Optimal step length for the Newton method near the minimum of a self-concordant function

Roland Hildebrand

Optimal step length for the Newton method near the minimum of a self-concordant function

Roland Hildebrand *

April 6, 2020

Abstract

The quadratic convergence region of the exact Newton method around the minimum of a self-concordant function makes up a fraction of the Dikin ellipsoid. Outside of this region, the Newton method has to be damped in order to ensure convergence. However, the available estimates of both the size of the convergence region and the step length to be used outside of it are based on conservative relations between the Hessians at different points and are hence sub-optimal. In this contribution we use methods of optimal control theory to compute the optimal step length of the Newton method on the class of self-concordant functions, as a function of the Newton decrement. With this step length quadratic convergence can be achieved on the whole Dikin ellipsoid. The exact bounds are expressed in terms of solutions of ordinary differential equations which cannot be integrated explicitly. As an application, the neighbourhood of the central path in which the iterates of path-following methods for conic programming are required to stay can be enlarged, enabling faster progress along the central path during each iteration and hence fewer iterations to achieve a given accuracy.

1 Introduction

The Newton method is a century-old second order method to find a zero of a vector field or a stationary point of a sufficiently smooth function. It is well-known that it is guaranteed to converge only in a neighbourhood of a solution, even on such well-behaved classes of objective functions as the self-concordant functions. In this neighbourhood the method converges quadratically. Outside the quadratic convergence region the step length has to be decreased to ensure convergence, leading to the *damped* Newton method. Several rules have been proposed to choose the step length on different classes of cost functions or vector fields. Among these are line searches or path searches until some pre-defined condition is met [1, 8], strategies imported from gradient descent methods [6, p. 37], and explicit formulas [7, p. 24].

These rules provide sufficient conditions for convergence, but they are optimal only if the order of convergence is concerned [6]. Given the fast convergence rate of the Newton method, the performance estimate of a single iterate is not important if the problem has to be solved to optimality with a starting point already in the quadratic convergence region. However, the situation is different if the Newton method is applied as it is in path-following methods for solving conic programs. Here we are not interested in finding the exact minimum, rather one or a few iterations are made before changing the cost function, thus moving the iterate again farther away from the minimum. In this regime the iterates permanently stay in the outskirts of the quadratic convergence region, and the overall progress is limited by the requirement that they do not drop out of it. Exact knowledge of the worst-case performance of a single iterate and of the boundaries of the convergence region would allow to significantly enlarge the step size when updating the cost function, thus boosting the overall convergence of the path-following method.

We shall consider the class of *self-concordant* functions. This class has been introduced in [7], because due to its invariance under affine coordinate transformations it is especially well suited for analysis in conjunction with the also affinely invariant Newton method. Self-concordant functions are locally strongly convex C^3 functions F which satisfy the inequality

$$|F'''(x)[h,h,h]| \leq 2(F''(x)[h,h])^{3/2}$$

*Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France (roland.hildebrand@univ-grenoble-alpes.fr).

for all x in the domain of definition and all h in the tangent space at x. The *Dikin ellipsoid* of F around a point $\hat{x}$ is given by the set

$$W(\hat{x}) = \left\{ x \mid ||x - \hat{x}||_{\hat{x}} = \sqrt{(x-\hat{x})^T F''(\hat{x})(x-\hat{x})} \leq 1 \right\}.$$

Here the distance between the points is measured in the local metric given by the Hessian $F''(\hat{x})$.

The basic iteration of a path-following method consists of a (damped) Newton step

$$x_{k+1} = x_k - \gamma_k (F''(x_k))^{-1} F'(x_k)$$

towards the minimum of a composite self-concordant function F, where $\gamma_k = 1$ for a full step and $\gamma_k \in (0,1)$ if the step is damped. This function is a sum of the self-concordant barrier used to describe the feasible set and the linear cost function of the conic problem multiplied with a positive weight. The function F hence depends on a scalar parameter, given by the weight, which is updated after one or a few Newton steps. The minima of the family of composite functions form a path, the so-called *central path* of the problem. The iterates must stay in a small neighbourhood of the central path to guarantee inclusion in the quadratic convergence region around the minimum. The parameter must be updated cautiously enough so that the iterate still remains inside the convergence region around the new minimum.

The estimate of the size of the quadratic convergence region and the performance of the Newton step is based on the following bound on the Hessian of a self-concordant function F at a point x near the current iterate x_k [7, Theorem 2.1.1] (see also [6, Theorem 5.1.7])

$$(1 - ||x - x_k||_{x_k})^2 F''(x_k) \preceq F''(x) \preceq (1 - ||x - x_k||_{x_k})^{-2} F''(x_k).$$

Here $||x - x_k||_k \in (0,1)$ is the distance between x and x_k as measured in the local metric at x_k. Based on this bound, a full Newton step can be made safely if its length

$$\begin{aligned} ||x_{k+1} - x_k||_{x_k} &= \sqrt{((F''(x_k))^{-1}F'(x_k))^T F''(x_k)((F''(x_k))^{-1}F'(x_k))} = \sqrt{F'(x_k)^T (F''(x_k))^{-1} F'(x_k)} \\ &= ||F'(x_k)||_{x_k} \end{aligned}$$

does not exceed $\lambda^* = \frac{3-\sqrt{5}}{2} \approx 0.3820$. With a little abuse of notion we shall say that the quadratic convergence region is guaranteed to contain an open ellipsoid of radius λ^* times the radius of the Dikin ellipsoid, and call λ^* a bound on the radius of the convergence region. The quantity $\rho(x_k) = ||F'(x_k)||_{x_k}$ is called the *Newton decrement*. At the next iterate, the decrement is guaranteed to be upper bounded by [6, Theorem 5.2.2]

$$\rho(x_{k+1}) \leq \left(\frac{\rho(x_k)}{1 - \rho(x_k)} \right)^2.$$

The bound λ^* is obtained by equating the upper bound on $\rho(x_{k+1})$ and $\rho(x_k)$. Thus if the first step is of length strictly smaller than λ^*, then the decrement is guaranteed to form a strictly decreasing sequence which tends to zero quadratically.

We now derive an estimate of how far we may advance the minimum of the current composite objective function along the central path by updating its parameter. It serves as a proxy for the overall performance of the path-following method. We have $\rho(x_k) - \rho(x_{k+1}) \geq \rho(x_k) - \left(\frac{\rho(x_k)}{1-\rho(x_k)} \right)^2$. The right-hand side of the inequality attains its maximum value at $\rho(x_k) = \rho^* \approx 0.2291$. Therefore, if we guarantee that the Newton decrement before the Newton step is not larger than ρ^*, we can be sure that it is not larger than $\rho_* = (\frac{\rho^*}{1-\rho^*})^2$ thereafter. We may then move the minimum by a distance of $\delta\rho \geq \rho^* - \rho_* \approx 0.1408$ in order to return to the initial state at the next iterate.

