A Rigorous Numerical Algorithm for Controllability

Fritz Colonius [*] — Tomasz Kapela [**]

[*] *Institut für Mathematik, Universität Augsburg, 86135 Augsburg, Germany*
e-mail: fritz.colonius@math.uni-augsburg.de

[**] *Institute of Computer Science, Jagiellonian University, Kraków, Poland*
e-mail: kapela@ii.uj.edu.pl

ABSTRACT. *The paper presents a rigorous numerical algorithm for controllability, based on the Krawczyk operator. For given two sets X and Y we check whether each point from the set X is connected by some controlled trajectory with each point in the set Y. Two examples are included.*

RÉSUMÉ. *A définir par la commande* \resume{...}

KEYWORDS: *controllability, rigorous numerics, Krawczyk operator, Takens-Bogdanov oscillator, escape equation*

MOTS-CLÉS : *A définir par la commande* \motscles{...}

1. Introduction

The purpose of this paper is to provide a rigorous numerical method that allows us to confirm controllability results. The notion of 'rigorous' in this context may need some explanation. It is very different from standard notions of convergence and, maybe, error estimates. It refers to a by now well established line of research which aims at computer assisted proofs for mathematical results. In particular, methods from interval arithmetics are used, see e.g. Neumair [Neu 90] or Jaulin et al. [JKDW 01], in order to take errors of floating point numerics into account. Thus the numerical results have the status of mathematically proven theorems. To the best of our knowledge, such algorithms for controllability have not yet been provided in the literature, with the exception of Marquardt [Mar 05]. Marquardt computes outer approximations to reachable sets using Fliess series together with subdivision methods. This generalizes the classical Lohner algorithm for higher order numerical solution of initial value problems for ordinary differential equations; see e.g. Zgliczyński [ZLo]. In contrast to this, we fix sets of initial and final values and ask for existence of a controlled trajectory connecting them.

We remark that the controllability problem often is considered as an optimal control problem: If two points can be connected by a trajectory, they can (under mild assumptions) also be connected by a time or energy optimal trajectory. The resulting optimal control problem leads to two-point boundary value problems, which then are solved numerically. While this yields a numerical approximation, it does not yield a rigorous proof confirming controllability (although one might try to apply rigorous numerics to the resulting two-point boundary value problem). One should not expect that rigorous numerical algorithms are competitive with respect to generality or effectivity to these nonrigorous methods. This is the price we have to pay for mathematical rigor.

We consider the control system

$$\dot{x}(t) = f(x(t), u(t)), \tag{1.1}$$

where $f : \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n$ is a C^1-function with respect to the first argument and $U \subset \mathbb{R}^m$ is a set of admissible control values, hence $u(t) \in U$ for all t. For simplicity, we assume that the relevant initial value problems with $x(0) = x_0$ and control function u are uniquely solvable with controlled trajectories $\varphi(t, x_0, u),\ t \in \mathbb{R}$.

We fix two subsets X and Y in the state space and ask, if we can connect each point in X to each point in Y by controlled trajectories, i.e., for all $x_0 \in X$ and $y_0 \in Y$ there are $T > 0$ and a control function u such that

$$\varphi(T, x_0, u) = y_0.$$

In order to simplify the problem, we consider only piecewise constant controls and assume that the number of control values is finite. Thus we switch between a finite number of autonomous ordinary differential equations and our main task is to compute the switching times. This can be reformulated as a problem to find a zero of a function.

We use the Krawczyk operator [Kra 69] to solve this problem; the numerical implementation is based on the package CAPD, see [Capd]. The CAPD package is developed by a group of mathematicians and computer scientists mainly from Jagiellonian University. Among other things it provides tools for rigorous ODE's integration using C^1-Lohner algorithm [ZLo] and various methods of set representations. The source code of the software performing proofs described in this article can be downloaded from the webpage [TK].

Good initial guesses are of fundamental importance for rigorous numerical proofs. Here we need initial guesses for the desired trajectory or, more precisely, for the piecewise constant control and the switching times. For simplicity we mainly consider the two dimensional case ($n = 2$) and assume that $u(t) \in \{u_-, u^+\} \subset \mathbb{R}$. We remark, that the restriction to two control values is not as restrictive as it may appear since in many applications it suffices to take only extremal values of the control range.

