

THE NESTEROV–TODD DIRECTION AND ITS RELATION TO WEIGHTED ANALYTIC CENTERS (FINAL REVISED EDITION)

Raphael Hauser

Oxford University Computing Laboratory

The subject of this report concerns differential-geometric properties of the Nesterov–Todd search direction for linear optimization over symmetric cones. In particular, we investigate the rescaled asymptotics of the associated flow near the central path. Our results imply that the Nesterov–Todd direction arises as the solution of a Newton system defined in terms of a certain transformation of the primal-dual feasible domain. This transformation has especially appealing properties which generalize the notion of weighted analytic centers for linear programming.

AMS (MOS) Subject classification: primary 90C25, 34E10, 34B15; secondary 90C60, 90C05, 90C20

Key words and phrases: Symmetric cones, self-scaled programming, Nesterov-Todd direction, weighted analytic centers, target map, interior point methods, Magnus series, orthogonal flow.

I am immensely grateful to Mike Todd for pointing me to the open problem that has led to this project, for his relentless and tremendous efforts in guiding my work and making me bring it to fruition, and generally for the many things I have learned from him.

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD
E-mail: hauser@comlab.oxford.ac.uk

January, 2003

Contents

1	Introduction	3
2	A Variational Analysis of the Nesterov–Todd Flow	10
3	Nesterov-Todd Directions in the Target Framework	13
4	Proof of Lemma 3.3	24
5	Proof of Lemma 3.4	29
6	Proof of Lemma 3.5	30
	References	35

1 Introduction

In recent years linear optimization over symmetric cones, or self-scaled programming, has become the accepted standard framework to treat linear, quadratic and semidefinite programming in a unified framework. This theory originated in the work of Nesterov–Todd [29, 30], Güler [9, 10], and Faybusovich [7, 8].

A powerful and popular family of methods for solving such problems numerically consists of primal-dual interior point methods (IPMs) based on the Nesterov–Todd (NT) search direction [29, 30]. As we will show below, the NT direction is defined in terms of a pseudo Newton system that arises from the Karush–Kuhn–Tucker (KKT) conditions of a perturbed problem. At first sight, this derivation seems like an ad-hoc fix to a difficulty that occurs because of a mismatch in the dimensions of the preimage and image spaces of these equations. Other, more straightforward approaches exist to overcome this problem, making the NT approach seem unnecessarily complicated. However, a more in-depth look reveals that the NT framework has compelling properties that further motivate it. Much of the intrinsic beauty of this approach is due to the fact that the so-called *scaling point* (see explanations below) relates it to the group action on symmetric cones, which are a particular type of homogeneous spaces. The NT approach is designed to be invariant under this group action which allows for the construction of algorithms that are invariant under exchanging an input problem with its dual (*primal-dual symmetry*) and under coordinate change (*scale-invariance*). Such methods are unaffected by special types of ill-conditioning associated with representing the problem geometry in a coordinate system.

In this paper we explore another interesting property of the NT approach: its relation to so-called *weighted analytic centers* and a transformation that is sometimes called the *target map*. This construction was first analyzed in the linear programming (LP) literature by Kojima–Misuno–Yoshise [22] who proved that the target map is a diffeomorphism between the primal-dual strictly feasible domain and the strictly positive orthant, and that this map rectifies the primal-dual central path, an object that plays an important role in guiding interior point methods to an optimal solution. The Kojima construction has been extremely useful in the LP literature, and many attempts have been made at generalizing it to other convex optimization problems, see further details below. The main result of this article shows that the NT-direction arises as the Newton direction defined in terms of a generalization of this construction to arbitrary symmetric cones, see Theorem 3.7, Section 3. Apart from linking the NT and target frameworks, Theorem 3.7 has the immediate implication that the NT process converges at a quadratic rate in a neighborhood of its attractor, and it contributes new techniques to the IPM theory by employing *Magnus series* and *orthogonal flows* as essential analytic tools in its proof. As a prerequisite for these arguments one needs a thorough variational understanding of the NT process: if the process is started at two nearby points, how far will the iterates deviate from one another after a fixed number of iterations, and how will this affect the direction from where the fixed point is approached? A wealth of related questions are answered in Section 2 which makes an independent contribution to the understanding of the NT process.

In our exposition we consider the following pair of convex programs in conic duality

$$\begin{aligned} (P) \quad & \inf \langle x; s_0 \rangle \\ & x \in (L + x_0) \cap \bar{K} \end{aligned} \qquad \begin{aligned} (D) \quad & \inf \langle x_0; s \rangle \\ & s \in (L^\perp + s_0) \cap \bar{K}^\sharp. \end{aligned} \tag{1.1}$$

Here E is a finite dimensional Euclidean space equipped with an inner product $\langle \cdot; \cdot \rangle$, L is a linear subspace of E , and L^\perp is its orthogonal complement. K is a convex cone which is open and has a pointed closure \bar{K} , that is, \bar{K} does not contain any whole lines. The points $x_0 \in K$ and $s_0 \in K^\sharp$ are fixed. The dual cone

$$K^\sharp := \{s \in E : \langle x; s \rangle > 0, \quad \forall x \in K\} \tag{1.2}$$

is also convex, open and with pointed closure \bar{K}^\sharp . Note that the definition of K^\sharp depends on the choice of the inner product $\langle \cdot; \cdot \rangle$ on E .

In the problems we consider, K belongs to a special family of cones which we shall now define: The *automorphism group* of an open convex cone K is the set of nonsingular linear maps $A : E \rightarrow E$ that map K onto K , that is, $\text{Aut}(K) := \{A \in \text{Gl}(E) : A(K) = K\}$. The cone K is called *homogeneous* if $\text{Aut}(K)$ acts transitively on K , that is, given arbitrary points $x, y \in K$, there exists a map $A \in \text{Aut}(K)$ such that $Ax = y$. K is called *self-dual* when the inner product $\langle \cdot; \cdot \rangle$ on E can be chosen so that $K^\sharp = K$, see (1.2). K is called a *symmetric cone* if it is both homogeneous and self-dual. In the sequel, we will always assume that E is endowed with an inner product under which $K = K^\sharp$. Symmetric cones arise in Jordan algebra theory as follows: a Euclidean Jordan algebra is a finite-dimensional real commutative algebra E endowed with a weakly associative multiplication with identity element e and an associative inner product. The set of invertible squares of a Euclidean Jordan algebra is a symmetric cone, and every symmetric cone can be represented in this form. Euclidean Jordan algebras and, by extension, symmetric cones have been algebraically classified, see Köcher [21] and the references therein. Every symmetric cone has a unique decomposition into a direct sum of elementary building blocks, so-called *irreducible* symmetric cones, of which there exist only five types. For a complete account of this theory, see Faraut–Korányi [6]. Three examples of symmetric cones are of particular interest to the optimization community: the positive orthant $K = \mathbb{R}_{++}^n$, which is in fact the direct sum of n irreducible symmetric cones consisting of open half-lines, the cone $K = S^{n \times n}(\mathbb{R})_{++}$ of $n \times n$ symmetric positive definite matrices with real coefficients, and the Lorentz cone $K = \left\{ \begin{pmatrix} \tau \\ x \end{pmatrix} \in \mathbb{R}^{n+1} : \tau > \|x\|_2 \right\}$, which is also called the second-order cone. The conic optimization problems associated with these cones are linear programming, semidefinite programming and second-order cone programming respectively. Considering more general symmetric cones, one can treat linear optimization problems with mixed linear, semidefinite and convex quadratic constraints in a single unified framework, see e.g. Todd–Toh–Tütüncü [36], Alizadeh–Schmieta [2] or Sturm [35].

In [29], Nesterov and Todd defined the concept of *self-scaled* barriers, a special class of self-concordant barrier functions whose Hessians form a transitive subset of the automorphism group of their domain of definition. Self-scaled barriers are well understood: Güler [9] and Nesterov–Todd [29] showed that K is the domain of definition of a self-scaled barrier if and only if K is a symmetric cone. Recently, Hauser [15, 13, 14], Schmieta

[33], Güler [11], Hauser–Güler [16] and Hauser–Lim [17] developed a classification theory showing that all self-scaled barriers defined on a symmetric cone with irreducible decomposition $K = K_1 \oplus \cdots \oplus K_p$ are of the form

$$F : x_1 \oplus \cdots \oplus x_p \mapsto c_0 + \sum_{i=1}^p c_i \ln \int_{K_i^\sharp} e^{-\langle x_i; s_i \rangle} ds_i, \quad (1.3)$$

where $c_i \geq 1$ ($i = 1, \dots, p$). The dual barrier is defined on K^\sharp as the Legendre–Fenchel transform $F_\sharp : s \mapsto \max\{-\langle x; s \rangle - F(x) : x \in K\}$. Under the self-dual embedding $K^\sharp \hookrightarrow K$ it is then the case that $F_\sharp(s) = F(x) + c$, where c is a constant. See [16] for a complete survey of self-scaled barriers and symmetric cones.

Using the barrier function F , most primal-dual interior point methods attack a sequence of unconstrained subproblems

$$\begin{aligned} (P_\mu) \quad & \inf \mu F(x) + \langle x; s_0 \rangle \\ & x \in (L + x_0) \cap K \end{aligned} \quad \begin{aligned} (D_\mu) \quad & \inf \mu F_\sharp(s) + \langle x_0; s \rangle \\ & s \in (L^\perp + s_0) \cap K^\sharp, \end{aligned} \quad (1.4)$$

for a monotone decreasing sequence of *barrier parameter* values $(\mu_k)_\mathbb{N} \rightarrow 0+$. Under the above made assumptions, (P_μ) and (D_μ) have unique optimal solutions for all $\mu > 0$. The KKT conditions are necessary and sufficient optimality conditions for these strictly convex problems, because the linear independence constraint qualification holds, see e.g. Borwein–Lewis [3]. The KKT conditions for (P_μ) are $\mu F'(x) + s_0 + z = 0$, $z \in L^\perp$ and $x \in L + x_0$. Moreover, $F(x) < \infty$ implies that $x \in K$. Setting $s = z + s_0$, we get

$$s = -\mu F'(x), \quad s \in L^\perp + s_0, \quad x \in L + x_0, \quad (1.5)$$

and it can be shown that $-\mu F'(x) \in K^\sharp$ and $-\mu F'_\sharp(-\mu F'(x)) = x$, see [29]. Therefore, the first equation in (1.5) can be reformulated as $x = -\mu F'_\sharp(s)$ and implies that $s \in K^\sharp$. This shows that the KKT conditions for (P_μ) and (D_μ) are equivalent, a property that is referred to as *primal-dual symmetry*. Since both problems are strictly convex, the solution pair $(x_\mu, s_\mu) \in K \times K^\sharp$ is unique. The path $\mu \mapsto (x_\mu, s_\mu)$ is called the *primal-dual central path* of (1.1). The paradigm of interior point methods is to follow the central path to the optimal solution of (1.1) that lies at its endpoint.

In the LP case (1.5) takes the form

$$s = \mu x^{-1}, \quad Ax = b, \quad s = c - A^T y, \quad (1.6)$$

where A is a matrix with nullspace L , b and c are vectors, (x, s, y) are the vectors of unknowns, x^{-1} is the componentwise inverse of x , and $s > 0$, $x > 0$ componentwise. Writing $X = \text{diag}(x)$ for the diagonal matrix with $X_{ii} = x_i$, and $e = (1, \dots, 1)^T$ for the vector of ones, the first equation in (1.6) can be rewritten as $\gamma(x, s) = \mu e$, where $\gamma(x, s) = Xs$. The definition of γ is primal-dual symmetric, because $Xs = Sx$ for $S = \text{diag}(s)$. The linearization of $\gamma(x, s) = \mu e$ yields the Newton system $Sd_x + Xd_s = \mu e - \gamma(x, s)$, $Ad_x = 0$, $d_s = -A^T d_y$, or, expressed in terms of γ ,

$$\frac{\partial}{\partial x} \gamma(x, s)[d_x] + \frac{\partial}{\partial s} \gamma(x, s)[d_s] = \mu e - \gamma(x, s), \quad d_s \in L^\perp, \quad d_x \in L. \quad (1.7)$$

In the main result of this article we show that the operator fields X and S and, by extension, the target map γ can be generalized so that the NT direction is defined as the *target direction* obtained as the solution of the Newton equation (1.7), see Theorem 3.7. This result yields a new motivation for the NT direction as a special case of a more general family of search directions with compelling properties described in the next paragraph. For a classical motivation of the NT approach see the last part of this introduction.

In the case of linear programming (1.7) defines the standard search direction for primal-dual IPMs. Kojima–Misuno–Yoshise [22] and Kojima–Megiddo–Noma–Yoshise [23] showed that γ is a diffeomorphism that transforms the primal-dual strictly feasible domain

$$\mathcal{F}(PD) := \{x \in \mathbb{R}^n : x > 0, Ax = b\} \times \{s \in \mathbb{R}^n : s > 0, \exists y \in \mathbb{R}^m \text{ s.t. } A^T y + s = c\}$$

into the positive orthant $\mathbb{R}_{++}^n := \{v \in \mathbb{R}^n : v > 0\}$. The primal-dual central path is rectified in the process, because $\gamma(x_\mu, s_\mu) = \mu e$. This makes it possible to monitor the progress of IPMs in the image space of γ , which is often called V -space. The paradigm of following the central path $\gamma^{-1}(\{\mu e : \mu > 0\})$ – also called the set of *analytic centers* – can therefore be relaxed and replaced by the new paradigm of following any ray $\gamma^{-1}(\{\mu v : \mu > 0\})$ where $v \in \mathbb{R}_{++}^n$. Points along such rays are called *weighted analytic centers*. It is possible to follow such rays by computing search directions based on the Newton equation $\frac{\partial}{\partial x} \gamma(x, s)[d_x] + \frac{\partial}{\partial s} \gamma(x, s)[d_s] = \mu v - \gamma(x, s)$. The V -space approach based on weighted analytic centers offers additional flexibility in the design of algorithms and conceptual simplicity in their analysis. This framework has therefore attracted a lot of interest in the IPM community. Several competing notions of V -space have been proposed both for LP and SDP, notions that are conceptually related but not equivalent: Jansen–Roos–Terlaky–Vial [20] and Roos–Terlaky–Vial [32] used the transformation $\gamma^{\frac{1}{2}}$ defined by the componentwise square-root of the γ defined above. They called $\gamma^{\frac{1}{2}}$ the *target map* and developed a theory of target-following algorithms for linear programming. By slight abuse of language we will call any V -space transformation a *target map* in the sequel. Monteiro–Pang [27], Sturm–Zhang [34], Monteiro–Zanjacomo [26], and Burer–Monteiro [4] all proposed V -space approaches for SDP that are based on slightly different target maps.

