

A dynamical systems approach to constrained minimization

Johannes Schropp

Konstanzer Schriften in Mathematik und Informatik

Nr. 89, Mai 1999

ISSN 1430–3558

Konstanzer Online-Publikations-System (KOPS)
URL: <http://www.ub.uni-konstanz.de/kops/volltexte/2006/2059/>

A dynamical systems approach to constrained minimization

J. Schropp *

Department of Mathematics and Computer Science

University of Konstanz

P.O. Box 5560

D-78434 Konstanz

Abstract

We present an ordinary differential equations approach to solve general smooth minimization problems including a convergence analysis. Generically often the procedure ends up at a point which fulfills sufficient conditions for a local minimum. This procedure will then be rewritten in the concept of differential algebraic equations which opens the route to an efficient implementation. Furthermore, we link this approach with the classical SQP-approach and apply both techniques onto two examples relevant in applications.

AMS Subject Classification: 34C35, 49B36, 65L05.

1 Introduction

Sometimes concrete modelling of naturally occurring phenomena leads to optimization problems. According to the properties of the underlying problem, the independent variables are restricted by constraints or not. In a formal description, the structure of these problems reads as follows.

Let $f : D \rightarrow \mathbb{R}$, $D \subset \mathbb{R}^N$ open, $g : D \rightarrow \mathbb{R}^l$, $0 \leq l \leq N$ and $k : D \rightarrow \mathbb{R}^p$, $p \geq 0$ be arbitrary smooth functions. With these triplet of functions we consider the following problem:

$$\text{Minimize } f(x) \text{ subject to } g(x) = 0 \text{ and } k(x) \geq 0. \quad (1.1)$$

*Email: johannes.schropp@uni-konstanz.de

Now, let

$$\Gamma := \{x \in D \mid g(x) = 0, k(x) \geq 0\} \quad (1.2)$$

denote the set of feasible points. It is well known from classical analysis that, under certain mild constraint qualifications on g and k , for a point $\bar{x} \in \Gamma$ the following first order conditions

$$\begin{aligned} \exists(\bar{\alpha}, \bar{\beta}) \in \mathbb{R}^{l+p} \text{ such that } 0 &= \nabla f(\bar{x}) - \sum_{i=1}^l \bar{\alpha}_i \nabla g_i(\bar{x}) - \sum_{i=1}^p \bar{\beta}_i \nabla k_i(\bar{x}), \\ 0 &= k(\bar{x})^T \bar{\beta}, \\ \bar{\beta}_i &\geq 0, \quad i = 1, \dots, p \end{aligned} \quad (1.3)$$

are necessary and, additionally, the second order condition

$$\begin{aligned} \nabla^2 f(\bar{x}) - \sum_{i=1}^l \bar{\alpha}_i \nabla^2 g_i(\bar{x}) - \sum_{i=1}^p \bar{\beta}_i \nabla^2 k_i(\bar{x}) \text{ positive definite on } N(Dg(\bar{x})) \cap \Lambda, \\ \Lambda := \{v \in \mathbb{R}^N \mid v^T \nabla k_i(\bar{x}) = 0, \text{ if } k_i(\bar{x}) = 0 \text{ and } \bar{\beta}_i > 0, \\ v^T \nabla k_i(\bar{x}) \geq 0, \text{ if } k_i(\bar{x}) = 0 \text{ and } \bar{\beta}_i = 0\} \end{aligned} \quad (1.4)$$

is sufficient to be a local solution of (1.1) (see, e.g., Gill et al. (1981), Ch. 3 or Spellucci (1993), Ch. 2.1). The conditions (1.3) are usually known as Kuhn-Tucker conditions and feasible points that fulfil (1.3) are called Kuhn-Tucker points (see Kuhn and Tucker (1951)).

Naturally, there are a lot of standard methods to solve (1.1), e.g., descent methods for unconstrained minimization problems, the Lagrangian method, projected gradient methods and sequential quadratic programming methods (SQP-methods) for constrained problems. If we are looking for a class of general minimization methods possessing pleasing global convergence properties, from the methods mentioned above, only the SQP-methods are appropriate (see, e.g., Spellucci (1993), Theorem 3.6.13 for a convergence theorem). On the other hand, as we will demonstrate in Section 4 and as it is shown in Brown and Bartholomew-Biggs (1987), (1989), these parameter dependent optimization methods sometimes run into trouble if applied to nonlinear problems.

One of the goals of the present paper is to show that, appropriate for nonlinear minimization problems may be a dynamical systems approach. The first step of that approach is to replace the minimization problem (1.1) by an initial value problem. Next, this ordinary differential equation is discretized by a time stepping method. To assure that finally the original minimization problem is solved, the solution flow of such an ordinary differential equation (ODE) should possess the following properties:

(P1) The longtime behaviour of its evolutions becomes stationary.

(P2) The equilibria \bar{x} fulfill $\bar{x} \in \Gamma$ and (1.3).

(P3) The stable stationary points satisfy additionally (1.4).

The first step, the replacement of (1.1) by an exact ODE, is well known in unconstrained minimization. Botsaris (1978), for example, shows that the equation

$$\dot{x} = -Q(x)^{-1}\nabla f(x), \quad Q(x) \text{ symmetric and positive definite}$$

fulfills (P1)-(P3). Moreover, Amaya (1987) presents a uniqueness convergence theorem for the same equation under an additional convexity assumption on the levelset.

Also for equality constrained problems, the dynamical systems approach is not unknown. For example, Brown and Bartholomew-Biggs (1987), (1989) suggest the parameter dependent differential equation

$$\dot{x} = -\frac{\partial}{\partial x}\Phi(x, r)$$

with the penalty function $\Phi(x, r) := f(x) + \frac{1}{r} \sum_{i=1}^l g_i(x)^2$ and compare this differential equation method with SQP-methods.

Under the assumption that the constraints are explicitly separable into independent and dependent variables, Botsaris (1979), (1981) presents a dynamical system appropriate to (1.1).

In addition, with the use of the Lagrangian function of (1.1), Tanabe (1974), Yamashita (1980) and Evtushenko & Zhadan (1994) suggest the following dynamical systems to replace the minimization problem (1.1) for $p = 0$. Under a regularity assumption concerning g with

$$Q(x) := Dg(x)^T(Dg(x)Dg(x)^T)^{-1}Dg(x) \quad (1.5)$$

one can find

$$\dot{x} = (I - Q(x))(-\nabla f(x)) \quad (1.6)$$

in Tanabe (1974) and

$$\dot{x} = (I - Q(x))(-\nabla f(x)) - Dg(x)^T(Dg(x)Dg(x)^T)^{-1}g(x) \quad (1.7)$$

in Yamashita (1980) (formula (5.2) with $s = 1$) and Evtushenko & Zhadan (1994).

