FULL NESTEROV-TODD STEP INTERIOR-POINT METHODS FOR SYMMETRIC OPTIMIZATION
G. GU *, M. ZANGIABADI †, AND C. ROOS*

Abstract. Some Jordan algebras were proved more than a decade ago to be an indispensable tool in the unified study of interior-point methods. By using it, we generalize the infeasible interior-point method for linear optimization of Roos [SIAM J. Optim., 16(4):1110–1136 (electronic), 2006] to symmetric optimization. This unifies the analysis for linear, second-order cone and semidefinite optimizations.

Key words. symmetric optimization, interior-point method, Nesterov-Todd step

AMS subject classifications. 17C99, 90C25, 90C51

1. Introduction. Jordan algebras were initially created in Quantum Mechanics, and they turned out to have a very large spectrum of applications. Indeed, some Jordan algebras were proved more than a decade ago to be an indispensable tool in the unified study of IPMs (Interior-Point Methods) for symmetric cones [4].

By using Jordan algebras, we generalize the IIPM (Infeasible IPM) for LO (Linear Optimization) of Roos [16] to symmetric optimization. This unifies the analysis for LO (linear optimization), SOCO (second-order cone optimization) and SDO (semidefinite optimization) problems. Since its analysis requires a quadratic convergence result for the feasible case, we first present a primal-dual (feasible) IPM with full NT-steps (Nesterov-Todd). Then we extend Roos’s IIPM for LO to symmetric cones. We prove that the complexity bound of our IIPM is $O(r \log(r/\varepsilon))$, where $r$ is the rank of the associated Euclidean Jordan algebra.

The paper is organized as follows. In Section 2 we briefly recall some properties of symmetric cones and its associated Euclidean Jordan algebras, focussing on what is needed in the rest of the paper. Then, in Section 3 we present a feasible IPM for symmetric cones, and in Section 4 our IIPM. Section 5 contains some conclusions and topics for further research.

2. Analysis on symmetric cones. This section offers an introduction to the theory of Jordan algebras, Euclidean Jordan algebras and symmetric cones. The presentation is mainly based on Faraut and Korányi [3]. After presenting the basic properties of Jordan algebras and Euclidean Jordan algebras, we state the following fundamental result due to M. Koecher and E.B. Vinberg [8]: the cone of squares in a Euclidean Jordan algebra is a symmetric cone, and every symmetric cone is obtained in this way. Based on these, we derive some more properties of Euclidean Jordan algebras and its associated symmetric cones as needed for the optimization techniques we present later. We restrict ourselves to finite-dimensional algebras over the real field $\mathbb{R}$. For omitted proofs, we refer to the given references and also to [1, 3, 7, 8, 9, 21].

2.1. Jordan algebras. In this subsection, we introduce Jordan algebras as well as some of their basic properties.

---

*Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, the Netherlands. e-mail: [g.gu, c.roos]@tudelft.nl
†Department of Mathematical Science, Shahrekord University, P.O. Box 115, Shahrekord, Iran. e-mail: m.zangiabadi@tudelft.nl
2.1.1. Definition. Definition 2.1 (Bilinear map). Let $J$ be a finite-dimensional vector space over $R$. A map $\circ : J \times J \rightarrow J$ is called bilinear if for all $x, y, z \in J$ and $\alpha, \beta \in R$:

(i) $(\alpha x + \beta y) \circ z = \alpha (x \circ z) + \beta (y \circ z)$;
(ii) $x \circ (\alpha y + \beta z) = \alpha (x \circ y) + \beta (x \circ z)$.

Definition 2.2 (R-algebra). A finite-dimensional vector space $J$ over $R$ is called an algebra over $R$ if a bilinear map from $J \times J$ into $J$ is defined.

Definition 2.3 (Jordan algebra). Let $J$ be a finite-dimensional R-algebra along with a bilinear map $\circ : J \times J \rightarrow J$. Then $(J, \circ)$ is called a Jordan algebra if for all $x, y \in J$:

Commutativity: $x \circ y = y \circ x$;

Jordan’s Axiom: $x \circ (x^2 \circ y) = x^2 \circ (y \circ x)$, where $x^2 = x \circ x$.

In the sequel, we always assume $(J, \circ)$ is a Jordan algebra, and simply denote as $J$. For an element $x \in J$, let $L(x) : J \mapsto J$ be the linear map defined by

$$L(x)y := x \circ y, \quad \text{for all } y \in J.$$ 

Then Jordan’s Axiom in Definition 2.3 means that the operators $L(x)$ and $L(x^2)$ commute.

We define $x^n, n \geq 2$ recursively by $x^n = x \circ x^{n-1}$. An algebra is said to be power associative if, for any $x$ in the algebra, $x^m \circ x^n = x^{m+n}$. This means that the subalgebra generated by $x$ is associative. Jordan algebras are not necessarily associative, but they are power associative (cf. Proposition II.1.2 of [3]).

2.1.2. Characteristic polynomial. Let $J$ be a Jordan algebra over $R$. An element $e \in J$ is said to be an identity element if

$$e \circ x = x \circ e = x, \quad \text{for all } x \in J.$$ 

Note that the identity element $e$ is unique. As if $e_1$ and $e_2$ are identity elements of $J$, then $e_1 = e_1 \circ e_2 = e_2$.

From now on, we always assume the existence of the identity element $e$. Let $R[X]$ denote the algebra over $R$ of polynomials in one variable with coefficients in $R$. For an element $x$ in $J$ we define

$$R[x] := \{ p(x) : p \in R[X] \}.$$

Since $J$ is a finite-dimensional vector space, for each $x \in J$, there exists a positive integer $k$ (bounded by the dimension of $J$) such that $e, x, x^2, \ldots, x^k$ are linearly dependent. This implies the existence of a nonzero polynomial $p \in R[X]$, such that $p(x) = 0$. If, in addition, this polynomial is monic (i.e., with the leading coefficient equal to 1) and of minimal degree, we call it the minimal polynomial of $x$.

The minimal polynomial of an element $x \in J$ is unique. As if $p_1$ and $p_2$ are two distinct minimal polynomials of $x$, their difference $p_1 - p_2$ vanishes in $x$ as well. Since $p_1$ and $p_2$ are monic and of the same degree, the degree of $p_1 - p_2$ is smaller than that of $p_1$ (or $p_2$). This contradicts the minimality of the degree of $p_1$ (or $p_2$).

We define the degree of an element $x \in J$, denoted as $\deg(x)$, as the degree of the minimal polynomial of $x$. Obviously, this number is bounded by the dimension of the vector space $J$. Moreover we define the rank of $J$ as

$$r := \max \{ \deg(x) : x \in J \},$$
which again bound by the dimension of \( J \). An element \( x \in J \) is called regular if \( \deg(x) = r \).

**Proposition 2.4** ([3, Proposition II.2.1]). The set of regular elements is open and dense in \( J \). There exist polynomials \( a_1, a_2, \ldots, a_r \) on \( J \) such that the minimal polynomial of every regular element \( x \) is given by

\[
f(\lambda; x) = \lambda^r - a_1(x)\lambda^{r-1} + a_2(x)\lambda^{r-2} + \cdots + (-1)^ra_r(x).
\]

The polynomials \( a_1, \ldots, a_r \) are unique and \( a_i \) is homogeneous of degree \( i \).

The polynomial \( f(\lambda; x) \) is called the characteristic polynomial of the regular element \( x \). Since the regular elements are dense in \( J \), by continuity we may extend the polynomials \( a_i(x) \) to all elements of \( J \) and consequently the characteristic polynomial. Note that the characteristic polynomial is a polynomial of degree \( r \) in \( \lambda \), where \( r \) is the rank of \( J \). Moreover, the minimal polynomial coincides with the characteristic polynomial for regular elements, but it divides the characteristic polynomial of non-regular elements.

The coefficient \( a_1(x) \) is called the trace of \( x \), denoted as \( \text{tr}(x) \). And the coefficient \( a_r(x) \) is called the determinant of \( x \), denoted as \( \text{det}(x) \).

**Proposition 2.5** ([3, Proposition II.4.3]). The symmetric bilinear form \( \text{tr}(x \circ y) \) is associative, i.e.,

\[
\text{tr}((x \circ y) \circ z) = \text{tr}(x \circ (y \circ z)), \quad \text{for all } x, y, z \in J.
\]

An element \( x \) is said to be invertible if there exists an element \( y \) in \( R[x] \) such that \( x \circ y = e \). Since \( R[x] \) is associative, \( y \) is unique. It is called the inverse of \( x \) and denoted by \( x^{-1} \).

**Proposition 2.6** ([3, Proposition II.2.3]). If \( L(x) \) is invertible, then \( x \) is invertible and \( x^{-1} = L(x)^{-1}e \).

**2.1.3. Quadratic representation.** Let \( J \) be a finite-dimensional Jordan algebra over \( R \) with the identity element \( e \). For \( x \in J \) we define

\[
P(x) = 2L(x)^2 - L(x^2),
\]

where \( L(x)^2 = L(x)L(x) \). The map \( P(\cdot) \) is called the quadratic representation of \( J \).

**Proposition 2.7** ([3, Proposition II.3.1]). An element \( x \in J \) is invertible if and only if \( P(x) \) is invertible. In this case,

\[
P(x)x^{-1} = x,
\]

\[
P(x)^{-1} = P(x^{-1}).
\]

**Proposition 2.8** ([3, Proposition II.3.3]). One has:

(i) The differential of the map \( x \mapsto x^{-1} \) is \( -P(x)^{-1} \).

(ii) If \( x \) and \( y \) are invertible, then \( P(x)y \) is invertible and

\[
(P(x)y)^{-1} = P(x^{-1})y^{-1}.
\]

(iii) For any elements \( x \) and \( y \):

\[
P(P(y)x) = P(y)P(x)P(y).
\]

In particular, the equation in (iii) of the above proposition is known as the fundamental formula (cf. [8, Chapter IV.1]).
2.2. Euclidean Jordan algebras and symmetric cones. Euclidean Jordan algebras (or formally real Jordan algebras) form a subclass of Jordan algebras. In this subsection, we recall their basic properties as well as their relation with symmetric cones.

2.2.1. Euclidean Jordan algebras. We consider finite-dimensional Jordan algebra over \( \mathbb{R} \) and assume the existence of the identity element \( e \). A Jordan algebra \( J \) over \( \mathbb{R} \) is said to be Euclidean if there exists a positive definite symmetric bilinear form on \( J \) which is associative; in other words, there exists an inner product denoted by \( \langle \cdot, \cdot \rangle \), such that

\[
\langle x \circ y, z \rangle = \langle x, y \circ z \rangle,
\]

for all \( x, y, z \in J \).

In the sequel, unless state otherwise, we always assume that \( J \) is a Euclidean Jordan algebra with the identity element \( e \). An element \( c \in J \) is said to be an idempotent if \( c^2 = c \). Two idempotents \( c_1 \) and \( c_2 \) are said to be orthogonal if \( c_1 \circ c_2 = 0 \). Since

\[
\langle c_1, c_2 \rangle = \langle c_1^2, c_2 \rangle = \langle c_1, c_1 \circ c_2 \rangle,
\]

orthogonal idempotents are orthogonal with respect to the inner product. Moreover, an idempotent is primitive if it is non-zero and cannot be written as the sum of two (necessarily orthogonal) non-zero idempotents. We say that \( \{c_1, \ldots, c_r\} \) is a complete system of orthogonal primitive idempotents, or Jordan frame, if each \( c_i \) is a primitive idempotent and

\[
c_i \circ c_j = 0, \quad i \neq j,
\]

\[
\sum_{i=1}^{r} c_i = e.
\]

Note that Jordan frames always contain \( r \) primitive idempotents, where \( r \) is the rank of \( J \).

Following, we recall a “spectral theorem”, which, in the case of the algebra of real symmetric matrices, specializes to the usual spectral theorem.