In this article, we use a V -space generalization that was independently developed both by Tunçel [39, 40] and Hauser [15], apart from the difference that the latter approach includes a differentiable structure which is needed to define an associated target map. This leads to the only generalization of the Kojima LP target map that inherits all of its essential properties. Let us now briefly describe this construction. The primal-dual strictly feasible domain of the general self-scaled programming problem pair (1.1) is given as

$$\mathcal{F}(PD) := (K \cap (L + x_0)) \times (K^\sharp \cap (L^\perp + s_0)). \quad (1.8)$$

Although the base space E is endowed with an inner product, we find it sometimes conceptually preferable to distinguish between E and its dual E^\sharp and think of this inner product as a bilinear form $\langle \cdot; \cdot \rangle : E \times E^\sharp \rightarrow \mathbb{R}$. Let V be a Euclidean space with

inner product $\langle \cdot; \cdot \rangle$ and dimension $\dim V = \dim E$. Let $e_V \in V$ be a fixed vector with $\|e_V\|_2 = \nu^{\frac{1}{2}}$, where $\nu := \sup\{\langle F''(x)^{-1}[-F'(x)]; -F'(x) \rangle : x \in K\}$ is the complexity parameter of the self-scaled barrier F , see [28] or [31]. The bilinear products $\langle \cdot; \cdot \rangle$ and $(\cdot; \cdot)$ define a notion of *adjoint* $\varphi^* : V \rightarrow E^\sharp$ of a linear operator $\varphi : E \rightarrow V$ via the usual requirement that $\langle x; \varphi^*(v) \rangle = (\varphi(x); v)$ for all $(x, v) \in E \times V$. Analogously, a notion of adjoint ψ^* exists for linear operators $\psi : E^\sharp \rightarrow V$. The gist in generalizing γ is to find appropriate generalizations of the operators X and S , defined as $X = \text{diag}(x)$, $S = \text{diag}(s)$ in the LP case. If we aim at preserving all the essential properties of γ from the LP framework, then the conditions we need to impose on X and S follow naturally from the NT equations (1.17) introduced further below: we must find sufficiently smooth operator fields $X : \mathcal{F}(PD) \rightarrow \text{Iso}(E^\sharp, V)$ and $S : \mathcal{F}(PD) \rightarrow \text{Iso}(E, V)$, such that for each $(x, s) \in \mathcal{F}(PD)$,

$$\begin{aligned} X^*(x, s) \circ X(x, s) &= F''^{-1}(x), \\ X(x, s) \circ F''(w(x, s)) &= S(x, s), \\ \text{and } X(x, s)[-F'(x)] &= e_V. \end{aligned} \tag{1.9}$$

The point $w(x, s)$ that appears in the second equation is the *scaling point* of x and s : Nesterov–Todd [29] showed that whenever K is a symmetric cone and F a self-scaled barrier for K then for all $x \in K$ and $s \in K^\sharp$ there exists a unique point $w(x, s) \in K$ such that $F''(w(x, s))[x] = s$. Our definition of X and S is primal-dual symmetric, because equations (1.9) are equivalent to their dual analogues. The following is an example of such a pair of operator fields: we endow E with an inner product under which K is self-dual: $E \simeq E^\sharp$, $K = K^\sharp$. This implies that there exists a unique $e \in K$ such that $F''(e) = \iota$, and this point also satisfies $\|e\|_2 = \nu^{\frac{1}{2}}$, see [29]. Let us choose $V = E$, $e_V = e$, $X(x, s) = F''^{-1}(w(e, -F'(x)))$ and $S(x, s) = X(x, s) \circ F''(w(x, s))$. Then (X, S) is a pair of operator fields that satisfy conditions (1.9), see [39, 15]. Note that X represents a square-root of $F''^{-1}(x)$, and S a square-root of $F''^{-1}(s)$, with respect to appropriately chosen coordinate systems. It can be shown that operator fields X and S that satisfy the conditions (1.9) can be constructed so that X depends only on the primal variables x and S only on the dual variables s if and only if K is the interior of a positive orthant, that is, only when (1.1) corresponds to the linear programming problem, see [15]. Thus, the general theory is necessarily more complicated than the LP case. Nevertheless, any pair of operator fields (X, S) that satisfy the conditions (1.9) defines a generalized target map via the assignment

$$\begin{aligned} \gamma : \mathcal{F}(PD) &\rightarrow V, \\ (x, s) &\mapsto X(x, s)[s] = S(x, s)[x]. \end{aligned} \tag{1.10}$$

This generalized target map inherits all properties of its LP version. This includes the rectification of the central path (note that substitution of (1.5) into the last equation of (1.9) shows that $\gamma(x_\mu, s_\mu) = \mu e_V$) and the transformation of the primal-dual strictly feasible domain into a cone isomorphic to K . The only weakening that can occur is that γ may be one-to-one only in a neighborhood of the central path, see Theorem 4.3.3, [15].

Let us conclude this introduction by presenting a classical motivation of the NT approach explained from a modern perspective. Recall that any self-scaled barrier F is of

the form (1.3). It is then the case that $-\mu F'(x) = \mu \bigoplus_{i=1}^p c_i x_i^{-1}$, or in the particular case where $c_i = 1$ for all i , $-\mu F'(x) = \mu x^{-1}$, where x^{-1} denotes the Jordan algebra inverse of x . The canonical way of solving the system of nonlinear equations (1.5) would appear to be as follows: a multiplication of the first equation with $\bigoplus_{i=1}^p c_i x_i$, using Jordan algebra multiplication, transforms the equations into primal-dual symmetric form:

$$\bigoplus_{i=1}^p c_i x_i s_i = \mu e, \quad s \in L^\perp + s_0, \quad x \in L + x_0, \quad (1.11)$$

Here $x_i s_i$ is the Jordan algebra product of x_i and s_i , and hence this is a member of E . One can then apply a damped Newton method to (1.11) and enforce the constraints $x \in K$, $s \in K^\sharp$ explicitly using line searches. Indeed, the approach we have just described leads to a family of algorithms which was first analyzed by Alizadeh–Haeberly–Overton [1] in the case of semidefinite programming, although their motivation for the method was different, see the explanations following (1.13) below. The generalization to symmetric cones and the interpretation of the method in the Jordan algebra setting is due to Faybusovich [8]. The work of Nesterov and Todd [29, 30], though later leading to the discovery of the connections between IPMs and Jordan algebras, was originally motivated by an earlier interpretation of the system (1.5): in the case of semidefinite programming (SDP) where K is the cone of $n \times n$ symmetric positive definite matrices $S^{n \times n}(\mathbb{R})_{++}$ Equation (1.5) takes the form

$$S = \mu X^{-1}, \quad \text{tr}(A_i(X - X_0)) = 0, \quad (i = 1, \dots, m), \quad S = S_0 - \sum_{i=1}^m y_i A_i, \quad (1.12)$$

where $S, X \in S^{n \times n}(\mathbb{R})_{++}$, $y \in \mathbb{R}^m$, $S_0, X_0 \in S^{n \times n}(\mathbb{R})_{++}$ are fixed positive definite symmetric matrices, and where $A_i \in S^{n \times n}(\mathbb{R})$ ($i = 1, \dots, m$) are $n \times n$ symmetric matrices. X, S , and y are the unknown variables. In this case we have $L = \{X \in S^{n \times n}(\mathbb{R}) : \text{tr}(A_i X) = 0\}$ and $L^\perp = \text{span}\{A_i : i = 1, \dots, m\}$. $S^{n \times n}(\mathbb{R})$ is a Euclidean Jordan algebra when endowed with the multiplication $(X, S) \mapsto \frac{1}{2}(XS + SX)$. Thus, if the term X^{-1} in the first equation of (1.12) is interpreted as the Jordan algebra inverse of X , then Jordan algebra multiplication with X yields the AHO equation

$$\frac{1}{2}(XS + SX) = \mu I. \quad (1.13)$$

However, X^{-1} is also the inverse of X under standard matrix multiplication. Matrix multiplication of the first equation in (1.12) by X then yields $XS = \mu I$. Note that XS is in general not symmetric. Therefore, the image space of this system is higher dimensional than the preimage space, which makes a direct application of Newton's method impossible. A wealth of fixes to this problem have been proposed. One solution is to apply the Gauss–Newton method instead of Newton's, see Kruk et.al. [24]. Most other solutions are based on symmetrizing the equation $XS = \mu I$, see Todd [38] for a survey. Equation (1.13) and the AHO approach were also originally motivated in this vein, [1].

One of the drawbacks of symmetrization is that the resulting search directions are not scale-invariant. Let us consider the AHO method as an example. For any fixed

$W \in S^{n \times n}(\mathbb{R})_{++}$, one can reformulate the primal SDP problem equivalently as follows:

$$\begin{aligned}
 (P) \quad & \min \operatorname{tr}(XS_0) \\
 \text{s.t.} \quad & \operatorname{tr}(A_i X) = b_i, \quad (i = 1, \dots, m) \\
 & X \succeq 0
 \end{aligned}
 \Leftrightarrow
 \begin{aligned}
 (\hat{P}) \quad & \min \operatorname{tr}(\hat{X}\hat{S}_0) \\
 \text{s.t.} \quad & \operatorname{tr}(\hat{A}_i \hat{X}) = \hat{b}_i, \quad (i = 1, \dots, m) \\
 & \hat{X} \succeq 0,
 \end{aligned}$$

where $\hat{X} = W^{-1}XW^{-1}$, $\hat{S}_0 = WS_0W$, $\hat{A}_i = WA_iW$ and $\hat{b}_i = b_i$. The dual problem has a corresponding reformulation with new dual variables $\hat{S} = WSW$, $\hat{y}_i = y_i$. The problem pairs $((P), (D))$ and $((\hat{P}), (\hat{D}))$ represent the same geometric problem represented in two different coordinate systems. A coordinate independent (scale-invariant) algorithm would move along sequences of points that correspond to one another via the same coordinate transformation when running on the problem inputs $((P), (D))$ and $((\hat{P}), (\hat{D}))$ respectively. But for this to be true, any search direction used by the algorithm would have to be scale-invariant too. However, the AHO equation for the rescaled variables $(\hat{X}\hat{S} + \hat{S}\hat{X})/2 = \mu I$ is equivalent to $XS + W^2SXW^{-2} = 2\mu I$ and generally leads to different Newton updates than (1.13). Other symmetrizations of $XS = \mu I$ lead to the same drawback. In order to overcome this defect, Nesterov and Todd took a different approach to symmetrization: multiplying the linearization of $XS = \mu I$ with X^{-1} , one gets

$$\Delta_S + X^{-1}\Delta_X S = \mu X^{-1} - S. \quad (1.14)$$

Note that $Z \mapsto X^{-1}ZS$ maps X to S . But likewise does the map $Z \mapsto W^{-1}ZW^{-1}$, where

$$W = X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2}, \quad (1.15)$$

and this map takes $S^{n \times n}(\mathbb{R})$ to $S^{n \times n}(\mathbb{R})$, whereas the map $Z \mapsto X^{-1}ZS$ does not. Thus, $Z \mapsto W^{-1}ZW^{-1}$ is a symmetrized version of $Z \mapsto X^{-1}ZS$. Replacing therefore (1.14) by $\Delta_S + W^{-1}\Delta_X W^{-1} = \mu X^{-1} - S$ and rewriting this equation in the form

$$\Delta_S + F''(W)[\Delta_X] = -\mu F'(X) - S, \quad (1.16)$$

where $F(Z) = -\ln \det Z$ is the ordinary logarithmic barrier function for the cone of positive definite symmetric matrices, one can check that the resulting search directions Δ_X, Δ_S are scale-invariant. This approach can be used on an arbitrary symmetric cone K endowed with an arbitrary self-scaled barrier F . Indeed, Nesterov–Todd [29] showed that every pair $(x, s) \in K \times K^\sharp$ defines a unique *scaling point* $w \in K$ such that $s = F''(w)[x]$. The NT direction (d_x, d_s) is then defined as the solution to the generalization of equation (1.16):

$$F''(w)d_x + d_s = -\mu F'(x) - s, \quad d_s \in L^\perp, \quad d_x \in L. \quad (1.17)$$

Various IPMs based on this search direction have been analyzed by Nesterov–Todd [29, 30], and variants of this method have been efficiently implemented by Toh–Todd–Tütüncü [37] and by Sturm [35].

2 A Variational Analysis of the Nesterov–Todd Flow

We will now develop a variational analysis of the NT direction field and the flow associated with it. Let X and S be fixed operator fields that satisfy the conditions of (1.9). The associated target map γ (see (1.10)) will serve as an essential tool in our analysis.

We start by placing the primal and dual problems from (1.1) and (1.4) in the setting of a single space: consider the vector space $Z := L \oplus L^\perp$, which has the same dimension as E and which we call the *primal-dual domain*. Let us consider the projections $\pi_L : Z \rightarrow L$ and $\pi_{L^\perp} = \iota - \pi_L$ of Z onto L and L^\perp along L^\perp and L respectively, where ι denotes the identity mapping. Since K is self-dual, there exists an element $e \in K$ such that $(F''(e))^{-1} = \iota$ is the canonical embedding $E^\sharp \hookrightarrow E$, see [29]. Therefore, we can endow Z with the inner product

$$(z_1; z_2) := \langle \pi_L z_2; F''(e) \pi_L z_1 \rangle + \langle F''^{-1}(e) \pi_{L^\perp} z_2; \pi_{L^\perp} z_1 \rangle.$$

Z thereby becomes a Euclidean space in which L and L^\perp are mutually orthogonal. The following coordinate transformation allows us to parametrize $\mathcal{F}(PD)$ (see (1.8)) with variables in Z :

$$\begin{aligned} x(z) &= x_0 + \pi_L z, & s(z) &= s_0 + \pi_{L^\perp} z, \\ z(x, s) &= (x - x_0) \oplus (s - s_0). \end{aligned}$$

Since both X and S are defined on $\mathcal{F}(PD)$, we can write

$$X(z) := X(x(z), s(z)), \quad S(z) := S(x(z), s(z)), \quad \gamma(z) := X(x(z), s(z)) [s(z)]$$

for $z \in \mathcal{F}(PD)$. It can easily be established that $z \mapsto F(x(z)) + F_\sharp(s(z))$ is a ν -self-concordant barrier for the convex open set $\mathcal{F}(PD)$, where ν is the common complexity parameter of F and F_\sharp , see [28] or [31]. Despite its quadratic appearance, the function $\text{gap}(z) := \langle x(z); s(z) \rangle$ is a linear functional on Z . Indeed, $\langle x(z) - x_0; s(z) - s_0 \rangle = 0$, so $\text{gap}(z) = \langle \pi_L z; s_0 \rangle + \langle x_0; \pi_{L^\perp} z \rangle + \langle x_0; s_0 \rangle$. This is the so-called *duality gap* of $x(z)$ and $s(z)$ and has the important property that

$$\begin{aligned} (PD) \quad & \inf \{ \text{gap}(z) : z \in \mathcal{F}(PD) \} \quad \text{and} \\ (PD_\mu) \quad & \min \{ \text{gap}(z) + \mu (F(x(z)) + F_\sharp(s(z))) \} \end{aligned}$$

are optimization problems that are equivalent to (1.1) and (1.4) respectively, see e.g. [15]. Thus, the primal-dual central path is the set of minimizers z_μ of (PD_μ) for all $\mu > 0$: $x(z_\mu) = x_\mu$, $s(z_\mu) = s_\mu$. The paradigm of the primal-dual framework is to reduce the duality gap to zero while maintaining feasibility.

We are now going to present a series of results which are proven in [15]. These are technical arguments that typically rely on propagating bounds via ODEs especially engineered to that aim. Though most of the properties described below are unsurprising in the sense that one would expect these from a good search direction, these results are new and not straightforward to prove, because the NT direction is implicitly defined with respect to axiomatic objects. All of these results play important roles in Section 3.