But, up to now, the differential equations approach has received only little attention in the optimization literature. To our opinion, this is due to the fact that the concrete ODE-codes mentioned are overall not competitive with an efficient SQP-implementation. Moreover, the connections of the differential equations suggested to any classical method in optimization is not clear. It is our objective to display how the efficiency of the ODE-approach can be improved and to show the close relationship between the SQP- and the

ODE-approach.

To be more precise, in the present paper, we suggest a family of dynamical systems slightly more general than (1.7). In contrast to Yamashita (1980) and Evtushenko & Zhadan (1994), a detailed comprehensive analysis about the behaviour of the solution flow and a convergence theorem is presented. This approach then will be rewritten in the concept of differential algebraic equations (DAEs) which opens a route for a more efficient implementation. Finally, we link the ODE- with the SQP-approach. It will be shown that the SQP-approach for equality constrained problems can be regarded as a variable step size Euler-Cauchy method applied to an appropriate dynamical system.

In Schropp (1998), we discretize our family of dynamical systems by applying arbitrary one- or linear multistep methods with constant step size. The reader will find a discrete analogon of our convergence theorem and a comparison of the properties of the discrete and the continuous solution flow. In addition, the convergence theorem for the discretization of the corresponding differential algebraic equation is under construction.

2 The main results

We consider an initial value problem

$$\dot{x} = F(x), x(0) = x_0, F \in C^1(\Omega, \mathbb{R}^N), \Omega \subset \mathbb{R}^N \text{ open} \quad (2.1)$$

with solution $\phi(t, x_0)$. The longtime behaviour of the evolution $\phi(t, x_0)$ can be characterized by the set

$$\omega(x_0) := \begin{cases} \emptyset, & \text{if } t^+(x_0) < \infty, \\ \{y \in \Omega \mid \exists (t_l)_{l \in \mathbb{N}}, t_l \rightarrow \infty \text{ such that } \phi(t_l, x_0) \rightarrow y \text{ as } l \rightarrow \infty\}, & \text{if } t^+(x_0) = \infty. \end{cases}$$

Here, $J(x_0) :=]t^-(x_0), t^+(x_0)[$ denotes the open maximal interval of existence of the solution $\phi(t, x_0)$. We make constant use of the general properties of the limit sets.

Lemma 2.1 *Let $\phi(t, x_0)$, $t \geq 0$ be a solution of (2.1) possessing compact closure in Ω . Then $\omega(x_0)$ is a nonempty, invariant, connected and compact subset of the phase space. Furthermore, the relation*

$$\text{dist}(\phi(t, x_0), \omega(x_0)) \rightarrow 0 \text{ as } t \rightarrow \infty$$

holds.

For a proof of Lemma 2.1 see, e.g., Hale (1980), Ch. I.8.

Our first step is to set up dynamical systems possessing the properties (P1)-(P3) for problem (1.1). We assume for the moment that there are no restrictions of inequality type and consider the minimization problem (1.1), $p = 0$ under the smoothness assumptions

$$f \in C^2(D, \mathbb{R}), \quad g \in C^2(D, \mathbb{R}^l). \quad (2.2)$$

For the constrained condition $g(x) = 0$ we make the following regularity assumption, which is standard in minimization theory.

(R) There is $\tau > 0$ such that every $v \in \mathbb{R}^l$, $\|v\|_2 \leq \tau$ is a regular value of g .

Let A be a smooth family of symmetric, positive definite $(l \times l)$ -matrices such that $B(x) := Dg(x)Dg(x)^T A(x)$ satisfies

$$\inf\{\mu_2(-B(x)) \mid x \in D, \|g(x)\|_2 \leq \tau\} \leq -\eta, \eta > 0. \quad (2.3)$$

Here, the logarithmic norm $\mu_2(C)$ of a matrix C is defined by

$$\mu_2(C) := \lim_{\delta \rightarrow 0} \frac{1}{\delta} (\|I + \delta C\|_2 - 1).$$

Under these assumptions, the augmented family of systems reads

$$\begin{aligned} \dot{x} = G(x) &:= (I - Q(x))(-\nabla f(x)) - Q(x)Dg(x)^T A(x)g(x) \\ &= (I - Q(x))(-\nabla f(x)) - Dg(x)^T A(x)g(x), \\ Q(x) &:= Dg(x)^T (Dg(x)Dg(x)^T)^{-1} Dg(x) \end{aligned} \quad (2.4)$$

for initial values

$$x(0) = x_0 \in \Omega := \{x \in D \mid \|g(x)\|_2 < \tau\}. \quad (2.5)$$

The reader may notice that the assumptions (R), (2.2) assure $G \in C^1(\Omega, \mathbb{R}^N)$. $Q(x)$, respectively, $I - Q(x)$ denotes the orthogonal projector onto the normal space $N_x M_v$, respectively, the tangential space $T_x M_v$ of the manifold $M_v := g^{-1}(v)$ at x . Roughly spoken, we obtain the vectorfield of equation (2.4) by taking the tangential part of $-\nabla f$ and the normal part of $-Dg^T A g$.

In general, the matrix family A can be obtained by solving the Liapunov equation

$$Dg(x)Dg(x)^T A(x) + A(x)^T Dg(x)Dg(x)^T = W(x) \quad (2.6)$$

for an arbitrary smooth positive definite matrix family W . The theory of Liapunov equations (see, e.g., Lancaster and Tismenetsky (1985), Ch. 13, Theorem 2) then ensures that $A(x)$ is well defined and positive definite. Moreover, an application of the implicit function theorem guarantees the smoothness of the matrix family A .

On the other hand, there are two natural choices of A which avoid solving equation (2.6). The first choice is $W(x) = 2I$ or $A(x) = (Dg(x)Dg(x)^T)^{-1}$. This leads us to (1.7). The second possibility is $W(x) = 2Dg(x)Dg(x)^T$, that is, $A(x) = I$. To ensure $\eta > 0$ in this case, e.g., the assumption $\Omega := \{x \in D \mid \|g(x)\|_2 < \tau\}$ bounded is sufficient.