This situation is different if instead of a full Newton step one applies a shorter one with a damping coefficient $\gamma = \frac{1}{1+\rho(x_k)}$ depending on the decrement at the actual iterate. Then one can guarantee that [6, Theorem 5.2.2]

$$\rho(x_{k+1}) \leq \frac{\rho(x_k)^2 (2 + \rho(x_k))}{1 + \rho(x_k)}.$$

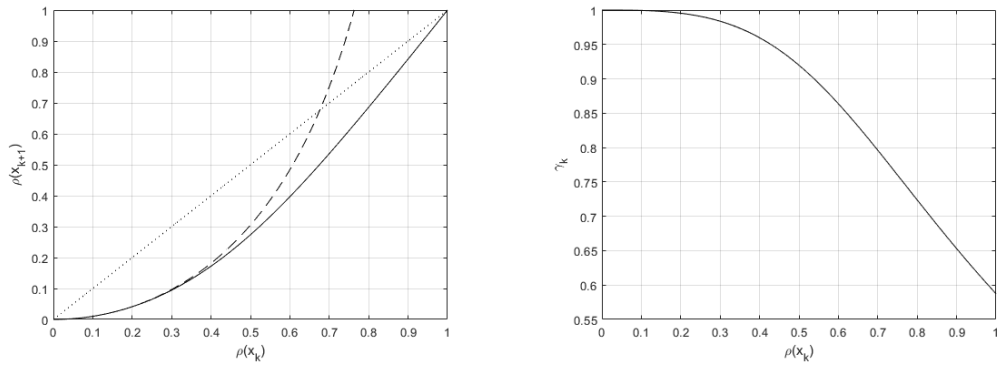

Figure 1: Upper bounds on the Newton decrement $\rho(x_{k+1})$ (left: dashed — full Newton step, solid — optimal damped Newton step) and optimal damping coefficient γ_k (right) as a function of the current Newton decrement $\rho(x_k)$.

Equating $\rho(x_{k+1})$ and $\rho(x_k)$ yields the improved bound $\lambda^* = \frac{\sqrt{5}-1}{2} \approx 0.6180$ on the radius of the convergence region. However, even if ρ^* increases to ≈ 0.2972, the bound on the difference $\delta\rho$ remains the same as in the case of a full Newton step.

In this contribution we perform an exact worst-case analysis of the performance of the Newton iterate by reformulating it as an optimal control problem. The performance is measured by the value $\rho(x_{k+1})$ of the Newton decrement at the next iterate. We show that for a full Newton step the radius of the convergence region, i.e., the set of initial points which certainly yield a strictly decreasing sequence of Newton decrements, is actually $\lambda^* \approx 0.6757$, while for a damped Newton step with optimal step size the quadratic convergence region covers the whole open Dikin ellipsoid, $\lambda^* = 1$. The values of ρ^* for the full and the optimal damped Newton step are given by ≈ 0.3943 and ≈ 0.4429, respectively. The respective bounds on $\delta\rho$ are given by ≈ 0.2184 and ≈ 0.2300, which improves the previously known bounds on $\delta\rho$ by more than 50%. The tight bound on $\rho(x_{k+1})$ and the optimal damping coefficient as a function of $\rho(x_k)$ are depicted in Fig. 1.

The idea of analyzing iterative algorithms by optimization techniques is not new. An exact worst-case analysis of gradient descent algorithms by semi-definite programming has been performed in [2, 9]. In these papers an arbitrary finite number of steps is analyzed, but the function classes are such that the resulting optimization problem is finite-dimensional.

In [3] a performance analysis for a single step of the Newton method on self-concordant functions is conducted. The class of self-concordant functions is, however, overbounded by a class of functions with Lipschitz-continuous Hessian, and the gradient at the next iterate is measured in the local norm of the previous iterate or in the (algorithmically inaccessible) local norm of the minimum. This yields also a finite-dimensional optimization problem. In [4] it is shown that the step size $\gamma_k = \frac{1}{1+\rho(x_k)}$ proposed in [7] maximizes a lower bound on the progress if the latter is measured in terms of the decrease of the function value. Using the techniques of this paper, one can show that this step length is actually optimal for this performance criterion[1].

In our case the properties of the class of self-concordant functions do not allow to obtain tight constraints on gradient and Hessian values at different points without taking into consideration all intermediate points, and the problem becomes infinite-dimensional. Our techniques, however, are borrowed from optimal control theory and nevertheless allow to obtain optimal bounds.

The remainder of the paper is structured as follows. In Section 2 we analyze the Newton iteration for self-concordant functions in one dimension. In this case the problem can be solved analytically. In Section 3 we generalize to an arbitrary number of dimensions. It turns out that the general case is no more difficult than the 2-dimensional one due to the rotational symmetry of the problem, and is described by the solutions of a Hamiltonian dynamical system in a 4-dimensional space. In Section 4 we analyze the solutions of the Hamiltonian system and derive the bound for the full Newton step. In Section 5 we minimize the bound

[1] Unpublished joint work with Anastasia S. Ivanova (Moscow Institute of Physics and Technology).

on the Newton decrement $\rho(x_{k+1})$ with respect to the step length in order to obtain the optimal damped Newton step. In turns out that in this case the system simplifies to an ordinary differential equation (ODE) on the plane. The optimal step length and the optimal bound on the decrement are then described in terms of solutions of this equation. In Section 6 we provide a game-theoretic interpretation of our results.

2 One-dimensional case

In this section we analyze the Newton iteration in one dimension. Given a damping coefficient and the value of the Newton decrement at the current iterate, we would like to find the maximal possible value of the decrement at the next iterate. We reformulate this optimization problem as an optimal control problem. The solution of this problem is found by presenting an analytic expression for the Bellman function.

Let $F : I \to \mathbb{R}$ be a self-concordant function on an interval, i.e., a C^3 function satisfying $F'' > 0$, $|F'''| \le 2(F'')^{3/2}$. Suppose the Newton decrement $\rho(x_k) = \sqrt{\frac{F'(x_k)^2}{F''(x_k)}}$ at some point $x_k \in I$ equals $a \in (0,1)$. Fix a constant $\gamma \in (0,1)$ and consider the damped Newton step

$$x_{k+1} = x_k - \gamma \frac{F'(x_k)}{F''(x_k)}.$$

The Newton decrement at the next iterate, which we suppose to lie in I, is given by $\rho(x_{k+1}) = \sqrt{\frac{F'(x_{k+1})^2}{F''(x_{k+1})}}$. Our goal in this section is to find the maximum of $\rho(x_{k+1})$ as a function of the parameters a, γ.

