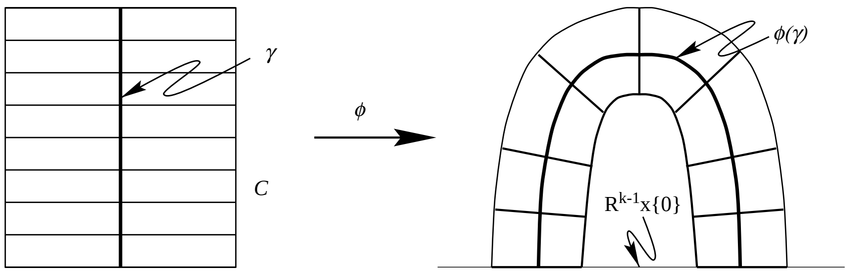

Figure 5. The U-shape construction.

well, where the images of $D(-1)$ and $D(1)$ lie in the plane $\mathbb{R}^{k-1} \times \{0\}$. By means of shifting and scaling we may assume that the image $\phi(C)$ is contained in a prescribed neighborhood of the origin of $\mathbb{R}^k$.

The constant metric around C induces, via the diffeomorphism ϕ, a Riemannian metric around $\phi(C)$, say $\mathcal{R}_\phi$. In an analogous way the function h defines a smooth function $h_\phi := h \circ \phi^{-1}$, defined around $\phi(C)$.

By the very construction, just as the image under a diffeomorphism, the (semi-)flow for h_ϕ starting in one of the germ neighborhoods $\phi(U_{\min})$ or $\phi(U_{\max})$ will reach the other one without touching the image of the cylindrical part of the boundary. In particular, the image $\phi(\gamma)$ is the common trajectory of an ascent and a descent flow for h_ϕ w.r.t. the metric $\mathcal{R}_\phi$, see Figure 5.

We shortly refer to $(\phi(C), h_\phi, \mathcal{R}_\phi)$ as a *U-shape two-cycle*.

2.3. THE IMPLANTATION STEP

Now suppose that $\dim(M) \geqslant 2$ and that at some point $\bar{x} \in M$ we have $J_0(\bar{x}) \neq \emptyset$. We will change a given pair $(f, \mathcal{R})$ in a neighborhood of $\bar{x}$ by successively implanting m U-shape two-cycles. It will be clear that all the assumptions regenerate; hence it suffices to consider the case $m = 1$, i.e. we merely present the implantation of just one U-shape two-cycle.

Since LICQ holds at $\bar{x}$, we may locally use the function values of the functions h_i, $i \in I$, g_j, $j \in J_0(\bar{x})$ as new coordinates. Hence M locally takes the form $\mathbb{R}^p \times \mathbb{H}^q$, where $q := |J_0(\bar{x})|$ and $p := n - |I| - q$. Note that $\mathbb{H}^q$ is the nonnegative orthant in $\mathbb{R}^q$. Consequently we may choose a point $\hat{x} \neq \bar{x}$ outside the set of local minima and maxima of $f|_M$ such that $|J_0(\hat{x})| = 1$. Now we choose local coordinates around $\hat{x}$ such that $\hat{x}$ is sent to the origin of $\mathbb{R}^k$ and M takes the form $\mathbb{R}^{k-1} \times \mathbb{H}$. We may assume that the latter open neighborhood U of the origin does not contain any local minimum or maximum of $f|_M$. On the other hand, our U-shape two-cycle may be assumed to be so small that $\phi(C) \subset U$. In an open neighborhood $V \subset U$ of $\phi(C)$ we replace the function f and the Riemannian metric $\mathcal{R}$ by h_ϕ and $\mathcal{R}_\phi$, respectively. In this way we have two "partial" functions and metrics with overlapping regions of definition: $(h_\phi, \mathcal{R}_\phi)$ on V and $(f, \mathcal{R})$ outside $\phi(C)$. By means of a suitable smooth partition of the unity we glue the partial functions and metrics together and we arrive at a modified function $\tilde{f}$ and Riemannian metric $\widetilde{\mathcal{R}}$ such that, on $\phi(C)$, they coincide with the ones of the U-shape two-cycle and, outside U, with the original ones. This implies, even for the glued model, that ascent semi-flows starting in $\phi(U_{\min})$ necessarily reach $\phi(U_{\max})$ and descent semi-flows starting in $\phi(U_{\max})$ reach $\phi(U_{\min})$.

Due to the partition of unity construction the functions $\widetilde{f}|_M$ and $(-\widetilde{f})|_M$ might have degenerate KKT-points. However, arbitrarily close to $\widetilde{f}$ (in the C^2-topology, see [4]), there are functions $\hat{f}$ which meet all the conditions in Assumption 1. Finally, taking $\hat{f}$ as objective function and $\widetilde{\mathcal{R}}$ as Riemannian metric, we obtain a min-max digraph $\Gamma(\hat{f}, \widetilde{\mathcal{R}})$ with at least one absorbing two-cycle.

This completes the proof of Theorem 4.

References

1. Jongen, H.Th. (1997), Zur Geometrie endlichdimensionaler nichtkonvexer Optimierungsaufgaben. *International Series of Numerical Mathematics* 36, 111–136.
2. Jongen, H.Th. and Sprekels, J. (1981), The index-k-stabilizing differential equation. *OR-Spektrum* 2, 223–225.
3. Jongen, H.Th. and Ruiz Jhones, A. (1999), Nonlinear optimization: On the min-max digraph and global smoothing. in: Ioffe, A., Reich S. and Shafrir, I. (eds.), *Calculus of Variations and Differential Equations*, *Research Notes in Mathematics Series*, Vol. 410, pp. 119–135. Chapman & Hall, Boca Raton.
4. Jongen, H.Th., Jonker, P. and Twilt, F. (2000), *Nonlinear Optimization in Finite Dimensions*, Kluwer Academic Publishers, Dordrecht.
5. Jongen, H.Th. and Stein, O. (2004), Constrained Global Optimization: Adaptive Gradient Flows. in: Floudas, C.A. and Pardalos, P.M. (eds.), *Frontiers in Global Optimization*, *Nonconvex Optimization and Its Applications*, (1994), Vol. 74, pp. 223–236. Kluwer Academic Publishers, Dordrecht.
6. Zank, H. (1994), personal communication.