In order to check our method, we study two (two-dimensional) systems where delicate controllability problems occur. The problem is if two given points in the state space can be connected by a trajectory corresponding to a control with one switching. The first system from Häckl and Schneider [HaSch 96] concerns the existence of control-homoclinic orbits. The second system from Gayer [Ga 04] is the escape equation and concerns bifurcation problems for control sets. In both examples we show rigorous results and compare them with the numerical results in the cited references. We hope that in the future it will be possible to do computations for problems with more than one switching, in state spaces of dimensions higher than two, and also with higher dimensional controls.

The contents of this paper are as follows:

Section 2 recalls the definition and properties of the Krawczyk operator. Section 3 shows how to use it for determination of the switching times and section 4 presents our algorithm. Here we put particular emphasis on describing how to find good initial guesses. The final Section 5 analyzes the announced examples. Note the very precise estimates needed here.

2. Interval Krawczyk Method

Our controllability result will be based on a transformation to a problem to find a zero of a function. A basic tool will be the Krawczyk operator. We cite the following result from Krawczyk [Kra 69].

Assume that:

- $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a C^1 function,
- $X \subset \mathbb{R}^n$ is an interval set,
- $\bar{x} \in X$
- $C \in \mathbb{R}^{n \times n}$ is a linear isomorphism.

Then the Krawczyk operator is given by

$$K(\bar{x}, X, F) := \bar{x} - CF(\bar{x}) + (Id - C[DF(X)])(X - \bar{x});$$

here $[DF(X)]$ denotes the interval hull of the derivative computed over the set X (this is the smallest product of intervals containing $DF(X)$).

Theorem 2.1. *With the assumptions and notation introduced above, the following holds:*

1) If $x^* \in X$ and $F(x^*) = 0$, then $x^* \in K(\bar{x}, X, F)$.

2) If $K(\bar{x}, X, F) \subset intX$, then there exists a unique $x^* \in X$ such that $F(x^*) = 0$.

3) If $K(\bar{x}, X, F) \cap X = \emptyset$, then $F(x) \neq 0$ for all $x \in X$.

Remark. To compute the Krawczyk operator we need the value of function F in the point $\bar{x}$ and the interval hull of its derivative $[DF]$ computed on the whole set X. Although C can be any non-degenerate matrix, the best choice is to take an approximation of the inverse of $DF(\bar{x})$.

Remark 2. To be rigorous all computations instead on double precision numbers are performed using interval arithmetics (for details see [Neu 90, Moo 66]) to include all possible errors coming from rounding, numerical method of ODE integration etc.

3. Searching for the switching times

This section shows how determine the switching times using the Krawczyk operator. It is convenient to introduce the following notation. Consider the autonomous equations in $\mathbb{R}^n$

$$\dot{x} = f_i(x) := f(x, u_i), \text{ where } i \in \{1, 2, \ldots, n\}.$$

generating flows $\varphi_i(t, x)$. Thus they correspond to the constant controls u_i. Since we assume that f is a C^1-function with respect to x, the flows are C^1.

Fix $x_0, y_0 \in \mathbb{R}^n$. We want to find times $t_1, t_2, ..., t_n \geq 0$ such that we can go from the point x_0 to the point y_0 following the first flow for the time t_1 and then the second flow for the time t_2, etc. This means that we are searching for solutions $(t_1, \ldots, t_n)$ of the equation

$$\varphi_n(t_n, \varphi_{n-1}(t_{n-1}, \ldots, \varphi_1(t_1, x_0)...)) = y_0,$$

which is the same as searching for a zero of the function $F_{x_0,y_0} : \mathbb{R}^n \to \mathbb{R}^n$ given by

$$F_{x_0,y_0}(t_1, \ldots, t_n) := \varphi_n(t_n, \varphi_{n-1}(t_{n-1}, \ldots, \varphi_1(t_1, x_0)...)) - y_0.$$

To this function we can apply the Interval Krawczyk method.