Let us fix a value $\mu > 0$ of the barrier parameter, and let us consider the corresponding NT direction which is defined as the solution (d_x, d_s) to the system (1.17). Using our parameterization in Z , we can define vector fields

$$\begin{aligned} d_x(z) &= d_x(x(z), s(z)), & d_s(z) &= d_s(x(z), s(z)), \\ d(z) &= (d_x(x(z), s(z)), d_s(x(z), s(z))), \end{aligned}$$

which are all in $C^\infty(\mathcal{F}(PD), Z)$, see [15]. The standard existence and uniqueness results for solutions of ordinary differential equation imply that $d(z)$ is the phase velocity field of a C^∞ maximal local flow $\varphi : W \rightarrow \mathcal{F}(PD)$, where $W \subset \mathbb{R} \times \mathcal{F}(PD)$ is an open set containing $\{0\} \times \mathcal{F}(PD)$, and $I_z := \{t : (t, z) \in W\}$ is the time interval over which the flux line through z is defined, see any textbook on differential topology, e.g., [5]. φ and d are then related as follows:

$$\begin{aligned} \frac{\partial}{\partial t} \varphi(0, z) &= d(z), & \forall z \in \mathcal{F}(PD), \\ \varphi(0, z) &= z, & \forall z \in \mathcal{F}(PD), \\ \varphi(t_1 + t_2, z) &= \varphi(t_2, \varphi(t_1, z)), \\ \forall z \in \mathcal{F}(PD), t_1, t_2 \in \mathbb{R} \text{ s.t. } (t_1 + t_2, z), (t_1, z) &\in W. \end{aligned} \tag{2.1}$$

Let us now investigate the global behaviour of the NT flow. The distance of $\varphi(t, z)$ from z_μ is best measured in the image under γ . Recall that $\gamma(z_\mu) = \mu e_V$. For all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$ and for all $t \in I_z$ we have

$$\|\mu e_V - \gamma(\varphi(t, z))\| = \|\mu e_V - \gamma(z)\| e^{-t}, \tag{2.2}$$

see Lemma 5.2.1 of [15]. The flux lines of φ extend to the point z_μ when moving in the positive time direction, and to the boundary of $\mathcal{F}(PD)$ or infinity when moving in negative time direction. In fact, for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$ there exists $l_z \in (-\infty, 0)$ such that $I_z = (l_z, +\infty)$, and

$$\lim_{t \rightarrow +\infty} \varphi(t, z) = z_\mu, \tag{2.3}$$

$$\lim_{t \rightarrow l_z} \left(\text{gap}(\varphi(t, z)) + \frac{\mu}{2} (F(x(\varphi(t, z))) + F_\sharp(s(\varphi(t, z)))) \right) = +\infty, \tag{2.4}$$

see Lemma 5.2.2 of [15]. If z is close enough to the central path, then the distance formula (2.2) provides an estimate for the corresponding distance in the preimage space: there exist real numbers $\delta > 0$ and $\sigma > 1$ such that for all $z \in B_\delta(z_\mu) \cap \mathcal{F}(PD) \setminus \{z_\mu\}$ and $t \in [0, +\infty)$,

$$\sigma^{-1} \|z_\mu - z\| e^{-t} \leq \|z_\mu - \varphi(t, z)\| \leq \sigma \|z_\mu - z\| e^{-t}, \tag{2.5}$$

see Lemma 5.2.3 of [15].

Next, we investigate the flux line $\varphi(t, z)$ through $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$. We are particularly interested in the effect caused at a later time when z is perturbed at time 0. First, we

note that the integral $\int_0^{+\infty} d(\varphi(t, z)) dt$ is absolutely convergent for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$, that is, the flux lines of φ are of bounded variation. This follows from Lemma 5.2.5 of [15], which shows that for $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$,

$$\|d(\varphi(t, z))\| \leq O(\|z_\mu - z\|) e^{-t}. \quad (2.6)$$

We derive from this inequality that

$$\int_0^{+\infty} \|d(\varphi(t, z))\| dt = O(\|z_\mu - z\|). \quad (2.7)$$

Lemma 5.2.6 of [15] then shows that the derivative of d is approximately the negative identity mapping in a neighborhood of z_μ :

$$d'(z) = -\text{id} + O(\|z_\mu - z\|). \quad (2.8)$$

The first order growth of perturbations in the initial value z can be described by the finite-time Lyapunov exponents λ_i of the linearized flow around the orbit $\varphi(t, z)$. In the case of the NT flow, all of these exponents satisfy $\lambda_i = -1 + O(\|z_\mu - z\|)$, as follows from the following inequality proven in Lemma 5.2.7 of [15]:

$$\|v\| e^{-t(1+O(\|z_\mu - z\|))} \leq \left\| \frac{\partial}{\partial z} \varphi(t, z)[v] \right\| \leq \|v\| e^{-t(1-O(\|z_\mu - z\|))}. \quad (2.9)$$

Lemma 5.2.7 of [15] also shows that

$$\left\| \frac{\partial}{\partial z} d(\varphi(t, z))[v] \right\| = (1 + O(\|z_\mu - z\|)) \left\| \frac{\partial}{\partial z} \varphi(t, z)[v] \right\|. \quad (2.10)$$

Together with (2.9) this implies that the integral $\int_0^{+\infty} \frac{\partial}{\partial z} d(\varphi(t, z))[\cdot] dt$ is absolutely convergent: for all $t \geq 0$ and $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$,

$$\int_0^{+\infty} \left\| \frac{\partial}{\partial z} d(\varphi(t, z))[\cdot] \right\| dt = 1 + O(\|z_\mu - z\|). \quad (2.11)$$

The second order variations are characterized in Lemma 5.2.8 of [15], which shows that for all $\epsilon > 0$ and $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$ close enough to z_μ ,

$$\left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[\cdot; \cdot] \right\| \leq 2\|d''(z_\mu)\| \frac{1+\epsilon}{1-\epsilon} e^{-t(1-\epsilon)}. \quad (2.12)$$

This equation implies that for all $t \geq 0$ and $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$,

$$\int_0^{+\infty} \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[\cdot; \cdot] \right\| dt \leq 2\|d''(z_\mu)\| (1 + O(\|z_\mu - z\|)), \quad (2.13)$$

and

$$\left\| \frac{\partial^2}{\partial z^2} \varphi(t, z)[v; w] \right\| = e^{-t(1-O(\|z_\mu - z\|))} O(1). \quad (2.14)$$

Finally, Lemma 5.2.9 of [15] shows that

$$\int_0^{+\infty} \frac{\partial}{\partial z} d(\varphi(t, z)) [\cdot] dt = -\iota, \quad (2.15)$$

$$\int_0^{+\infty} \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) [\cdot; \cdot] dt = 0. \quad (2.16)$$

We conclude our variational analysis by investigating the limiting behaviour of the directions from which flux lines approach z_μ . Lemma 5.2.10 of [15] shows that for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$, the corresponding flux line has a limiting direction, because

$$\lim_{t \rightarrow +\infty} e^t d(\varphi(t, z)) \quad (2.17)$$

exists. Moreover, Lemma 5.2.11 of [15] shows that

$$\lim_{t \rightarrow +\infty} e^t (d(\varphi(t, z)) - (z_\mu - \varphi(t, z))) = 0. \quad (2.18)$$

As one would expect, the NT flow is strictly contracting in a neighborhood of z_μ . In fact, there exists a radius $\delta > 0$ such that for all $z \in B_\delta(z_\mu)$,

$$\|z_\mu - \varphi(t, z)\| < \|z_\mu - z\| \quad \forall t > 0, \quad (2.19)$$

see Lemma 5.2.12 of [15]. And finally, Lemmas 5.2.13 and 5.2.14 of [15] show that for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$, the limit

$$\lim_{t \rightarrow +\infty} e^t \frac{\partial}{\partial t} (e^t d(\varphi(t, z))) \quad (2.20)$$

exists, and that

$$\begin{aligned} \frac{\partial}{\partial z} d(\varphi(t, z)) &= e^{-t} (-\iota + O(\|z_\mu - z\|)), \\ \frac{\partial}{\partial z} \varphi(t, z) &= e^{-t} (\iota + O(\|z_\mu - z\|)). \end{aligned} \quad (2.21)$$

3 Nesterov-Todd Directions in the Target Framework

In this section we will continue to use the primal-dual framework introduced in Section 2 and analyze the NT direction defined by a fixed point z_μ on the primal-dual central path, that is, we consider the vector field $d(z) : \mathcal{F}(PD) \rightarrow Z$ that solves the system of equations

$$F''((w(x(z)), s(z))) \pi_L d(z) + \pi_{L^\perp} d(z) = -\mu F'(x(z)) - s(z), \quad (3.1)$$

c.f. (1.17). Our goal is to construct a pair of C^2 operator fields (X, S) that satisfy the conditions (1.9), and such that the associated target map γ (see (1.10)) has the property that the NT direction satisfies the Newton equation

$$\gamma'(z)d(z) = \mu e_V - \gamma(z) \quad (3.2)$$

for all $z \in \mathcal{F}(PD)$. In other words, we will prove that the NT direction is a special case of a target direction, see Theorem 3.7.

Before we start the construction of (X, S) , let us further explore the difference between the systems (3.1) and (3.2). Multiplying (3.1) by $X(z)$, we get $S(z)\pi_L d(z) + X(z)\pi_{L^\perp} d(z) = \mu e_V - \gamma(z)$, which can be written as

$$M(z)d(z) = \mu e_V - \gamma(z), \quad (3.3)$$

where $M \in C^2(\mathcal{F}(PD), \mathcal{L}(Z, V))$ is the operator field $M(z) : d \mapsto S(z)\pi_L d + X(z)\pi_{L^\perp} d$.

On the other hand, for all $z \in \mathcal{F}(PD)$, and for fixed orthogonal bases on $E \simeq E^\sharp$ and V , a linear operator $F''^{-\frac{1}{2}}(x(z)) : E^\sharp \rightarrow V$ is well-defined with respect to these bases by the unique positive definite symmetric square-root of the matrix that represents $F''^{-1}(x(z)) : E^\sharp \rightarrow E$ with respect to the basis on E . Likewise, $F_\sharp''^{-\frac{1}{2}}(s(z)) \in \mathcal{L}(E, V)$ is well-defined. It can then be shown (see Chapter 3 of [15]) that there exist C^2 operator fields $\Omega_x, \Omega_s : \mathcal{F}(PD) \rightarrow \mathcal{O}(V)$ such that

$$\begin{aligned} X(z) &= \Omega_x(z)F''^{-\frac{1}{2}}(x(z)), \\ S(z) &= \Omega_s(z)F_\sharp''^{-\frac{1}{2}}(s(z)), \end{aligned} \quad (3.4)$$

and then (1.9) implies that

$$\Omega_s(z) = \Omega_x(z)F''^{-\frac{1}{2}}(x(z))F''(w(x(z), s(z)))F_\sharp''^{\frac{1}{2}*}(s(z)).$$

In (3.4), $\mathcal{O}(V)$ denotes the set of orthogonal transformations of V , endowed with the canonical differentiable structure that turns it into a differentiable manifold and a topological group. This is an example of a Lie group (see e.g. [18]), and we call it the *orthogonal group* of V . Now, applying the product rule in the computation of $\gamma'(z)$ and splitting the left hand side of (3.2) into parts, we get

$$M(z)d(z) + R(z)d(z) = \mu e_V - \gamma(z), \quad (3.5)$$

where $R \in C^2(\mathcal{F}(PD), \mathcal{L}(Z, V))$ is the operator field

$$R(z) : d \mapsto (\Omega'_s(z)[\pi_L d] \circ \Omega_s^*(z) + \Omega'_x(z)[\pi_{L^\perp} d] \circ \Omega_x^*(z)) \gamma(z).$$

Therefore, the NT equation (3.3) and the target equation (3.5) differ only in the term $R(z)d(z)$, and for (3.2) to hold we need to construct the operator fields (X, S) such that

$$R(z)d(z) \equiv 0. \quad (3.6)$$

Proposition 4.1.9, [15] shows that for all $z \in \mathcal{F}(PD)$, $M(z)$ is nonsingular, $\dim(\ker R(z)) \geq 2$ and $\text{im } R(z) \subseteq \text{span}\{e_V, \gamma(z)\}^\perp$. Moreover, if z lies on the central path,

then $R(z) = 0$. Since $R(z)$ has a nontrivial kernel, the requirement (3.6) is not a priori impossible to satisfy. Ideally, we would like to construct (X, S) such that $d(z) \in \ker R(z)$ for all $z \in \mathcal{F}(PD)$ and for the NT direction fields arising from *all* possible values of $\mu > 0$ *simultaneously*. A necessary and sufficient condition for this to be true would be that this requirement can be satisfied for only *two different values* of μ simultaneously (see [15]). Again, this requirement is not a priori impossible to satisfy because $\dim(\ker R(z)) \geq 2$. However, the difficulties of proving that such a pair of operator fields (X, S) exists seem rather extraordinary and we restrict our analysis to the NT field corresponding to a fixed value of $\mu > 0$ throughout.