An important observation is now that, with $u(t) = g(\phi(t, x_0))$, $\alpha(t) = r(\phi(t, x_0))$ and

$r(x) = (Dg(x)Dg(x)^T)^{-1}Dg(x)\nabla f(x) - A(x)g(x)$ the initial value problem (2.4), (2.5) is equivalent to the differential algebraic equation (DAE)

$$\begin{aligned}\dot{x} &= -\nabla f(x) + Dg(x)^T\alpha, \\ \dot{u} &= -Dg(x)Dg(x)^TA(x)u, \\ 0 &= g(x) - u\end{aligned}\tag{2.7}$$

with the consistent initial values

$$x(0) = x_0 \in \Omega, \quad u(0) = g(x_0), \quad \alpha(0) = r(x_0)\tag{2.8}$$

on the existence interval $]t^-(x_0), t^+(x_0)[$. In addition, the regularity assumption (R) ensures that the DAE (2.7) possesses differential index 2 (for a definition see, e.g., Hairer et al. (1991), 476-478). Thus we call $(\bar{x}, g(\bar{x}), r(\bar{x}))$ an equilibrium, respectively, a stable equilibrium of (2.7), if \bar{x} is a stationary point, respectively, a stable stationary point of the differential equation (2.4).

The properties of the equilibria of (2.4) or x -components of time independent solutions of (2.7) are characterized in the following lemma.

Lemma 2.2 *Suppose the regularity condition (R), the smoothness condition (2.2) and $\eta > 0$ hold for the constrained minimization problem (1.1) with $p = 0$. In addition, let $\bar{x} [(\bar{x}, g(\bar{x}), r(\bar{x}))]$ be an equilibrium of (2.4) [(2.7)]. Then, \bar{x} satisfies $g(\bar{x}) = 0$, (1.3) for $p = 0$. If $\bar{x} [(\bar{x}, g(\bar{x}), r(\bar{x}))]$ is hyperbolic and stable for (2.4) [(2.7)], \bar{x} satisfies additionally (1.4) for $p = 0$.*

Here, \bar{x} is a hyperbolic equilibrium of (2.1) if $F(\bar{x}) = 0$ and $\text{Re}(\lambda) \neq 0$ for all eigenvalues $\lambda \in \mathbb{C}$ of $DF(\bar{x})$. The description of the longtime behaviour of evolutions of (2.4) and (2.7) reads as follows.

Theorem 2.3 *Suppose the regularity condition (R), the smoothness condition (2.2) and $\eta > 0$ hold for the constrained minimization problem (1.1) with $p = 0$ and consider the initial value problem (2.4), (2.5), respectively, the DAE (2.7), (2.8). We assume that these equations possess merely finitely many equilibria. Then, M_0 is a closed, invariant, globally attractive set for equation (2.4). Every bounded solution $\phi(t, x_0)$ of (2.4), respectively, $(\phi(t, x_0), g(\phi(t, x_0)), r(\phi(t, x_0)))$ of (2.7) exists for $t \geq 0$ and converges towards a steady state as $t \rightarrow \infty$.*

In the case of a general minimization problem (1.1) we assume $k \in C^2(D, \mathbb{R}^p)$. We apply a standard procedure (see, e.g., Spellucci (1993), Remark 1.2.3) and reduce the original problem to a smooth equality constrained problem. We introduce the so called slack variables y_1, \dots, y_p , the diagonal matrix $\text{diag}(y_1, \dots, y_p)$ and the C^2 -functions

$$\begin{aligned}\bar{f}(x, y) &:= f(x), \\ \bar{g}(x, y) &:= \begin{pmatrix} g(x) \\ k(x) - \text{diag}(y_1, \dots, y_p)y \end{pmatrix}\end{aligned}\tag{2.9}$$

which are defined on $D \times \mathbb{R}^p$ and consider the following minimization problem:

$$\text{Minimize } \bar{f}(x, y) \text{ subject to } \bar{g}(x, y) = 0. \quad (2.10)$$

The hidden difficulty with this formally elegant shift from (1.1) to (2.10) is, that it increases the number of Kuhn-Tucker points. Since (2.10) is again a smooth problem of the form (1.1) with $p = 0$, assuming the regularity condition (R) for $z = (x, y)$ the $(N + p)$ -dimensional dynamical system reads

$$\begin{aligned} \dot{z} &= H(z) := (I - \bar{Q}(z))(-\nabla \bar{f}(z)) - D\bar{g}(z)^T \bar{A}(z) \bar{g}(z), \\ \bar{Q}(z) &:= D\bar{g}(z)^T (D\bar{g}(z) D\bar{g}(z)^T)^{-1} D\bar{g}(z), \\ z(0) &= z_0 \in \Sigma := \{(x, y) \in D \times \mathbb{R}^p \mid \|\bar{g}(x, y)\|_2 < \tau\}. \end{aligned} \quad (2.11)$$

With $u = (v, w)$, $\bar{r} = (D\bar{g} D\bar{g}^T)^{-1} D\bar{g} \nabla \bar{f} - \bar{A} \bar{g}$ equation (2.11) is equivalent to the index 2 DAE

$$\begin{aligned} \dot{x} &= -\nabla f(x) + Dg(x)^T \alpha + Dk(x)^T \beta, \quad x(0) = x_0, \\ \dot{y} &= -2 \operatorname{diag}(y) \beta, \quad y(0) = y_0, \\ \begin{pmatrix} \dot{v} \\ \dot{w} \end{pmatrix} &= -D\bar{g}(x, y) D\bar{g}(x, y)^T \bar{A}(x, y) \begin{pmatrix} v \\ w \end{pmatrix}, \quad (v(0), w(0)) = \bar{g}(x_0, y_0), \\ 0 &= \begin{pmatrix} g(x) - v \\ (k(x) - \operatorname{diag}(y)y) - w \end{pmatrix}, \quad (\alpha(0), \beta(0)) = \bar{r}(x_0, y_0). \end{aligned} \quad (2.12)$$

Under the appropriate topological assumptions on Σ , Theorem 2.3 states that the long-time behaviour of the evolutions of (2.11), (2.12) is stationary. In addition, for the x -components of the equilibria of these systems the following characterization holds.

Lemma 2.4 *Suppose the regularity condition (R) and the smoothness condition (2.2) hold for the minimization problem (1.1). Then, for every equilibrium of (2.11) or (2.12) the x -component \bar{x} satisfies $g(\bar{x}) = 0$, $k(\bar{x}) \geq 0$ and the first two statements in (1.3). For every hyperbolic, stable equilibrium of (2.11), respectively, (2.12), the x -component \bar{x} satisfies additionally $\bar{\beta} \geq 0$ and (1.4).*

Now, let us justify the introduction of slack variables. Lemma 2.4 and Theorem 2.3 ensure that the ODE (2.11) and the DAE (2.12) are equivalent to the minimization problem (1.1) up to the fact that stationary points automatically do not fulfil $\bar{\beta} \geq 0$. This problem, which arises from the introduction of slack variables, has an elegant solution in the ODE- and DAE-approach. It is well known that, generically often, evolutions of dynamical systems with property (P1) saddle down at stable equilibria. But according to Lemma 2.4 these equilibria fulfil the sufficient conditions (P2) and (P3) for a local solution of the corresponding minimization problem (1.1). This means, that formally, (2.11), (2.12) may possess stationary points whose x -components are no Kuhn-Tucker

points, but this does not matter the forward dynamics of almost every trajectory.