Theorem 3.1. *If for some interval set $T \subset \mathbb{R}^n_+$ and a point $t_0 \in T$ the rigorously computed Krawczyk operator $K(t_0, T, F_{x_0,y_0})$ is contained in the interior of T, then there exists a trajectory joining the points x_0 and y_0. Explicitly, there exist times $(t_1, \dots, t_n)$ such that $(t_1, \dots, t_n) \in T$ and $y_0 = \varphi_n(t_n, \varphi_{n-1}(t_{n-1}, \dots, \varphi_1(t_1, x_0)...))$.*

Proof. For an application of Theorem 2.1 we have to show that F is a C^1-map. This follows, since by assumption the flows are C^1 and, by the chain rule, the partial derivatives $\frac{\partial F}{\partial t_i}$ exist and are continuous:

$$\frac{\partial F}{\partial t_i} = (\varphi_n)_* \dots (\varphi_{i+1})_* f(\varphi_i(t_i, \varphi_{i-1}(t_{i-1}, ..., \varphi_1(t_1, x_0)...)));$$

here the linearized flows $(\varphi_j)_*$ are evaluated as

$$(\varphi_j)_* = (\varphi_j)_* (t_{j+1}, f(\varphi_j(t_j, \varphi(t_{j-1}, \dots, \varphi(t_1, x_0)...)))).$$

□

We can easily generalize this result by replacing the initial and final points by interval sets.

Let X and Y be two interval sets in $\mathbb{R}^n$ and define

$$\bar{F}_{X,Y}(t_1, \dots, t_n) = \varphi_n(t_n, \varphi(t_{n-1}, \dots, \varphi(t_1, X)...)) - Y.$$

Then we have $F_{x_0,y_0}(t_1, \dots, t_n) \in \bar{F}_{X,Y}(t_1, \dots, t_n)$ for each $(x_0, y_0) \in X \times Y$.

Corollary 3.2. *If for some interval set $T \subset \mathbb{R}^n_+$ and a point $t_0 \in T$ the rigorously computed Krawczyk operator $K(t_0, T, \bar{F}_{X,Y})$ is contained in the interior of T, then for each $(x_0, y_0) \in X \times Y$ there exists a trajectory from the point x_0 to y_0.*

Proof. Let $(x_0, y_0) \in X \times Y$. Then $K(t_0, T, F_{x_0,y_0}) \subset K(t_0, T, \bar{F}_{X,Y}) \subset intT$. Theorem 2.1 yields the existence of a point $t^* = (t^*_1, \dots, t^*_n) \in T$ such that $F_{x_0,y_0}(t^*) = 0$. Hence from the point x_0 we can reach the point y_0 by first going for time t^*_1 with the first flow, then with second one for time t^*_2, etc. □

Remark. In order to obtain a function F on $\mathbb{R}^n$, we need that the number of switching times is equal to the dimension of the state space. If more switching times are needed, we have to split up the problem into subproblems. In practice, however, this will require good estimates for the intermediate end points.

4. Algorithm isReachable

From now on we restrict our attention to two-dimensional case ($n = 2$). For two given sets X and Y if algorithm **isReachable** returns **true** then each point in the set Y can be reached from each point in the set X. On the other hand if this algorithm returns **false** we can not conclude that there are points in Y that are not reachable from some points in X.

INPUT

– $\varphi_1(t,x)$, $\varphi_2(t,x)$ - two flows,

– X - an interval set of "starting" points,

– Y - an interval set, whose reachability we want to check,

OUTPUT

– **true** if we succeeded to show that every point in Y can be reached from each point in X,

– **false** otherwise.

We return **true** if and only if for some interval set T and $t_0 \in T$ the rigorously computed Krawczyk operator $K(t_0, T, \bar{F}_{X,Y})$ is subset of $intT$. For this we need good approximations of the times T and $t_0 \in T$. On one hand the set T should be as small as possible to enable rigorous computation and to provide a good approximation of the Krawczyk operator. On the other hand it should be also big enough to guarantee that all "switching" times are included in T and that the values of the Krawczyk operator, even if overestimated, form a subset of the interior of T.

In the first step of the algorithm we want to approximate the "switching" times and define a "good" set T. There are several possibilities:

– we may already have very good approximations from analytic knowledge or from numerical simulations and we want to define the set T manually.

– we may have only rough estimates and we want to improve them.

– we do not have any initial guesses and we need an algorithm to find estimates and to define the set T.

In the first case we just use those approximations. A possible method for the third case is to perform numerical simulation of some forward trajectories from the set X and some backwards trajectories from the set Y and search for approximated times of its intersections. If the used approximation method fails and we do not have any initial guesses for the set T then we return **false**.

The main part of this algorithm is to compute rigorously the Krawczyk operator and to check a suitable inclusion. We use the C^1-Lohner algorithm (see [ZLo]) implemented in the CAPD package [Capd], which provides rigorous bounds for trajectories and derivatives of flows. We can perform computations using various values of parameters such as the time step and the order of the Taylor method. We can also try to increase or decrease the set T.