Equations (1.9) and (3.4) show that any two pairs of operator fields (X, S) and (\tilde{X}, \tilde{S}) must be related to each other via a C^2 operator field $\Omega^* : \mathcal{F}(PD) \rightarrow \mathcal{O}(V, e_V) := \{\theta \in \mathcal{O}(V) : \theta^* e_V = e_V\}$ as follows:

$$X(z) = \Omega^*(z)\tilde{X}(z), \quad S(z) = \Omega^*(z)\tilde{S}(z).$$

This means that for our construction of a pair of operator fields (X, S) that satisfy the requirement (3.6), we can start with an arbitrary known pair of operator fields (\tilde{X}, \tilde{S}) that satisfy the conditions (1.9), e.g., the example of Section 1 for which $\tilde{X}, \tilde{S} \in C^\infty$, and then we must construct a C^2 operator field $\Omega^* : \mathcal{F}(PD) \rightarrow \mathcal{O}(V, e_V)$ such that $(X, S) = (\Omega^* \circ \tilde{X}, \Omega^* \circ \tilde{S})$ satisfies (3.6). We adopt the adjoint notation Ω^* for later convenience. Let us denote the operator fields R, Ω_x and Ω_s associated with (X, S) and (\tilde{X}, \tilde{S}) respectively by $R(z)$ and $\tilde{R}(z)$, $\Omega_x(z)$ and $\tilde{\Omega}_x(z)$, $\Omega_s(z)$ and $\tilde{\Omega}_s(z)$ respectively. Likewise, let us write $\gamma(z)$ and $\tilde{\gamma}(z)$ respectively for the associated target map. Then $\Omega_x(z) = \Omega^*(z)\tilde{\Omega}_x(z)$, $\Omega_s(z) = \Omega^*(z)\tilde{\Omega}_s(z)$, and

$$R(z)[d(z)] = \Omega^{*'}[d(z)]\tilde{\gamma}(z) + \Omega^*(z) \left(\tilde{\Omega}_s'(z)[\pi_L d]\tilde{\Omega}_s^*(z) + \tilde{\Omega}_x'(z)[\pi_{L^\perp} d]\tilde{\Omega}_x^*(z) \right) \tilde{\gamma}(z).$$

Therefore, the condition (3.6) is equivalent to

$$(\Omega^{*'}[d(z)])[\tilde{\gamma}] = -\Omega^* \left(\tilde{\Omega}_s'[\pi_L d(z)]\tilde{\Omega}_s^* + \tilde{\Omega}_x'[\pi_{L^\perp} d(z)]\tilde{\Omega}_x^* \right) [\tilde{\gamma}(z)] \quad (3.7)$$

for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$. For $z = z_\mu$ we don't need to make any assumptions, because $R(z_\mu) = \tilde{R}(z_\mu) = 0$, as remarked above. However, for specificity, we require that $\Omega^*(z_\mu) = \iota$ be the identity map. Moreover, we strengthen the condition (3.7) by dropping the multiplication with $\tilde{\gamma}(z)$. Taking adjoints and using $\Omega\Omega^{*'} + \Omega'\Omega^* = 0$, the requirement becomes finding a C^2 operator field $\Omega : \mathcal{F}(PD) \rightarrow \mathcal{O}(V, e_V)$ such that

$$\begin{aligned} \Omega(z_\mu) &= \iota, \\ \Omega'(z)[d(z)] &= \left(\tilde{\Omega}_s'(z)[\pi_L d(z)]\tilde{\Omega}_s^*(z) + \tilde{\Omega}_x'(z)[\pi_{L^\perp} d(z)]\tilde{\Omega}_x^*(z) \right) \Omega(z) \quad \forall z \in \mathcal{F}(PD) \setminus \{z_\mu\}. \end{aligned} \quad (3.8)$$

Note that (3.8) constitutes a boundary value problem: this is a partial differential equation for an operator valued function $z \mapsto \Omega(z) \in \mathcal{O}(V, e_V)$ with domain of definition $\mathcal{F}(PD) \setminus \{z_\mu\}$ and with the requirement that the boundary condition $\Omega(z_\mu) = \iota$ be satisfied at the isolated boundary point z_μ . Thus, for the purposes of showing the existence of

(X, S) that satisfy (3.6), it suffices to show that the boundary value problem (3.8) has a C^2 solution which can be extended in a twice continuously differentiable manner at the boundary point z_μ . Indeed, we are going to show that the boundary value problem (3.8) has a unique solution, and that its extension to z_μ is C^2 , see Theorem 3.7. Showing the last property is the technically most difficult part of the proof.

Lemma 3.1. *The boundary value problem (3.8) has a solution that can be extended in a twice continuously differentiable manner at z_μ if and only if it has such a solution in a neighborhood of z_μ .*

Proof. The only if part is of course trivially true. Let us therefore assume that there exists an open ball $B_\delta(z_\mu) \subset \mathcal{F}(PD)$ and a mapping $\check{\Omega} \in C^2(B_\delta(z_\mu), O(V, e_V))$ such that

$$\check{\Omega}'(z)[d(z)] = \left(\check{\Omega}'_s(z)[\pi_L d(z)]\check{\Omega}_s^*(z) + \check{\Omega}'_x(z)[\pi_{L^\perp} d(z)]\check{\Omega}_x^*(z) \right) \check{\Omega}(z) \quad \forall z \in B_\delta(z_\mu) \setminus \{z_\mu\}.$$

Consider the following boundary value problem:

$$\begin{aligned} \Omega(z) &= \check{\Omega}(z) \quad \forall z \in \partial B_{\delta/2}(z_\mu), \\ \Omega'(z)[d(z)] &= \left(\check{\Omega}'_s(z)[\pi_L d(z)]\check{\Omega}_s^*(z) + \check{\Omega}'_x(z)[\pi_{L^\perp} d(z)]\check{\Omega}_x^*(z) \right) \Omega(z) \quad \forall z \in \mathcal{F}(PD) \setminus \bar{B}_{\delta/2}(z_\mu). \end{aligned} \quad (3.9)$$

For any $z \in \partial B_{\delta/2}(z_\mu)$, the standard existence and uniqueness theorems for solutions of ordinary differential equations can be implied to show that there exists a unique function $\Omega(\varphi(t, z))$ that satisfies (3.9) for all points $\varphi(t, z)$ on the interval $t \in (l_z, 0]$ (see Section 2 for notation). The standard theorems on the smooth dependence of solutions of ODEs on parameters also imply that $\Omega(\varphi(t, z))$ varies in a C^2 fashion as a function of z . The required Ω is then obtained for all $z \in \mathcal{F}(PD) \setminus B_{\delta/2}(z_\mu)$ by setting $\Omega(z) := \Omega(\varphi(t, \hat{z}))$ where (t, \hat{z}) is the unique point in $\mathbb{R}_- \times \partial B_{\delta/2}(z_\mu)$ such that $z = \varphi(t, \hat{z})$. It follows from the arguments above that the extension Ω is unique and coincides with $\check{\Omega}$ on the intersection of their domains of definition. For a more detailed proof, see Lemma 3.11 of [15]. \square

Notational Convention 3.2. In the remainder of the present and subsequent sections the following shorthand notation will often be employed, where v_i are vectors:

$$\begin{aligned} \tilde{\Omega}'[\pi v_1]\tilde{\Omega}^* &:= \tilde{\Omega}'_s[\pi_L v_1]\tilde{\Omega}_s^* + \tilde{\Omega}'_x[\pi_{L^\perp} v_1]\tilde{\Omega}_x^*, \\ \tilde{\Omega}'[\pi v_1]\tilde{\Omega}^{*'}[v_2] &:= \tilde{\Omega}'_s[\pi_L v_1]\tilde{\Omega}_s^{*'}[v_2] + \tilde{\Omega}'_x[\pi_{L^\perp} v_1]\tilde{\Omega}_x^{*'}[v_2], \\ \tilde{\Omega}''[\pi v_1, v_2]\tilde{\Omega}^* &:= \tilde{\Omega}''_s[\pi_L v_1, v_2]\tilde{\Omega}_s^* + \tilde{\Omega}''_x[\pi_{L^\perp} v_1, v_2]\tilde{\Omega}_x^*, \\ \tilde{\Omega}'[\pi v_1]\tilde{\Omega}^{*''}[v_2, v_3] &:= \tilde{\Omega}'_s[\pi_L v_1]\tilde{\Omega}_s^{*''}[v_2, v_3] + \tilde{\Omega}'_x[\pi_{L^\perp} v_1]\tilde{\Omega}_x^{*''}[v_2, v_3], \\ \tilde{\Omega}''[\pi v_1, v_2]\tilde{\Omega}^{*'}[v_3] &:= \tilde{\Omega}''_s[\pi_L v_1, v_2]\tilde{\Omega}_s^{*'}[v_3] + \tilde{\Omega}''_x[\pi_{L^\perp} v_1, v_2]\tilde{\Omega}_x^{*'}[v_3], \\ \tilde{\Omega}'''[\pi v_1, v_2, v_3]\tilde{\Omega}^* &:= \tilde{\Omega}'''_s[\pi_L v_1, v_2, v_3]\tilde{\Omega}_s^* + \tilde{\Omega}'''_x[\pi_{L^\perp} v_1, v_2, v_3]\tilde{\Omega}_x^*. \end{aligned}$$

We will henceforth concentrate on the problem of showing the existence and uniqueness of a solution of (3.8) which is locally defined around z_μ and C^2 extendable there. For

$z \in \mathcal{F}(PD) \setminus \{z_\mu\}$, let us consider the following coordinate change for the time parameter of the flux line $\varphi(t, z)$:

$$r(t, z) = \begin{cases} e^{-t} & \text{if } t \in (l_z, +\infty), \\ 0 & \text{if } t = +\infty, \end{cases}$$

where $l_z < 0$ is defined as in (2.4). Then $[0, 1] \subset \text{im}(r)$ for all $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$. We write $t(r, z)$ for the inverse of $r(t, z)$ and

$$\psi(r, z) := \begin{cases} \varphi(t(r, z), z) & \text{if } r > 0, \\ z_\mu & \text{if } r = 0. \end{cases}$$

We claim that $\psi \in C^1$ for any fixed $z \in \mathcal{F}(PD) \setminus \{z_\mu\}$. In fact, it follows from (2.3) that ψ is continuous. Moreover, (2.17) shows that the limit

$$\lim_{r \rightarrow 0} \frac{\partial}{\partial r} \psi(r, z) = \lim_{r \rightarrow 0} \frac{\partial}{\partial t} \varphi(t(r, z), z) \frac{\partial}{\partial r} t(r, z) = \lim_{t \rightarrow +\infty} \frac{d(\varphi(t, z))}{-r(t, z)} = - \lim_{t \rightarrow +\infty} e^t d(\varphi(t, z))$$

exists.

Suppose that Ω is a local solution to (3.8), defined on $B_\delta(z_\mu)$, where $\delta > 0$ is chosen small enough for (2.19) to be true. Then the function

$$y(r, z) = \Omega(\psi(r, z))$$

is well-defined on $[0, 1] \times B_\delta(z_\mu)$, and for $r > 0$ we have

$$\begin{aligned} \frac{\partial}{\partial r} y(r, z) &= \Omega'(\psi(r, z)) \left[\frac{\partial}{\partial r} \psi(r, z) \right] \\ &= \Omega'(\psi(r, z)) \left[\frac{\partial}{\partial t} \varphi(t(r, z), z) \frac{\partial}{\partial r} t(r, z) \right] \\ &= -\Omega'(\psi(r, z)) \left[\frac{d(\psi(r, z))}{r} \right] \\ &\stackrel{(3.8)}{=} -\tilde{\Omega}'(\psi(r, z)) \left[\pi \frac{d(\psi(r, z))}{r} \right] \tilde{\Omega}^*(\psi(r, z)) \Omega(\psi(r, z)) \\ &= a(r, z) y(r, z), \end{aligned} \tag{3.10}$$

where the mapping

$$a(r, z) = -\tilde{\Omega}'(\psi(r, z)) \left[\pi \frac{d(\psi(r, z))}{r} \right] \tilde{\Omega}^*(\psi(r, z)) \tag{3.11}$$

is defined on $(0, 1] \times B_\delta(z_\mu)$.

$O(V, e_V)$ is a closed subgroup of the Lie group $O(V)$, and it is therefore a Lie group itself, see e.g. [5]. Moreover, since $\Omega(z) \in O(V, e_V)$ for all $z \in B_\delta(z_\mu)$, we have

$$\frac{\partial}{\partial r} y(r, z) \in T_{y(r, z)} O(V, e_V), \tag{3.12}$$

where $T_y \mathcal{O}(V, e_V)$ denotes the tangent space of $\mathcal{O}(V, e_V)$ at y . It is a trivial fact from the theory of Lie groups that $(T_y \mathcal{O}(V, e_V))y^{-1} = T_i \mathcal{O}(V, e_V)$, where ι is the identity mapping, that is, ι is the multiplicative neutral element of $\mathcal{O}(V, e_V)$. $T_i \mathcal{O}(V, e_V)$, henceforth denoted by $\mathfrak{o}(V, e_V)$, consists of the set of skew-adjoint endomorphisms of V that contain e_V in their kernel, that is,

$$v \in \mathfrak{o}(V, e_V) \Leftrightarrow v \in \text{End}(V), \quad v^* = -v, \quad ve_V = 0.$$

This characterization shows that the following *commutator* operation is well-defined:

$$\begin{aligned} [\cdot, \cdot] : \mathfrak{o}(V, e_V) \times \mathfrak{o}(V, e_V) &\rightarrow \mathfrak{o}(V, e_V), \\ [u, v] &\mapsto uv - vu. \end{aligned}$$

When $\mathfrak{o}(V, e_V)$ is endowed with this operation, it becomes a *Lie algebra*, see e.g. [18]. This is called the *Lie algebra associated with* the Lie group $\mathcal{O}(V, e_V)$.

Equations (3.10) and (3.12) show that $a \in C^1((0, 1] \times B_\delta(z_\mu), \mathfrak{o}(V, e_V))$ (see (3.11)). We claim that for fixed $z \in B_\delta(z_\mu)$, a can be extended to $[0, 1] \times \{z\}$ in a C^1 fashion. In fact, $\tilde{\Omega}' \in C^1$ and (2.17) shows that

$$\lim_{r \rightarrow 0} \frac{d(\psi(r, z))}{r} = \lim_{t \rightarrow +\infty} e^t d(\varphi(t, z))$$

exists. This proves that a can be continuously extended at $(0, z)$. On the other hand, (2.20) shows that

$$\lim_{r \rightarrow 0} \frac{\partial}{\partial r} \left(\frac{d(\psi(r, z))}{r} \right) = \lim_{t \rightarrow +\infty} \frac{\partial}{\partial t} (e^t d(\varphi(t, z))) \frac{\partial t(r, z)}{\partial r} = - \lim_{t \rightarrow +\infty} e^t \frac{\partial}{\partial t} (e^t d(\varphi(t, z)))$$

exists. Together with $\tilde{\Omega} \in C^2$ and $\psi \in C^\infty$ this proves that the extension of a is continuously differentiable with respect to r at $(0, z)$.

In summary, we have shown that if a local solution Ω to (3.8) exists, then $y(r, z) = \Omega(\psi(r, z))$ must satisfy the differential equation

$$\begin{aligned} y(0, z) &= \iota, \\ \frac{\partial}{\partial r} y(r, z) &= a(r, z)y(r, z) \quad (r \in [0, 1]), (z \in B_\delta(z_\mu)), \end{aligned} \tag{3.13}$$

where a is defined as in (3.11) and continuously extended at $r = 0$. If (3.13) has a unique solution and if we can integrate this equation then we know Ω along the characteristic $\psi(\cdot, z)$. In particular, since this characteristic flows through z , $\Omega(z)$ is uniquely determined. Thus, if (3.8) has a local solution, then (3.13) provides a *mechanism to find this solution* explicitly. On the other hand, if (3.13) has a unique solution for all $z \in B_\delta(z_\mu) \setminus \{z_\mu\}$ then $\Omega(z) := y(1, z)$ satisfies (3.8) for all $z \in B_\delta(z_\mu) \setminus \{z_\mu\}$. In Lemma 3.4 we will prove that this is indeed the case. In Lemma 3.5 we will then prove that $\Omega(z_\mu) = \iota$ extends this solution in a twice continuously differentiable manner at z_μ . This proves the existence and uniqueness of a local solution for (3.8), and together with Lemma 3.1 this constitutes a proof of Theorem 3.7.