**Theorem 2.9 (Spectral theorem, cf. [3, Theorem III.1.2])**. Suppose \( J \) has rank \( r \). Then for \( x \in J \) there exist a Jordan frame \( c_1, c_2, \ldots, c_r \) and real numbers \( \lambda_1, \lambda_2, \ldots, \lambda_r \) such that

\[
x = \sum_{i=1}^{r} \lambda_i c_i.
\]

The numbers \( \lambda_i \) (with their multiplicities) are uniquely determined by \( x \). Furthermore,

\[
\text{tr}(x) = \sum_{i} \lambda_i, \quad \det(x) = \prod_{i=1}^{r} \lambda_i,
\]

More generally,

\[
a_k(x) = \sum_{1 \leq i_1 < \ldots < i_k \leq r} \lambda_{i_1} \ldots \lambda_{i_k},
\]

where \( a_k(1 \leq k \leq r) \) is the polynomial defined in Proposition 2.4.
In fact, the above \( \lambda_1, \ldots, \lambda_r \), are exactly the roots of the characteristic polynomial \( f(\lambda; x) \). Normally, to express their dependence on \( x \), they are denoted as \( \lambda_1(x), \ldots, \lambda_r(x) \), or simply as a vector \( \lambda(x) \in \mathbb{R}^r \). We call them the eigenvalues (or spectral values) of \( x \). Moreover, we denote the largest eigenvalue of \( x \) as \( \lambda_{\text{max}}(x) \), and analogously the smallest as \( \lambda_{\text{min}}(x) \). Note that since \( e \) has eigenvalue 1, with multiplicity \( r \), it follows that \( \text{tr}(e) = r \) and \( \det(e) = 1 \).

Now it is possible to extend the definition of any real valued, continuous univariate function \( f(\cdot) \) to elements of a Euclidean Jordan algebra, using eigenvalues:

\[
f(x) := f(\lambda_1)c_1 + \cdots + f(\lambda_k)c_k.
\]

 Particularly we have:

**Inverse:** \( x^{-1} := \lambda_1^{-1}c_1 + \cdots + \lambda_r^{-1}c_r \), whenever all \( \lambda_i \neq 0 \) and undefined otherwise;

**Square root:** \( x^{1/2} := \lambda_1^{1/2}c_1 + \cdots + \lambda_r^{1/2}c_r \), whenever all \( \lambda_i \geq 0 \) and undefined otherwise;

**Square:** \( x^2 := \lambda_1^2c_1 + \cdots + \lambda_r^2c_r \).

For each \( i = 1, \ldots, r \), we have \( c_i \in \mathbb{R}[x] \) (cf. [3, Theorem III.1.1]). Moreover \((\lambda_1^{-1}c_1 + \cdots + \lambda_r^{-1}c_r)\circ x = e \). Hence, the expression of \( x^{-1} \) here, using a spectral decomposition of \( x \), coincides with the algebraic definition of inverse in Subsection 2.1.2. In the same way it follows that \( x^{1/2} \circ x^{1/2} = x \) and \( x^2 = x \circ x \).

**Theorem 2.10** (cf. [3, Theorem III.1.5]). Let \( \mathcal{J} \) be a Jordan algebra over \( \mathbb{R} \) with the identity element \( e \). The following properties are equivalent.

(i) \( \mathcal{J} \) is a Euclidean Jordan algebra.

(ii) The symmetric bilinear form \( \text{tr}(x \circ y) \) is positive definite.

The above proposition implies that if \( \mathcal{J} \) is a Euclidean Jordan algebra, then \( \text{tr}(x \circ y) \) is an inner product. In the sequel, \( \langle x, y \rangle \) will always denote this inner product, and we refer to it as the trace inner product.

The norm induced by the above inner product is named as the Frobenius norm, which is given by

\[
\|x\|_F := \sqrt{\langle x, x \rangle}.
\]

In fact, the above norm can also be obtained by eigenvalues. By Theorem 2.9, \( x \in \mathcal{J} \) has a spectral decomposition \( x = \sum_{i=1}^r \lambda_i(x)c_i \), and \( x^2 = \sum_{i=1}^r \lambda_i^2(x)c_i \). Hence

\[
\|x\|_F = \sqrt{\langle x, x \rangle} = \sqrt{\text{tr}(x^2)} = \sqrt{\sum_{i=1}^r \lambda_i^2(x)}.
\]

Moreover the operator \( L(x) \) and \( P(x) \) are self-adjoint with respect to this inner product (cf. Proposition 2.5 and the definition of the quadratic representation, respectively).

**2.2.2. Symmetric cones.** In this subsection, we recall some definitions concerning symmetric cones.

**Definition 2.11** (Convex set). A set \( \mathcal{K} \) is convex if for any \( x, y \in \mathcal{K} \) and any \( \alpha \) with \( 0 \leq \alpha \leq 1 \), we have \( \alpha x + (1-\alpha)y \in \mathcal{K} \).

**Definition 2.12** (Cone). A set \( \mathcal{K} \) is called a cone if for every \( x \in \mathcal{K} \) and \( \alpha \geq 0 \), we have \( \alpha x \in \mathcal{K} \).

Therefore, a set \( \mathcal{K} \) is called a convex cone if it is convex and a cone, which means that for any \( x, y \in \mathcal{K} \) and \( \alpha, \beta \geq 0 \), we have \( \alpha x + \beta y \in \mathcal{K} \).
Definition 2.13 (Dual cone). Let \( \mathcal{K} \subseteq \mathcal{J} \) be a cone. The set
\[
\mathcal{K}^* := \{ y \in \mathcal{J} : \langle x, y \rangle \geq 0, \text{ for all } x \in \mathcal{K} \}
\]
is called the dual cone of \( \mathcal{K} \).

As the name suggests, \( \mathcal{K}^* \) is a cone, and is always convex, even when the original cone is not. If cone \( \mathcal{K} \) and its dual \( \mathcal{K}^* \) coincide, we say that \( \mathcal{K} \) is self-dual. In particular, this implies that \( \mathcal{K} \) has a nonempty interior and does not contain any straight line (i.e., it is pointed).

Definition 2.14 (Homogeneity). The convex cone \( \mathcal{K} \) is said to be homogeneous if for every pair \( x, y \in \text{int} \mathcal{K} \), there exists an invertible linear operator \( g \) for which \( g \mathcal{K} = \mathcal{K} \) and \( gx = y \).

In fact, the above linear operator \( g \) is an automorphism of the cone \( \mathcal{K} \), i.e., \( g \in \text{Aut}(\mathcal{K}) \), which is defined later in subsection 2.2.5.

Definition 2.15 (Symmetric cone). The convex cone \( \mathcal{K} \) is said to be symmetric if it is self-dual and homogeneous.

In [3] the self-dual cone and hence the symmetric cone are defined to be open. Here, we follow the definition used by the optimization community (cf. e.g. [19, Definition 1]). This minor difference will not affect too much.

2.2.3. One-to-one correspondence. In this subsection, we recall a fundamental result which establishes the one-to-one correspondence between (cones of squares of) Euclidean Jordan algebras and symmetric cones.

Let \( \mathcal{J} \) be a Euclidean Jordan algebra. We define the cone of squares \( \mathcal{K}(\mathcal{J}) \) of \( \mathcal{J} \) as
\[
\mathcal{K}(\mathcal{J}) := \{ x^2 : x \in \mathcal{J} \}.
\]

The following theorem brings together some major properties of the cone of squares in a Euclidean Jordan algebra.

Theorem 2.16 (cf. [3, Theorem III.2.1, Proposition III.2.2]). Let \( \mathcal{J} \) be a Euclidean Jordan algebra, then \( \mathcal{K}(\mathcal{J}) \) is a symmetric cone, and is the set of elements \( x \) in \( \mathcal{J} \) for which \( L(x) \) is positive semidefinite. Furthermore, if \( x \) is invertible, then
\[
P(x) \text{int} \mathcal{K} = \text{int} \mathcal{K}.
\]

The above theorem indicates that the cone of squares in a Euclidean Jordan algebra is a symmetric cone. Conversely, given any symmetric cone in a Euclidean space, one may define a Euclidean Jordan algebra such that the given cone is its cone of squares (cf. [3, Theorem III.3.1]). Therefore we have the following Jordan algebraic characterization of symmetric cones.

Theorem 2.17 (cf. [3, Section III.2–5]). A cone is symmetric if and only if it is the cone of squares of some Euclidean Jordan algebra.

Because of the above one-to-one correspondence, the notions of cone of squares in a Euclidean Jordan algebra and symmetric cone are equivalent. In the sequel, unless stated otherwise, we denote them the same, simply as \( \mathcal{K} \).

Proposition 2.18 (cf. [4, Theorem 5.13]). Given \( x \in \text{int} \mathcal{K} \), we have
\[
\langle x, s \rangle > 0, \quad \text{for all } s \in \mathcal{K} \setminus \{0\}.
\]
We associate with the (proper) cone $K$ the partial order defined by

$$x \succeq_K y \iff x - y \in K.$$ 

We also write $y \preceq_K x$ for $x \succeq_K y$. Similarly, we define an associated strict partial order by

$$x \succ_K y \iff x - y \in \text{int } K,$$

and write $y \prec_K x$ for $x \succ_K y$.

2.2.4. Simple Jordan algebras. Another consequence of the one-to-one correspondence between Euclidean Jordan algebras and symmetric cones is the unique decomposition of every symmetric cone into a direct product of irreducible ones. As usual, we start with some definitions.

Definition 2.19 (Ideal). Let $J$ be an $\mathbb{R}$-algebra. An ideal $I$ of $J$ is a vector subspace of $J$ such that for every $x \in I$ and every $y \in J$ the elements $x \circ y$ and $y \circ x$ belong to $I$.

Definition 2.20 (Simple algebra). An $\mathbb{R}$-algebra $J$ is simple if it contains only two ideals, namely $\{0\}$ and $J$ (trivial ideals).

Proposition 2.21 (cf. [3, Proposition III.4.4]). If $J$ is a Euclidean Jordan algebra, then it is, in a unique way, a direct sum of simple ideals.

A symmetric cone $K$ in a Euclidean space $J$ is said to be irreducible if there do not exist non-trivial subspaces $J_1$, $J_2$, and symmetric cones $K_1 \subset J_1$, $K_2 \subset J_2$, such that $J$ is the direct sum of $J_1$ and $J_2$, and $K$ the direct sum of $K_1$ and $K_2$.

Proposition 2.22 (cf. [3, Proposition III.4.5]). Any symmetric cone $K$ is, in a unique way, the direct product of irreducible symmetric cones.

The following theorem states that there are five kinds of simple Euclidean Jordan algebras and correspondingly five kinds of irreducible symmetric cones.

Theorem 2.23 (cf. [3, Chapter V]). Let $J$ be a simple Euclidean Jordan algebra. Then $J$ is isomorphic to one of the following algebras.

(i) The algebra in space $\mathbb{R}^{n+1}$ with Jordan multiplication defined as

$$x \circ y = (x^T y; x_0 y_0 + y_0 x_0),$$

where $x := (x_0; \bar{x})$ and $y := (y_0; \bar{y})$ with $x_0, y_0 \in \mathbb{R}$ and $\bar{x}, \bar{y} \in \mathbb{R}^n$.

(ii) The algebra of real symmetric matrices with Jordan multiplication defined as

$$X \circ Y = (XY + YX)/2.$$

(iii) The algebra of complex Hermitian matrices with Jordan multiplication defined as in (ii).

(iv) The algebra of quaternion Hermitian matrices, with Jordan multiplication defined as in (ii).

(v) The algebra of $3 \times 3$ octonion Hermitian matrices with Jordan multiplication defined as in (ii).

2.2.5. Automorphisms. Let $J$ be a Euclidean Jordan algebra and $K$ its cone of squares (or its associated symmetric cone). In this section, we recall definitions and some properties of automorphisms of $J$ as well as $K$.

We denote henceforth the set of all invertible linear maps from $J$ into itself by $\text{GL}(J)$.
Definition 2.24 (Automorphism of $\mathcal{J}$). A map $g \in \mathrm{GL}(\mathcal{J})$ is called an automorphism of $\mathcal{J}$ if for every $x$ and $y$ in $\mathcal{J}$, we have $g(x \circ y) = gx \circ gy$, or equivalently, $gL(x)g^{-1} = L(gx)$. The set of automorphisms of $\mathcal{J}$ is denoted as $\text{Aut}(\mathcal{J})$.

Definition 2.25 (Automorphism of $\mathcal{K}$). A map $g \in \mathrm{GL}(\mathcal{J})$ is called an automorphism of $\mathcal{K}$ if $g\mathcal{K} = \mathcal{K}$. The set of automorphisms of $\mathcal{K}$ is denoted as $\text{Aut}(\mathcal{K})$.