First of all, we may use the affine invariance of the problem setup to make some simplifications. We move the current iterate to the origin, $x_k = 0$, normalize the Hessian at the initial point to $F''(0) = 1$, and possibly flip the real axis to achieve $F'(0) = -a < 0$. Then we get $x_{k+1} = a\gamma$. Introducing the functions $h = F''(x)$, $g = F'(x)$, we obtain the optimal control problem

$$g' = h, \qquad h' = 2uh^{3/2}, \qquad u \in U = [-1, 1]$$

with initial conditions

$$g(0) = -a, \qquad h(0) = 1$$

and objective function

$$\sqrt{\frac{g(a\gamma)^2}{h(a\gamma)}} \to \sup.$$

Replacing the state variable g by $y = h^{-1/2}g$ and the independent variable x by $t = h^{1/2} \cdot (x - a\gamma)$, we obtain $\frac{dt}{dx} = h^{1/2} \cdot (1 + ut)$ and the problem becomes

$$\dot{y} = \frac{1 - uy}{1 + ut}, \qquad u \in U = [-1, 1]$$

with initial conditions $y(-a\gamma) = -a$ and objective function $|y(0)| \to \sup$. The variable h becomes disconnected from the relevant part of the dynamics and can be discarded. If the control is bang-bang, i.e., u is piece-wise constant with values in $\{-1, 1\}$, then the dynamics can be integrated explicitly, with solutions

$$-u \log|1 - uy| + const = u \log|1 + ut| \qquad \Rightarrow \quad (1 - uy)(1 + ut) = const.$$

Since the objective function depends only on the end-point $y(0)$, the Bellman function $B(t, y)$ of the problem satisfies the boundary condition $B(0, y) = |y|$ and is constant along the trajectories. The Bellman equation becomes

$$\max_{u \in [-1,1]} \frac{dB}{dt} = \max_{u \in [-1,1]} \left(\frac{\partial B}{\partial y} \frac{1 - uy}{1 + ut} + \frac{\partial B}{\partial t} \right) = 0.$$

After a bit of calculation one obtains the solution

$$B(t, y) = \begin{cases} -y + t + ty, & y \le \frac{2(-1+\sqrt{1+t^3})}{t^2}, \\ 4 - y + t - ty - 4\sqrt{(1-y)(1+t)}, & \frac{2(-1+\sqrt{1+t^3})}{t^2} \le y \le -t, \\ y - t - ty, & y \ge -t \end{cases}$$

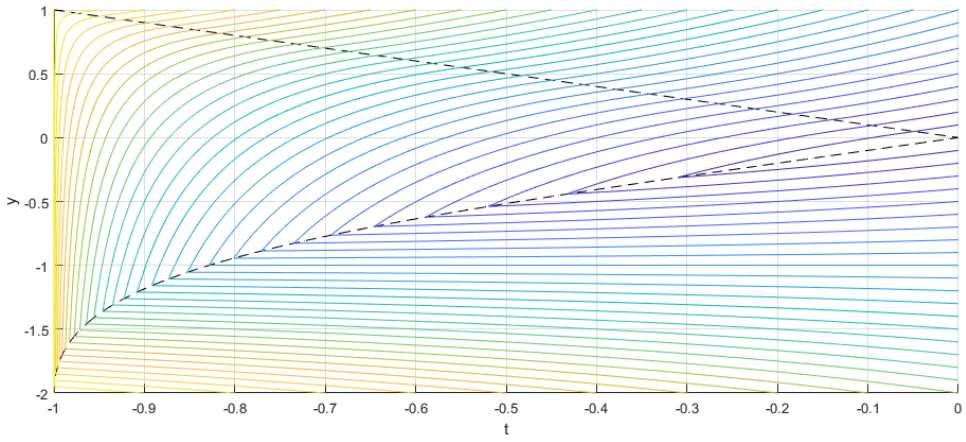

Figure 2: Optimal synthesis of the control problem modeling the one-dimensional case. The switching curve and the dispersion curve are dashed. The level curves of the Bellman function are at the same time the optimal trajectories of the system.

on the domain $(t, y) \in [-1, 0] \times \mathbb{R}$. The curve $y = -t$ is a switching curve, where the control u switches from $+1$ to -1. The curve $y = \frac{2(-1+\sqrt{1+t^3})}{t^2}$ is a dispersion curve, there are two optimal trajectories with controls $u = \pm 1$ emanating from the points of this curve. The optimal synthesis of the system is depicted in Fig. 2.

Let us now consider the result for the full Newton step. In this case $\gamma = 1$, and at the initial point is given by $(t, y) = (-a, -a)$. It hence lies between the dispersion curve and the switching curve, yielding the upper bound

$$\rho(x_{k+1}) \leq B(-a, -a) = 4 - a^2 - 4\sqrt{1 - a^2} = 4 - \rho(x_k)^2 - 4\sqrt{1 - \rho(x_k)^2}.$$

Setting this bound equal to a and resolving with respect to a, we obtain $\mathrm{Roots}(\lambda^3 + 2\lambda^2 + 9\lambda - 8) \approx 0.7282$ for the radius of the convergence region λ^*.

For the damped Newton step, we have to minimize $B(t, -a)$ with respect to t for given a. The minimum is given by the intersection point $(\frac{2(1-\sqrt{1+a^3})}{a^2}, -a)$ of the line $y = -a$ with the dispersion curve. Hence the optimal damping coefficient is given by $\gamma_k = \frac{2(\sqrt{1+\rho(x_k)^3}-1)}{\rho(x_k)^3}$. The corresponding upper bound on the Newton decrement is given by

$$\rho(x_{k+1}) \leq a + \frac{2(1 - \sqrt{1 + a^3})}{a^2}(1 - a) = \frac{2(1 - \rho(x_k))(1 - \sqrt{1 + \rho(x_k)^3}) + \rho(x_k)^3}{\rho(x_k)^2}.$$

In this case we have $\rho(x_{k+1}) < \rho(x_k)$ for every $\rho(x_k) \in (0, 1)$, and the convergence region covers the whole open Dikin ellipsoid.

3 Reduction to a control problem in the general case

In this section we perform a similar analysis for the case of self-concordant functions F defined on n-dimensional domains. We reduce the problem to an optimal control problem in two state space dimensions. For this problem the Bellman function cannot be presented in closed form, however, and we have to employ Pontryagins maximum principle to solve it. This results in a Hamiltonian system in 4-dimensional extended phase space.

Introduce the vector-valued variable $g = F'$ and the matrix-valued variable $h = F'' = WW^T$, where W is the lower-triangular factor of h with positive diagonal. Let further $\mathcal{P}$ be the set of homogeneous cubic polynomials $p(x) = \sum_{i,j,k=1}^{n} p_{ijk} x_i x_j x_k$ which are bounded by 1 on the unit sphere $S^{n-1} \subset \mathbb{R}^n$. This is a

compact convex set. The self-concordance condition then expresses the third derivatives of F in the form

$$\frac{\partial^3 F}{\partial x_i \partial x_j \partial x_k} = 2 \sum_{r,s,t=1}^{n} W_{ir} W_{js} W_{kt} p_{rst}, \qquad p \in \mathcal{P}.$$

We shall need also the projection

$$\mathcal{U} = \{U \mid \exists\, p \in \mathcal{P} :\ U_{ij} = p_{ij1}\}$$

of $\mathcal{P}$. The set $\mathcal{U}$ is a compact convex subset of the space of real symmetric matrices $\mathcal{S}^n$. Clearly it is overbounded by the set $\mathcal{U}' = \{U \mid -I \preceq U \preceq I\}$. We shall also introduce the sets of lower-triangular matrices

$$\mathcal{V} = \left\{V \mid \frac{V+V^T}{2} \in \mathcal{U}\right\}, \qquad \mathcal{V}' = \left\{V \mid \frac{V+V^T}{2} \in \mathcal{U}'\right\}$$

which are also compact and convex.