In particular, we can compute solutions of (1.1) by integrating equation (2.11) or (2.12) with a numerical integration method. The qualitative features of the discretized version of equation (2.11) with an arbitrary one- or linear multistep method are analyzed in Schropp (1998). There we show that, for sufficiently small stepsize, the resulting discrete dynamical system inherits all decisive properties from the continuous differential equation. An analogous result for the discretization of the DAE (2.12) is under construction.

Another interesting feature of the dynamical systems approach is that by Lemma 2.4 no extra work is necessary to determine whether the computed point is a solution of (1.1) or merely a saddle point. In addition, $\bar{g}^{-1}(0)$ is a globally attractive set for equation (2.11). Thus, it is not necessary that the x -component of the starting point is feasible. Nevertheless, from the efficiency point of view, a disadvantage of adding slack variables is that with every inequality constraint the dimension of the resulting dynamical system (2.11) increases by one.

From the way of introduction, the reader might think that the ODE- or the DAE-approach is completely different from the classical SQP-approach. Our aim here to show is that, despite the motivation for both methods is completely different, the resulting algorithms are very similar. To be more precise, we show that the SQP-method for solving a minimization problem (1.1) with $p = 0$ can be regarded as variable step size Euler-Cauchy method applied to an ordinary differential equation whose solution flow satisfies (P1)-(P3). In the SQP-approach for equality constrained problems one solves the parameter dependent free problem defined by the objective function

$$\Phi(x, \alpha) := f(x) + \sum_{i=1}^l \alpha_i |g_i(x)|, \quad \alpha = (\alpha_1, \dots, \alpha_l) > 0 \text{ sufficiently big} \quad (2.13)$$

using a descent method $x_{n+1} = x_n + \gamma_n p_n$, $\gamma_n > 0$. The descent direction p_n is computed by solving the following quadratic subproblem:

$$\begin{aligned} \text{Minimize } f_n(p) &:= f(x_n) + \nabla f(x_n)^T p + \frac{1}{2} p^T \Delta(x_n) p, \quad \Delta(x) \text{ positive definite} \\ \text{subject to } g_n(p) &:= g(x_n) + Dg(x_n)p = 0. \end{aligned} \quad (2.14)$$

Next we compare the SQP- with the ODE-approach.

Lemma 2.5 *Assume the regularity assumption (R) holds. Then, the SQP-method (2.13), (2.14) applied to a minimization problem (1.1) with $p = 0$ is equivalent to a variable step size Euler-Cauchy method applied to $\dot{x} = C(x)^{-1}G(x)$, $x(0) = x_0$ with G from (2.4) and $C(x) := (I - Q(x))\Delta(x) + Dg(x)^T A(x) Dg(x)$. Moreover, the solution flow of $\dot{x} = C(x)^{-1}G(x)$ satisfies (P1)-(P3).*

Let us draw the conclusions from Lemma 2.5. Since the equation $C(x)p = G(x)$ is not solved explicitly, the SQP-approach works very efficient. Nevertheless, Euler-Cauchy is

an explicit time stepping method. Thus, there may arise trouble, if $\dot{x} = G(x)$ is a stiff dynamical system. In this case, the SQP-preconditioner $C(x)^{-1}$ must be designed to avoid stiffness of $\dot{x} = C(x)^{-1}G(x)$. So, there might be advantages for implicit time stepping methods if preconditioning via $C(x)$ fails.

3 Proof of the main results

Proof of Lemma 2.2: Let $\bar{x} \in \Omega$ be a stationary point of (2.4), respectively, $(\bar{x}, g(\bar{x}), r(\bar{x}))$ be a time independent solution of (2.7). Equation (2.7) directly implies $g(\bar{x}) = 0$ and $\nabla f(\bar{x}) = Dg(\bar{x})^T \bar{\alpha}$ with $\bar{\alpha} = r(\bar{x})$.

Now, let $\bar{x} \in \Omega$ be a hyperbolic stable equilibrium of (2.4), that is, $G(\bar{x}) = 0$ and $\text{Re}(\lambda) < 0$ for every $\lambda \in \sigma(DG(\bar{x}))$. Using the chain- and product-rules, we can calculate

$$\begin{aligned} DG(\bar{x}) &= -(I - Q(\bar{x}))L(\bar{x}) - Dg(\bar{x})^T A(\bar{x})Dg(\bar{x}), \\ L(x) &= \nabla^2 f(x) - \sum_{i=1}^l \bar{\alpha}_i \nabla^2 g_i(x). \end{aligned}$$

Let $V = (v_1, \dots, v_{N-l}) \in \mathbb{R}^{N, N-l}$ be an orthonormal basis of $R(I - Q(\bar{x}))$, let $W = (w_1, \dots, w_l) \in \mathbb{R}^{N, l}$ be an orthonormal basis of $R(Dg(\bar{x})^T)$ and let $S = (V, W) \in \mathbb{R}^{N, N}$. With $Q(\bar{x})V = 0$, $(I - Q(\bar{x}))W = 0$ the matrix $S^{-1}DG(\bar{x})S$ computes to

$$S^{-1}DG(\bar{x})S = \begin{pmatrix} -V^T L(\bar{x})V & -V^T L(\bar{x})W \\ 0 & -W^T Dg(\bar{x})^T A(\bar{x})Dg(\bar{x})W \end{pmatrix}$$

and the spectral relation

$$\sigma(DG(\bar{x})) = \sigma(-V^T L(\bar{x})V) \cup \sigma(-W^T Dg(\bar{x})^T A(\bar{x})Dg(\bar{x})W)$$

follows. Since $W^T Dg(\bar{x})^T A(\bar{x})Dg(\bar{x})W$ is an inertia transformation of $A(\bar{x})$, we obtain

$$\langle \alpha, -V^T L(\bar{x})V \alpha \rangle = -\langle V \alpha, L(\bar{x})V \alpha \rangle < 0 \quad \forall \alpha \in \mathbb{R}^{N-l} \setminus \{0\}$$

for a hyperbolic stable equilibrium \bar{x} of equation (2.4). This implies (1.4) for $p = 0$, since $R(V) = N(Dg(\bar{x}))$.