For a fixed $z \in B_\delta(z_\mu)$, (3.13) is a linear ordinary differential equation evolving on the Lie group $O(V, e_V)$ and is driven by the operator $a(\cdot, z) \in C^1([0, 1], \mathfrak{o}(V, e_V))$. This type of initial value problem was studied by Hausdorff [12] for general Lie groups G and their associated Lie algebras \mathfrak{g} . Substituting $G = O(V, e_V)$, it follows from this theory that there exists a number $r^* > 0$ and a function $\sigma(\cdot, z) \in C^1([0, r^*], \mathfrak{o}(V, e_V))$ such that $y(r, z) = \exp(\sigma(r, z))$ is the unique solution of (3.13) on $r \in [0, r^*]$, where \exp is the matrix exponential, and where $\sigma(\cdot, z)$ satisfies the *dexpinv* equation

$$\begin{aligned} \sigma(0, z) &= 0, \\ \frac{\partial}{\partial r} \sigma(r, z) &= \sum_{m=0}^{\infty} h_m \text{ad}^m(a(r, z), \sigma(r, z)), \quad (r \in [0, r^*]). \end{aligned} \quad (3.14)$$

In (3.14), h_m is the m -th Taylor series coefficient of the function $h : \mathbb{C} \rightarrow \mathbb{C}$,

$$h(w) = \frac{w}{e^w - 1} + w \quad (3.15)$$

expanded around $w = 0$, and the ad -operator ad^k is recursively defined as follows: $\text{ad}^0(v, u) = v$, and $\text{ad}^k(v, u) = [\text{ad}^{k-1}(v, u), u]$ for $k \in \mathbb{N}$, where $[\cdot, \cdot]$ denotes the commutator operator defined above. Using Picard-Lindelöf iteration it is possible to explicitly determine more and more terms of a series development for the solution of (3.14). Magnus [25] derived the first four terms of this series:

$$\begin{aligned} \sigma(r, z) &= \int_0^r a(\kappa, z) d\kappa + \frac{1}{2} \int_0^r \left[a(\kappa, z), \int_0^\kappa a(\xi, z) d\xi \right] d\kappa \\ &\quad + \frac{1}{4} \int_0^r \left[a(\kappa, z), \int_0^\kappa \left[a(\xi, z), \int_0^\xi a(\eta, z) d\eta \right] d\xi \right] d\kappa \\ &\quad + \frac{1}{12} \int_0^r \left[\left[a(\kappa, z), \int_0^\kappa a(\xi, z) d\xi \right], \int_0^\kappa a(\eta, z) d\eta \right] d\kappa + \dots \end{aligned} \quad (3.16)$$

The general term of this series was characterized by Iserles-Nørsett [19]. We now describe their construction for the special case where $G = O(V, e_V)$ that applies to our problem. Consider the set of functions $\mathcal{E} \subset \mathcal{F}([0, r^*], \mathfrak{o}(V, e_V))$ for which membership is defined by recursively applying the following rules: $a(\cdot, z) \in \mathcal{E}$, and if $p, q \in \mathcal{E}$ then $r \mapsto [p(r), \int_0^r q(\kappa) d\kappa] \in \mathcal{E}$. It would be difficult to work with \mathcal{E} without a proper indexing system. This is most elegantly achieved by use of rooted trees. We recursively apply the following rules: the map $a(\cdot, z)$ is associated with the tree τ_0 consisting of a single node,

$$a(r, z) \sim \tau_0 = \bullet,$$

and if $p(r), q(r) \in \mathcal{E}$ are associated with the trees $\tau^{[1]}$ and $\tau^{[2]}$ respectively, then the mapping

$$r \mapsto \left[p(r), \int_0^r q(\kappa) d\kappa \right] \in \mathcal{E}$$

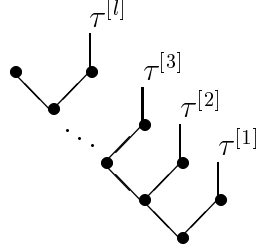
is associated with the tree obtained by appending a new root to $\tau^{[2]}$ and joining the resulting tree with $\tau^{[1]}$ via a new root on the left:

$$\left[p(r), \int_0^r q(\kappa) d\kappa \right] \sim \begin{array}{c} \tau^{[1]} \quad \tau^{[2]} \\ \diagdown \quad \diagup \\ \bullet \end{array} \quad (3.17)$$

We denote the set of trees that can be obtained in this fashion by \mathcal{T} , and we denote the member of \mathcal{E} associated with τ by $H_\tau(\cdot, z)$. By \mathcal{T}_k we denote the set of members of \mathcal{T} that contain k nodes, and we say that these trees are of *order* k . An induction argument shows that all trees in \mathcal{T} are of order $3k + 1$ for some $k \in \mathbb{N}_0$. Iserles–Nørsett [19] proved that

$$\#\mathcal{T}_{3k+1} = \frac{(2k)!}{k!(k+1)!} \quad \forall k \in \mathbb{N}_0. \quad (3.18)$$

Each $\tau \in \mathcal{T}_k$ can be written uniquely in the form



for some trees $\tau^{[1]}, \dots, \tau^{[l]} \in \mathcal{T}$ of order strictly less than k . We write $\tau = \mathcal{R}(\tau^{[1]}, \dots, \tau^{[l]})$ to express this relationship. For later convenience, let us denote the tree $\mathcal{R}(\tau_0)$ by τ_1 . With this notation it is possible to define a sequence of numbers $(\alpha_\tau)_\tau$ by recursively applying the following rules: $\alpha_{\tau_0} = h_0$, and if all $\alpha_{\tau^{[i]}}$ are defined for $(i = 1, \dots, l)$, then $\alpha_{\mathcal{R}(\tau^{[1]}, \dots, \tau^{[l]})} = h_l \prod_{i=1}^l \alpha_{\tau^{[i]}}$, where h_l is defined with respect to (3.15) as above. Note that since the function h has a convergence radius strictly greater than 1, we have

$$|\alpha_\tau| \leq 1 \quad (3.19)$$

for all τ of sufficiently high order (actually for all $\tau \in \mathcal{T}$). It follows from the results of Iserles–Nørsett [19] that the general term in the series (3.16) is $\alpha_\tau \int_0^r H_\tau(\kappa, z) d\kappa$, that is, the solution to the dexpinv equation (3.14) is given by the *Magnus series*

$$\sigma(r, z) = \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^r H_\tau(\kappa, z) d\kappa \quad (3.20)$$

on the interval $[0, \tilde{r})$ where both this series and its termwise derivative converge absolutely. Moreover, the solution of the initial value problem (3.13) is given by $y(r, z) = \exp(\sigma(r, z))$ on the interval $[0, r^*]$ on which $\exp(\sigma(r, z))$ is defined and $r^* \leq \tilde{r}$. We will see below that $r^* \geq 1$.

We will now express the Magnus series (3.20) in terms of the parameters (t, z) instead of (r, z) . For each $\tau \in \mathcal{T}$ we can define a function $L_\tau(t, z) : [0, +\infty) \times B_\delta(z_\mu) \rightarrow \mathfrak{o}(V, e_V)$ by recursively applying the following rules:

(i) $L_{\tau_0}(t, z) := -\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))]\tilde{\Omega}^*(\varphi(t, z)).$

(ii) If τ is the tree defined in (3.17) then

$$L_\tau(t, z) := \left[L_{\tau[1]}(t, z), \int_t^{+\infty} L_{\tau[2]}(\theta, z) d\theta \right].$$

The functions defined above then satisfy

$$L_\tau(t, z) = e^{-t} H_\tau(e^{-t}, z) \quad (3.21)$$

for all $\tau \in \mathcal{T}$. This can easily be seen via induction.

Now note that if $\sigma(\cdot, z)$ is expressible by the Magnus series (3.20), then $\Omega(z) = \Omega(\varphi(0, z)) = y(1, z) = \exp(\varsigma(z))$, where

$$\begin{aligned} \varsigma(z) &:= \sigma(1, z) = \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^1 H_\tau(\kappa, z) d\kappa = \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_{+\infty}^0 \frac{\partial r(t, z)}{\partial t} H_\tau(r(t, z), z) dt \\ &= \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^{+\infty} L_\tau(t, z) dt. \end{aligned} \quad (3.22)$$

So far we have treated z as a fixed parameter, but (3.22) now shows that the free variable t disappears when taking the integral. We therefore obtain an explicit series representation for $\Omega(z)$ as a function of z , now considered as a variable. Let us endow $\mathfrak{o}(V, e_V)$ with the usual operator matrix norm. Recall that we chose δ small enough for (2.19) to be true. For our further analysis we need to restrict the neighborhood around z_μ even further. The results of Section 2 and the fact that $\tilde{\Omega} \in C^\infty$ imply that it is possible to choose $Q > 0$ large enough and $\varrho > 0$ small enough so that the following inequalities are satisfied for all $z \in B_\varrho(z_\mu) \setminus \{z_\mu\}$ and $t \geq 0$:

$$\|\tilde{\Omega}'(z)[\pi \cdot] \tilde{\Omega}^*(z)\| \leq Q, \quad (3.23)$$

$$\|\tilde{\Omega}''(z)[\pi \cdot; \cdot] \tilde{\Omega}^*(z)\|, \|\tilde{\Omega}'(z)[\pi \cdot] \tilde{\Omega}^{*'}(z)[\cdot]\| \leq Q, \quad (3.24)$$

$$\|\tilde{\Omega}'''(z)[\pi \cdot; \cdot; \cdot] \tilde{\Omega}^*(z)\|, \|\tilde{\Omega}''(z)[\pi \cdot; \cdot] \tilde{\Omega}^{*'}(z)[\cdot]\|, \|\tilde{\Omega}'(z)[\pi \cdot] \tilde{\Omega}^{*''}(z)[\cdot; \cdot]\| \leq Q, \quad (3.25)$$

$$\|d(\varphi(t, z))\| \leq Q \|z_\mu - z\| e^{-t}, \text{ see (2.5) and (2.18),} \quad (3.26)$$

$$\|z_\mu - \varphi(t, z)\| \leq \|z_\mu - z\|, \text{ see (2.19),} \quad (3.27)$$

$$\left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| \leq e^{-\frac{t}{2}}, \text{ see (2.9),} \quad (3.28)$$

$$\left\| \frac{\partial^2}{\partial z^2} \varphi(t, z) \right\| \leq Q, \text{ see (2.14),} \quad (3.29)$$

$$\left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) \right\| \leq Q e^{-\frac{t}{2}}, \text{ see (2.12),} \quad (3.30)$$

$$\int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| dt \leq 1 + Q \|z_\mu - z\|, \text{ see (2.11),} \quad (3.31)$$

$$\int_0^\infty \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) \right\| dt \leq Q(1 + Q \|z_\mu - z\|), \text{ see (2.13).} \quad (3.32)$$

Lemma 3.3. *If $Q, \varrho > 0$ are chosen so that (3.23)-(3.32) hold true, then for all $k \in \mathbb{N}_0$, $\tau \in \mathcal{T}_{3k+1}$ and $z \in B_\varrho(z_\mu) \setminus \{z_\mu\}$ the following inequalities hold true:*

$$\begin{aligned} i) \quad & \int_0^\infty \|L_\tau(t, z)\| dt \leq \|z_\mu - z\|^{k+1} Q^{2k+2} 2^k, \\ ii) \quad & \int_0^\infty \left\| \frac{\partial}{\partial z} L_\tau(t, z) \right\| dt \leq \|z_\mu - z\|^k Q^{2k+1} 2^{2k} + \|z_\mu - z\|^{k+1} Q^{2k+2} 3^{2k+1}, \\ iii) \quad & \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_\tau(t, z) \right\| dt \leq \|z_\mu - z\|^{k-1} Q^{2k} \chi(k) + \|z_\mu - z\|^k Q^{2k+1} (3^{2k+2} + Q 3^{2k}) \\ & \quad + \|z_\mu - z\|^{k+1} Q^{2k+2} (3^{2k+3} + Q 3^{2k+1}), \end{aligned}$$

where

$$\chi(k) = \begin{cases} 0 & \text{if } k = 0, \\ 2^{2(k+1)} & \text{if } k \geq 1. \end{cases}$$

We will prove this lemma in Section 4.

Lemma 3.4. *There exists a radius $\rho > 0$ such that the mapping $\varsigma(z)$ defined in (3.22) is well defined and twice continuously differentiable on $B_\rho(z_\mu) \setminus \{z_\mu\}$. Moreover,*

$$\begin{aligned} i) \quad & \varsigma(z) = O(\|z_\mu - z\|), \\ ii) \quad & \frac{\partial}{\partial z} \varsigma(z) = \int_0^\infty \frac{\partial}{\partial z} L_{\tau_0}(t, z) dt + O(\|z_\mu - z\|), \\ iii) \quad & \frac{\partial^2}{\partial z^2} \varsigma(z) = \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_0}(t, z) dt + \frac{1}{2} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_1}(t, z) dt + O(\|z_\mu - z\|). \end{aligned}$$

We will prove this lemma in Section 5.

Lemma 3.5. *Let $\rho > 0$ be chosen as in Lemma 3.4 and let $\varsigma(z)$ be continuously extended at z_μ , that is, $\varsigma(z_\mu) = 0$. Then $\varsigma \in C^2(B_\rho(z_\mu), \mathfrak{o}(V, e_V))$. In particular, the derivatives at z_μ are given as follows: for all $v, w \in Z$,*

$$\varsigma'(z_\mu)[w] = \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu), \quad (3.33)$$

$$\begin{aligned} \varsigma''(z_\mu)[w; v] &= \frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi v; w] \tilde{\Omega}^*(z_\mu) + \frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi w; v] \tilde{\Omega}^*(z_\mu) \\ &\quad + \frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^{*'}(z_\mu)[v] + \frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^{*'}(z_\mu)[w]. \end{aligned} \quad (3.34)$$

We will prove this lemma in Section 6.

Theorem 3.6 (Local Solution).

Let $\rho > 0$ and $\varsigma(z)$ be chosen as in Lemma 3.5. Then $\Omega(z) = \exp(\varsigma(z))$ is a twice continuously differentiable solution of (3.8) defined on $B_\rho(z_\mu)$. The continuous derivatives at the boundary point z_μ are given as follows: for all $v, w \in Z$,

$$\Omega'(z_\mu)[w] = \varsigma'(z_\mu)[w], \quad (3.35)$$

$$\Omega''(z_\mu)[w; v] = \varsigma''(z_\mu)[w; v] + \frac{1}{2} \varsigma'(z_\mu)[w] \varsigma'(z_\mu)[v] + \frac{1}{2} \varsigma'(z_\mu)[v] \varsigma'(z_\mu)[w] \quad (3.36)$$

Proof. The first statement is clear from Lemma 3.5, the fact that the exponential mapping is analytic and the developments that led to equation (3.22). In order to prove the second statement, note that at the origin the first and second derivatives of the matrix exponential are as follows: for all $V, W \in M_{n \times n}(\mathbb{R})$,

$$\exp'(0)[V] = V, \quad (3.37)$$

$$\exp''(0)[W; V] = \frac{1}{2}(WV + VW). \quad (3.38)$$

Since $\mathfrak{o}(V, e_V)$ is a matrix Lie algebra, it is therefore the case that for all $v, w \in Z$,

$$\Omega'(z_\mu)[w] = \exp'(\varsigma(z_\mu))[\varsigma'(z_\mu)[w]] = \exp'(0)[\varsigma'(z_\mu)[w]] \stackrel{(3.37)}{=} \varsigma'(z_\mu)[w],$$

and

$$\begin{aligned} \Omega''(z_\mu)[w; v] &= \exp'(\varsigma(z_\mu))[\varsigma''(z_\mu)[w; v]] + \exp''(\varsigma(z_\mu))[\varsigma'(z_\mu)[w]; \varsigma'(z_\mu)[v]] \\ &= \exp'(0)[\varsigma''(z_\mu)[w; v]] + \exp''(0)[\varsigma'(z_\mu)[w]; \varsigma'(z_\mu)[v]] \\ &\stackrel{(3.38)}{=} \varsigma''(z_\mu)[w; v] + \frac{1}{2} \varsigma'(z_\mu)[w] \varsigma'(z_\mu)[v] + \frac{1}{2} \varsigma'(z_\mu)[v] \varsigma'(z_\mu)[w]. \end{aligned}$$

□

Theorem 3.7 (NT and Target).