If one of these attempts succeeds (which means that values of the Krawczyk operator are contained in the interior of the set T) then we return **true**. If all of them fail we return **false**.

5. Examples

The first example is the Takens-Bogdanov oscillator, where Häckl and Schneider [HaSch 96] (compare also Colonius and Kliemann [CK 00]) could analytically confirm the existence of a control range such that two given points can be connected by a trajectory with one switching (while constant controls do not allow this). Furthermore, they established a nonrigorous numerical example of such a control range. The second one is the escape equation. Here, for controllability problems, Gayer [Ga 04] could establish numerical estimates for the required size of the control range.

5.1. *The Takens-Bogdanov oscillator*

The controlled Takens-Bogdanov oscillator is given by

$$\begin{aligned} &\dot{x} = y \\ &\dot{y} = \lambda_1 + \lambda_2 x + x^2 + xy + u(t), \\ &u(t) \in [-\rho, \rho], \end{aligned}$$

with real parameters $\lambda_1, \lambda_2, \rho$. It is known (cp. [CK 00, Section 9.4] for the following assertions) that for parameters (λ_1, λ_2) in a certain subset $S_2 \subset \mathbb{R}^2$ the uncontrolled system has a stable fixed point and a hyperbolic fixed point without homoclinic orbit. For small $\rho > 0$ there are two simply connected control sets (i.e., maximal subsets of complete controllability), one around each of the equilibria. These control sets contain all stable and all hyperbolic equilibria, respectively, corresponding to constant controls in $[-\rho, \rho]$. For the parameter values in S_2 given by

$$\lambda_1 = -0.3,\ \lambda_2 = -1,$$

one can prove that there is $\rho > 0$ such that the control set containing the hyperbolic equilibrium is not simply connected and surrounds the other control set. A numerical computation for the control range given by $\rho = 0.05$ shows this behavior; it is reproduced in Figure 1: There are control homoclinic orbits (i.e. trajectories connecting hyperbolic equilibria), while it is known that for constant controls $u \in [-\rho, \rho]$ there is no homoclinic orbit.

Our goal is to prove that for these parameter values there is a control heteroclinic orbit.

We will show that there exists a trajectory joining the point $x_0 = (1.26, 0.0)$, which is the hyperbolic fixed point for $u(t) = -0.0276$ (it is close to the hyperbolic fixed point for $u(t) = -\rho$) and the fixed point $y_0 = (1.207106781, 0.0)$ for $u(t) = \rho$. This trajectory goes around the invariant control set.

To show this we will use the following two control values: first $u_1(t) = \rho = -0.05$ and then $u_2(t) = 0.04$. This determines the two vector fields