Using affine invariance, we may achieve the normalization $x_k = \mathbf{0}$, $g(\mathbf{0}) = -ae_1$, $h(\mathbf{0}) = W(\mathbf{0}) = I$, $x_{k+1} = a\gamma e_1$. Here $a \in (0,1)$ is the Newton decrement $\rho(x_k)$, $\gamma \in (0,1]$ the damping coefficient, and e_1 the first canonical basis vector. We consider the evolution of the variables g, h, W only on the line segment joining x_k and x_{k+1}, and may hence pass to a scalar variable $\tau \in [0, a\gamma]$, such that $x(\tau) = \tau e_1$ and $g(\tau) := g(x(\tau))$, $h(\tau) := h(x(\tau))$. The dynamics of the resulting control system can be written as

$$\frac{dg}{d\tau} = he_1, \qquad \frac{dh_{ij}}{d\tau} = 2 \sum_{r,s,t=1}^{n} W_{ir} W_{js} W_{1t} p_{rst} = 2W_{11} \sum_{r,s=1}^{n} W_{ir} W_{js} U_{rs}, \qquad p \in \mathcal{P},\ U \in \mathcal{U}.$$

It follows that

$$\frac{dh}{d\tau} = \frac{dW}{d\tau} W^T + W \frac{dW^T}{d\tau} = 2W_{11} WUW^T \quad \Rightarrow \quad W^{-1}\frac{dW}{d\tau} + \frac{dW^T}{d\tau} W^{-T} = 2W_{11} U, \qquad U \in \mathcal{U}.$$

Since $W^{-1}\frac{dW}{d\tau}$ is lower-triangular, and $\frac{dW^T}{d\tau}W^{-T}$ is its transpose, we finally obtain

$$\frac{dW}{d\tau} = W_{11} WV, \qquad V \in \mathcal{V}.$$

We now replace g by the variable $y = W^{-1}g$ and introduce a new independent variable $t = W_{11} \cdot (\tau - a\gamma)$. This variable then evolves in the interval $t \in [-a\gamma, 0]$, and we have $\frac{dt}{d\tau} = W_{11}^2 V_{11} \cdot (\tau - a\gamma) + W_{11} = W_{11} \cdot (tV_{11} + 1)$. The dynamics of the system becomes

$$\dot{y} = \frac{1}{W_{11}(tV_{11}+1)}(-W^{-1}(W_{11}WV)W^{-1}g + W^{-1}he_1) = \frac{-Vy + e_1}{tV_{11}+1}, \qquad V \in \mathcal{V}$$

with initial condition $y(-a\gamma) = -ae_1$ and objective function $\rho(x_{k+1}) = \sqrt{g^T(x_{k+1})h^{-1}(x_{k+1})g(x_{k+1})} = ||y(0)|| \to \sup$. The matrix-valued variable W becomes disconnected and can be discarded.

Let us apply Pontryagins maximum principle [5] to this optimal control problem. Introduce an adjoint vector-valued variable p, then the Pontryagin function and the Hamiltonian of the system are given by

$$\mathcal{H}(t, y, p, V) = \frac{\langle p, -Vy + e_1\rangle}{tV_{11}+1}, \qquad H(t, y, p) = \max_{V \in \mathcal{V}} \frac{\langle p, -Vy + e_1\rangle}{tV_{11}+1},$$

respectively. The transversality condition is non-trivial at the end-point $t = 0$ and states that $p(0)$ equals the gradient $\frac{\partial ||y(0)||}{\partial y(0)} = \frac{y(0)}{||y(0)||}$ of the objective function.

Note that the problem setting is invariant with respect to orthogonal transformations of $\mathbb{R}^n$ which leave the distinguished vector e_1 invariant. Suppose that at some point on the trajectory the vectors y, p are located in a plane containing e_1. Then at this point the derivatives of the Pontryagin function in the orthogonal directions vanish due to symmetry, and the derivatives of y, p are also contained in this plane. Therefore these variables will remain in this plane along the whole trajectory. We may hence assume without

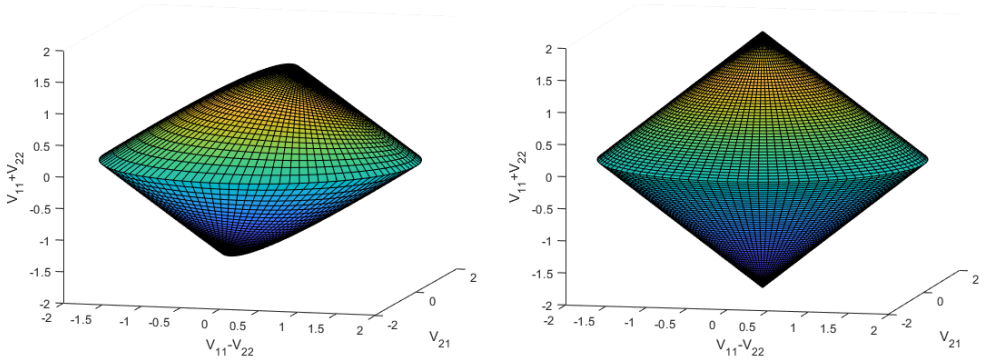

Figure 3: True set $\mathcal{V}$ of controls V (left) and overbounding set $\mathcal{V}'$ (right). The sharp circular edge of both bodies is the circle $\mathcal{C}$.

loss of generality that y, p are contained in the plane spanned by the basis vectors e_1, e_2, or equivalently, that the dimension n equals 2.