We say that a linear map is orthogonal if $g^* = g^{-1}$. The set of orthogonal automorphisms of $\mathcal{K}$ is denoted as $\text{OAut}(\mathcal{K})$. So we have

$$\text{OAut}(\mathcal{K}) = \{ g \in \text{Aut}(\mathcal{K}) : g^* = g^{-1} \}.$$ 

We would like to stress that $\text{Aut}(\mathcal{J}) \neq \text{Aut}(\mathcal{K})$. For example, an element $g$ of $\text{Aut}(\mathcal{K})$ may not satisfy $g(x \circ y) = gx \circ gy$. Since $\mathcal{K}$ is not closed under the Jordan product (which implies that $x \circ y$ maybe not in $\mathcal{K}$), $g(x \circ y)$ is not well defined.

Proposition 2.26 (cf. [3, Proposition II.4.2]). The trace and the determinant are invariant under $\text{Aut}(\mathcal{J})$.

The next theorem establishes a connection between the automorphism group of a Euclidean Jordan algebra and the orthogonal automorphism group of its associated symmetric cone.

Proposition 2.27 (cf. [21, Theorem 2.8.4]). We have

$$\text{Aut}(\mathcal{J}) = \text{OAut}(\mathcal{K}).$$

Proposition 2.28 (cf. [3, Proposition IV.2.5]). Let $\mathcal{J}$ be a simple Euclidean Jordan algebra. If $\{c_1, \ldots, c_r\}$ and $\{d_1, \ldots, d_r\}$ are two Jordan frames, then there exists an automorphism $g$ in $\text{Aut}(\mathcal{J})$ such that

$$gc_i = d_i, \quad \text{for all } 1 \leq i \leq r.$$ 

2.3. More algebraic properties. In this section we recall or derive some more properties of Euclidean Jordan algebras and their associated symmetric cones. These results play a key role in our analysis of optimization techniques for symmetric cones. Recall that we always assume $\mathcal{K}$ is a symmetric cone (or equivalently the cone of squares of some Euclidean Jordan algebra).

2.3.1. NT-scaling. When defining the search direction in our algorithms, we need a rescaling of the space in which the symmetric cone lives. In this subsection, we show the existence and uniqueness of a scaling point $w$ corresponding to any points $x, s \in \text{int} \mathcal{K}$, such that $P(w)$ takes $s$ into $x$. This was done for self-scaled cones in [13, 14]. Later, Faybusovich [5] derived it in the framework of Euclidean Jordan algebras.

In order to introduce the scaling, we need the following proposition.

Proposition 2.29. If $x \in \mathcal{K}$, then $x^{1/2}$ is well defined and $P(x^{1/2}) = P(x)^{1/2}$.

Proof. As $x$ is in the cone of squares, all its eigenvalues are nonnegative. Hence $x^{1/2}$ is well defined. By the definition of the quadratic representation, we have

$$P(x^{1/2})e = \left(2L(x^{1/2})^2 - L(x)\right)e = x.$$ 

Therefore, by the fundamental formula (cf. Proposition 2.8),

$$P(x) = P(P(x^{1/2})e) = P(x^{1/2})P(e)P(x^{1/2}) = P(x^{1/2})^2.$$
Moreover, since \( P(x^{1/2}) \) is self-adjoint, \( P(x)^{1/2} = P(x^{1/2}) \).

**Proposition 2.30** (NT-scaling, cf. [5, Lemma 3.2]). Given \( x, s \in \text{int} \mathcal{K} \), there exists a unique \( w \in \text{int} \mathcal{K} \) such that

\[
x = P(w)s.
\]

Moreover,

\[
w := P(x)^{1/2}(P(x)^{1/2} s)^{-1/2} = P(s)^{-1/2}(P(s)^{1/2} x)^{1/2}.
\]

We call \( w \) the scaling point of \( x \) and \( s \) (in this order).

**2.3.2. Similarity.** Recall that two matrices \( X \) and \( S \) are similar if they share the same set of eigenvalues; in this case, we write \( X \sim S \). Analogously, we say that two elements \( x \) and \( s \) in \( \mathcal{F} \) are similar, denoted as \( x \sim s \), if and only if \( x \) and \( s \) share the same set of eigenvalues. For more details we refer to [19, 20, 21].

**Proposition 2.31** ([19, Proposition 19]). Two elements \( x \) and \( s \) of a Euclidean Jordan algebra are similar if and only if \( L(x) \) and \( L(s) \) are similar.

**Proposition 2.32** ([19, Corollary 20]). Let \( x \) and \( s \) be two elements in \( \text{int} \mathcal{K} \). Then \( x \) and \( s \) are similar if and only if \( P(x) \) and \( P(s) \) are similar.

**Proposition 2.33** ([19, Proposition 21]). Let \( x \) and \( s \) be two elements in \( \text{int} \mathcal{K} \), then \( P(x)^{1/2}s \) and \( P(s)^{1/2}x \) are similar.

Following are two important generalizations. Because of their importance we include the proofs.

**Lemma 2.34** ([19, Proposition 21]). Let \( x, s, u \in \text{int} \mathcal{K} \). Defining \( \tilde{x} = P(u)x \) and \( \tilde{s} = P(u^{-1})s \), then one has

\[
P(\tilde{x}^{1/2})\tilde{s} \sim P(x^{1/2})s.
\]

**Proof.** By the fundamental formula,

\[
P(P(\tilde{x}^{1/2})\tilde{s}) = P(\tilde{x}^{1/2})P(\tilde{s})P(\tilde{x}^{1/2}) \sim P(\tilde{x})P(\tilde{s}).
\]

Similarly \( P(P(x^{1/2})s) \sim P(x)P(s) \). Since \( P(\tilde{x}^{1/2})\tilde{s}, P(x^{1/2})s \in \text{int} \mathcal{K} \) (cf. Theorem 2.16), by Proposition 2.32, it suffices to show that \( P(\tilde{x})P(\tilde{s}) \sim P(x)P(s) \). Using the fundamental formula again, we obtain

\[
P(\tilde{x})P(\tilde{s}) = P(P(u)x)P(P(u^{-1})s)
= P(u)P(x)P(u^{-1})P(s)P(u^{-1})
\sim P(x)P(s).
\]

Hence the proof is complete.

**Lemma 2.35** (cf. [21, Proposition 3.2.4]). Let \( x, s \in \text{int} \mathcal{K} \), and \( w \) the scaling point of \( x \) and \( s \), then

\[
(P(x)^{1/2}s)^{1/2} \sim P(w^{1/2})s.
\]

**Proof.** From Theorem 2.16 and Proposition 2.29, it readily follows that

\[
(P(x)^{1/2}s)^{1/2} \in \text{int} \mathcal{K}, \quad P(w^{1/2})s \in \text{int} \mathcal{K}.
\]
Then Proposition 2.32 implies that the statement is equivalent to
\[ P((P(x^{1/2})s)^{1/2}) \sim P(P(w^{1/2})s). \]
Since we have, by the fundamental formula,
\[ P(P(w^{1/2})s) = P(w)^{1/2}P(s)P(w)^{1/2} \sim P(w)P(s), \]
the statement is also equivalent to
\[ P((P(x^{1/2})s)^{1/2}) \sim P(w)P(s). \]

From Proposition 2.30, \( w = P(x^{1/2})(P(x^{1/2})s)^{-1/2} \). Then by the fundamental formula and Proposition 2.29
\[ P(w) = P(P(x^{1/2})(P(x^{1/2})s)^{-1/2}) = P(x^{1/2})P(P(x^{1/2})s)^{-1/2}P(x^{1/2}). \]
Hence, by substitution and using the fundamental formula again, we derive
\[
P(w)P(s) = P(x^{1/2})P(P(x^{1/2})s)^{-1/2}P(x^{1/2})P(s)
\sim P((P(x^{1/2})s)^{-1/2}P(x^{1/2})P(s))P(x^{1/2})
= P(P(x^{1/2})s)^{1/2}.
\]
The proposition follows. \( \square \)

2.3.3. Inequalities. To analyze our algorithms, we need some more inequalities, which are presented in the sequel.

**Lemma 2.36 (Cauchy-Schwarz inequality).** Let \( x, s \in J \), then
\[ \langle x, s \rangle \leq \|x\|_F \|s\|_F. \]

**Proof.** As the inequality is trivially true in the case \( s = 0 \), we may assume \( \langle s, s \rangle \) is nonzero. Let \( \alpha \in \mathbb{R} \), then
\[
0 \leq \|x - \alpha s\|_F^2 = \langle x - \alpha s, x - \alpha s \rangle = \langle x, x \rangle - 2\alpha \langle x, s \rangle + \alpha^2 \langle s, s \rangle.
\]
The above expression is valid for any \( \alpha \), which implies
\[ 4\langle x, s \rangle^2 - 4\langle s, s \rangle \langle x, x \rangle \leq 0, \]
and from this the lemma follows. \( \square \)

**Lemma 2.37.** Let \( x, s \in J \), then
\[ \|x \circ s\|_F \leq \frac{1}{2}\|x^2 + s^2\|_F. \]

**Proof.** Since
\[
x^2 + s^2 + 2x \circ s = (x + s)^2 \in K,
\]
\[
x^2 + s^2 - 2x \circ s = (x - s)^2 \in K,
\]
then we have
\[\langle x^2 + s^2 + 2x \circ s, x^2 + s^2 - 2x \circ s \rangle \geq 0,\]
which is equivalent to
\[\langle x^2 + s^2, x^2 + s^2 \rangle - 4\langle x \circ s, x \circ s \rangle \geq 0,\]
or equivalent to
\[\|x^2 + s^2\|_F^2 - 4\|x \circ s\|_F^2 \geq 0.\]
This implies the lemma.

**Lemma 2.38.** Let \(x \in K\), then
\[
\text{tr}(x^2) \leq \text{tr}(x^2).
\]

**Proof.** Since \(x \in K\), we have \(\lambda(x) \geq 0\). Therefore
\[
\text{tr}(x^2) = \sum_{i=1}^{r} \lambda_i(x)^2 \leq \left( \sum_{i=1}^{r} \lambda_i(x) \right)^2 = \text{tr}(x^2).
\]
This proves the lemma.

**Lemma 2.39.** Let \(x \in J\), then
\[
\|x^2\|_F \leq \|x\|^2.
\]

**Proof.** Use the definition of Frobenius norm, we have
\[
\|x^2\|_F^2 = \text{tr}((x^2)^2) \leq \text{tr}(x^2)^2 = (\|x\|^2)^2,
\]
where the inequality follows from Lemma 2.38. As both \(\|x^2\|_F\) and \(\|x\|_F\) are nonnegative, the lemma follows.

**Lemma 2.40.** Let \(J\) be a Euclidean Jordan algebra, and \(x, s \in J\) with \(\langle x, s \rangle = 0\). Then one has
\[
(i) \quad -\frac{1}{4} \|x + s\|^2_F \leq \lambda_{\text{max}}((x + s)^2 - (x - s)^2) \leq \frac{1}{4} \|x + s\|^2_F e;
(ii) \quad \|x \circ s\|_F \leq \frac{1}{\sqrt{2}} \|x + s\|^2_F.
\]

**Proof.** We write
\[
x \circ s = \frac{1}{4} ((x + s)^2 - (x - s)^2).
\]
Since \((x + s)^2 \in K\), we have
\[
x \circ s + \frac{1}{4} (x - s)^2 \in K.
\]
Using
\[
(x - s)^2 \leq \lambda_{\text{max}}((x - s)^2) e \leq \|x - s\|^2_F e,
\]
it follows that
\[ x \circ s + \frac{1}{4} \| x - s \|_F^2 e \in \mathcal{K}, \]
which means that \(-\frac{1}{4} \| x - s \|_F^2 e \preceq_{\mathcal{K}} x \circ s\). In the similar way one derives that \(x \circ s \preceq_{\mathcal{K}} \frac{1}{4} \| x + s \|_F^2 e\). Since \((x, s) = 0\), \(\| x - s \|_F = \| x + s \|_F\), part (i) of the lemma follows.