$$f_1(x, y) = (y, x^2 + xy - x - 0.35),$$

$$f_2(x, y) = (y, x^2 + xy - x - 0.26).$$

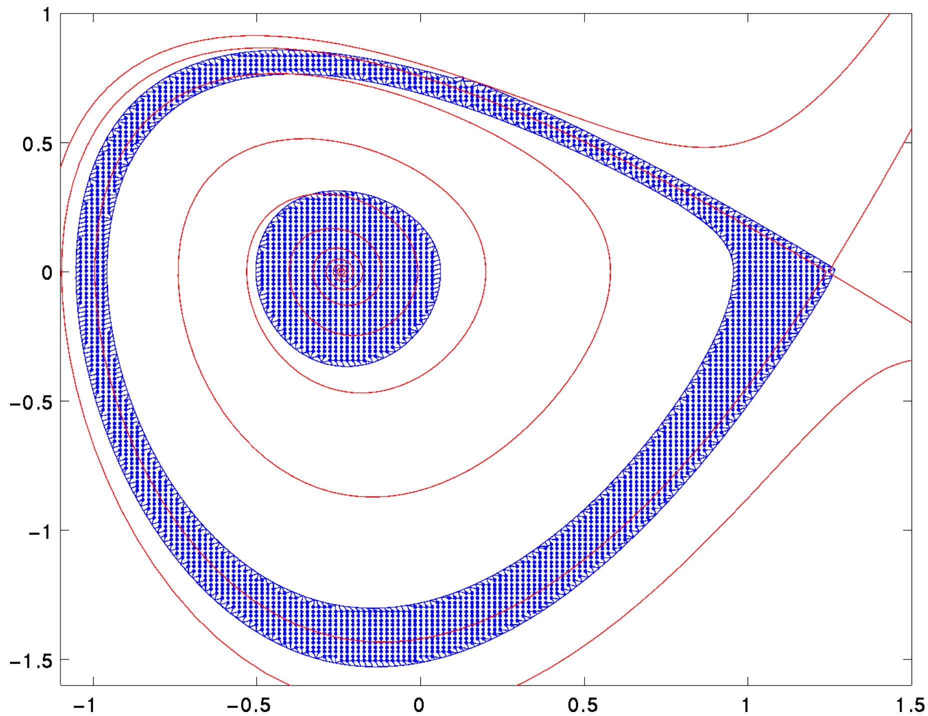

Figure 1. *Numerical simulations of control sets for the Takens-Bogdanov oscillator with* $\lambda_1 = -0.3$, $\lambda_2 = -1$, $\rho = 0.05$

Our initial guesses for the times are $t_0 = (t_1, t_2) = (2.9039196889, 10.559934)$. We set $T = [t_1 - \delta_1, t_1 + \delta_1] \times [t_2 - \delta_2, t_2 + \delta_2]$ where $\delta_1 = 2.0 * 10^{-9}$ and $\delta_2 = 2.0 * 10^{-3}$. In the numerical simulation we obtain two intersection points of the forward trajectory from x_0 and the backward trajectory from y_0. We choose the point with the smaller time t_1 and then we start several times our approximation procedure with decreasing time steps.

To compute the Krawczyk operator we use the C^1 Lohner algorithm with time step equal to 0.01 and a 6th order Taylor method. We obtain these values by performing computations with different time steps and orders and then comparing diameters of computed Krawczyk operators.

Finally, we obtain that

$$K(t_0, T, F_{x_0,y_0}) \subset [2.9039196882987688, 2.9039196895134038] \times$$
$$[10.558743378108383, 10.561125701344494] \subset \mathrm{int} T.$$

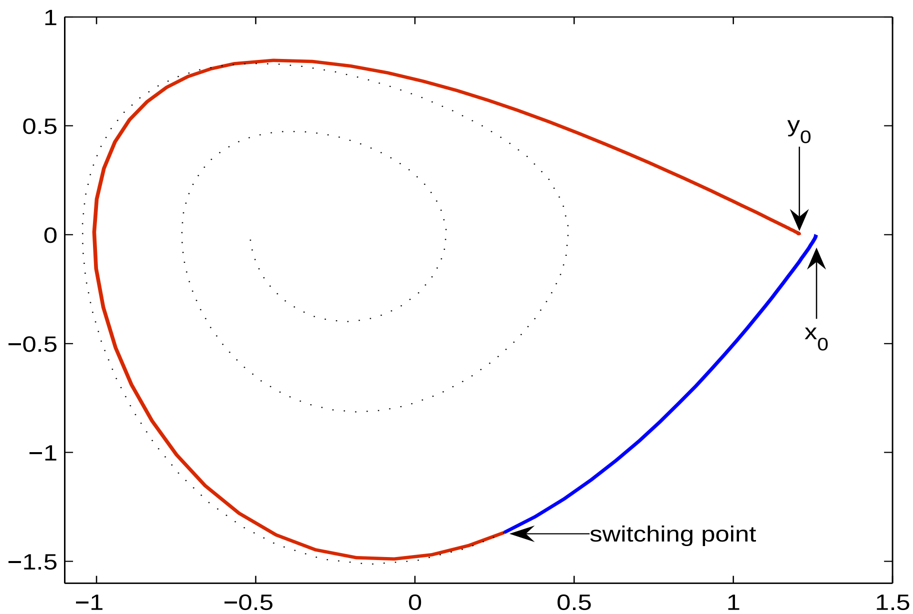

Figure 2. *Controlled homoclinic trajectory for the Takens-Bogdanov oscillator with* $\lambda_1 = -0.3$, $\lambda_2 = -1$, $\rho = 0.05$

From Theorem 3.1 it follows that there exists a trajectory joining the points x_0 and y_0.