Then the set $\mathcal{P}$ is given by the set of bi-variate homogeneous cubic polynomials which are bounded by 1 on the unit circle, and can be expressed via the semi-definite representable set of nonnegative univariate trigonometric polynomials. This allows to obtain an explicit description of the set $\mathcal{V}$. Its boundary is given by the matrices

$$V = \pm\frac{1}{2\cos^3\xi}\begin{pmatrix}\cos\phi(3\cos^2\xi - \cos^2\phi) & 0\\ 2\sin\phi(\sin^2\xi - \sin^2\phi) & \cos\phi(\cos^2\xi - \sin^2\phi)\end{pmatrix}$$

with $\xi \in [0, \frac{\pi}{3}]$, $|\phi| \le \xi$. Recall that the set $\mathcal{V}$ is overbounded by the set $\mathcal{V}'$ of lower-triangular matrices (see Fig. 3). Both sets share the circle

$$\mathcal{C} = \left\{ V = \begin{pmatrix}\cos\phi & 0\\ 2\sin\phi & -\cos\phi\end{pmatrix} \mid \phi \in [-\pi, \pi] \right\}.$$

Note that the level sets of the Pontryagin function $\mathcal{H}$ are planes, because $\mathcal{H}$ is a fractional-linear function of V. Hence the maximum of $\mathcal{H}$ over a compact convex set is attained at an extreme point of this set. The maximum of $\mathcal{H}$ the over the circle $\mathcal{C}$ can be computed explicitly and is given by the expression

$$\max_{V\in\mathcal{C}} \mathcal{H} = \frac{p_1 + (p_1y_1 - p_2y_2)t + \sqrt{(p_1y_1 - p_2y_2 + p_1t)^2 + 4p_2^2y_1^2(1-t^2)}}{1-t^2}.$$

Besides $\mathcal{C}$, the set $\mathcal{V}'$ has the extreme points $V = \pm I$, on which $\mathcal{H}$ evaluates to $\frac{p_1 \mp (p_1y_1 + p_2y_2)}{1\pm t}$. Hence we have $\max_{V\in\mathcal{V}'} \mathcal{H} = \max_{V\in\mathcal{C}} \mathcal{H}$ if

$$\sqrt{(p_1y_1 - p_2y_2 + p_1t)^2 + 4p_2^2y_1^2(1-t^2)} \ge -p_1t - p_1y_1 - p_2y_2 + 2p_2y_2t,$$

$$\sqrt{(p_1y_1 - p_2y_2 + p_1t)^2 + 4p_2^2y_1^2(1-t^2)} \ge p_1t + p_1y_1 + p_2y_2 + 2p_2y_2t.$$

These conditions then also imply $\max_{V\in\mathcal{V}} \mathcal{H} = \max_{V\in\mathcal{C}} \mathcal{H}$ and hence

$$H(t, y, p) = \frac{p_1 + (p_1y_1 - p_2y_2)t + \sqrt{(p_1y_1 - p_2y_2 + p_1t)^2 + 4p_2^2y_1^2(1-t^2)}}{1-t^2}. \tag{1}$$

It turns out a posteriori that the conditions are satisfied on the relevant solutions, and by virtue of the necessity of Pontryagins maximum principle for optimality we obtain the following result.

Theorem 3.1. *Let $a, \gamma \in (0,1)$ be given. Then the upper bound on the Newton decrement $\rho(x_{k+1})$ after a damped Newton step with damping coefficient γ and initial value of the decrement $\rho(x_k) = a$ is given by the norm $||y(0)||$, where $(y(t), p(t))$, $y = (y_1, y_2)$, $p = (p_1, p_2)$, $t \in [-a\gamma, 0]$, is a solution of the Hamiltonian system defined by (1) and satisfying the boundary conditions $p(0) = \frac{y(0)}{||y(0)||}$, $y(-a\gamma) = -ae_1$.* □

In the next section we shall analyze the qualitative behaviour of the solutions of this Hamiltonian system.

4 Bound for the full Newton step

In this section we analyze the Hamiltonian system obtained in the previous section. It turns out that for small enough damping coefficients the trajectories corresponding to the 1-dimensional solution obtained in Section 2 are optimal. If the initial point of the trajectory lies beyond a critical curve in the (t, y_1)-plane, however, the second variable y_2 is no more identically zero. This is the case for the full Newton step. The corresponding bound on the Newton decrement can be obtained numerically by integration of the Hamiltonian system.

As in the 1-dimensional case, the Bellman function $B(t, y)$ of the problem is constant on the trajectories of the Hamiltonian system and satisfies the boundary condition $B(0, y) = ||y||$. In order to construct it, we have to integrate the system in backward time with initial condition $p(0) = \frac{y(0)}{||y(0)||}$, launching a trajectory from every point of the y-plane.

In general, the projections on y-space of trajectories launched from different points may eventually intersect. In this case the trajectory with the maximal value of the Bellman function along it is retained. Therefore trajectories cease to be optimal at dispersion surfaces where they meet other trajectories with the same value of the Bellman function. In our case the plane $y_2 = 0$ acts as a dispersion surface by virtue of the symmetry $(y_1, y_2, p_1, p_2) \mapsto (y_1, -y_2, p_1, -p_2)$ of the system. Indeed, a trajectory launched from a point with $y_2(0) \neq 0$ and hitting the plane $y_2 = 0$ at some time will meet there with its image under the symmetry, which necessarily has the same value of the Bellman function. Thus the trajectories which are relevant for Theorem 3.1 are those which either completely evolve on the plane $y_2 = 0$, or which do not cross this plane on the time interval $(-a\gamma, 0]$.

The first kind of trajectories correspond to the 1-dimensional system considered in Section 2 and are depicted in Fig. 2. In the regions between the dashed curves and the vertical axis they are given by

$$y(t) = \frac{c+t}{1-t} e_1, \quad p(t) = \frac{c(1-t)}{|c|} e_1, \tag{2}$$

where $c = y_1(0)$ is a parameter.

We now perturb trajectory (2) by launching it from a nearby point $y(0) = (c, \epsilon)$, and consider its evolution up to first order in ϵ. This can be done by solving the linearized ODE. The right-hand side of the Hamiltonian system is given by $(\dot{y}, \dot{p}) = (\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial y})$. Hence the coefficient matrix of the linearized system is given by

$$\begin{pmatrix} \frac{\partial^2 H}{\partial p \partial y} & \frac{\partial^2 H}{\partial p^2} \\ -\frac{\partial^2 H}{\partial y^2} & -\frac{\partial^2 H}{\partial y \partial p} \end{pmatrix} = \begin{pmatrix} \frac{1}{1-t} & 0 & 0 & 0 \\ 0 & -\frac{1}{1-t} & 0 & \frac{4y_1^2}{p_1(t+y_1)} \\ 0 & 0 & -\frac{1}{1-t} & 0 \\ 0 & 0 & 0 & \frac{1}{1-t} \end{pmatrix} = \begin{pmatrix} \frac{1}{1-t} & 0 & 0 & 0 \\ 0 & -\frac{1}{1-t} & 0 & \frac{4|c|(c+t)^2}{c(c+2t-t^2)(1-t)^2} \\ 0 & 0 & -\frac{1}{1-t} & 0 \\ 0 & 0 & 0 & \frac{1}{1-t} \end{pmatrix}.$$