For the proof of part (ii), we derive as follows,
\[
\| x \circ s \|_F^2 = \left\| \frac{1}{4} ((x + s)^2 - (x - s)^2) \right\|_F^2 = \frac{1}{16} \text{tr} \left( ((x + s)^2 - (x - s)^2)^2 \right)
= \frac{1}{16} \left[ \text{tr} \left( (x + s)^4 \right) + \text{tr} \left( (x - s)^4 \right) - 2 \text{tr} \left( (x + s)^2 \circ (x - s)^2 \right) \right].
\]
Since \((x + s)^2\) and \((x - s)^2\) belong to \(\mathcal{K}\), the trace of their product is nonnegative. Thus we obtain
\[
\| x \circ s \|_F^2 \leq \frac{1}{16} \left[ \| (x + s)^4 \|_F^2 + \| (x - s)^4 \|_F^2 \right].
\]
Using Lemma 2.39 and \(\| x + s \|_F = \| x - s \|_F\) again, we get
\[
\| x \circ s \|_F^2 \leq \frac{1}{16} \left[ \| x + s \|_F^4 + \| x - s \|_F^4 \right] = \frac{1}{8} \| x + s \|_F^4.
\]
This implies part (ii) of the lemma. Hence the proof of the lemma is complete. \(\square\)

**Lemma 2.41.** Let \(x \in \mathcal{J}\) and \(s \in \mathcal{K}\). Then
\[ \lambda_{\min}(x) \text{tr}(s) \leq \text{tr}(x \circ s) \leq \lambda_{\max}(x) \text{tr}(s). \]

**Proof.** For any \(x \in \mathcal{J}\) we have \(\lambda_{\max}(x)e - x \in \mathcal{K}\). Since also \(s \in \mathcal{K}\), it follows that
\[ \text{tr}(\lambda_{\max}(x)e - x \circ s) \geq 0. \]
Hence the second inequality in the lemma follows by writing
\[ \text{tr}(x \circ s) \leq \text{tr}(\lambda_{\max}(x)e \circ s) = \lambda_{\max}(x) \text{tr}(s). \]
The proof of the first inequality goes in the similar way. \(\square\)

**Lemma 2.42.** If \(x \circ s \in \text{int} \mathcal{K}\), then \(\text{det}(x) \neq 0\).

**Proof.** By Theorem 2.9, the element \(x\) can be write as
\[ x = \sum_{i=1}^{r} \lambda_i(x)c_i, \]
where \(\{c_1, \ldots, c_r\}\) is a Jordan frame. Suppose \(\det(x) = 0\), then there must be an integer \(k\) with \(1 \leq k \leq r\), such that \(\lambda_k(x) = 0\). Since \(x \circ s \in \text{int} \mathcal{K}\) and \(c_k \in \mathcal{K} \setminus \{0\}\), by Proposition 2.18,
\[ 0 < \langle x \circ s, c_k \rangle = \langle s, x \circ c_k \rangle = 0. \]
This contradiction completes the proof. \(\square\)
Lemma 2.43 (cf. [15, Lemma 2.9]). Given $x \in \text{int} \mathcal{K}$, we have
\[ \|x - x^{-1}\|_F \leq \frac{\|x^2 - e\|_F}{\lambda_{\text{min}}(x)}. \]

Lemma 2.44 (cf. [19, Lemma 30]). Let $x, s \in \text{int} \mathcal{K}$, then the following relations hold.
(i) $\|P(x)^{1/2}s - e\|_F \leq \|x \circ s - e\|_F$;
(ii) $\lambda_{\text{min}}(P(x)^{1/2}s) \geq \lambda_{\text{min}}(x \circ s)$.

3. A full NT-step feasible IPM. In this section we present a full NT-step feasible IPM and its analysis. The main result of this section, Theorem 3.5, will be used later on, when dealing with the main purpose of this paper, a full NT-step infeasible IPM.

3.1. The symmetric optimization problem. We introduce the symmetric optimization problems in this subsection, and refer to [19] for some more details.

Let $J$ be a Euclidean Jordan algebra with dimension $n, \text{rank } r$, and cone of squares (or its associated symmetric cone) $K$. Consider the primal-dual pair of symmetric optimization problems
\[
\min \{ \langle c, x \rangle : Ax = b, \; x \in K \} \quad (\text{CP})
\]
and
\[
\max \{ b^T y : A^T y + s = c, \; y \in \mathbb{R}^m, \; s \in K \}. \quad (\text{CD})
\]
Here $c$ and the rows of $A$ lies in $J$, and $b \in \mathbb{R}^m$. Moreover, assume that $a_i$ is the $i$-th row of $A$, then $Ax = b$ means that
\[ \langle a_i, x \rangle = b_i, \; \text{for each } i = 1, \ldots, m, \]
while $A^T y + s = c$ means
\[ \sum_{i=1}^m y_i a_i + s = c. \]

We say that $x$ is in the null space of $A$, if $x \in J$ and $Ax = 0$, and $s$ in the row space of $A$ (or column space of $A^T$), if $s = A^T y$ for some $y \in \mathbb{R}^m$. Moreover, we say that $x$ and $s$ are orthogonal with respect to the trace inner product if $\text{tr}(x \circ s) = 0$ (or equivalently $\langle x, s \rangle = 0$). Note that if $x$ is in the null space of $A$ and $s$ in the row space of $A$, then
\[ \langle x, s \rangle = \langle x, A^T y \rangle = (Ax)^T y = 0, \]
i.e., they are orthogonal (with respect to the trace inner product).

3.2. Conic duality. We call (CP) feasible if there exists $x \in \mathcal{K}$ such that $Ax = b$, and strictly feasible, if in addition, $x \in \text{int} \mathcal{K}$. Similarly, we call (CD) feasible if there exists $(y, s) \in \mathbb{R}^m \times \mathcal{K}$ such that $A^T y + s = c$, and strictly feasible, if in addition $s \in \text{int} \mathcal{K}$.

Let $x$ and $(y, s)$ be a primal-dual feasible pair, i.e., a pair comprised of feasible solutions to (CP) and (CD). Then
\[ \langle c, x \rangle - b^T y = \langle A^T y + s, x \rangle - b^T y = \langle x, s \rangle \geq 0, \]
where \( (x, s) \) is called the duality gap.

If the primal problem (CP) is strictly feasible and below bounded, then the dual (CD) is solvable and the optimal values in the problems coincide. Similarly, if the dual (CD) is strictly feasible and above bounded, then the primal (CP) is solvable and the optimal values coincide. If both of the problems are strictly feasible, then both of them are solvable, and the optimal values coincide [2, Theorem 2.4.1].

3.3. The central path. In the sequel of the current section, we always assume that both (CP) and (CD) satisfy the IPC (Interior-Point Condition), i.e., both (CP) and (CD) are strictly feasible. Then finding an optimal solution of (CP) and (CD) is equivalent to solving the following system (cf. [4, 19]).

\[
\begin{align*}
Ax &= b, \quad x \in K, \\
A^Ty + s &= c, \quad s \in K, \\
x \circ s &= 0.
\end{align*}
\] (3.1)

The basic idea of primal-dual IPMs is to replace the third equation in (3.1), the so-called complementarity condition for (CP) and (CD), by the parameterized equation \( x \circ s = \mu e \), with \( \mu > 0 \). Thus we consider the system

\[
\begin{align*}
Ax &= b, \quad x \in K, \\
A^Ty + s &= c, \quad s \in K, \\
x \circ s &= \mu e.
\end{align*}
\] (3.2)

For each \( \mu > 0 \) the parameterized system (3.2) has a unique solution \( (x(\mu), y(\mu), s(\mu)) \), and we call \( x(\mu) \) and \( (y(\mu), s(\mu)) \) the \( \mu \)-center of (CP) and (CD), respectively. The set of \( \mu \)-centers (with \( \mu \) running through all positive real numbers) gives a homotopy path, which is called the central path of (CP) and (CD). If \( \mu \to 0 \), then the limit of the central path exists and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (CP) and (CD) [4].

3.4. The NT-search direction. The natural way to define a search direction is to follow the Newton approach and to linearize the third equation in (3.2). This leads to the system

\[
\begin{align*}
A \Delta x &= 0, \\
A^T \Delta y + \Delta s &= 0, \\
x \circ \Delta s + s \circ \Delta x &= \mu e - x \circ s.
\end{align*}
\] (3.3)

Due to the fact that \( x \) and \( s \) do not operator commute in general, i.e., \( L(x)L(s) \neq L(s)L(x) \), this system not always has a unique solution. In particular, for semidefinite optimization, the above system defines the AHO direction, which is not necessarily unique [4, 11]. It is now well known that this difficulty can be solved by applying a scaling scheme [19]. This goes as follows. Let \( u \in \text{int} \ K \). Then we have

\[ x \circ s = \mu e \quad \iff \quad P(u)x \circ P(u^{-1})s = \mu e. \]

Since \( x, s \in \text{int} \ K \), this is an easy consequence of Proposition 2.8 (ii), as becomes clear when using that \( x \circ s = \mu e \) holds if and only if \( x = \mu s^{-1} \). Now replacing the third
equation in (3.3) by \( P(u)x \circ P(u^{-1})s = \mu e \), and then applying Newtons method, we obtain the system
\[
A\Delta x = 0,
\]
\[
A^T \Delta y + \Delta s = 0,
\]
\[
P(u)x \circ P(u^{-1})\Delta s + P(u^{-1})s \circ P(u)\Delta x = \mu e - P(u)x \circ P(u^{-1})s.
\]
By choosing \( u \) appropriately (in a subclass of the Monteiro-Zhang family called the commutative class [12, 19, 23]) this system can be used to define search directions. In the literature the following choices are well known: \( u = s^{1/2} \), \( u = x^{-1/2} \), and \( u = w^{-1/2} \), where \( w \) is the NT-scaling point of \( x \) and \( s \). The first two choices lead to the so-called \( xs \)-direction and \( sz \)-direction, respectively [18, 19]. In this paper we focus on the third choice, which gives rise to the NT-direction. For that case we define
\[
v := \frac{P(w)^{-1/2}x}{\sqrt{\mu}} \quad \left[ = \frac{P(w)^{1/2}s}{\sqrt{\mu}} \right],
\]
and
\[
d_x := \frac{P(w)^{-1/2}\Delta x}{\sqrt{\mu}}, \quad d_s := \frac{P(w)^{1/2}\Delta s}{\sqrt{\mu}}.
\]
This enables us to rewrite the system (3.4) as follows:
\[
\sqrt{\mu}AP(w)^{1/2}d_x = 0,
\]
\[
\left( \sqrt{\mu}AP(w)^{1/2} \right)^T \frac{\Delta y}{\mu} + d_s = 0,
\]
\[
d_x + d_s = v^{-1} - v.
\]
That substitution of (3.5) and (3.6) into the first two equations of (3.4) yields (3.7) and (3.8) is easy to verify. It is less obvious that the third equation in (3.4) yields (3.9). By substitution we get, after dividing both sides by \( \mu \), \( v \circ (d_x + d_s) = e - v^2 \). This can be written as \( L(v)(d_x + d_s) = e - v^2 \). After multiplying of both sides from the left with \( L(v)^{-1} \), while using \( L(v)^{-1}e = v^{-1} \) and \( L(v)^{-1}v^2 = v \), we obtain (3.9). It easily follows that the above system has unique solution. Since (3.7) requires that \( d_x \) belongs to the null space of \( \sqrt{\mu}AP(w)^{1/2} \), and (3.8) that \( d_s \) belongs to the row space of \( \sqrt{\mu}AP(w)^{1/2} \), it follows that system (3.7)–(3.9) determines \( d_x \) and \( d_s \) uniquely as the (mutually orthogonal with respect to the trace inner product) components of the vector \( v^{-1} - v \) in these two spaces. From (3.9) and the orthogonality of \( d_x \) and \( d_s \) we obtain
\[
\|d_x\|_F^2 + \|d_s\|_F^2 = \|d_x + d_s\|^2_F = \|v^{-1} - v\|^2_F.
\]
Therefore the displacements \( d_x, d_s \) (and since \( \sqrt{\mu}AP(w)^{1/2} \) has full row rank, also \( \Delta y \)) are zero if and only if \( v^{-1} - v = 0 \). In this case it easily follows that \( v = e \), and that this implies that \( x \) and \( (y,s) \) coincide with the respective \( \mu \)-centers. To get the search directions \( \Delta x \) and \( \Delta s \) in the original space we simply transform the scaled search directions back to the \( x \)- and \( s \)-space by using (3.6):
\[
\Delta x = \sqrt{\mu}P(w)^{1/2}d_x,
\]
\[
\Delta s = \sqrt{\mu}P(w)^{-1/2}d_s.
\]
The new iterates are obtained by taking a full step, as follows.

\[
\begin{align*}
x^+ &= x + \Delta x, \\
y^+ &= y + \Delta y, \\
z^+ &= z + \Delta z.
\end{align*}
\] (3.12)

3.5. Proximity measure. In the analysis of the algorithm, we need a measure for the distance of the iterates \((x, y, s)\) to the current \(\mu\)-center \((x(\mu), y(\mu), s(\mu))\). The aim of this section is to present such a measure and to show how it depends on the eigenvalues of the vector \(v\).