Proof of Theorem 2.3: Let $x_0 \in \Omega$ be an initial value for equation (2.4), respectively, $(x_0, g(x_0), r(x_0))$ be a consistent initial value for the DAE (2.7). Due to (2.7) the function $g(\phi(\cdot, x_0))$: $J(x_0) \rightarrow \mathbb{R}^l$ solves the initial value problem

$$\begin{aligned} \dot{u} &= -Dg(\phi(t, x_0))Dg(\phi(t, x_0))^T A(\phi(t, x_0))u = -B(\phi(t, x_0))u, \\ u(0) &= g(x_0). \end{aligned}$$

Using (2.3) and Theorem 5.1.3 in Strehmel and Weiner (1995) with $v = 0$, we obtain

$$\|g(\phi(t, x_0))\|_2 \leq \|g(x_0)\|_2 \exp(-\eta t), \quad t \in [0, t^+(x_0)[, \quad x_0 \in \Omega. \quad (3.1)$$

Formula (3.1) ensures that for every $x_0 \in \Omega$ the closed set $C(x_0) := \{x \in \Omega \mid \|g(x)\|_2 \leq \|g(x_0)\|_2\}$ is positive invariant. Thus, $t^+(x_0) = \infty$ and $\emptyset \neq \omega(x_0)$ follow for every bounded solution. In addition, the set $M_0 = g^{-1}(0)$ is a closed, invariant, globally attractive subset of Ω . Hence, we obtain $\omega(x_0) \subset M_0 \forall x_0 \in \Omega$.

Next we characterize the dynamics of equation (2.4) restricted to the subset M_0 . For $x_1 \in M_0$ we can compute

$$\frac{\partial}{\partial t} f(\phi(t, x_1)) = - \|G(\phi(t, x_1))\|_2^2 \leq 0, \quad t \in [0, t^+(x_1)],$$

that is, f is a Lyapunov function of the dynamical system (2.4) with phase space M_0 . To characterize $\omega(x_1)$ for $x_1 \in M_0$ we use the invariance principle of La Salle (La Salle (1976), Ch. 2, Theorem 6.4). This yields

$$G(\tilde{x}) = 0 \quad \forall \tilde{x} \in \omega(x_1). \quad (3.2)$$

Since (2.4) possesses only finitely many equilibria, formula (3.2) assures

$$\omega(M_0) := \overline{\cup_{x \in M_0} \omega(x)} = \{\bar{x}_1, \dots, \bar{x}_r\} \quad \text{for some } r \in \mathbb{N}.$$

Now, our aim is to apply the ω -reduction Theorem 2.1 of Schropp (1996) to ensure

$$\omega(x_0) \subset \omega(M_0) = \{\bar{x}_1, \dots, \bar{x}_r\} \quad \forall x_0 \in \Omega. \quad (3.3)$$

Therefore we have to show that the sets $\{\bar{x}_i\}$, $i = 1, \dots, r$ are isolated invariant in the sense of Conley (1976), Ch. I and that equation (2.4) possesses no closed cycle in M_0 .

The reader may recall that $\{\bar{x}_i\}$ is called isolated invariant if a neighbourhood V of \bar{x}_i exists such that $\{\bar{x}_i\}$ is the maximal invariant set of the underlying dynamical system with phase space V .

Let V be a neighbourhood of \bar{x}_i and let $W \subset V$ be invariant. Then (3.1) implies $W \subset (M_0 \cap V)$. Since f is a Lyapunov function of equation (2.4) with phase space M_0 , we obtain $W = \{x_i\}$ for V sufficiently small.

Now, suppose that $\gamma(x_0) := \{\phi(t, x_0) \mid t \in \mathbb{R}\}$, $G(x_0) \neq 0$ is a connecting orbit from \bar{x}_i to \bar{x}_j in M_0 . The Lyapunov property then ensures $f(\bar{x}_i) > f(\bar{x}_j)$ and, hence, a closed cycle of (2.4) in M_0 is impossible. In particular, Theorem 2.1 in Schropp (1996) is applicable and (3.3) is shown.

The properties of ω -limit sets (see Lemma 2.1) then imply

$$\omega(x_0) = \{\bar{x}_{i_0}\} \quad \text{for some } i_0 \in \{1, \dots, r\},$$

that is, $\lim_{t \rightarrow \infty} \phi(t, x_0) = \bar{x}_{i_0}$ and Theorem 2.3 is proved.

Proof of Lemma 2.4: Let $\bar{z} = (\bar{x}, \bar{y})$ be an equilibrium of equation (2.11). The DAE (2.12) then directly shows $g(\bar{x}) = 0$, $k(\bar{x}) \geq 0$ as well as

$$\begin{aligned} \nabla f(\bar{x}) &= Dg(\bar{x})^T \bar{\alpha} + Dk(\bar{x})^T \bar{\beta}, \\ 0 &= -2 \operatorname{diag}(\bar{y}_1, \dots, \bar{y}_p) \bar{\beta} \end{aligned} \quad (3.4)$$

and the first two statements in (1.3) are verified.

Now, let $\bar{z} = (\bar{x}, \bar{y})$ be a hyperbolic stable equilibrium of equation (2.11). Differentiation of H yields

$$\begin{aligned} DH(\bar{z}) &= -(I - \bar{Q}(\bar{z}))\bar{L}(\bar{z}) - D\bar{g}(\bar{z})^T \bar{A}(\bar{z}) D\bar{g}(\bar{z}), \\ \bar{L}(\bar{z}) &= \begin{pmatrix} \tilde{L}(\bar{x}) & 0 \\ 0 & 2 \operatorname{diag}(\bar{\beta}_1, \dots, \bar{\beta}_p) \end{pmatrix}, \\ \tilde{L}(\bar{x}) &:= \nabla^2 f(\bar{x}) - \sum_{i=1}^l \bar{\alpha}_i \nabla^2 g_i(\bar{x}) - \sum_{i=1}^p \bar{\beta}_i \nabla^2 k_i(\bar{x}) \end{aligned}$$

and the application of Lemma 2.2 guarantees that $\bar{L}(\bar{z})$ is positive definite on $N(D\bar{g}(\bar{z}))$. Next we show $\bar{\beta}_i \geq 0$.

In the case $\bar{y}_i \neq 0$ formula (3.4) shows everything.