There exists a unique twice continuously differentiable operator field $\Omega : \mathcal{F}(PD) \rightarrow \mathcal{O}(V, e_V)$ that solves the boundary value problem (3.8). Moreover, the NT direction solves the Newton system (3.2) when the target map γ is defined with respect to $X(z) = \Omega^(z) \circ X(z)$.*

Proof. The existence of Ω follows from Theorem 3.6 and Lemma 3.1. Moreover, the local solution constructed in Theorem 3.6 is unique because $y(r, z) = \exp(\sigma(r, z))$ is the unique local solution of (3.13), by virtue of Hausdorff's theory of the dexpinv equation [12]. Lemma 3.1 shows that there is a unique extension of this local solution to all of $\mathcal{F}(PD)$. Finally, (3.8) was explicitly designed so as to render the remaining claims true. \square

4 Proof of Lemma 3.3

Proof. We use induction over k . For $k = 0$ we have $\tau_0 = \bullet$, and then we can check claims i), ii), and iii) of Lemma 3.3 as follows:

$$\begin{aligned}
 i) \quad & \int_0^\infty \|L_{\tau_0}(t, z)\| dt = \int_0^\infty \|\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))]\tilde{\Omega}^*(\varphi(t, z))\| dt \\
 & \stackrel{(3.23), (3.27)}{\leq} \int_0^\infty Q \|d(\varphi(t, z))\| dt \stackrel{(3.26)}{\leq} Q \int_0^\infty \|z_\mu - z\| Q e^{-t} dt \\
 & = 2^0 Q^{0+2} \|z_\mu - z\|^{0+1}.
 \end{aligned}$$

$$\begin{aligned}
 ii) \quad & \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau_0}(t, z) \right\| dt \leq \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
 & + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
 & + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))]\tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
 & \stackrel{(3.23), (3.24), (3.27)}{\leq} Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt + Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| dt \\
 & + Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
 & \stackrel{(3.26), (3.28), (3.31)}{\leq} 2Q^2 \|z_\mu - z\| + Q(1 + Q\|z_\mu - z\|) \\
 & = \|z_\mu - z\|^0 Q^{0+1} 2^0 + \|z_\mu - z\|^1 Q^{0+2} 3^{0+1}.
 \end{aligned}$$

$$\begin{aligned}
& iii) \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_{\tau_0}(t, z) \right\| dt \\
& \leq \int_0^\infty \left\| \tilde{\Omega}'''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial^2}{\partial z^2} \varphi(t, z) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}'^*(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)) \right] \tilde{\Omega}'^*(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) \right] \tilde{\Omega}'^*(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)) \right] \tilde{\Omega}'^*(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^{*''}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z); \frac{\partial}{\partial z} \varphi(t, z) \right] \right\| dt \\
& \quad + \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}'^*(\varphi(t, z)) \left[\frac{\partial^2}{\partial z^2} \varphi(t, z) \right] \right\| dt
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(3.23)-(3.25),(3.27)}{\leq} Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 dt \\
& + Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt + Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial^2}{\partial z^2} \varphi(t, z) \right\| dt \\
& + Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 dt + Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& + Q \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) \right\| dt + Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& + Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 dt + Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& + Q \int_0^\infty \|d(t, z)\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 dt + Q \int_0^\infty \|d(\varphi(t, z))\| \left\| \frac{\partial^2}{\partial z^2} \varphi(t, z) \right\| dt \\
& \stackrel{(3.28),(3.29)}{\leq} (4 + 2Q)Q \int_0^\infty \|d(\varphi(t, z))\| dt \\
& \quad + 4Q \int_0^\infty \left\| \frac{\partial}{\partial z} d(\varphi(t, z)) \right\| dt + Q \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) \right\| dt \\
& \stackrel{(3.26),(3.31),(3.32)}{\leq} (4 + 2Q)Q^2 \|z_\mu - z\| + 4Q(1 + Q\|z_\mu - z\|) + Q^2(1 + Q\|z_\mu - z\|) \\
& \leq \|z_\mu - z\|^{-1} Q^0 0 + \|z_\mu - z\|^0 Q^1 (3^2 + Q3^0) + \|z_\mu - z\|^1 Q^2 (3^3 + Q3^1).
\end{aligned}$$

This completes the base case. In order to prove the induction step, let $k \geq 1$ and suppose the lemma holds true for all $\tau \in \mathcal{T}_{3i+1}$, ($i = 0, \dots, k-1$). Let $\tau \in \mathcal{T}_{3k+1}$. Because of the recursive definition of \mathcal{T} there exist an integer $l < k$ and two oriented rooted trees $\tau^{[1]} \in \mathcal{T}_{3l+1}$ and $\tau^{[2]} \in \mathcal{T}_{3(k-l-1)+1}$ such that τ is related to $\tau^{[1]}$ and $\tau^{[2]}$ as in (3.17). Therefore, assuming that statements i), ii) and iii) of the lemma hold for $\tau^{[1]}$ and $\tau^{[2]}$, the following arguments show that they hold for τ too:

$$\begin{aligned}
i) \quad & \int_0^\infty \|L_\tau(t, z)\| dt = \int_0^\infty \left\| \left[L_{\tau^{[1]}}(t, z), \int_t^\infty L_{\tau^{[2]}}(\xi, z) d\xi \right] \right\| dt \\
& \leq 2 \int_0^\infty \|L_{\tau^{[1]}}(t, z)\| \int_t^\infty \|L_{\tau^{[2]}}(\xi, z)\| d\xi dt \\
& \stackrel{(2.1)}{=} 2 \int_0^\infty \|L_{\tau^{[1]}}(t, z)\| \int_0^\infty \|L_{\tau^{[2]}}(\xi, \varphi(t, z))\| d\xi dt \\
& \stackrel{i)}{\leq} \int_0^\infty \|L_{\tau^{[1]}}(t, z)\| 2^{k-l} Q^{2k-2l} \|z_\mu - \varphi(t, z)\|^{k-l} dt \\
& \stackrel{(3.27)}{\leq} 2^{k-l} Q^{2k-2l} \|z_\mu - z\|^{k-l} \int_0^\infty \|L_{\tau^{[1]}}(t, z)\| dt \\
& \stackrel{i)}{\leq} 2^{k-l} Q^{2k-2l} \|z_\mu - z\|^{k-l} 2^l Q^{2l+2} \|z_\mu - z\|^{l+1} = 2^k Q^{2k+2} \|z_\mu - z\|^{k+1}.
\end{aligned}$$

$$\begin{aligned}
ii) \quad & \int_0^\infty \left\| \frac{\partial}{\partial z} L_\tau(t, z) \right\| dt \leq \int_0^\infty \left\| \left[\frac{\partial}{\partial z} L_{\tau[1]}(t, z), \int_t^\infty L_{\tau[2]}(\xi, z) d\xi \right] \right\| dt \\
& + \int_0^\infty \left\| \left[L_{\tau[1]}(t, z); \int_t^\infty \frac{\partial}{\partial z} L_{\tau[2]}(\xi, z) d\xi \right] \right\| dt \\
& \stackrel{(2.1)}{\leq} 2 \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau[1]}(t, z) \right\| \int_0^\infty \|L_{\tau[2]}(\xi, \varphi(t, z))\| d\xi dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \int_t^\infty \left\| \frac{\partial}{\partial \tilde{z}} L_{\tau[2]}(\xi, \tilde{z}) \Big|_{\tilde{z}=\varphi(t, z)} \left[\frac{\partial}{\partial z} \varphi(t, z) \right] \right\| d\xi dt \\
& \stackrel{i), ii)}{\leq} 2^{k-l} Q^{2k-2l} \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau[1]}(t, z) \right\| \|z_\mu - \varphi(t, z)\|^{k-l} dt \\
& + 2 \cdot 2^{2k-2l-2} Q^{2k-2l-1} \int_0^\infty \|L_{\tau[1]}(t, z)\| \|z_\mu - \varphi(t, z)\|^{k-l-1} \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& + 2 \cdot 3^{2k-2l-1} Q^{2k-2l} \int_0^\infty \|L_{\tau[1]}(t, z)\| \|z_\mu - \varphi(t, z)\|^{k-l} \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& \stackrel{i), ii), (3.27), (3.28)}{\leq} 2^{k-l} Q^{2k-2l} \|z_\mu - z\|^{k-l} \{ 2^{2l} Q^{2l+1} \|z_\mu - z\|^l \\
& + 3^{2l+1} Q^{2l+2} \|z_\mu - z\|^{l+1} \} + 2 \{ 2^{2k-2l-2} Q^{2k-2l-1} \|z_\mu - z\|^{k-l-1} \\
& + 3^{2k-2l-1} Q^{2k-2l} \|z_\mu - z\|^{k-l} \} 2^l Q^{2l+2} \|z_\mu - z\|^{l+1} \\
& = \|z_\mu - z\|^k \cdot Q^{2k+1} (2^{k-l} 2^{2l} + 2^{l+1} 2^{2k-2l-2}) \\
& + \|z_\mu - z\|^{k+1} Q^{2k+2} (2^{k-l} 3^{2l+1} + 2^{l+1} 3^{2k-2l-1}) \\
& \leq \|z_\mu - z\|^k Q^{2k+1} 2^{2k} + \|z_\mu - z\|^{k+1} Q^{2k+2} 3^{2k+1}.
\end{aligned}$$

$$\begin{aligned}
iii) \quad & \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_\tau(t, z) \right\| dt = \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} \left[L_{\tau[1]}(t, z), \int_t^\infty L_{\tau[2]}(\xi, z) d\xi \right] \right\| dt \\
& \leq 2 \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_{\tau[1]}(t, z) \right\| \int_t^\infty \|L_{\tau[2]}(\xi, z)\| d\xi dt \\
& + 4 \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau[1]}(t, z) \right\| \int_t^\infty \left\| \frac{\partial}{\partial z} L_{\tau[2]}(\xi, z) \right\| d\xi dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \int_t^\infty \left\| \frac{\partial^2}{\partial z^2} L_{\tau[2]}(\xi, z) \right\| d\xi dt \\
& \stackrel{(2.1)}{\leq} 2 \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_{\tau[1]}(t, z) \right\| \int_0^\infty \|L_{\tau[2]}(\xi, \varphi(t, z))\| d\xi dt \\
& + 4 \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau[1]}(t, z) \right\| \int_0^\infty \left\| \frac{\partial}{\partial \tilde{z}} L_{\tau[2]}(\xi, \tilde{z}) \Big|_{\tilde{z}=\varphi(t, z)} \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| d\xi dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \int_0^\infty \left\| \frac{\partial^2}{\partial \tilde{z}^2} L_{\tau[2]}(\xi, \tilde{z}) \Big|_{\tilde{z}=\varphi(t, z)} \right\| \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 d\xi dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \int_0^\infty \left\| \frac{\partial}{\partial \tilde{z}} L_{\tau[2]}(\xi, \tilde{z}) \Big|_{\tilde{z}=\varphi(t, z)} \right\| \left\| \frac{\partial^2}{\partial z^2} \varphi(t, z) \right\| d\xi dt
\end{aligned}$$

$$\begin{aligned}
& \stackrel{i),ii),iii)}{\leq} 2 \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_{\tau[1]}(t, z) \right\| 2^{k-l-1} Q^{2k-2l} \|z_\mu - \varphi(t, z)\|^{k-l} dt \\
& + 4 \int_0^\infty \left\| \frac{\partial}{\partial z} L_{\tau[1]}(t, z) \right\| \left\{ 2^{2k-2l-2} Q^{2k-2l-1} \|z_\mu - \varphi(t, z)\|^{k-l-1} \right. \\
& + 3^{2k-2l-1} Q^{2k-2l} \|z_\mu - \varphi(t, z)\|^{k-l} \left. \right\} \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\| dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \left\{ \chi(k-l-1) Q^{2k-2l-2} \|z_\mu - \varphi(t, z)\|^{k-l-2} \right. \\
& + (3^{2k-2l} + Q 3^{2k-2l-2}) Q^{2k-2l-1} \|z_\mu - \varphi(t, z)\|^{k-l-1} \\
& + (3^{2k-2l+1} + Q 3^{2k-2l-1}) Q^{2k-2l} \|z_\mu - \varphi(t, z)\|^{k-l} \left. \right\} \left\| \frac{\partial}{\partial z} \varphi(t, z) \right\|^2 dt \\
& + 2 \int_0^\infty \|L_{\tau[1]}(t, z)\| \left\{ 2^{2k-2l-2} Q^{2k-2l-1} \|z_\mu - \varphi(t, z)\|^{k-l-1} \right. \\
& + 3^{2k-2l-1} Q^{2k-2l} \|z_\mu - \varphi(t, z)\|^{k-l} \left. \right\} \left\| \frac{\partial^2}{\partial z^2} \varphi(t, z) \right\| dt
\end{aligned}$$

$$\begin{aligned}
& \stackrel{(3.27)-(3.29),i),ii),iii)}{\leq} 2 \left\{ \chi(l) Q^{2l} \|z_\mu - z\|^{l-1} + (3^{2l+2} + Q 3^{2l}) Q^{2l+1} \|z_\mu - z\|^l \right. \\
& + (3^{2l+3} + Q 3^{2l+1}) Q^{2l+2} \|z_\mu - z\|^{l+1} \left. \right\} 2^{k-l-1} Q^{2k-2l} \|z_\mu - z\|^{k-l} \\
& + 4 \left\{ 2^{2l} Q^{2l+1} \|z_\mu - z\|^l + 3^{2l+1} Q^{2l+2} \|z_\mu - z\|^{l+1} \right\} \cdot \\
& \cdot \left\{ 2^{2k-2l-2} Q^{2k-2l-1} \|z_\mu - z\|^{k-l-1} + 3^{2k-2l-1} Q^{2k-2l} \|z_\mu - z\|^{k-l} \right\} \\
& + 2^{l+1} Q^{2l+2} \|z_\mu - z\|^{l+1} \left\{ \chi(k-l-1) Q^{2k-2l-2} \|z_\mu - z\|^{k-l-2} \right. \\
& + (3^{2k-2l} + Q 3^{2k-2l-2}) Q^{2k-2l-1} \|z_\mu - z\|^{k-l-1} \\
& + (3^{2k-2l+1} + Q 3^{2k-2l-1}) Q^{2k-2l} \|z_\mu - z\|^{k-l} \left. \right\} \cdot 1 \\
& + 2^{l+1} Q^{2l+2} \|z_\mu - z\|^{l+1} \left\{ 2^{2k-2l-2} Q^{2k-2l-1} \|z_\mu - z\|^{k-l-1} \right. \\
& + 3^{2k-2l-1} Q^{2k-2l} \|z_\mu - z\|^{k-l} \left. \right\} Q \\
& = \|z_\mu - z\|^{k-1} Q^{2k} \left\{ 2^{k-l} \chi(l) + 2^{2k} + 2^{l+1} \chi(k-l-1) \right\} \\
& + \|z_\mu - z\|^k Q^{2k+1} \left\{ 2^{k-l} 3^{2l+2} + 2^{2l+2} 3^{2k-2l-1} + 2^{2k-2l} 3^{2l+1} \right. \\
& + 2^{l+1} 3^{2k-2l} + Q(2^{k-l} 3^{2l} + 2^{l+1} 3^{2k-2l-2} + 2^{2k-l-1}) \left. \right\} \\
& + \|z_\mu - z\|^{k+1} Q^{2k+2} \left\{ 2^{k-l} 3^{2l+3} + 2^2 3^{2k} + 2^{l+1} 3^{2k-2l+1} \right. \\
& + Q(2^{k-l} 3^{2l+1} + 2^{l+1} 3^{2k-2l-1} + 2^{l+1} 3^{2k-2l-1}) \left. \right\}.
\end{aligned}$$