We also performed computations with the same parameters for small interval sets X and Y around the points x_0 and y_0 (we took radius 10^{-14}) and we obtained

$$K(t_0, T, F_{X,Y}) \subset [2.903919687831086, 2.9039196899790425]\times$$
$$[10.557940607196224, 10.5619284723529] \subset \mathrm{int}T.$$

Theorem 3.2 implies the existence of trajectories from each point in the set X to each point in Y. In this example, the sets X and Y are very small because we are very close to the hyperbolic point.

In both cases the computation times were approximately 1 second.

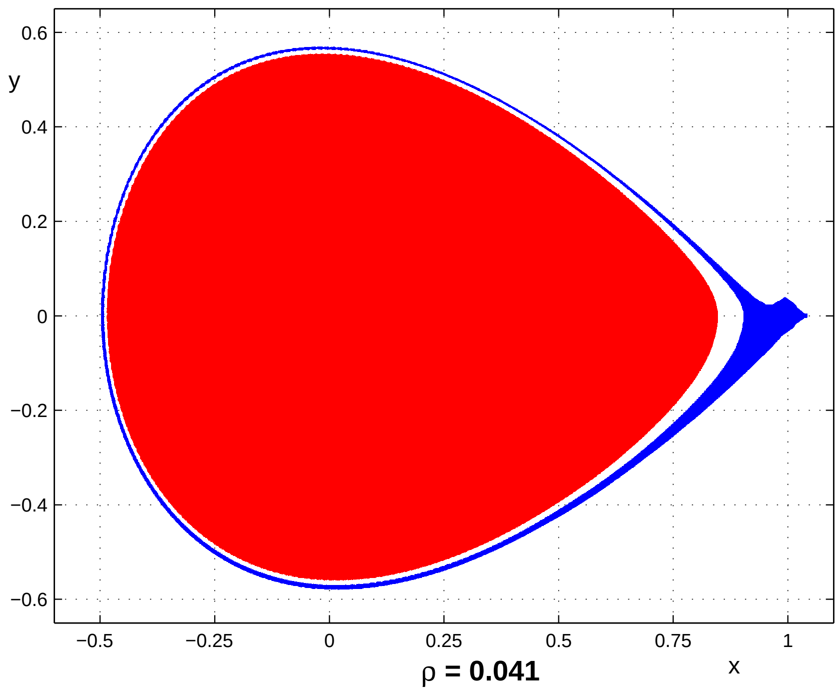

Figure 3. *Simulations of control sets for the controlled escape equation with* $\rho = 0.041$.

5.2. *Escape equation*

The controlled escape equation is given by

$$\begin{aligned}\dot{x} &= y\\ \dot{y} &= \lambda y + x^2 - x + u(t),\\ u(t) &\in [-\rho, \rho]\end{aligned}$$

where λ and ρ are real parameters.

Figures 3 and 5 show simulations of control sets for $\lambda = -0.1$ and two values of ρ due to T. Gayer (for more details see [Ga 04]). For $\rho = 0.041$ there are two control sets, a variant control set surrounds an invariant one. For $\rho = 0.04$ we also have two control sets, but the variant control set is simple connected.

For $\lambda = -0.1$ and $\rho = -0.041$ we will prove the existence of a periodic orbit (see Figure 4), going around the invariant control set. For $\rho = 0.040$ we proved that the simulation is not exact and that for this parameter there still exists a controlled trajectory starting and ending in the variant control set and going around the invariant control set (see Figure 6).

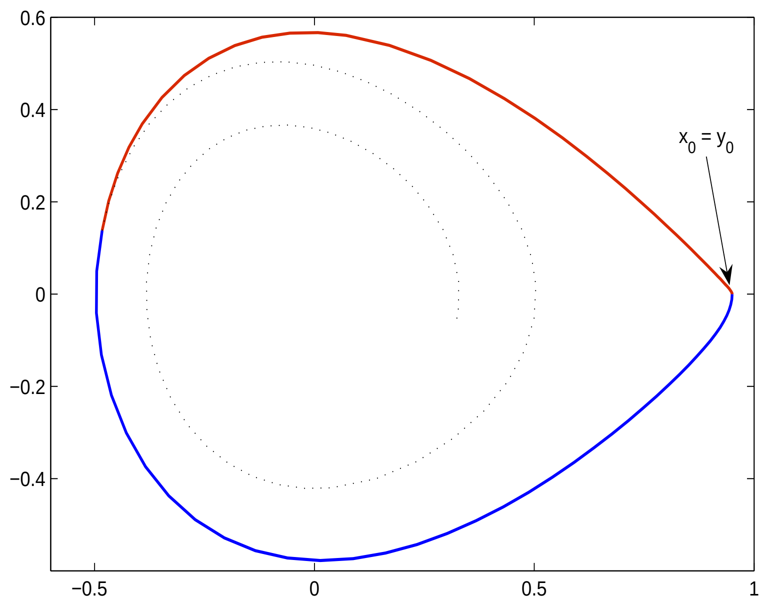

Figure 4. *Controlled periodic orbit for the controlled escape equation with* $\rho = 0.041$