Now, let $\bar{y}_i = 0$ hold for some $i \in \{1, \dots, p\}$. Then we see $(0, e_i) \in N(D\bar{g}(\bar{z}))$. Here e_i denotes the i -th unit vector. Thus, we obtain $0 < (0, e_i)^T \bar{L}(\bar{x}, \bar{y})(0, e_i) = 2\bar{\beta}_i$.

With Λ from (1.4) it remains to show $v^T \tilde{L}(\bar{x})v > 0 \quad \forall v \in (N(Dg(\bar{x})) \cap \Lambda) \setminus \{0\}$.

Let $\hat{v} \in (N(Dg(\bar{x})) \cap \Lambda) \setminus \{0\}$. With $\hat{w} = (\hat{w}_1, \dots, \hat{w}_p) \in \mathbb{R}^p$,

$$\hat{w}_j = \begin{cases} 0, & \text{if } \bar{y}_j = 0, \\ \frac{1}{2}(\bar{y}_j)^{-1} Dk_j(\bar{x})\hat{v}, & \text{if } \bar{y}_j \neq 0, \end{cases} \quad j = 1, \dots, p$$

the relation $(\hat{v}, \hat{w}) \in N(D\bar{g}(\bar{x}, \bar{y})) \setminus \{0\}$ holds and $0 < (\hat{v}, \hat{w})^T \bar{L}(\bar{x}, \bar{y})(\hat{v}, \hat{w}) = \hat{v}^T \tilde{L}(\bar{x})\hat{v}$ follows. This proves Lemma 2.4.

Proof of Lemma 2.5: To compute the descent direction p_n , we set up the Lagrangian equations for the quadratic subproblem (2.14). Next we use the regularity assumption (R) to eliminate the Lagrangian multiplier variables. Finally, with G from (2.4) we end up with the equation

$$C(x)p := ((I - Q(x))\Delta(x) + Dg(x)^T A(x) Dg(x))p = G(x) \quad (3.5)$$

for $x = x_n$ to determine the descent direction p_n .

To show that $C(x)$ is invertible, let $V = (v_1, \dots, v_{N-l}) \in \mathbb{R}^{N, N-l}$ be an orthonormal base of $R(I - Q(x))$, let $W = (w_1, \dots, w_l) \in \mathbb{R}^{N, l}$ be an orthonormal base of $R(Dg(x)^T)$ and $S = (V, W) \in \mathbb{R}^{N, N}$. We calculate

$$S^{-1}C(x)S = \begin{pmatrix} V^T \Delta(x) V & V^T \Delta(x) W \\ 0 & W^T Dg(x)^T A(x) Dg(x) W \end{pmatrix} \quad (3.6)$$

and see that $C(x)$ is invertible.

It remains to show that the solution flow of $\dot{x} = C(x)^{-1}G(x)$ fulfills (P1)-(P3).

First we remark that condition (P2) is obvious by Lemma 2.2. Now let \bar{x} be a stationary point of $\dot{x} = C(x)^{-1}G(x)$. An easy calculation shows

$$\frac{d}{dx}(C(x)^{-1}G(x)) \Big|_{x=\bar{x}} = C(\bar{x})^{-1}DG(\bar{x}).$$

To determine the spectrum of $C(\bar{x})^{-1}DG(\bar{x})$, we compute

$$S^{-1}C(\bar{x})^{-1}DG(\bar{x})S = \begin{pmatrix} -(V^T\Delta(\bar{x})V)^{-1}V^TL(\bar{x})V & * \\ 0 & -I \end{pmatrix}$$

with a not specified entry up right. In particular, $(V^T\Delta(\bar{x})V)^{-1} \in \mathbb{R}^{N-l, N-l}$ is a symmetric, positive definite matrix and

$$\sigma((V^T\Delta(\bar{x})V)^{-1}V^TL(\bar{x})V) = \sigma(((V^T\Delta(\bar{x})V)^{-1/2})^TV^TL(\bar{x})V(V^T\Delta(\bar{x})V)^{-1/2})$$

holds. Thus, $(V^T\Delta(\bar{x})V)^{-1}V^TL(\bar{x})V$ and $V^TL(\bar{x})V$ have the same inertia of eigenvalues and Lemma 2.2 guarantees (P3).

To analyze the long time behaviour of the solution flow $\psi(t, x_0)$ of $\dot{x} = C(x)^{-1}G(x)$, the reader may notice

$$\frac{\partial}{\partial t}g(\psi(t, x_0)) = Dg(\psi(t, x_0))C(\psi(t, x_0))^{-1}G(\psi(t, x_0)) = -g(\psi(t, x_0)). \quad (3.7)$$

Formula (3.7) ensures, that $g^{-1}(0)$ is an invariant and globally attractive subset of the phase space. For $x_1 \in g^{-1}(0)$ we can compute

$$\begin{aligned} \frac{\partial}{\partial t}f(\psi(t, x_1)) &= \nabla f(\psi(t, x_1))^TC(\psi(t, x_1))^{-1}G(\psi(t, x_1)) \\ &= -G(\psi(t, x_1))^TC(\psi(t, x_1))^{-1}G(\psi(t, x_1)). \end{aligned}$$

Now, let $x \in g^{-1}(0)$ arbitrary, and let $G(x) = S(\alpha, 0)$, $\alpha \in \mathbb{R}^{N-l}$ with $S = (V, W)$. Using (3.6) we obtain

$$-G(x)^TC(x)^{-1}G(x) = -\alpha^T(V^T\Delta(x)V)^{-1}\alpha \begin{cases} \leq 0 \quad \forall \alpha \in \mathbb{R}^{N-l}, \\ = 0 \quad \longleftrightarrow \quad \alpha = 0. \end{cases}$$

Hence, $\frac{\partial}{\partial t}f(\psi(t, x_1)) \leq 0$ and $\frac{\partial}{\partial t}f(\psi(t, x_1)) = 0 \longleftrightarrow G(\psi(t, x_1)) = 0$ follows. Thus we can adapt the proof of Theorem 2.3 to this situation and guarantee (P1).

4 Numerical implementations and applications

In this section we compare the dynamical systems approach with the classical minimization methods. For unconstrained minimization problems we refer the reader to Schropp (1995), (1997).