But note that

$$\begin{aligned}
2^{k-l} \chi(l) + 2^{2k} + 2^{l+1} \chi(k-l-1) & \leq \begin{cases} 2^{2k} + 2 \cdot 2^{2k} & \text{if } l = 0, \\ 2^{2l+2} 2^{k-l} + 2^{2k} + 2^{l+1} 2^{2k-2l} & \text{if } l \geq 1, \end{cases} \\
& \leq \begin{cases} 2^{2k+2} & \text{if } l = 0, \\ 2^{2k+1} + 2^{2k} + 2^{2k} & \text{if } l \geq 1, \end{cases} \leq \begin{cases} 2^{2k+2} & \text{if } l = 0, \\ 2^{2k+2} & \text{if } l \geq 1. \end{cases}
\end{aligned}$$

Moreover,

$$\begin{aligned}
& 2^{k-l}3^{2l+2} + 2^{2l+2}3^{2k-2l-1} + 2^{2k-2l}3^{2l+1} + 2^{l+1}3^{2k-2l} \\
& \leq \begin{cases} 3^{2k} \{ (2/9)^k \cdot 9 + 2 + (2/3)^k \cdot 3 + 2 \} & \text{if } l = 0, \\ 3^{2k+1} + 3^{2k+1} + 2^2 3^{2k-2l-2} 3^{2l+1} + 2^2 3^{l-1} 3^{2k-2l} & \text{if } l \geq 1, \end{cases} \\
& \leq \begin{cases} 3^{2k} \cdot 8 & \text{if } l = 0, \\ 3^{2k+1} + 3^{2k+1} + 8 \cdot 3^{2k-1} & \text{if } l \geq 1, \end{cases} \leq \begin{cases} 3^{2k+2} & \text{if } l = 0, \\ 3^{2k+2} & \text{if } l \geq 1, \end{cases}
\end{aligned}$$

And finally,

$$\begin{aligned}
& 2^{k-l}3^{2l} + 2^{l+1}3^{2k-2l-2} + 2^{2k-l-1} \leq 3 \cdot 3^{2k-1} = 3^{2k}, \\
& 2^{k-l}3^{2l+3} + 2^2 3^{2k} + 2^{l+1}3^{2k-2l+1} \leq 3 \cdot 3^{2k+2} = 3^{2k+3}, \\
& 2^{k-l}3^{2l+1} + 2^{l+1}3^{2k-2l-1} + 2^{l+1}3^{2k-2l-1} \leq 3 \cdot 3^{2k} = 3^{2k+1}.
\end{aligned}$$

This concludes the proof. \square

5 Proof of Lemma 3.4

Proof. Let Q and $\varrho > 0$ be chosen so as to render (3.23)–(3.32) true. Note that in condition (3.26) we made the implicit assumption that $Q \geq \sigma$ and $\varrho < \delta$, where σ and δ are as in (2.5). Let $\tau \in \mathcal{T}_{3k+1}$. It follows from Lemma 3.3 that all of the integrals $\int_0^\infty L_\tau(t, z)dt$, $\int_0^\infty \frac{\partial}{\partial z} L_\tau(t, z)dt$ and $\int_0^\infty \frac{\partial^2}{\partial z^2} L_\tau(t, z)dt$ exist and converge absolutely. Therefore,

$$\begin{aligned}
& \frac{\partial}{\partial z} \int_0^\infty L_\tau(t, z)dt = \int_0^\infty \frac{\partial}{\partial z} L_\tau(t, z)dt, \\
& \frac{\partial^2}{\partial z^2} \int_0^\infty L_\tau(\varphi(t, z))dt = \int_0^\infty \frac{\partial^2}{\partial z^2} L_\tau(\varphi(t, z))dt
\end{aligned}$$

for all $z \in B_\varrho(z_\mu) \setminus \{z_\mu\}$ and for all $\tau \in \mathcal{T}$.

Lemma 3.3 also implies that there exists a radius $\rho \in (0, \varrho)$ such that for $z \in B_\rho(z_\mu) \setminus \{z_\mu\}$ and $k \in \mathbb{N}$ the following inequalities hold true:

$$\int_0^\infty \|L_\tau(t, z)\|dt < 4^{-(k+1)}, \quad \int_0^\infty \left\| \frac{\partial}{\partial z} L_\tau(t, z) \right\|dt < 4^{-k}, \quad \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} L_\tau(t, z) \right\|dt < 4^{-(k-1)}. \quad (5.1)$$

Let $\Theta(z) := \sum_{k=0}^\infty (\#\mathcal{T}_{3k+1})z^k$ be the generating function of the sequence $(\#\mathcal{T}_{3k+1})_{k \in \mathbb{N}_0}$. It follows from (3.18) that

$$\Theta(z) = \frac{1 - \sqrt{1 - 4z}}{2z} = \sum_{k=0}^\infty \frac{(2k)!}{k!(k+1)!} z^k.$$

Since this function is analytic in $B_{1/4}(0) \subset \mathbb{C}$, equations (3.19) and (5.1) imply that the series

$$\sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^\infty L_\tau(t, z)dt, \quad \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^\infty \frac{\partial}{\partial z} L_\tau(t, z)dt, \quad \sum_{\tau \in \mathcal{T}} \alpha_\tau \int_0^\infty \frac{\partial^2}{\partial z^2} L_\tau(t, z)dt$$

all converge absolutely in $B_\rho(z_\mu) \setminus \{z_\mu\}$ and equal $\varsigma(z)$, $\varsigma'(z)$ and $\varsigma''(z)$ respectively. The claims i), ii) and iii) of the lemma now follow from Lemma 3.3 and the facts that $\mathcal{T}_1 = \{\tau_0\}$ and $\mathcal{T}_3 = \{\tau_1\}$ are singletons, and that $\alpha_{\tau_0} = 1$ and $\alpha_{\tau_1} = 1/2$, see (3.16). \square

6 Proof of Lemma 3.5

Proof. We already know from Lemma 3.4 that ς is twice continuously differentiable on $B_\rho(z_\mu) \setminus \{z_\mu\}$. In order to prove twofold continuous differentiability at the point z_μ it suffices to show that $\lim_{z \rightarrow z_\mu} \varsigma(z)$, $\lim_{z \rightarrow z_\mu} \varsigma'(z)$ and $\lim_{z \rightarrow z_\mu} \varsigma''(z)$ exist. Lemma 3.4 shows that

$$\lim_{z \rightarrow z_\mu} \varsigma(z) = 0, \quad (6.1)$$

$$\lim_{z \rightarrow z_\mu} \varsigma'(z) = \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial}{\partial z} L_{\tau_0}(t, z) dt, \quad \text{and} \quad (6.1)$$

$$\lim_{z \rightarrow z_\mu} \varsigma''(z) = \lim_{z \rightarrow z_\mu} \left(\int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_0}(t, z) dt + \frac{1}{2} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_1}(t, z) dt \right). \quad (6.2)$$

Therefore, all we need to show is that the limits on the right hand sides of (6.1) and (6.2) exist, and that these equal the right hand sides of (3.33) and (3.34).

Let us first show this for $\varsigma'(z_\mu)$. For all $w \in Z$ we have

$$\begin{aligned} \lim_{z \rightarrow z_\mu} \varsigma'(z)[w] &= - \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w] dt \\ &= - \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \end{aligned} \quad (6.3)$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z)) [w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \quad (6.4)$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi[w] \right] dt, \quad (6.5)$$

so long as all the limits in (6.3)–(6.5) exist. Using the fact that $\tilde{\Omega} \in C^\infty$, we can compute these limits as follows:

$$\begin{aligned} \lim_{z \rightarrow z_\mu} \left\| \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \right\| \\ \leq \lim_{z \rightarrow z_\mu} \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi \cdot; \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| \|d(\varphi(t, z))\| dt \\ \stackrel{(2.3), (2.9)}{\leq} \|\tilde{\Omega}''(z_\mu) [\pi \cdot; \cdot] \tilde{\Omega}^*(z_\mu)\| \|w\| \lim_{z \rightarrow z_\mu} \int_0^\infty \|d(\varphi(t, z))\| dt \stackrel{(2.6)}{=} 0, \end{aligned} \quad (6.6)$$

and likewise,

$$\lim_{z \rightarrow z_\mu} \left\| \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[w] \right] dt \right\| = 0.$$

Finally,

$$\begin{aligned}
& \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \\
& \stackrel{(2.3)}{=} \tilde{\Omega}'(z_\mu) \left[\pi \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial}{\partial z} d(\varphi(t, z))[w] dt \right] \tilde{\Omega}^*(z_\mu) \\
& \stackrel{(2.15)}{=} -\tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu),
\end{aligned}$$

Therefore, the limit $\lim_{z \rightarrow z_\mu} \varsigma'(z)[w] = \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu)$ exists and equation (3.33) holds true.

Let us now consider $\varsigma''(z_\mu)$. For all $v, w \in Z$ we have

$$\begin{aligned}
& \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_0}(t, z)[v; w] dt \\
& = - \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial^2}{\partial z^2} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [v, w] dt \\
& = - \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z)[w]; \frac{\partial}{\partial z} \varphi(t, z)[v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt
\end{aligned} \tag{6.7}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[v]; \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \tag{6.8}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial^2}{\partial z^2} \varphi(t, z)[w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \tag{6.9}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[v] \right] dt \tag{6.10}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[w]; \frac{\partial}{\partial z} \varphi(t, z)[v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \tag{6.11}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \tag{6.12}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[w] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[v] \right] dt \tag{6.13}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z)[v] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[w] \right] dt \tag{6.14}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[v] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[w] \right] dt \tag{6.15}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^{*''}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[w]; \frac{\partial}{\partial z} \varphi(t, z)[v] \right] dt \tag{6.16}$$

$$- \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial^2}{\partial z^2} \varphi(t, z)[w; v] \right] dt, \tag{6.17}$$

so long as the limits (6.7)–(6.17) exist.

Expressions (6.7), (6.10), (6.14) and (6.16) can be shown to be equal to zero in much the same way as (6.6), that is by relying on the smoothness of $\tilde{\Omega}$, and on (2.3) and (2.9). Likewise, (6.9) and (6.17) are equal to zero. The argument is almost identical for both expressions, and we will show it only for (6.9):

$$\begin{aligned} & \lim_{z \rightarrow z_\mu} \left\| \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial^2}{\partial z^2} \varphi(t, z)[w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \right\| \\ & \leq \lim_{z \rightarrow z_\mu} \int_0^\infty \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi \cdot; \frac{\partial^2}{\partial z^2} \varphi(t, z)[w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| \|d(\varphi(t, z))\| dt \\ & \stackrel{(2.3), (3.29)}{\leq} \|\tilde{\Omega}''(z_\mu)[\pi \cdot; \cdot] \circ \tilde{\Omega}^*(z_\mu)\| Q \|w\| \|v\| \lim_{z \rightarrow z_\mu} \int_0^\infty \|d(\varphi(t, z))\| dt \stackrel{(2.7)}{=} 0. \end{aligned}$$

Next, we show that expression (6.12) is equal to zero:

$$\begin{aligned} & \lim_{z \rightarrow z_\mu} \left\| \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \right\| \\ & \leq \lim_{z \rightarrow z_\mu} \int_0^\infty \left\| \tilde{\Omega}'(\varphi(t, z))[\pi \cdot] \tilde{\Omega}^*(\varphi(t, z)) - \tilde{\Omega}'(z_\mu)[\pi \cdot] \tilde{\Omega}^*(z_\mu) \right\| \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[w; v] \right\| dt \\ & + \lim_{z \rightarrow z_\mu} \left\| \tilde{\Omega}'(z_\mu) \left[\pi \left(\int_0^\infty \frac{\partial^2}{\partial z^2} d(\varphi(t, z))[w; v] dt \right) \right] \tilde{\Omega}^*(z_\mu) \right\| \stackrel{(2.3), (2.13), (2.16)}{=} 0. \end{aligned}$$

Let us next take the limit in expression (6.8):

$$\begin{aligned} & \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[v]; \frac{\partial}{\partial z} \varphi[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \\ & = \lim_{z \rightarrow z_\mu} \int_0^\infty -e^{-2t} \tilde{\Omega}''(\varphi(t, z)) \left[-\pi e^t \frac{\partial}{\partial z} d(\varphi(t, z))[v]; e^t \frac{\partial}{\partial z} \varphi(t, z)[w] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \\ & \stackrel{(2.21)}{=} \tilde{\Omega}''(z_\mu)[\pi v; w] \tilde{\Omega}^*(z_\mu) \int_0^\infty -e^{-2t} dt = -\frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi v; w] \tilde{\Omega}^*(z_\mu). \end{aligned}$$

Similar arguments can be applied to expressions (6.11), (6.13) and (6.15), yielding

$$\begin{aligned} & \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}''(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[w]; \frac{\partial}{\partial z} \varphi(t, z)[v] \right] \tilde{\Omega}^*(\varphi(t, z)) dt \\ & = -\frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi w; v] \tilde{\Omega}^*(z_\mu), \\ & \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[w] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[v] \right] dt \\ & = -\frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^{*'}(z_\mu)[v], \\ & \lim_{z \rightarrow z_\mu} \int_0^\infty \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial}{\partial z} d(\varphi(t, z))[v] \right] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z)[w] \right] dt \\ & = -\frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^{*'}(z_\mu)[w]. \end{aligned}$$

In summary we get that for all $v, w \in Z$,

$$\begin{aligned} \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_0}(t, z)[w; v] dt &= \frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi v; w] \tilde{\Omega}^*(z_\mu) + \frac{1}{2} \tilde{\Omega}''(z_\mu)[\pi w; v] \tilde{\Omega}^*(z_\mu) \\ &\quad + \frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^{*'}(z_\mu)[v] + \frac{1}{2} \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^{*'}(z_\mu)[w]. \end{aligned} \quad (6.18)$$

This is the first term in the right-hand side of (6.2), and we have shown that it is equal to the right-hand side of equation (3.34).