We set $x_0 = y_0 = (0.95, 0)$, this is a point near the leftmost hyperbolic equilibrium. The extremal control values determine the two vector fields

$$f_1(x, y) = (y, -0.1y + x^2 - x - 0.041),$$
$$f_2(x, y) = (y, -0.1y + x^2 - x + 0.041).$$

An approximation of the switching times yields $t_1 = 5.18658693445$ and $t_2 = 7.18515699298$. Again we take $T = [t_1 - \delta_1, t_1 + \delta_1] \times [t_2 - \delta_2, t_2 + \delta_2]$ to be the interval ball around our time approximation, with $\delta_1 = 2.0 * 10^{-9}$ and $\delta_2 = 2.0 * 10^{-3}$. The first approximation brings us two intersection points between the forward trajectory from x_0 and the backward trajectory of y_0, but this time we take the point with larger t_1.

We check that for parameters: time step equal to 0.0425 and the 6th order Taylor method, the values of the Krawczyk operator are contained in interior of the set T. Hence Theorem 3.1 implies the existence of a trajectory joining the points x_0 and $y_0 = x_0$. Thus a periodic orbit is obtained (cp. Figure 4).

Now consider the case $\rho = 0.040$. In order to prove that the variant control set is not simply connected we take two hyperbolic fixed points $x_0 = (1.0385, 0.0)$

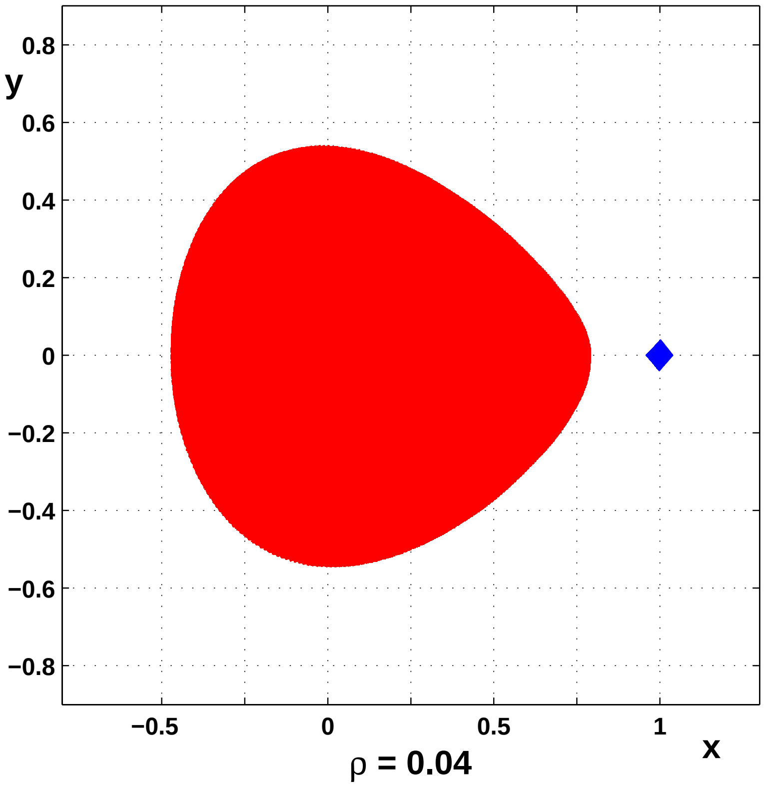

Figure 5. *Simulations of control sets for the controlled escape equation with* $\rho = 0.040$.