In the constraint case, we start by analyzing the different approaches structurally. Due to Lemma 2.5, a concrete SQP-realization can be regarded as variable step size Euler-Cauchy method applied to a certain ODE. In the ODE-approach we have to integrate the suggested ODE (2.4) for $t \rightarrow \infty$. Thus, it seems quite natural to select an implicit A or $A(\alpha)$ -stable time stepping method and use their facilities to make huge step sizes. The

main difference between a time stepping method applied to the ODE (2.4) and an SQP method is the determination of the step length. Time stepping methods adjust the step length appropriate to the local error, whereas SQP methods estimate the step length geometrically by analyzing the surface defined by the penalty function Φ from (2.13). This may allow bigger step sizes and lead to more efficiency. In addition, the evaluation of the right hand side of the ODE (2.4) seems to be costly, because a $(l \times l)$ -system of linear equations must be solved. Another advantage of the SQP-method is that it works directly with $f, g, \nabla f, Dg$ and, hence, one can exploit the structures of these functions. On the other hand, by changing the local error parameter of a concrete time stepping method, in the ODE case the user is able to balance out the criteria efficiency and reliability.

At this point, the DAE-approach (2.7) may be of interest. Obviously, for $A(x) = (Dg(x)Dg(x)^T)^{-1}$ this approach directly deals with $f, g, \nabla f$ and Dg and the computation of the projector Q is superfluous. Moreover, this efficient evaluation of the right hand side of (2.7) is combined with the reliability of the ODE-approach. Nevertheless, the dimension of (2.7) is big. One possibility to reduce the dimension of the DAE (2.7) with $A(x) = (Dg(x)Dg(x)^T)^{-1}$ is to use the index 1 formulation with eliminated u -variables. This yields to

$$\begin{aligned} \dot{x} &= -\nabla f(x) + Dg(x)^T \alpha, \\ 0 &= Dg(x)(-\nabla f(x) + Dg(x)^T \alpha) + g(x) \end{aligned} \tag{4.1}$$

with consistent initial values

$$x(0) = x_0 \in \Omega, \quad \alpha(0) = r(x_0). \tag{4.2}$$

Our first test example is the computation of the steady states of a symmetric hydrostatic skeleton with N segments and M parts between which no exchange of fluidum is possible. The general unsymmetric model was developed by Beyn, Wadepuhl (1989). If we denote the total energy of the skeleton in a state x by $E_{tot}(x)$, the actual volume of the i -th part K_i by $V_{K_i}(x)$ and the prescribed total volume with $V_{tot}^{K_i}$, the mathematical description of the possible steady states of the symmetric skeleton reads as follows:

$$\text{Minimize } E_{tot}(x) \text{ subject to } V^{K_i}(x) - V_{tot}^{K_i} = 0, \quad i = 1, \dots, M. \tag{4.3}$$

Here, a state $x = (b_1, c_1, \dots, b_N, c_N, b_{N+1}) \in \mathbb{R}^{2N+1}$ has the total energy

$$E_{tot}(x) = 4 \left(\sum_{j=1}^{N+1} P_{\epsilon_{R_j}}(b_j, \beta_j) + \sum_{j=1}^N P_{\epsilon_{L_j}}(c_j, \gamma_j) \right),$$

provided the single muscle energy is defined by

$$P_\epsilon(L, \alpha) = \int_{L_0}^L F_\epsilon(l, \alpha) dl$$

with the force law

$$F_\epsilon(L, \alpha) = \frac{\alpha}{\epsilon} \tan(\epsilon(L - L_0)), \quad L_0 - \frac{\pi}{2\epsilon} < L < L_0 + \frac{\pi}{2\epsilon}.$$

For our actual computations we chose the parameters $L_0 = \alpha = 1$, $\beta_j = \gamma_j = 1 \ \forall j$, $\epsilon_{L_j} = \epsilon_{R_j} = .15 \ \forall j$. Additionally, for $N = 21$ and $M = 3$ we put the segments 1-7, 8-14 and 15-21 to one part together and require $V^{K_i}(x) - 1000/3 = 0$ for $i = 1, 2, 3$. In this parameter specification the problem (4.3) has two local solutions and one saddle point.

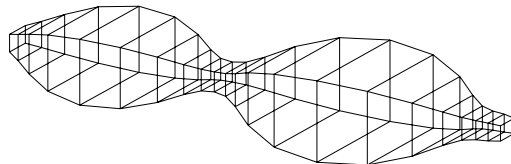


Figure 1: Solution of (4.3) for $N = 21$, $M = 3$

Now we solve (4.3) with the dynamical systems, the differential algebraic and the SQP-approach. To be more precise, we take the NAG-routine E04UCF as efficient SQP realization and we apply NAG-routine D02NGF onto the initial value problem (2.4), (2.5) with $A(x) = I$ or $A(x) = (Dg(x)Dg(x)^T)^{-1}$ or onto (4.1), (4.2) in the DAE case. The NAG-routine D02NGF is driven with the option, that the nonlinear systems of equations in every time step of the integration are solved with functional iterations instead of Newtons method. All codes are supplied with subroutines for calculating f , g and their first derivatives. Then our results are as follows.

For a lot of initial values the SQP-routine was able to solve (4.3). Nevertheless, for initial values x_0 possessing the symmetry $(x_0)_i = (x_0)_{2N+1-i}$, $i = 1, \dots, 2N + 1$ the iterates of the SQP-routine used converge towards a saddle instead of a solution. Moreover, the SQP-routine claims via failure parameter $IFAIL = 0$ that the computed point is a local solution of (4.3).

If we solve (4.3) with the NAG-routine D02NGF applied onto (2.4), (2.5) or (4.1), (4.2) our results decisively depend on the error parameter $ATOL$ and $RTOL$. With $ATOL = RTOL = 10^{-4}$ the results are the same as in the SQP-case. On the other hand, with $ATOL = RTOL = 10^{-8}$ the dynamical systems and the DAE-approach were able to solve (4.3) for arbitrary initial values. Depending on the initial value, the SQP-method and the DAE-approach with $ATOL = RTOL = 10^{-4}$ need between 0.5 and 2 seconds, whereas the dynamical systems approach with the same tolerance needs 2 to 10 seconds on a HP-workstation 715/64. Moreover, with the smaller tolerance $ATOL = RTOL = 10^{-8}$ the ODE- and the DAE-approach need about double the time than with coarse tolerance.