The proximity measure that we are going to use is defined as follows:

\[
\delta(x, s; \mu) \equiv \delta(v) := \frac{1}{2} \|v - v^{-1}\|_F,
\] (3.13)

where \(v\) is defined in (3.5). It follows that

\[
4\delta^2(v) = \|v - v^{-1}\|_F^2 = \text{tr}(v^2) + \text{tr}(v^{-2}) - 2 \text{tr}(e),
\] (3.14)

which expresses \(\delta^2(v)\) in the eigenvalues of \(v^2\) and its inverse.


**Algorithm 3.1** A full NT-step feasible IPM

**Input:**
- accuracy parameter \(\varepsilon > 0\);
- barrier update parameter \(\theta, 0 < \theta < 1\);
- strictly feasible triple \((x^0, y^0, s^0)\) such that \(\text{tr}(x^0 \circ s^0) = \mu^0\text{tr}(e)\) and \(\delta(x^0, s^0; \mu^0) \leq 1/2\).

**begin**
\[
x := x^0; \quad y := y^0; \quad s := s^0; \quad \mu := \mu^0;
\]
**while** \(\text{tr}(x \circ s) \geq \varepsilon\)
\[
\mu\text{-update:} \quad \mu := (1 - \theta)\mu.
\]
\[
\text{NT-step:} \quad (x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s);
\]
**end**

3.7. The analysis of the NT-step.

3.7.1. Feasibility. Our aim is to find a condition that guarantees feasibility of the iterates after a full NT-step. As before, let \(x, s \in \text{int} \mathcal{K}, \mu > 0\) and let \(w\) be the scaling point of \(x\) and \(s\). Using (3.5), (3.6) and (3.12), we obtain

\[
\begin{align*}
x^+ &= x + \Delta x = \sqrt{\mu}P(w)^{1/2}(v + d_x), \\
(\Delta x, \Delta y, \Delta s) &= (\Delta x, \Delta y, \Delta s).
\end{align*}
\] (3.15)

Since \(P(w)^{1/2}\) and its inverse \(P(w)^{-1/2}\) are automorphisms of \(\text{int} \mathcal{K}\) (cf. Theorem 2.16), \(x^+\) and \(s^+\) will belong to \(\text{int} \mathcal{K}\) if and only if \(v + d_x\) and \(v + d_s\) belong to \(\text{int} \mathcal{K}\). For the
Since (3.9) that lemma. proof of our main result in this subsection, which is Lemma 3.2, we need the following lemma.

**Lemma 3.1.** If \( \delta(v) \leq 1 \) then \( e + d_x \circ d_s \in \mathcal{K} \). Moreover, if \( \delta(v) < 1 \) then \( e + d_x \circ d_s \in \text{int} \mathcal{K} \).

**Proof.** Since \( d_x \) and \( d_s \) are orthogonal with respect to the trace inner product, Lemma 2.40 implies that the absolute values of the eigenvalues of \( d_x \circ d_s \) do not exceed \( \frac{1}{2} \|d_x + d_s\|_F^2 \). In addition, it follows from (3.9) and (3.13) that

\[
\|d_x + d_s\|_F^2 = \|v - v^{-1}\|_F^2 = 4\delta(v)^2.
\]

Hence, the absolute values of the eigenvalues of \( d_x \circ d_s \) do not exceed \( \delta(v)^2 \). This implies that \( 1 - \delta(v)^2 \) is a lower bound for the eigenvalues of \( e + d_x \circ d_s \). Hence, if \( \delta(v) \leq 1 \) then \( e + d_x \circ d_s \in \mathcal{K} \) and if \( \delta(v) < 1 \) then \( e + d_x \circ d_s \in \text{int} \mathcal{K} \). This proves the lemma. \( \square \)

**Lemma 3.2.** The full NT-step is feasible if \( \delta(v) \leq 1 \) and strictly feasible if \( \delta(v) < 1 \).

**Proof.** We introduce a step length \( \alpha \) with \( 0 \leq \alpha \leq 1 \), and define

\[
v^0_x = v + \alpha d_x, \\
v^0_s = v + \alpha d_s.
\]

We then have \( v^0_x = v, v^1_x = v + d_x \), and similarly \( v^0_s = v, v^1_s = v + d_s \). It follows from (3.9) that

\[
v^0_x \circ v^0_s = (v + \alpha d_x) \circ (v + \alpha d_s) = v^2 + \alpha v \circ (d_x + d_s) + \alpha^2 d_x \circ d_s
\]

\[
= v^2 + \alpha v \circ (v^{-1} - v) + \alpha^2 d_x \circ d_s = (1 - \alpha)v^2 + \alpha e + \alpha^2 d_x \circ d_s.
\]

Since \( \delta(v) \leq 1 \), Lemma 3.1 implies that \( d_x \circ d_s \succeq \mathcal{K} - e \). Substitution gives

\[
v^0_x \circ v^0_s \succeq \mathcal{K} (1 - \alpha)v^2 + \alpha e - \alpha^2 e = (1 - \alpha)(v^2 + \alpha e).
\]

If \( 0 \leq \alpha < 1 \), the last vector belongs to \( \text{int} \mathcal{K} \), i.e., we then have

\[
v^0_x \circ v^0_s \succeq \mathcal{K} 0, \quad \text{for } \alpha \in [0, 1).
\]

By Lemma 2.42, we derive that \( \det(v^0_x) \) and \( \det(v^0_s) \) do not vanish for \( \alpha \in [0, 1) \). Since

\[
\det(v^0_x) = \det(v^0_s) = \det(v) > 0,
\]

by continuity, \( \det(v^0_x) \) and \( \det(v^0_s) \) stay positive for all \( \alpha \in [0, 1) \). Again by continuity, we also have that \( \det(v^1_x) \) and \( \det(v^1_s) \) are nonnegative. This proves that if \( \delta(v) \leq 1 \) then \( v + d_x \in \mathcal{K} \) and \( v + d_s \in \mathcal{K} \).

For \( \delta(v) < 1 \), we have by Lemma 3.1, \( d_x \circ d_s \succeq \mathcal{K} - e \) and similar arguments imply that \( \det(v^0_x) \) and \( \det(v^0_s) \) do not vanish for \( \alpha \in [0, 1] \), whence \( v + d_x \in \text{int} \mathcal{K} \) and \( v + d_s \in \text{int} \mathcal{K} \). This proves the lemma. \( \square \)

**Lemma 3.3.** Let \( x, s \in \text{int} \mathcal{K} \) and \( \mu > 0 \), then \( \langle x^+, s^+ \rangle = \mu \text{tr}(e) \).

**Proof.** Due to (3.15) we may write

\[
\langle x^+, s^+ \rangle = (\sqrt{\mu}P(w)^{1/2}(v + d_x), \sqrt{\mu}P(w)^{-1/2}(v + d_s)) = \mu(v + d_x, v + d_s).
\]
Using (3.9) we obtain
\[
\langle v + d_x, v + d_s \rangle = \langle v, v \rangle + \langle v, d_x + d_s \rangle + \langle d_x, d_s \rangle \\
= \langle v, v \rangle + \langle v, v^{-1} - v \rangle + \langle d_x, d_s \rangle \\
= \text{tr}(e) + \langle d_x, d_s \rangle.
\]
Since \(d_x\) and \(d_s\) are orthogonal with respect to the trace inner product, the lemma follows.

3.7.2. Quadratic convergence. In this subsection we prove quadratic convergence to the target point \((x(\mu), s(\mu))\) when taking full NT-steps. According to (3.5), the \(v\)-vector after the step is given by:
\[
v^+ := \frac{P(w^+)^{-1/2} x^+}{\sqrt{\mu}} = \frac{P(w^+)^{1/2} s^+}{\sqrt{\mu}}, \tag{3.16}
\]
where \(w^+\) is the scaling point of \(x^+\) and \(s^+\).

**Lemma 3.4 ([21, Proposition 5.9.3]).** One has
\[
v^+ \sim \left( P(v + d_x)^{1/2}(v + d_s) \right)^{1/2}.
\]

**Proof.** It readily follows from (3.16) and Lemma 2.35 that
\[
\sqrt{\mu} v^+ = P(w^+)^{1/2} s^+ \sim \left( P(x^+)^{1/2} s^+ \right)^{1/2}.
\]
Due to (3.15) and Lemma 2.34, we may write
\[
P(x^+)^{1/2} s^+ = \mu P \left( P(w)^{1/2}(v + d_x) \right)^{1/2} P(w)^{-1/2}(v + d_s) \\
\sim \mu P(v + d_x)^{1/2}(v + d_s).
\]
From this the lemma follows.

**Theorem 3.5.** If \(\delta := \delta(v) < 1\), then the full NT-step is strictly feasible and
\[
\delta(v^+) \leq \frac{\delta^2}{2(1 - \delta^2)}.
\]

**Proof.** Since \(\delta := \delta(v) < 1\), from Lemma 3.2 and its proof, we have
\[
v + d_x, v + d_s, (v + d_x) \circ (v + d_s) \in \text{int} K.
\]
For the moment, we denote
\[
u := P(v + d_x)^{1/2}(v + d_s), \\
\bar{u} := (v + d_x) \circ (v + d_s).
\]
Then, it follows from Lemma 3.4 that \(v^+ \sim u^{1/2}\). Therefore
\[
2\delta(v^+) = \|v^+ - (v^+)^{-1}\|_F = \|u^{1/2} - u^{-1/2}\|_F.
\]
By applying Lemma 2.43, we obtain
\[ 2\delta(v^+) = \|u^{1/2} - u^{-1/2}\|_F \leq \frac{\|u - e\|_F}{\lambda_{\min}(u^{1/2})} = \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}}. \]

In addition, we derive, by Lemma 2.44,
\[ 2\delta(v^+) \leq \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}} \leq \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}}. \]

As it follows from (3.9) that
\[ \bar{u} = (v + d_x) \circ (v + d_s) = v^2 + v \circ (d_x + d_s) + d_x \circ d_s \]
\[ = v^2 + v \circ (v^{-1} - v) + d_x \circ d_s = e + d_x \circ d_s, \]
substitution gives
\[ 2\delta(v^+) \leq \frac{\|d_x \circ d_s\|_F}{\lambda_{\min}(e + d_x \circ d_s)^{1/2}} = \frac{\|d_x \circ d_s\|_F}{[1 + \lambda_{\min}(d_x \circ d_s)]^{1/2}}. \]

Yet we apply Lemma 2.40. Part (i) of this lemma implies that \( \delta^2 \) is an upper bound for \( \|\lambda(d_x \circ d_s)\|_\infty \), as we already established in the proof of Lemma 3.1. Also using part (ii) of Lemma 2.40 we may now write
\[ 2\delta(v^+) \leq \frac{\|d_x \circ d_s\|_F}{[1 + \lambda_{\min}(d_x \circ d_s)]^{1/2}} \leq \frac{\sqrt{2}\delta^2}{\sqrt{1 - \delta^2}}, \]
which implies the lemma. \( \Box \)

As a result, the following corollary follows trivially.

**Corollary 3.6.** If \( \delta(v) \leq 1/\sqrt{2}, \) then the full NT-step is strictly feasible and \( \delta(v^+) \leq \delta(v)^2. \)

### 3.8. Updating the barrier parameter \( \mu. \)
In this section we establish a simple relation for our proximity measure just before and after a \( \mu \)-update.

**Lemma 3.7.** Let \( x, s \in \text{int} K \), \( \text{tr}(x \circ s) = \mu \text{tr}(e) \), and \( \delta := \delta(x, s; \mu) \). If \( \mu^+ = (1 - \theta)\mu \) for some \( 0 < \theta < 1 \), then
\[ \delta(x, s; \mu^+)^2 = \frac{\theta^2 \text{tr}(e)}{4(1 - \theta)} + (1 - \theta)\delta^2. \]

**Proof.** When updating \( \mu \) to \( \mu^+ \), the vector \( v \) is divided by the factor \( \sqrt{1 - \theta}. \)

Hence we may write
\[ 4\delta(x, s; \mu^+)^2 = \left\| \frac{v}{\sqrt{1 - \theta}} - \sqrt{1 - \theta}v^{-1} \right\|_F^2 = \left\| \frac{\theta v}{\sqrt{1 - \theta}} + \sqrt{1 - \theta}(v - v^{-1}) \right\|_F^2. \]

Yet we observe that the vectors \( v \) and \( v - v^{-1} \) are orthogonal with respect to the trace inner product. This is due to \( \text{tr}(v^2) = \text{tr}(e) \), as it readily follows from (3.5) that \( \mu \text{tr}(v^2) = \text{tr}(x \circ s) \). Hence we have
\[ \text{tr}(v \circ (v - v^{-1})) = \text{tr}(v^2 - e) = \text{tr}(v^2) - \text{tr}(e) = 0. \]

Therefore we may proceed as follows,
\[ 4\delta(x, s; \mu^+)^2 = \frac{\theta^2}{1 - \theta} \|v\|_F^2 + (1 - \theta)\|v - v^{-1}\|_F^2 = \frac{\theta^2}{1 - \theta} \text{tr}(e) + 4(1 - \theta)\delta^2. \]

This implies the lemma. \( \Box \)
3.9. Iteration bound. We conclude this section with an iteration bound for the Algorithm 3.1. Because the quadratic convergence lemma (i.e., Theorem 3.5) and, when we replace \( \text{tr}(e) \) by \( n \), the lemma describing the effect of a barrier parameter update (i.e., Lemma 3.7) are exactly the same as in [16] (cf. [16, Lemma 2.2] and [16, Lemma 2.2]), and also after a full NT-step the target value of the duality gap is attained, we can use the same arguments as in [16] to prove the following result.