Here the first relation is obtained by setting $y_2 = p_2 = 0$ in the partial derivatives of H, the second one by inserting the values (2). The linearized system has to be integrated with initial condition $\frac{\partial (c, \epsilon, \frac{c}{\sqrt{c^2+\epsilon^2}}, \frac{\epsilon}{\sqrt{c^2+\epsilon^2}})}{\partial \epsilon}|_{\epsilon=0} = (0, 1, 0, \frac{1}{|c|})$ at $t = 0$. It has the solution

$$\delta y(t) = \delta y_2 e_2, \qquad \delta p(t) = \frac{1}{|c|(1-t)} e_2,$$

where the scalar function δy_2 is a solution of the ODE

$$\frac{d(\delta y_2)}{dt} = -\frac{1}{1-t}\delta y_2 + \frac{4(c+t)^2}{c(c+2t-t^2)(1-t)^3}$$

with initial condition $\delta y_2(0) = 1$. Integrating the ODE, we obtain

$$\begin{aligned} \delta y_2(t) = {} & \frac{-(c^2+5c+16)(1-t)}{3c(c+1)} + \frac{4(c+2)}{c(c+1)} - \frac{4}{c(1-t)} + \frac{4(c+1)}{3c(1-t)^2} + \frac{4(1-t)\log\frac{c(1-t)^2}{c+2t-t^2}}{c(c+1)} \\ & + \frac{2(c+2)(1-t)\log\frac{(c+2t-t^2)(\sqrt{c+1}+1)^2}{c(\sqrt{c+1}+1-t)^2}}{c(c+1)^{3/2}} \end{aligned}$$

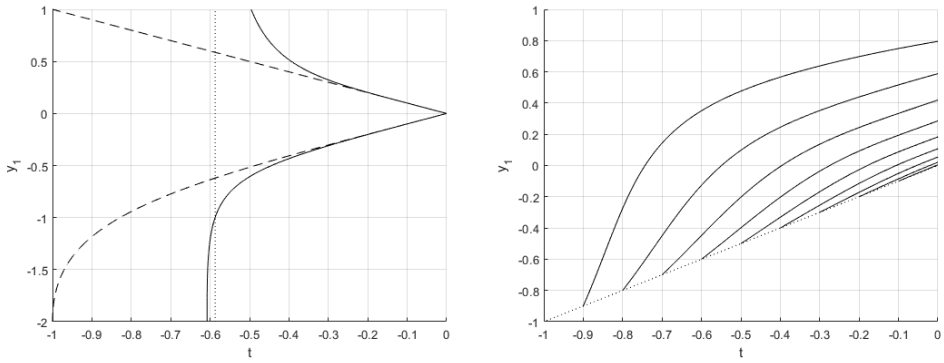

Figure 4: Left: Critical curve marking the limit of optimality of the 1-dimensional solution (solid). For comparison the switching curve and the dispersion curve of the 1-dimensional optimal synthesis are also depicted (dashed). For $|y| \to +\infty$ the critical curve tends to the line $t = 1 - 2^{2/3}$ (dotted). Right: Projections on the (t, y_1)-plane of the trajectories corresponding to a full Newton step for different initial values a of the Newton decrement. The dotted line is the locus of the initial points.

for $c > -1$, $\delta y_2 = \frac{4}{3(1-t)^2} - \frac{1-t}{3}$ for $c = -1$, and

$$\delta y_2(t) = \frac{-(c^2+5c+16)(1-t)}{3c(c+1)} + \frac{4(c+2)}{c(c+1)} - \frac{4}{c(1-t)} + \frac{4(c+1)}{3c(1-t)^2} + \frac{4(1-t)\log\frac{c(1-t)^2}{c+2t-t^2}}{c(c+1)} + \frac{4(c+2)(1-t)\left(\arctan\frac{1}{\sqrt{-1-c}} - \arctan\frac{1-t}{\sqrt{-1-c}}\right)}{c(-1-c)^{3/2}}$$

for $c < -1$.

Setting the variable δy_2 to zero, we obtain a critical value of t on trajectory (2) beyond which other trajectories of the system start to intersect the $y_2 = 0$ plane. At this point trajectory (2) ceases to be optimal. The ensemble of these critical points for all values $c \in \mathbb{R}$ forms a curve in the (t, y_1)-plane which marks the limit of optimality of the synthesis obtained in Section 2. In order to obtain an expression for this curve, we have to use (2) to express c as a function of t, y_1. Inserting this expression into the relation $\delta y_2 = 0$ yields a relation between t and y_1.

For $c > -1$ this relation is given by

$$(-Y^4T^6 + 4Y^4 - 3Y^2T^4 - 12y_1t)Y + 24T^2Y\log T + 6T(YT-1)^2\log\frac{Y-T}{YT-1} + 6T(YT+1)^2\log\frac{YT+1}{Y+T} = 0,$$

where $Y = \sqrt{1+y_1}$, $T = \sqrt{1-t}$. For $c = -1$ we obtain the point $(t^*, -1)$ with $t^* = 1 - 2^{2/3} \approx -0.5874$, and for $c < -1$ we get

$$(-Y^4T^3 + 4Y^4 + 3Y^2T^3 - 12T^2y_1t) + 12T^3\left(2\log T + \frac{(Y^2-1)(\arctan\frac{1}{Y} - \arctan\frac{T}{Y})}{Y} + \log\frac{Y^2+1}{Y^2+T^2}\right) = 0,$$

where $Y = \sqrt{-(1+y_1)(1-t)}$, $T = 1-t$.

In particular, both the switching curve and the dispersion curve lie beyond the critical curve and are not part of the optimal synthesis (see Fig. 4, left). The trajectories with initial condition $y(-a\gamma) = -ae_1$ corresponding to a point $(t, y_1) = (-a\gamma, -a)$ beyond the critical curve can be computed numerically. To each such initial condition there correspond two solutions which are taken to each other by the symmetry $(y_1, y_2, p_1, p_2) \mapsto (y_1, -y_2, p_1, -p_2)$. In Fig. 4, right, the trajectories corresponding to the value $\gamma = 1$ (full Newton step) and different a are depicted. The resulting bound on the decrement after the iteration is depicted in Fig. 1, left. A numerical analysis of the solutions yields the following results.

Numerical values: The radius of the convergence region of the exact Newton method equals $\lambda^* \approx 0.6757$. This means, if the Newton method is launched from an initial point x_0 with decrement $\rho(x_0) < \lambda^*$, then it

is guaranteed to converge to the nearby minimum. If at the iterate x_k we have $\rho(x_k) \le \rho^* \approx 0.3943$, then $\rho(x_{k+1}) \le \rho_* \approx 0.1758$. These values maximize the lower bound on the difference $\rho(x_k) - \rho(x_{k+1})$, which evaluates to $\delta\rho \approx 0.2184$. Applied to path-following methods, ρ^* gives the radius of the tube around the central path which the iterates should not leave, while $\delta\rho$ is a lower bound on the length of the step we may move the target point along the central path at each iteration. (The actual values are a bit different due to second order effects, but these details are not subject of this paper.)

5 Optimal step length and bounds for the damped Newton step

In this section we minimize the upper bound on $\rho(x_{k+1})$ with respect to the damping coefficient γ for fixed initial values of the decrement $\rho(x_k) = a$. The minimizer of this problem yields the optimal damping coefficient for the Newton iterate which leads to the largest guaranteed decrease of the decrement.