Let us now show that the second term in the right-hand side of (6.2) is equal to zero: for all $v, w \in Z$,

$$\begin{aligned} &\lim_{z \rightarrow z_\mu} \frac{1}{2} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_1}(t, z)[w; v] dt \\ &= \frac{1}{2} \lim_{z \rightarrow z_\mu} \int_0^\infty \frac{\partial^2}{\partial z^2} \left[\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)), \right. \\ &\quad \left. \int_t^\infty \tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) d\theta \right] [w; v] dt \\ &= \frac{1}{2} \lim_{z \rightarrow z_\mu} \int_0^\infty \left[\frac{\partial^2}{\partial z^2} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w; v], \right. \\ &\quad \left. \int_t^\infty \tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) d\theta \right] dt \quad (6.19) \\ &\quad + \frac{1}{2} \lim_{z \rightarrow z_\mu} \int_0^\infty \left[\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)), \right. \\ &\quad \left. \int_t^\infty \frac{\partial^2}{\partial z^2} \left(\tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) \right) [w; v] d\theta \right] dt \quad (6.20) \end{aligned}$$

$$\begin{aligned} &+ \frac{1}{2} \lim_{z \rightarrow z_\mu} \int_0^\infty \left[\frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w], \right. \\ &\quad \left. \int_t^\infty \frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) \right) [v] d\theta \right] dt \quad (6.21) \end{aligned}$$

$$\begin{aligned} &+ \frac{1}{2} \lim_{z \rightarrow z_\mu} \int_0^\infty \left[\frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [v], \right. \\ &\quad \left. \int_t^\infty \frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) \right) [w] d\theta \right] dt \quad (6.22) \end{aligned}$$

as long as all these limits exist. Note that for all z sufficiently close to z_μ ,

$$\begin{aligned} & \left\| \int_t^\infty \tilde{\Omega}'(\varphi(\theta, z)) [\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) d\theta \right\| \\ & \leq \int_t^\infty \|\tilde{\Omega}'(\varphi(\theta, z)) [\pi \cdot] \tilde{\Omega}^*(\varphi(\theta, z))\| \|d(\varphi(\theta, z))\| d\theta \\ & \stackrel{(3.26), (3.27)}{=} O(1) \|\tilde{\Omega}'(z_\mu) [\pi \cdot] \tilde{\Omega}^*(z_\mu)\| \int_t^\infty \|z_\mu - z\| e^{-\theta} d\theta = O(1) \|z_\mu - z\| e^{-t}. \end{aligned} \quad (6.23)$$

Likewise, still for z close enough to z_μ ,

$$\left\| \frac{\partial^2}{\partial z^2} \left(\tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) \right\| = O(1) e^{-\frac{t}{2}}. \quad (6.24)$$

In fact, this follows from our analysis of expressions (6.7)–(6.11) and (6.13)–(6.17), equations (2.6) and (2.10), and from the following revised analysis of expression (6.12):

$$\begin{aligned} & \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) [w; v] \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| \\ & = O(1) \|\tilde{\Omega}'(z_\mu) [\pi \cdot] \tilde{\Omega}^*(z_\mu)\| \left\| \frac{\partial^2}{\partial z^2} d(\varphi(t, z)) [w; v] \right\| \\ & \stackrel{(3.30)}{=} O(1) \|\tilde{\Omega}'(z_\mu) [\pi \cdot] \tilde{\Omega}^*(z_\mu)\| \|w\| \|v\| e^{-\frac{t}{2}}. \end{aligned}$$

(6.23) and (6.24) imply that expression (6.19) is zero, because its norm can be bounded from above by

$$\begin{aligned} & \lim_{z \rightarrow z_\mu} \int_0^\infty \left\| \frac{\partial^2}{\partial z^2} \left(\tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w; v] \right\| \\ & \quad \cdot \left\| \int_t^\infty \tilde{\Omega}'(\varphi(\theta, z)) [\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) d\theta \right\| dt \\ & \stackrel{(6.24), (6.23)}{\leq} \lim_{z \rightarrow z_\mu} O(1) \|z - z_\mu\| \int_0^\infty e^{-\frac{3t}{2}} dt = 0, \end{aligned}$$

Likewise, expression (6.20) equals zero, because its norm can be bounded similarly.

Finally, it remains to analyze expressions (6.21) and (6.22). Note that

$$\begin{aligned} & \left\| e^t \frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w] + \tilde{\Omega}'(z_\mu) [\pi w] \tilde{\Omega}^*(z_\mu) \right\| \\ & \leq e^t \left\| \tilde{\Omega}''(\varphi(t, z)) \left[\pi d(\varphi(t, z)); \frac{\partial}{\partial z} \varphi(t, z) [w] \right] \tilde{\Omega}^*(\varphi(t, z)) \right\| \\ & \quad + \left\| \tilde{\Omega}'(\varphi(t, z)) \left[\pi e^t \frac{\partial}{\partial z} d(\varphi(t, z)) [w] \right] \tilde{\Omega}^*(\varphi(t, z)) - \tilde{\Omega}'(z_\mu) \left[\pi e^t \frac{\partial}{\partial z} d(\varphi(t, z)) [w] \right] \tilde{\Omega}^*(z_\mu) \right\| \\ & \quad + \|\tilde{\Omega}'(z_\mu) [\pi \cdot] \tilde{\Omega}^*(z_\mu)\| \left\| e^t \frac{\partial}{\partial z} d(\varphi(t, z)) [w] + w \right\| \\ & \quad + e^t \left\| \tilde{\Omega}'(\varphi(t, z)) [\pi d(\varphi(t, z))] \tilde{\Omega}^{*'}(\varphi(t, z)) \left[\frac{\partial}{\partial z} \varphi(t, z) [w] \right] \right\| \end{aligned} \quad (6.25)$$

$$\begin{aligned}
& \stackrel{(2.6),(2.9),2.21}{\leq} O(\|z - z_\mu\|) e^{-\frac{t}{2}} \|w\| \\
& + \|\tilde{\Omega}'(\varphi(t, z))[\pi \cdot] \tilde{\Omega}^*(\varphi(t, z)) - \tilde{\Omega}'(z_\mu)[\pi \cdot] \tilde{\Omega}^*(z_\mu)\| \left\| e^t \frac{\partial}{\partial z} d(\varphi(t, z))[w] \right\| \\
& + \|\tilde{\Omega}'(z_\mu)[\pi \cdot] \tilde{\Omega}^*(z_\mu)\| O(\|z - z_\mu\|) + O(\|z - z_\mu\|) e^{-\frac{t}{2}} \|w\| \\
& \stackrel{(3.27),(2.21)}{=} O(\|z - z_\mu\|). \tag{6.26}
\end{aligned}$$

Equation (6.26) implies that

$$\frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w] = -e^{-t} \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu) + e^{-t} O(\|z - z_\mu\|). \tag{6.27}$$

This finally allows us to compute expression (6.21):

$$\begin{aligned}
& \lim_{z \rightarrow z_\mu} \int_0^\infty \left[\frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(t, z))[\pi d(\varphi(t, z))] \tilde{\Omega}^*(\varphi(t, z)) \right) [w] \right. \\
& \quad \left. \int_t^\infty \frac{\partial}{\partial z} \left(\tilde{\Omega}'(\varphi(\theta, z))[\pi d(\varphi(\theta, z))] \tilde{\Omega}^*(\varphi(\theta, z)) \right) [v] d\theta \right] dt \\
& \stackrel{(6.27)}{=} \lim_{z \rightarrow z_\mu} \int_0^\infty \left[-e^{-t} \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu) + e^{-t} O(\|z - z_\mu\|) \right. \\
& \quad \left. \int_t^\infty -e^{-\theta} \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^*(z_\mu) + e^{-\theta} O(\|z - z_\mu\|) d\theta \right] dt \\
& = \lim_{z \rightarrow z_\mu} \left[\tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu), \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^*(z_\mu) \right] \int_0^\infty e^{-t} \int_t^\infty e^{-\theta} d\theta dt \\
& \quad + \lim_{z \rightarrow z_\mu} O(\|z - z_\mu\|) \\
& = \frac{1}{2} \left[\tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu), \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^*(z_\mu) \right].
\end{aligned}$$

This implies of course that (6.22) equals

$$\frac{1}{2} \left[\tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^*(z_\mu), \tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu) \right] = -\frac{1}{2} \left[\tilde{\Omega}'(z_\mu)[\pi w] \tilde{\Omega}^*(z_\mu), \tilde{\Omega}'(z_\mu)[\pi v] \tilde{\Omega}^*(z_\mu) \right].$$

Expressions (6.21) and (6.22) thus cancel each other out, and we have

$$\lim_{z \rightarrow z_\mu} \frac{1}{2} \int_0^\infty \frac{\partial^2}{\partial z^2} L_{\tau_1}(t, z)[w, v] dt = 0. \tag{6.28}$$

Substituting (6.18) and (6.28) in (6.2) we find that $\lim_{z \rightarrow z_\mu} \zeta''(z)$ exists and (3.34) holds true. \square

References

- [1] F. Alizadeh, J.A. Haeberly, and M. Overton. Primal-dual interior-point methods for semidefinite programming: convergence rates, stability and numerical results. *SIAM J. Optim.* 8:746–768, 1998.
- [2] F. Alizadeh and S. Schmieta. Optimization with Semidefinite, Quadratic and Linear Constraints. RUTCOR, RRR Report 23-97, November 1997, 640 Bartholomew Rd, Piscataway NJ 08854.
- [3] J. Borwein and A. Lewis. *Convex Analysis and Nonlinear Optimization*. CMS Books in Mathematics, Springer, New York, 2000.
- [4] S. Burer and R.D.C. Monteiro. A General Framework for Establishing Polynomial Convergence of Long-Step Methods for Semidefinite Programming. Working paper, School of ISyE, Georgia Tech, USA, August 1999.
- [5] T. Bröcker and K. Jänich. *Einführung in die Differentialtopologie*. Heidelberger Taschenbücher Band 143. Springer Verlag, Berlin Heidelberg New York, 1990.
- [6] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Oxford Mathematical Monograph, Oxford University Press, New York, USA, 1994.
- [7] L. Faybusovich. Jordan algebras, Symmetric cones and Interior-point methods. *Technical Report*, Department of Mathematics, University of Notre Dame, IN, October 1995.
- [8] L. Faybusovich. Linear systems in Jordan algebras and primal-dual interior-point algorithms. *Journal of Computational and Applied Mathematics*, 86:149–175, 1997.
- [9] O. Güler. Private communication to Y. E. Nesterov and M. J. Todd, April 1994.
- [10] O. Güler. Barrier functions in interior-point methods. *Mathematics of Operations Research*, 21:860–885, 1996.
- [11] O. Güler, Interior-point methods on symmetric and homogeneous cones, Presentation, *17th International Symposium on Mathematical Programming*, Atlanta, Georgia, 7–11 August, 2000.
- [12] F. Hausdorff. Die symbolische Exponentialformel in der Gruppentheorie. *Berichte der Sächsischen Akademie der Wissenschaften (Math. Phys. Klasse)*, 58:19-48, 1906.
- [13] R.A. Hauser. Self-Scaled Barrier Functions: Decomposition and Classification. Numerical Analysis Report DAMTP 1999/NA13, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW.
- [14] R.A. Hauser. Self-scaled Barriers for Semidefinite Programming. Numerical Analysis Report DAMTP 2000/NA02, Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge, England CB3 9EW.
- [15] R.A. Hauser. *On Search Directions for Self-Scaled Conic Programming*. Ph.D. thesis. School of Operations Research and Industrial Engineering, Cornell University. January 2000.

- [16] R.A. Hauser and O. Güler. Self-scaled barrier functions on symmetric cones and their classification. *Found. Comput. Math.*, 2:121–143, 2002.
- [17] R.A. Hauser and Y. Lim. Self-scaled barriers for irreducible symmetric cones. *SIAM J. Optim.*, 12:715–723, 2002.
- [18] J. Hilgert and K. H. Neeb. *Lie Gruppen und Lie Algebren*. Vieweg Verlag, Braunschweig/Wiesbaden, Germany, 1991.
- [19] A. Iserles and S. P. Nørsett. On the solution of linear differential equations in Lie groups. In C.J. Budd and A. Iserles (eds.), *Geometric Integration: Numerical Solution of Differential Equations on Manifolds*, vol.357 of *Philosophical Transactions of the Royal Society A*, pages 983–1020, London Mathematical Society, 1999.
- [20] B. Jansen, C. Roos, T. Terlaky and J. P. Vial. Long-step primal-dual target-following algorithms for linear programming. *Zeitschrift für Operations Research - Mathematical Methods of Operations Research*, 44:11–30 (1996).
- [21] M. Koecher. *The Minnesota Notes on Jordan Algebras and their Applications*. Lecture Notes in Mathematics 1710, Springer-Verlag, New York, 1999.
- [22] M. Kojima, S. Mizuno and A. Yoshise. An $O(\sqrt{n}L)$ iteration potential reduction algorithm for linear complementarity problems. *Mathematical Programming*, 50:331–342, 1991.
- [23] M. Kojima, N. Megiddo, T. Noma and A. Yoshise. *A Unified Approach to Interior Point Algorithms for Linear Complementarity Problems*. Springer Verlag, Berlin, 1991.
- [24] S. Kruk, M. Muramatsu, F. Rendl, R.J. Vanderbei, and H. Wolkowicz. The Gauss-Newton Direction in Semidefinite Programming. *Optim. Methods Softw.* 15:1–28, 2001.
- [25] W. Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure & Appl. Maths*, 7:649–673, 1954.
- [26] R.D.C. Monteiro and P. Zanjacomo. General Interior-Point Maps and Existence of Weighted Paths for Nonlinear Semidefinite Complementarity Problems. Working paper, School of ISyE, Georgia Tech, USA, April 1998.
- [27] R.D.C. Monteiro and J.-S. Pang. On two interior point mappings for nonlinear semidefinite complementarity problems. *Mathematics of Operations Research*, 23:39–60, 1998.
- [28] Yu. E. Nesterov and A. S. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms*. SIAM Publications. SIAM, Philadelphia, USA, 1994.
- [29] Yu. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22:1–42, 1997.
- [30] Yu. E. Nesterov and M. J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM Journal on Optimization*, 8:324–364, 1998.
- [31] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization*. SIAM Publications, Philadelphia, 2001.

- [32] C. Roos, T. Terlaky and J.-Ph. Vial. *Theory and Algorithms for Linear Optimization*. Wiley-Interscience, John Wiley & Sons Ltd, West Sussex, England, 1997.
- [33] S. H. Schmieta. Complete classification of self-scaled barrier functions. *Technical Report*, Department of IEOR, Columbia University, New York, NY 10027, July 2000.
- [34] J. F. Sturm and S. Zhang. On weighted centers for semidefinite programming. Technical Report 9636/A, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1996.
- [35] J. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Interior point methods. Optim. Methods Softw.*, 11/12:625–653, 1999.
- [36] M.J. Todd, K.C. Toh and R.H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *SIAM Journal on Optimization*, 8:769–796, (1998).
- [37] K.C. Toh, M.J. Todd, and R.H. Tütüncü. SDPT3 – a MATLAB software package for semidefinite programming, version 1.3. *Interior point methods. Optim. Methods Softw.*, 11/12 (1999), no. 1-4, 545–581.
- [38] M.J. Todd. A study of search directions in primal-dual interior-point methods for semidefinite programming. *Interior point methods. Optim. Methods Softw.*, 11/12:1–46, 1999.
- [39] L. Tunçel. Primal-dual symmetry and scale invariance of interior-point algorithms for convex optimization. *Mathematics of Operations Research*, 23, 1998.
- [40] L. Tunçel. Generalization of primal-dual interior-point methods to convex optimization problems in conic form. *Found. Comput. Math.*, 1:229–254, 2001.