**Theorem 3.8.** If \( \theta = 1/\sqrt{2r} \), with \( r = \text{tr}(e) \) the rank of the associated Euclidean Jordan algebra, then the number of iterations of the feasible path-following algorithm with full NT-steps does not exceed 

\[
\sqrt{2r} \log \frac{\mu_0^2 r}{\epsilon}.
\]

4. A full NT-step Infeasible IPM. A feasible triple \((x, y, s)\) is called an \( \epsilon \)-solution of (CP) and (CD) if the norms of the residual vectors \( b - Ax \) and \( c - A^T y - s \) do not exceed \( \epsilon \), and also \( \text{tr}(x \circ s) \leq \epsilon \). In this section we present an infeasible-start algorithm that generates an \( \epsilon \)-solution of (CP) and (CD), if it exists, or establish that no such solution exists.

4.1. Perturbed problems. We assume (CP) and (CD) have an optimal solution \((x^*, y^*, s^*)\) with vanishing duality gap, i.e., \( \text{tr}(x^* \circ s^*) = 0 \). As has become usual for infeasible IPMs we start the algorithm with a triple \((x_0, y_0, s_0)\) and \( \mu_0 > 0 \) such that

\[
x_0 = \zeta e, \quad y_0 = 0, \quad s_0 = \zeta e, \quad \mu_0 = \zeta^2, \quad (4.1)
\]

where \( \zeta \) is a (positive) number such that

\[
x^* + s^* \preceq_K \zeta e. \quad (4.2)
\]

The algorithm generates an \( \epsilon \)-solution of (CP) and (CD), or it establishes that there do not exist optimal solutions with vanishing duality gap satisfying (4.2). The initial values of the primal and dual residual vectors are denoted as \( r_b^0 \) and \( r_c^0 \), respectively. So we have

\[
r_b^0 = b - Ax_0, \quad r_c^0 = c - A^T y_0 - s_0.
\]

In general we have \( r_b^0 \neq 0 \) and \( r_c^0 \neq 0 \). In other words, the initial iterates are not feasible. The iterates generated by the algorithm will (in general) be infeasible for (CP) and (CD) as well, but they will be feasible for perturbed versions of (CP) and (CD) that we introduce in the following.

For any \( \nu \) with \( 0 \leq \nu \leq 1 \) we consider the perturbed problem \((\text{CP}_\nu)\), defined by

\[
\min \left\{ (c - \nu r_c^0)^T x : b - Ax = \nu r_b^0, \ x \in K \right\}, \quad \text{(CP}_\nu)
\]

and its dual problem \((\text{CD}_\nu)\), which is given by

\[
\min \left\{ (b - \nu r_b^0)^T y : c - A^T y - s = \nu r_c^0, \ s \in K \right\}. \quad \text{(CD}_\nu)
\]

Note that these problem are defined in such a way that if \((x, y, s)\) is feasible for \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\) then the residual vectors for the given triple \((x, y, s)\) with respect to the original problems (CP) and (CD) are \( \nu r_b^0 \) and \( \nu r_c^0 \), respectively.
If \( \nu = 1 \) then \( x = x^0 \) yields a strictly feasible solution of \((\text{CP}_\nu)\), and \((y, s) = (y^0, s^0)\), a strictly feasible solution of \((\text{CD}_\nu)\). This means that if \( \nu = 1 \) then \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\) satisfy the IPC.

**Lemma 4.1** (cf. [22, Theorem 5.13]). Let \((\text{CP})\) and \((\text{CD})\) be feasible and \(0 < \nu \leq 1\). Then the perturbed problems \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\) satisfy the IPC.

**Proof.** Let \( \bar{x} \) be a feasible solution of \((\text{CP})\) and \((\bar{y}, \bar{s})\) a feasible solution of \((\text{CD})\), i.e., \(Ax = b\) with \( \bar{x} \in \mathcal{K} \) and \(A^T \bar{y} + \bar{s} = c\) with \( \bar{s} \in \mathcal{K} \). Consider

\[
\begin{align*}
x &= (1 - \nu)\bar{x} + \nu x^0, \\
y &= (1 - \nu)\bar{y} + \nu y^0, \\
s &= (1 - \nu)\bar{s} + \nu s^0.
\end{align*}
\]

Since \( x \) is the sum of the vectors \( (1 - \nu)\bar{x} \in \mathcal{K} \) and \( \nu x^0 \in \text{int} \mathcal{K} \), we have \( x \in \text{int} \mathcal{K} \). Moreover

\[
b - Ax = b - A[(1 - \nu)\bar{x} + \nu x^0] = b - (1 - \nu)b - \nu Ax^0 = \nu(b - Ax^0) = \nu r^0,
\]

showing that \( x \) is strictly feasible for \((\text{CP}_\nu)\). In precisely the same way one shows that \((y, s)\) is strictly feasible for \((\text{CD}_\nu)\). Thus we have shown that \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\) satisfy the IPC. \( \square \)

It should be mentioned that this kind of perturbed problems have been studied first in [10], and later also in [6].

4.2. Central path of the perturbed problems. Let \((\text{CP})\) and \((\text{CD})\) be feasible and \(0 < \nu \leq 1\). Then Lemma 4.1 implies that the problems \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\) satisfy the IPC, and therefore their central paths exist. This means that for every \( \mu > 0 \) the system

\[
\begin{align*}
b - Ax &= \nu r^0, & x &\in \mathcal{K}, \\
c - A^T y - s &= \nu r^0, & s &\in \mathcal{K}, \\
x \circ s &= \mu e
\end{align*}
\]

has a unique solution. This solution is denoted as \((x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))\). These are the \( \mu \)-centers of the perturbed problems \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\). In the sequel the parameters \( \mu \) and \( \nu \) will always be in a one-to-one correspondence, according to

\[
\mu = \nu \mu^0 = \nu \zeta^2.
\]

Therefore, we feel free to omit one parameter and denote \((x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))\) simply as \((x(\nu), y(\nu), s(\nu))\).

Due to the choice of the initial iterates, according to (4.1), we have \(x^0 \circ s^0 = \mu^0 e\). Hence \(x^0\) is the \( \mu^0 \)-center of the perturbed problem \((\text{CP}_1)\) and \((y^0, s^0)\) the \( \mu^0 \)-center of the perturbed problem \((\text{CD}_1)\). In other words, \((x(1), y(1), s(1)) = (x^0, y^0, s^0)\).

4.3. The full NT-step infeasible IPM algorithm. We just established that if \( \nu = 1 \) and \( \mu = \mu^0 \), then \( x = x^0 \) and \((y, s) = (y^0, s^0)\) are the \( \mu \)-center of \((\text{CP}_\nu)\) and \((\text{CD}_\nu)\), respectively. These are our initial iterates.

We measure proximity to the \( \mu \)-center of the perturbed problems by the quantity \( \delta(x, s; \mu) \) as defined in (3.13). So, initially we have \( \delta(x, s; \mu) = 0 \). In the sequel we assume that at the start of each iteration, just before the \( \mu \)-update, \( \delta(x, s; \mu) \) is smaller than or equal to a (small) threshold value \( \tau > 0 \). Since we then have
δ(x, s; μ) = 0, this condition is certainly satisfied at the start of the first iteration, and also \( \text{tr}(x \circ s) = \mu^0 \text{tr}(e) \).

Now we describe one (main) iteration of our algorithm. Suppose that for some \( \nu \in (0, 1] \) we have \( x, \) and \( (y, s) \) satisfying the feasibility conditions (4.3) and (4.4) for \( \mu = \nu \mu^0 \), and such that \( \text{tr}(x \circ s) = \mu \text{tr}(e) \) and \( \delta(x, s; \mu) \leq \tau \). We reduce \( \nu \) to \( \nu + = (1 - \theta) \nu \), with \( \theta \in (0, 1) \), and find new iterates \( x^+ \) and \( (y^+, s^+) \) that satisfy (4.3) and (4.4), with \( \nu \) replaced by \( \nu^+ + = (1 - \theta) \nu^+ \) and \( \mu \) by \( \mu^+ = \nu^+ \mu^0 = (1 - \theta) \mu \), and such that \( \text{tr}(x^+ \circ s^+) = \mu^+ \text{tr}(e) \) and \( \delta(x^+, s^+; \mu^+) \leq \tau \).

One (main) iteration consists of a feasibility step and a few centering steps. The feasibility step serves to get iterates \( (x^f, y^f, s^f) \) that are strictly feasible for (CP\( \nu^+ \)) and (CD\( \nu^+ \)), and such that \( \delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2} \). In other words, \( (x^f, y^f, s^f) \) belongs to the quadratic convergence neighborhood of the \( \mu^+ \)-center of (CP\( \nu^+ \)) and (CD\( \nu^+ \)). Hence, because the NT-step is quadratically convergent in that region, a few centering steps, starting at \( (x^f, y^f, s^f) \) and targeting at the \( \mu^+ \)-center of (CP\( \nu^+ \)) and (CD\( \nu^+ \)) will generate iterates \( (x^+, y^+, s^+) \) that are feasible for (CP\( \nu^+ \)) and (CD\( \nu^+ \)) and that satisfy \( \text{tr}(x^+ \circ s^+) = \mu^+ \text{tr}(e) \) and \( \delta(x^+, s^+; \mu^+) \leq \tau \).

A formal description of the algorithm is given in Algorithm 4.1. Recall that after each iteration the residuals and the duality gap are reduced by the factor \( (1 - \theta) \). The algorithm stops if the norms of the residuals and the duality gap are less than the accuracy parameter \( \varepsilon \).

Algorithm 4.1 A full NT-step infeasible IPM for CO

**Input:**
- accuracy parameter \( \varepsilon > 0 \);
- barrier update parameter \( \theta \), \( 0 < \theta < 1 \);
- threshold parameter \( \tau > 0 \);
- initialization parameter \( \zeta > 0 \).

**begin**
- \( x := \zeta e; \ y := 0; \ s := \zeta e; \ \mu := \mu^0 = \zeta^2; \ \nu := 1; \)
- **while** \( \max(\text{tr}(x \circ s), \|b - Ax\|, \|c - A^Ty - s\|_F) \geq \varepsilon \)
  - feasibility step:
    - \( (x, y, s) := (x, y, s) + (\Delta f x, \Delta f y, \Delta f s); \)
  - update of \( \mu \) and \( \nu \):
    - \( \mu := (1 - \theta) \mu; \ \nu := (1 - \theta) \nu; \)
  - centering steps:
    - **while** \( \delta(x, s; \mu) \geq \tau \)
      - \( (x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s). \)
- endwhile
- endwhile
**end**

4.4. **Analysis of the feasibility step.** In this subsection, we define and analyze the feasibility step. This is the most difficult part of the analysis. In essence we follow the same chain of arguments as in [16], but at several places the analysis is more tight and also more elegant.