Technically, releasing γ is equivalent to releasing the left end of the time interval on which the trajectory of the Hamiltonian system evolves, while leaving the initial state fixed. It is well known that the partial derivative with respect to time of the Bellman function, i.e., the objective achieved by the trajectory, equals the value of the Hamiltonian [5]. Therefore we look for trajectories with starting points lying on the surface $H = 0$.

Let us first evaluate H on the critical curve, more precisely on its arc between the lines $y_1 = -1$ and $y_1 = 0$. There we have $p_1 < 0$, $y_1 + t < 0$. Setting $y_2 = 0$, $p_2 = 0$, we obtain $H = \frac{p_1(1+y_1)}{1-t}$. Thus for $a < 1$ we have $H < 0$ and hence the optimal initial time instant t is strictly smaller than the time instant defined by the critical curve. For $a = 1$, or equivalently $y_1 = -1$, the trajectory of the 1-dimensional system is optimal. As a consequence, for $a \to 1$ the optimal damping coefficient tends to $2^{2/3} - 1 \approx 0.5874$.

Lemma 5.1. *The hyper-surface $H = 0$ is integral for the Hamiltonian system defined by* (1).

Proof. One easily computes

$$\dot{H} = \frac{\partial H}{\partial t} = H \cdot \frac{p_1 y_1 - p_2 y_2 + p_1 t + t\sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}}{\sqrt{(p_1 y_1 - p_2 y_2 + p_1 t)^2 + 4p_2^2 y_1^2 (1 - t^2)}(1 - t^2)},$$

and hence if $H = 0$ somewhere on a trajectory, then $H \equiv 0$ everywhere on the trajectory. □

In particular, it follows that at the end-point $t = 0$ of the trajectory we also have $H = 0$. From (1) we then obtain by virtue of the transversality conditions $p(0) = \frac{y(0)}{||y(0)||}$ that

$$H(0) = p_1 + \sqrt{(p_1 y_1 - p_2 y_2)^2 + 4p_2^2 y_1^2} = \frac{y_1 + y_1^2 + y_2^2}{||y(0)||} = 0.$$

The locus of the end-points in y-space is hence given by the circle $(y_1 + \frac{1}{2})^2 + y_2^2 = \frac{1}{4}$.

From now on we assume without loss of generality that $y_2 \ge 0$ on the trajectory of the system.

Using the homogeneity of the dynamics with respect to the adjoint variable p we may eliminate this variable altogether. On the surface $H = 0$ we have

$$(y_1^2 - 1)p_1^2 - 2y_1 y_2 p_1 p_2 + (4y_1^2 + y_2^2)p_2^2 = 0, \qquad \frac{p_1}{p_2} = \frac{y_1 y_2 + \sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}{y_1^2 - 1},$$

and consequently

$$\dot{y}_1 = \frac{-p_1 y_1^2 + p_2 y_2 y_1 + p_1}{p_1 + p_1 y_1 t - p_2 y_2 t} = \frac{\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}(-y_2(y_1 + t) + (y_1 t + 1)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2})}{4y_1^4 t^2 + 8y_1^3 t + 4y_1^2 - y_2^2 t^2 + y_2^2},$$

$$\dot{y}_2 = -\frac{4p_2 y_1^2 - p_1 y_1 y_2 + p_2 y_2^2}{p_1 + p_1 y_1 t - p_2 y_2 t} = \frac{\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}(4ty_1^3 + 4y_1^2 + y_2^2 - ty_2\sqrt{-4y_1^4 + 4y_1^2 + y_2^2})}{4y_1^4 t^2 + 8y_1^3 t + 4y_1^2 - y_2^2 t^2 + y_2^2}.$$

A closer look reveals that the quotient of the two derivatives does not depend on t, and we obtain a planar dynamical system defined by the scalar ODE

$$\frac{dy_2}{dy_1} = \frac{\sqrt{4y_1^2(1 - y_1^2) + y_2^2} + y_1 y_2}{1 - y_1^2}. \tag{3}$$

We obtain the following result.

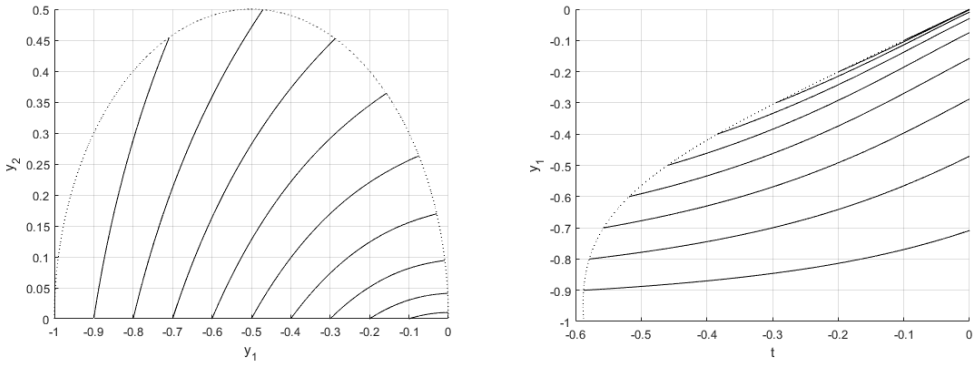

Figure 5: Left: Solution curves of ODE (3) between the line $y_2 = 0$ and the circle $(y_1 + \frac{1}{2})^2 + y_2^2 = \frac{1}{4}$ in the upper half-plane. Right: Solutions of ODE (4) on the curves depicted on the left part of the figure. The dotted curves are the locus of the end-points of the trajectories.

Theorem 5.2. *Let $a \in (0,1)$ be given, and let σ be the solution curve of ODE (3) through the point $y_0 = (-a, 0)$. Then the upper bound on the Newton decrement $\rho(x_{k+1})$ after a damped Newton step with optimal damping coefficient and initial value $\rho(x_k) = a$ of the decrement is given by $||y^*|| = \sqrt{-y_1^*}$, where $y^* = (y_1^*, y_2^*)$ is the intersection point of the curve σ with the circle centered on $(-\frac{1}{2}, 0)$ and with radius $\frac{1}{2}$ in the upper half-plane $y_2 > 0$.* □

The Riemann surfaces corresponding to the solutions of ODE (3) in the complex plane have an infinite number of quadratic ramification points, and the equation is not integrable in closed form. The optimal bound on $\rho(x_{k+1})$ as a function of $\rho(x_k)$ can be computed numerically and is depicted in Fig. 1, left.