Our second example is the optimization of an alkylation process in the chemotechnical industry. A detailed description of that chemical process can be found in Bracken

and Mc Cormick (1968), Ch. 4. The concrete mathematical model then reads as follows.
Minimize

$$f(x) := 5.04x_1 + 0.035x_2 + 10x_3 + 3.36x_5 - 0.063x_4x_7 \quad (4.4)$$

subject to the constraints

$$\begin{aligned} 1.22x_4 - x_1 - x_5 &= 0, \\ \frac{98000x_3}{x_4x_9 + 1000x_3} - x_6 &= 0, \\ \frac{x_2 + x_5}{x_1} - x_8 &= 0, \\ 35.82 - 0.222x_{10} - bx_9 &\geq 0, \quad b = 0.9, \\ -133 + 3x_7 - ax_{10} &\geq 0, \quad a = 0.99, \\ -35.82 + 0.222x_{10} + \frac{1}{b}x_9 &\geq 0, \\ 133 - 3x_7 + \frac{1}{a}x_{10} &\geq 0, \\ 1.12x_1 + 0.13167x_1x_8 - 0.00667x_1x_8^2 - ax_4 &\geq 0, \\ 57.425 + 1.098x_8 - 0.038x_8^2 + 0.325x_6 - ax_7 &\geq 0, \\ -1.12x_1 - 0.13167x_1x_8 + 0.00667x_1x_8^2 + \frac{1}{a}x_4 &\geq 0, \\ -57.425 - 1.098x_8 + 0.038x_8^2 - 0.325x_6 + \frac{1}{a}x_7 &\geq 0, \end{aligned} \quad (4.5)$$

$$\begin{aligned} 0 \leq x_1 \leq 2000, \quad 85 \leq x_6 \leq 93, \\ 0 \leq x_2 \leq 16000, \quad 90 \leq x_7 \leq 95, \\ 0 \leq x_3 \leq 120, \quad 3 \leq x_8 \leq 12, \\ 0 \leq x_4 \leq 5000, \quad 1.2 \leq x_9 \leq 4, \\ 0 \leq x_5 \leq 2000, \quad 145 \leq x_{10} \leq 162. \end{aligned} \quad (4.6)$$

This is a problem with a few equality and a big number of inequality constraints. It possesses the solution

$$\bar{x} = (1698.1, 15818, 54.103, 3031.2, 2000, 90.115, 95, 10.493, 1.561, 153.53)$$

(see also Spellucci (1993), p.374). Now we solve problem (4.4)-(4.6) with the SQP-, the ODE- and the DAE-approach. It turns out that this problem has no hidden difficulty and so our three methods are able to solve this problem appropriately for all suitable initial values. In addition, a coarse tolerance for *ATOL* and *RTOL* is enough in the ODE- and DAE-approach. Depending on the initial values, the SQP-method needs between 1 and 2, the ODE-approach needs between 45 and 90 and the DAE-approach needs about 2 to 4 seconds. Thus, there are small efficiency advantages for the SQP-method in comparison

with the DAE-approach, but huge advantages if we compare with the ODE-approach. The reason for this is that the 20 bounds (4.6) on the independent variables are treated very efficient in the SQP-case, efficient in the DAE-case but the ODE-approach cannot take advantage of the special structure of these constraints.

Summarizing our results, we can say that SQP-methods are valuable and efficient tools to solve optimization problems, whereas the ODE-approach may be more reliable. From the efficiency point of view, the ODE-approach seems to be more appropriate for nonlinear problems with only a few number of constraints. But the DAE-approach for minimization problems combines efficiency with reliability in a prosperous way.

References

Amaya, J.A. (1987): A differential equations approach to function minimization, *Rev. Mat. Apl.* **9**, 1-7.

Beyn, W.J., Wadepuhl, M. (1989): Computer Simulation of the Hydrostatic Skeleton, *The Physical Equivalent, Mathematics and Application to Wormlike Forms*, *J. theor. Biol.* **136**, 379-402.

Botsaris, C.A. (1978): Differential gradient methods, *J. Math. Anal. Appl.* **63**, 177-198.

Botsaris, C.A. (1979): A Newton type curvilinear search method for constrained optimization, *J. Math. Anal. Appl.* **69**, 372-397.

Botsaris, C.A. (1981): A class of differential descent methods for constrained optimization, *J. Math. Anal. Appl.* **79**, 96-112.

Bracken, J., Mc Cormick, G.P. (1968): *Selected Applications of Nonlinear Programming*, Wiley, New York.

Brown, A.A., Bartholomew-Biggs, M.C. (1987): ODE versus SQP methods for constrained optimization, Technical report No. 179, The Hatfield polytechnic optimization centre.

Brown, A.A., Bartholomew-Biggs, M.C. (1989): ODE versus SQP methods for constrained optimization, *J.O.T.A.* **62**, 371-386.

Conley, C. (1976): *Isolated invariant sets and the morse index*, CBMS Regional Conference Series in Mathematics, Vol. **38**.

Coppel, W.A. (1965): *Stability and asymptotic behaviour of differential equations*, D.C Heath and Company, Boston Englewood Chicago.

- Evtushenko, Y.G., Zhadan V.G. (1994): Stable barrier-projection and barrier-Newton methods in nonlinear programming, *Optimization Methods and Software* **3**, 237-256.
- Gill, P.E., Murray, W., Wright, M.H. (1981): *Practical Optimization*, Academic Press, London.
- Hairer, E., Wanner, G. (1991): *Solving ordinary differential equations II*, Springer Verlag.
- Hale, J.K. (1980): *Ordinary differential equations*, second edition, Wiley-Interscience, New York.
- Kuhn, H.W., Tucker A.W. (1951): Nonlinear Programming, Proc. 2, Berkeley Symp. Math. Stat. Prob., Univ. Cal. Press, 481-492.
- Lancaster, P., Tismenetsky, M. (1985): *The theory of matrices*, second edition, Academic Press, New York London.
- La Salle, J.P. (1976): *The stability of dynamical systems*, Regional Conference Series in applied Mathematics **25**, SIAM.
- Schropp, J. (1995): Using dynamical systems methods to solve minimization problems, *Applied Numerical Mathematics* **18**, 321-335.
- Schropp, J. (1996): A reduction principle for ω -limit sets, *Z. angew. Math. Mech.* **76**, 349-356.
- Schropp, J. (1997): A note on minimization problems and multistep methods, *Numer. Math.* **78**, 87-101.
- Schropp, J. (1998): One- and Multistep procedures for constrained minimization problems, to appear in *IMA Journal of Numerical Analysis*.
- Spellucci, P. (1993): *Numerische Verfahren der nichtlinearen Optimierung*, Birkhäuser Verlag, Basel Boston Berlin.
- Strehmel, K., Weiner, R. (1995): *Numerik gewöhnlicher Differentialgleichungen*, B. G. Teubner, Stuttgart.
- Tanabe, K. (1974): An algorithm for constrained maximization in nonlinear programming, *Journal of the Operations Research Society of Japan* **17**, 184-201.
- Yamashita, H. (1980): A differential equations approach to nonlinear programming, *Math. Prog.* **18**, 155-168.