4.4.1. **Definition.** We describe the feasibility step in detail. The analysis will follow in subsequent sections. Suppose we have strictly feasible iterates \( (x, y, s) \) for
(CP) and (CD). This means that $(x, y, s)$ satisfies (4.3) and (4.4), with $\mu = \nu\zeta^2$. We need displacements $\Delta f x$, $\Delta f y$ and $\Delta f s$ such that

$$
\begin{align*}
x^f &:= x + \Delta f x, \\
y^f &:= y + \Delta f y, \\
s^f &:= s + \Delta f s,
\end{align*}
$$

(4.6)

are feasible for (CP) and (CD). One may easily verify that $(x^f, y^f, s^f)$ satisfies (4.3) and (4.4), with $\nu$ replaced by $\nu^+$ and $\mu$ by $\mu^+ = \nu^+ \mu^0 = (1 - \theta)\mu$, only if the first two equations in the following system are satisfied.

$$
A\Delta f x = \theta \nu r^0_x, \\
A^T \Delta f y + \Delta f s = \theta \nu r^0_s, \\
P(u)x \circ P(u^{-1})\Delta f s + P(u^{-1})s \circ P(u)\Delta f x = (1 - \theta)\mu e - P(u)x \circ P(u^{-1})s.
$$

(4.7)

The third equation is inspired by the third equation in the system (3.4) that we used to define search directions for the feasible case, except that we target at the $\mu^+$-centers of (CP) and (CD). As in the feasible case, we use the NT-scaling scheme to guarantee that the above system has a unique solution. So we take $w = w^{-1/2}$, where $w$ is the NT-scaling point of $x$ and $s$. Then the third equation becomes

$$
P(w)^{-1/2}x \circ P(w)^{1/2} \Delta f s + P(w)^{1/2} s \circ P(w)^{-1/2} \Delta f x = (1 - \theta)\mu e - P(w)^{-1/2}x \circ P(w)^{1/2} s
$$

(4.8)

Due to this choice of $w$ the coefficient matrix of the resulting system is exactly the same as in the feasible case, and hence it defines the feasibility step uniquely.

By its definition, after the feasibility step the iterates satisfy the affine equations in (4.3) and (4.4), with $\nu$ replaced by $\nu^+$. The hard part in the analysis will be to guarantee that $(x^f, s^f) \in \text{int } K$ and to guarantee that the new iterates satisfy $\delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2}$.

Let $(x, y, s)$ denote the iterates at the start of an iteration with $\text{tr}(x \circ s) = \mu \text{tr}(e)$ and $\delta(x, s; \mu) \leq \tau$. Recall that at the start of the first iteration this is certainly true, because $\text{tr}(x^0 \circ s^0) = \mu^0 \text{tr}(e)$ and $\delta(x^0, s^0; \mu^0) = 0$.

We scale the search directions, just as we did in the feasible case (cf. (3.6)), by defining

$$
d_x^f := \frac{P(w)^{-1/2} \Delta f x}{\sqrt{\mu}}, \quad d_s^f := \frac{P(w)^{1/2} \Delta f s}{\sqrt{\mu}}.
$$

(4.9)

with $w$ denoting the scaling point of $x$ and $s$, as defined in Proposition 2.30. With the vector $v$ as defined before (cf. (3.5)), the equation (4.8) can be restated as

$$
\mu v \circ (d_x^f + d_s^f) = (1 - \theta)\mu e - \mu v^2.
$$

By multiplying both sides of this equation from left with $\mu^{-1} L(v)^{-1}$ this equation gets the form

$$
d_x^f + d_s^f = (1 - \theta) v^{-1} - v.
$$
Thus we arrive at the following system for the scaled search directions in the feasibility step:

\[
\begin{align*}
\sqrt{\mu}AP(w)^{1/2}d_x^f &= \theta v r_0^k, \\
\left(\sqrt{\mu}AP(w)^{1/2}\right)^T \frac{\Delta f}{\mu} + d_x^s &= \frac{1}{\sqrt{\mu}} \theta v P(w)^{1/2}r_0^k, \\
d_x^f + d_x^s &= (1 - \theta)v^{-1} - v.
\end{align*}
\] (4.10)

To get the search directions \(\Delta f x\) and \(\Delta f s\) in the original \(x\)- and \(s\)-space we use (4.9), which gives

\[
\begin{align*}
\Delta f x &= \sqrt{\mu}P(w)^{1/2}d_x^f, \\
\Delta f s &= \sqrt{\mu}P(w)^{-1/2}d_x^s.
\end{align*}
\]

The new iterates are obtained by taking a full step, as given by (4.6). Hence we have

\[
\begin{align*}
x^f &= x + \Delta f x = \sqrt{\mu}P(w)^{1/2}(v + d_x^f), \\
s^f &= s + \Delta f s = \sqrt{\mu}P(w)^{-1/2}(v + d_x^s).
\end{align*}
\] (4.11)

From the third equation in (4.10) we derive that

\[
(v + d_x^f) \circ (v + d_x^f) = v^2 + v \circ [(1 - \theta)v^{-1} - v] + d_x^f \circ d_x^f = (1 - \theta)e + d_x^f \circ d_x^f.
\] (4.12)

As we mentioned before the analysis of the algorithm as presented below is much more difficult than in the feasible case. The main reason for this is that the scaled search directions \(d_x^f\) and \(d_x^s\) are not (necessarily) orthogonal (with respect to the trace inner product).

**4.4.2. Feasibility.** By the same arguments as in Subsection 3.7 it follows from (4.11) that \(x^f\) and \(s^f\) are strictly feasible if and only if \(v + d_x^f\) and \(v + d_x^s\) belong to \(\text{int } K\). Using this we have the following result.

**Lemma 4.2.** The iterates \((x^f, y^f, s^f)\) are feasible if

\[
(1 - \theta)e + d_x^f \circ d_x^f \in \mathcal{K},
\]

and strictly feasible if

\[
(1 - \theta)e + d_x^f \circ d_x^f \in \text{int } \mathcal{K}.
\]

**Proof.** Just as in the proof of Lemma 3.2 we introduce a step length \(\alpha\) with \(0 \leq \alpha \leq 1\), and define

\[
\begin{align*}
v_0^x &= v, \\
v_0^s &= v.
\end{align*}
\]

We then have \(v_0^x = v, v_0^s = v + d_x^f\), and similarly \(v_0^s = v, v_0^s = v + d_x^s\). From the third equation in (4.10), i.e., \(d_x^f + d_x^s = (1 - \theta)v^{-1} + v\), it follows that

\[
\begin{align*}
v_0^x \circ v_0^s &= (v + \alpha d_x^f) \circ (v + \alpha d_x^f) = v^2 + \alpha v \circ (d_x^f \circ d_x^f) + \alpha^2 d_x^f \circ d_x^f \\
&= v^2 + \alpha v \circ [(1 - \theta)v^{-1} - v] + \alpha^2 d_x^f \circ d_x^f \\
&= (1 - \theta)v^2 + \alpha(1 - \theta)e + \alpha^2 d_x^f \circ d_x^f.
\end{align*}
\]
Thus we obtain

$$v^0_n v^0_n \succeq \mathcal{K} (1 - \alpha) v^2 + \alpha (1 - \theta) e - \alpha^2 (1 - \theta) e = (1 - \alpha) (v^2 + \alpha (1 - \theta) e).$$

Since $v^2 \in \text{int} \mathcal{K}$, we have $v^2 + \alpha (1 - \theta) e \in \text{int} \mathcal{K}$. Hence,

$$v^0_n v^0_n \succeq \mathcal{K} (1 - \alpha) (v^2 + \alpha (1 - \theta) e) \succ 0, \quad \text{for } \alpha \in [0,1).$$

By Lemma 2.42, it follows that $\det(v^0_n v^0_n)$ and $\det(v^0_n v^0_n)$ do not vanish for $\alpha \in [0,1)$. Since

$$\det(v^0_n v^0_n) = \det(v^0_n v^0_n) = \det(v) > 0,$$

by continuity, $\det(v^0_n v^0_n)$ and $\det(v^0_n v^0_n)$ stay positive for all $\alpha \in [0,1)$. Again by continuity, we also have that $\det(v^0_n v^0_n)$ and $\det(v^0_n v^0_n)$ are nonnegative. This proves that if $(1 - \theta) e + d^f_L \circ d^f_R \in \mathcal{K}$, then $v + d^f_e \in \mathcal{K}$ and $v + d^f_s \in \mathcal{K}$, i.e., iterates $(x^f, y^f, s^f)$ are feasible.

If $(1 - \theta) e + d^f_L \circ d^f_R \in \text{int} \mathcal{K}$, or equivalently $d^f_L \circ d^f_R \succ 0$, similar arguments imply that $\det(v^0_n v^0_n)$ and $\det(v^0_n v^0_n)$ do not vanish for $\alpha \in [0,1]$, whence $v + d^f_e \in \text{int} \mathcal{K}$ and $v + d^f_s \in \text{int} \mathcal{K}$. This proves the lemma. \(\Box\)

It is clear from the above lemma that the feasibility of the iterates $(x^f, y^f, s^f)$ highly depends on the eigenvalues of the vector $d^f_L \circ d^f_R$.

### 4.4.3. Proximity.

We proceed by deriving an upper bound for $\delta(x^f, s^f; \mu^+)$.

Let $w^f$ be the scaling point of $x^f$ and $s^f$. When denoting the $v$-vector after the feasibility step, with respect to the $\mu^+$-center, as $v^f$, according to (3.5) this vector is given by

$$v^f := \frac{P(w^f)^{-1/2} x^f}{\sqrt{\mu(1 - \theta)}} \left[ \frac{P(w^f)^{1/2} s^f}{\sqrt{\mu(1 - \theta)}} \right]. \quad (4.13)$$

**Lemma 4.3.** One has

$$\sqrt{1 - \theta} v^f \sim \left[ P(v + d^f_L)^{1/2}(v + d^f_R) \right]^{1/2}.$$

**Proof.** It follows from (4.13) and Lemma 2.35 that

$$\sqrt{\mu(1 - \theta)} v^f = P(w^f)^{1/2} s^f \sim (P(x^f)^{1/2} s^f)^{1/2}.$$  

Due to (4.11) and Lemma 2.34, we may write

$$P(x^f)^{1/2} s^f = \mu P(w^f)^{1/2}(v + d^f_L)^{1/2} P(w)^{-1/2}(v + d^f_R) \sim \mu P(v + d^f_L)^{1/2}(v + d^f_R).$$

Thus we obtain

$$\sqrt{\mu(1 - \theta)} v^f \sim \sqrt{\mu} \left[ P(v + d^f_L)^{1/2}(v + d^f_R) \right]^{1/2}.$$

From this the lemma follows. \(\Box\)

The above lemma implies that

$$(v^f)^2 \sim \left[ \frac{v + d^f_L}{\sqrt{1 - \theta}} \right]^{1/2} \left[ \frac{v + d^f_R}{\sqrt{1 - \theta}} \right]. \quad (4.14)$$
In the sequel we denote $\delta(x', s'; u^+)$ also shortly by $\delta(v')$.

**Lemma 4.4.** If $\|\lambda(d_x' \circ d_s')\|_\infty < 1 - \theta$, then

$$4\delta(v')^2 \leq \frac{\|d_s' \circ d_x'\|_F^2}{1 - \frac{\|\lambda(d_s' \circ d_x')\|_\infty}{1 - \theta}}.$$  

**Proof.** Since $\|\lambda(d_x' \circ d_s')\|_\infty < 1 - \theta$, from Lemma 4.2 and its proof, we have $v + d_x', v + d_s', (v + d_x') \circ (v + d_s') \in \text{int} \mathcal{K}$.

For the moment, we denote

$$\delta = \frac{\|d_x' \circ d_s'\|_F}{1 - \frac{\|\lambda(d_x' \circ d_s')\|_\infty}{1 - \theta}}.$$  

Then, as $v' \sim u'^2$ (cf. (4.14)), we have

$$2\delta(v') = \|v' - (v')^{-1}\|_F = \|u'^2 - u^{-1/2}\|_F.$$  

By applying Lemma 2.43, we obtain

$$2\delta(v')^2 \leq \frac{\|u - e\|_F}{\lambda_{\min}(u'^2)} = \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}}.$$  

In addition, we derive, by Lemma 2.44,

$$2\delta(v') \leq \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}} \leq \frac{\|u - e\|_F}{\lambda_{\min}(u)^{1/2}}.$$  

As it follows from the third equation of (4.10) that

$$(1 - \theta)\bar{u} = (v + d_x') \circ (v + d_s') = v^2 + v \circ (d_x + d_s) + d_x' \circ d_s'$$

$$= v^2 + v \circ ((1 - \theta)v^{-1} - v) + d_x' \circ d_s' = (1 - \theta)e + d_x' \circ d_s',$$

substitution gives

$$2\delta(v') \leq \frac{\|d_s' \circ d_x'\|_F}{\lambda_{\min}(e + d_x' \circ d_s')} = \frac{\|d_s' \circ d_x'\|_F}{\lambda_{\min}(1 - \theta)^{1/2}}.$$  

The lemma follows. $\square$

From the definition of the Frobenius norm, Lemma 2.37, Lemma 2.39, we have

$$\|\lambda(d_x' \circ d_s')\|_\infty \leq \|d_x' \circ d_s'\|_F \leq \frac{1}{2} \left(\|d_x'\|^2 + \|d_s'\|^2\right)$$

$$\leq \frac{1}{2} \left(\|d_x'\|^2 + \|d_s'\|^2\right) \leq \frac{1}{2} \left(\|d_x'\|^2 + \|d_s'\|^2\right).$$  

(4.15)

Substitution of the above inequality into the inequality of Lemma 4.4 yields that

$$4\delta(v')^2 \leq \frac{\|d_x' \circ d_s'\|_F^2}{1 - \frac{\|\lambda(d_x' \circ d_s')\|_\infty}{1 - \theta}} \leq \frac{1}{1 - \frac{\|\lambda(d_x' \circ d_s')\|_\infty}{1 - \theta}}.$$  

(4.16)

Now, we have derived an upper bound for $\delta(v')$, but in terms of $\|d_x'\|^2 + \|d_s'\|^2$. To proceed, we need an upper bound for $\|d_x'\|^2 + \|d_s'\|^2$. 