In order to obtain the value of the optimal damping coefficient one also has to integrate the linear differential equation

$$\begin{aligned} \frac{dt}{dy_1} &= \frac{4y_1^4t^2 + 8y_1^3t + 4y_1^2 - y_2^2t^2 + y_2^2}{\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}(-y_2(y_1 + t) + (y_1t + 1)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2})} \\ &= \frac{y_2(y_1 + t) + (y_1t + 1)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}{(1 - y_1^2)\sqrt{-4y_1^4 + 4y_1^2 + y_2^2}}. \end{aligned} \tag{4}$$

Theorem 5.3. *Let $a \in (0,1)$ be given, and let σ be the solution curve of ODE (3) through the point $y_0 = (-a, 0)$, intersecting the circle $(y_1 + \frac{1}{2})^2 + y_2^2 = \frac{1}{4}$ in the point y^* in the upper half-plane $y_2 > 0$. Then the optimal damping coefficient γ for the Newton step with initial value $\rho(x_k) = a$ of the decrement is given by the value of $t(y_0)$, where $t(y)$ is the solution of ODE (4) along the curve σ with initial value $t(y^*) = 0$.* □

In order to compute the optimal value of γ one hence has first to integrate ODE (3) from y_0 to y^* and then ODE (4) back from y^* to y_0. The result is depicted on Fig. 1, right. The solution curves of the ODEs are depicted in Fig. 5.

Numerical values: If the optimal damping coefficient is applied, then the Newton decrement is guaranteed to decrease whenever its current value is smaller than 1. The radius of the convergence region hence equals $\lambda^* = 1$. If at the iterate x_k we have $\rho(x_k) \le \rho^* \approx 0.4429$, then $\rho(x_{k+1}) \le \rho_* \approx 0.2129$. These values maximize the lower bound on the difference $\rho(x_k) - \rho(x_{k+1})$, which evaluates to $\delta\rho \approx 0.2300$.

The upper bound on $\rho(x_{k+1})$ is actually quite close to $\rho(x_k)^2$. An asymptotic analysis of ODE (3) for small values of a yields the expansion

$$\max \rho(x_{k+1}) = \rho(x_k)^2 - \frac{1}{4}\rho(x_k)^4 \log \rho(x_k) + \left(\frac{\log 2}{2} - \frac{1}{16}\right) \rho(x_k)^4 + o(\rho(x_k)^5).$$

On the whole interval $[0, 1]$ we have

$$\max \rho(x_{k+1}) \le \rho(x_k)^2 - 0.5552 \cdot \rho(x_k)^4 \cdot \log \rho(x_k)$$

with a maximal difference of $7 \cdot 10^{-3}$.

An asymptotic analysis of ODE (4) for small values of a leads to the expansion $\gamma = 1 - \frac{a^3}{2} + \frac{a^4}{4} + O(a^5 \log a)$ of the optimal damping coefficient. Unlike the situation in one dimension, in multiple dimensions the upper bound on the Newton decrement is a smooth function of the damping coefficient. Therefore a small deviation from the optimal value of γ will result only in an increase of the upper bound on $\rho(x_{k+1})$ by a term of second order.

6 Conclusion

We first furnish an interpretation of the results. Let us imagine the worst-case behaviour of the self-concordant function on the segment between the iterates x_k, x_{k+1} as the response of an adversarial player to our choice of the damping coefficient. The goal of this player is to maximize the Newton decrement $\rho(x_{k+1})$ at the next iterate. Since the control which is at the disposal of the adversarial player affects the third derivative of the function, we can roughly assume that he plays with the acceleration of the gradient.

First consider the case when the function is defined on an interval. Here the adversarial player has two different options. One is to maximally decelerate the gradient in order to prevent it from reaching zero at the end-point of the interval. This strategy will pay off more if we choose a smaller step length. The other strategy is to first maximally accelerate the gradient, in order to give it enough velocity to overshoot. At some point, corresponding to the crossing of the switching curve in Fig. 2, the gradient is again decelerated by decreasing the Hessian, because the effect of a smaller denominator F'' in the objective function overweighs the effect of a larger gradient F', which enters in the numerator. This strategy pays off more if we choose a larger step length. Our optimal strategy will therefore be to choose that value of the damping coefficient which results in the same objective for both strategies of the adversarial player, i.e., we choose the initial point $(-a\gamma, -a)$ on the dispersion curve in Fig. 2.

In the case of a multi-dimensional domain of definition the adversarial player has more options. In addition to acceleration or deceleration of the gradient in the direction of movement, he may boost it in a perpendicular direction. Here he may choose this perpendicular direction arbitrarily, but once it is chosen, it is optimal to keep the acceleration vector in the plane spanned by the direction of movement and this particular direction. If the damping coefficient is large enough, more precisely if it corresponds to an initial point (t, y_1) beyond the critical curve in Fig. 4, left, the optimal strategy of the adversarial player is then indeed a mixture of boosts in the parallel and the perpendicular direction. Here the parallel component may be an acceleration or a deceleration, but the perpendicular component is always increased. For smaller damping coefficients the optimal strategy is a pure deceleration of the gradient.

The optimal value γ of the damping coefficient can be computed numerically by integrating two scalar ODEs, one of which is linear. For small values a of the decrement at the current iterate the correction with respect to the value 1 of the full Newton step is cubic in a.

The obtained results enable to tune the parameters in the machinery of path-following methods to achieve larger steps along the central path. The optimal choices of these parameters may also be obtained as solutions to optimization problems, but these are beyond the scope of this paper. Besides this, they may serve in general to optimize the performance of the Newton method on self-concordant functions in the vicinity of a minimum.

Acknowledgments

The paper was written in coronavirus quarantine at Moscow Institute of Physics and Technology. The author would like to thank the medical staff for having provided appropriate working conditions.

References

[1] Oleg P. Burdakov. Some globally convergent modifications of Newton's method for solving systems of linear equations. *Soviet Math. Dokl.*, 22(2):376–379, 1980.

[2] Etienne de Klerk, François Glineur, and Adrien Taylor. On the worst-case complexity of the gradient method with exact line search for smooth strongly convex functions. *Optimization Letters*, 11:1185–1199, 2017.

[3] Etienne de Klerk, François Glineur, and Adrien Taylor. Worst-case convergence analysis of gradient and Newton methods through semidefinite programming performance estimation. arXiv 1709.05191, 2017.

[4] Wenbo Gao and Donald Goldfarb. Quasi-newton methods: superlinear convergence without line searches for self-concordant functions. *Optim. Method Softw.*, 34(1):194–217, 2019.

[5] R.V. Gamkrelidze L.S. Pontryagin, V.G. Boltyanskii and E.F. Mischchenko. *The mathematical theory of optimal processes*. Wiley, New York, London, 1962.

[6] Yurii Nesterov. *Lectures on Convex Optimization*, volume 137 of *Springer Optimization and its Applications*. Springer, 2nd edition, 2018.

[7] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point Polynomial Algorithms in Convex Programming*, volume 13 of *SIAM Stud. Appl. Math.* SIAM, Philadelphia, 1994.

[8] Daniel Ralph. Global convergence of damped Newton's method for nonsmooth equations via the path search. *Math. Oper. Res.*, 19(2):352–389, 1994.

[9] Adrien B. Taylor, Julien M. Hendrickx, and François Glineur. Exact worst-case performance of first-order methods for composite convex optimization. *SIAM J. Optim.*, 27(3):1283–1313, 2017.