and $y_0 = (0.9582575695, 0.0)$ which correspond to controls $u(t) = -0.03998225$ and $u(t) = 0.040$, respectively, and we show that there is a controlled trajectory joining these points going around the invariant control set. Our simulations indicate that for this control range there does not exist a controlled periodic orbit with only one switching. In order to reach the point y_0 in a finite time we cannot use the maximal value of the control. Instead we use the controls $u_1 = -0.04$ and $u_2 = 0.0399$. Then the approximation procedure yields $t_1 = 13.769247218$ and $t_2 = 11.400673878$. Using time step 0.0825 and a 10th order of the Taylor method we showed that the interval Krawczyk operator is contained in the interior of the set $T = [t_1 - \delta_1, t_1 + \delta_1] \times [t_2 - \delta_2, t_2 + \delta_2]$ where $\delta_1 = 1.0 * 10^{-8}$ and $\delta_2 = 2.0 * 10^{-3}$. Hence the existence of a trajectory joining the points x_0 and y_0 follows from Theorem 3.1.

Acknowledgement. Tomasz Kapela was supported by the Marie Curie Multi-partner project Control Training Site in the form of a fellowship, contract number

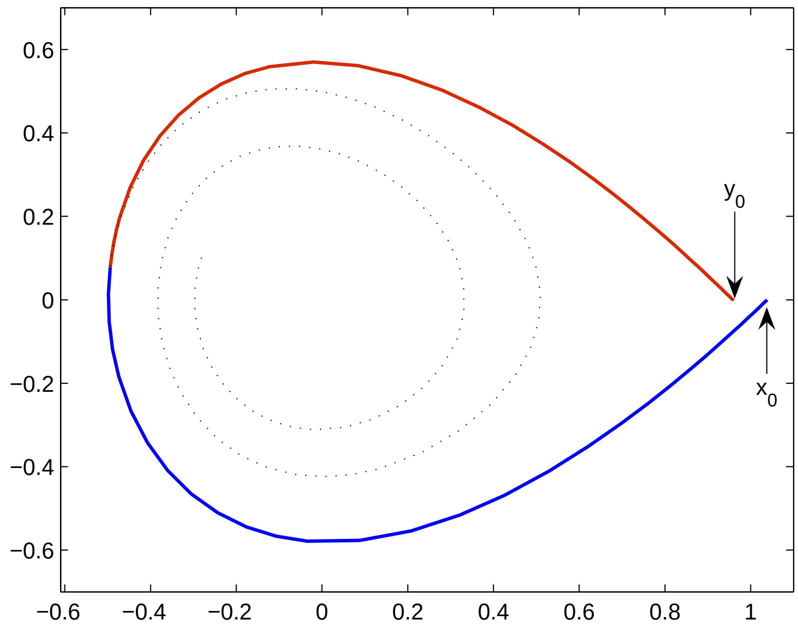

Figure 6. *Controlled homoclinic orbit for the controlled escape equation with* $\rho = 0.040$

HPMT-GH-01-00278-104. We thank Tobias Gayer for the permission to use Figures 3 and 5 from [Ga 04].

6. References

[Capd] CAPD - Computer assisted proofs in dynamics, a package for rigorous numerics, *http://capd.wsb-nlu.edu.pl/*

[CK 00] F. Colonius, W. Kliemann, The Dynamics of Control, Birkhäuser 2000.

[Ga 04] T. Gayer, *Control sets and their boundaries under parameter variation*, J. Differential Equations 201(2004), 177-200.

[HaSch 96] G. Häckl, K. Schneider, *Controllability near Takens-Bogdanov points*, J. Dyn. Control Systems 2(1996), 585-598.

[JKDW 01] J. Jaulin, M. Kiefer, O. Didrit, and E. Walter, Applied Interval Analysis, Springer-Verlag, London 2001.

[Kra 69] R. Krawczyk, *Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken*, Computing 4(1969), 187–201.

[Mar 05] A. Marquardt, Rigorous Numerical Enclosures for Control Affine Systems, Dissertation, Universität Augsburg, Augsburg 2005. Available online: http://www.opus-bayern.de/uni-augsburg/volltexte/2006/198/

[Moo 66] R.E. Moore, *Interval Analysis.* Prentice Hall, Englewood Cliffs, N.J., 1966

[Neu 90] A. Neumeier, Interval methods for systems of equations. Cambridge University Press, 1990.

[TK] T. Kapela Homepage, *http://www.ii.uj.edu.pl/~kapela/*

[ZLo] P. Zgliczyński, C^1-*Lohner algorithm*, Foundations of Computational Mathematics, (2002) 2:429–465