4.4.4. Upper bound for $\|d_f^e\|_F^2 + \|d_f^c\|_F^2$. To obtain an upper bound for $\|d_f^e\|_F^2 + \|d_f^c\|_F^2$ is the subject of this subsection. In subsequent subsections this will enable us to find a default value for $\theta$.

For the moment, let us define

$$r_b := \theta \nu r_b^0, \quad r_c := \theta \nu r_c^0, \quad \bar{r} := (1 - \theta)u^{-1} - v. \quad (4.17)$$

With $\xi = -\frac{\mu \theta}{\mu}$, it follows from (4.10), by eliminating $d^e_f$,

$$\sqrt{\mu} AP(w)^{1/2} d^e_f = r_b, \quad (4.18)$$

$$\sqrt{\mu} P(w)^{1/2} A^T \xi + d^e_f = \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c. \quad (4.19)$$

By multiplying both sides of (4.19) from the left with $\sqrt{\mu} AP(w)^{1/2}$ and using (4.18) it follows that

$$\mu AP(w) A^T \xi + r_b = \sqrt{\mu} AP(w)^{1/2} \left( \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c \right).$$

Therefore,

$$\xi = \frac{1}{\mu} (AP(w) A^T)^{-1} \left[ \sqrt{\mu} AP(w)^{1/2} \left( \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c \right) - r_b \right].$$

Substitution into (4.19) gives

$$d_f^e = \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c$$

$$- \frac{1}{\sqrt{\mu}} P(w)^{1/2} A^T (AP(w) A^T)^{-1} \left[ \sqrt{\mu} AP(w)^{1/2} \left( \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c \right) - r_b \right]$$

$$= \left[ I - P(w)^{1/2} A^T (AP(w) A^T)^{-1} A P(w)^{1/2} \right] \left( \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c \right)$$

$$+ \frac{1}{\sqrt{\mu}} P(w)^{1/2} A^T (AP(w) A^T)^{-1} r_b.$$

To simplify notation we denote

$$\bar{P} = P(w)^{1/2} A^T (AP(w) A^T)^{-1} A P(w)^{1/2}.$$

Note that $\bar{P}$ is (the matrix of) the orthogonal projection (with respect to the trace inner product) to the row space of the matrix $AP(w)^{1/2}$. We now may write

$$d_f^e = [I - \bar{P}] \left( \bar{r} - \frac{1}{\sqrt{\mu}} P(w)^{1/2} r_c \right) + \frac{1}{\sqrt{\mu}} P(w)^{1/2} A^T (AP(w) A^T)^{-1} r_b.$$

Let $(\bar{x}, \bar{y}, \bar{s})$ be such that $A \bar{x} = b$ and $A^T \bar{y} + \bar{s} = c$. Then we may write

$$r_b = \theta \nu r_b^0 = \theta \nu (b - Ax^0) = \theta \nu A(\bar{x} - \bar{x}^0),$$

$$r_c = \theta \nu r_c^0 = \theta \nu (c - A^T \bar{y}^0 - s^0) = \theta \nu (A^T (\bar{y} - \bar{y}^0) + \bar{s} - s^0).$$
Thus we obtain
\[ d^f_x = [I - \hat{P}] \left( \bar{r} - \frac{\theta \nu}{\sqrt{\mu}} P(w)^{1/2} (A^T \bar{y} - \bar{s} - s^0) \right) + \frac{\theta \nu}{\sqrt{\mu}} \hat{P} P(w)^{-1/2}(\bar{x} - x^0). \]

Since \( I - \hat{P} \) is the orthogonal projection to the null space of \( AP(w)^{1/2} \) we have
\[ [I - \hat{P}] P(w)^{1/2} A^T (\bar{y} - y^0) = 0. \]

Hence it follows that
\[ d^f_x = [I - \hat{P}] \left( \bar{r} - \frac{\theta \nu}{\sqrt{\mu}} P(w)^{1/2} (\bar{s} - s^0) \right) + \frac{\theta \nu}{\sqrt{\mu}} \hat{P} P(w)^{-1/2}(\bar{x} - x^0). \]

To proceed we further simplify the notation by defining
\[ u^x = \frac{\theta \nu}{\sqrt{\mu}} P(w)^{-1/2}(\bar{x} - x^0), \quad u^s = \frac{\theta \nu}{\sqrt{\mu}} P(w)^{1/2}(\bar{s} - s^0). \]  

Then we may write
\[ d^f_x = [I - \hat{P}](\bar{r} - u^x) + \hat{P} u^x. \]

For \( d^f_s \) we obtain, by using the third equation of (4.10) and the definition (4.17) of \( \bar{r} \),
\[ d^f_s = \bar{r} - d^f_x = \bar{r} - [I - \hat{P}] \bar{r} + [I - \hat{P}] u^s - \hat{P} u^s = [I - \hat{P}] u^s + \hat{P}(\bar{r} - u^x). \]

We denote \( [I - \hat{P}] \bar{r} = \bar{r}_1 \) and \( \hat{P} \bar{r} = \bar{r}_2 \), and use similar notations for the projection of \( u^x \) and \( u^s \). Then from the above expressions for \( d^f_x \) and \( d^f_s \) we derive that
\[ d^f_x = \bar{r}_1 - u^1 + u^2_s, \quad d^f_s = u^1 + \bar{r}_2 - u^2_s. \]

Therefore, using the orthogonality (with respect to the trace inner product) of vectors with different subscripts, we may write
\[
\|d^f_x\|_F^2 + \|d^f_s\|_F^2 = \|\bar{r}_1 - u^1\|_F^2 + \|u^2_s\|_F^2 + \|u^1_s\|_F^2 + \|\bar{r}_2 - u^2_s\|_F^2 \\
= \|\bar{r}_1\|_F^2 + \|u^1_s\|_F^2 + 2\langle \bar{r}_1, u^1_s \rangle + \|u^2_s\|_F^2 + \|\bar{r}_2\|_F^2 + \|u^2_s\|_F^2 - 2\langle \bar{r}_2, u^2_s \rangle \\
= \|\bar{r}\|_F^2 + 2\|u^2_s\|_F^2 + 2\|u^1_s\|_F^2 - 2\langle \bar{r}_1, u^1_s \rangle - 2\langle \bar{r}_2, u^2_s \rangle.
\]

Further by the Cauchy-Schwartz inequality (cf. Lemma 2.36), and the properties of orthogonal projection, we obtain
\[
\|d^f_x\|_F^2 + \|d^f_s\|_F^2 \leq \|\bar{r}\|_F^2 + 2\|u^2_s\|_F^2 + 2\|u^1_s\|_F^2 + \|\bar{r}_1\|_F \|u^1_s\|_F + 2\|\bar{r}_2\|_F \|u^2_s\|_F \\
\leq \|\bar{r}\|_F^2 + 2\|u^2_s\|_F^2 + 2\|u^1_s\|_F^2 + \|\bar{r}_1\|_F^2 + \|u^1_s\|_F^2 + \|\bar{r}_2\|_F^2 + \|u^2_s\|_F^2 \\
\leq 2\|\bar{r}\|_F^2 + 3 \left( \|u^2_s\|_F^2 + \|u^1_s\|_F^2 \right). \tag{4.21}
\]

Since \( v \) and \( v^{-1} - v \) are orthogonal (with respect to the trace inner product) and \( \|v\|_F \geq \text{tr}(e) \), we have
\[
\|\bar{r}\|_F^2 = \|((1 - \theta)v^{-1} - v)\|_F^2 = \|((1 - \theta)(v^{-1} - v) - \theta v)\|_F^2 \\
= (1 - \theta)^2 \|v^{-1} - v\|_F^2 + \theta^2 \|v\|_F^2 = 4(1 - \theta)^2 \delta(v)^2 + \theta^2 r. \tag{4.22}
\]
where \( r = \text{tr}(e) \) is the rank of the associated Euclidean Jordan algebra. Due to (4.20) we have

\[
\|u^r\|_F^2 + \|u^s\|_F^2 = \frac{\theta^2 \nu^2}{\mu} \left( \left\| P(w)^{-1/2}(\bar{x} - x^0) \right\|_F^2 + \left\| P(w)^{1/2}(\bar{s} - s^0) \right\|_F^2 \right). \tag{4.23}
\]

Let \((x^*, y^*, s^*)\) be optimal solutions satisfying (4.2). It follows that \(Ax^* = b\) and \(ATy^* + s^* = c\). Therefore we may take \(\bar{x} = x^*\), \(\bar{y} = y^*\) and \(\bar{s} = s^*\). Since \(x^*\) is feasible for (CP) we have \(x^* \succeq_K 0\). Also \(s^* \succeq_K 0\). Hence we have \(0 \succeq_K x^* \succeq_K x^* + s^* \succeq_K \zeta e\), or equivalently \(0 \succeq_K \bar{x} \succeq_K \zeta e\). In a similar way we derive that \(0 \succeq_K \bar{s} \succeq_K \zeta e\). It therefore follows that

\[0 \succeq_K x^0 - \bar{x} \succeq_K \zeta e, \quad 0 \succeq_K s^0 - \bar{s} \succeq_K \zeta e.\]

We first consider the term \(\left\| P(w)^{-1/2}(\bar{x} - x^0) \right\|_F^2\). Using that \(P(w)^{1/2}\) is self-adjoint with respect to the inner product and \(P(w)e = w^2\), we may write

\[
\left\| P(w)^{-1/2}(\bar{x} - x^0) \right\|_F^2 = \left\| P(w)^{-1/2}(x^0 - \bar{x}) \right\|_F^2 = \langle P(w)^{-1}(x^0 - \bar{x}), x^0 - \bar{x} \rangle
\]

\[
= \langle P(w)^{-1}(x^0 - \bar{x}), P(w)^{-1}(x^0 - \bar{x})e - (x^0 - \bar{x}) \rangle
\]

\[
\leq \langle P(w)^{-1}(x^0 - \bar{x}), P(w)^{-1}(x^0 - \bar{x})e \rangle = \zeta \langle P(w)^{-1}e, x^0 - \bar{x} \rangle
\]

\[
= \zeta \langle P(w)^{-1}e, P(w)^{-1}e - (x^0 - \bar{x}) \rangle \leq \zeta^2 \text{tr}(w^2).
\]

In the same way it follows that

\[
\left\| P(w)^{1/2}(\bar{s} - s^0) \right\|_F^2 \leq \zeta^2 \text{tr}(w^2).
\]

Substitution of the last two inequalities into (4.23) gives

\[
\left\| u^r \right\|_F^2 + \left\| u^s \right\|_F^2 = \frac{\theta^2 \nu^2 \zeta^2}{\mu} \text{tr}(w^2 + w^{-2}).
\]

By using \(\mu = \nu \mu_0 = \nu \zeta^2\) we derive

\[
\left\| u^r \right\|_F^2 + \left\| u^s \right\|_F^2 = \frac{\theta^2 \nu}{\mu} \text{tr}(w^2 + w^{-2}).
\]

**Lemma 4.5.** One has

\[
\text{tr}(w^2 + w^{-2}) \leq \frac{\text{tr}(x + s)^2}{\mu \lambda_{\text{min}}(e)^2}.
\]

**Proof.** For the moment, let \(u := (P(x^{1/2})s)^{-1/2}\). Then, by Proposition 2.30, \(w = P(x^{1/2})u\). Using that \(P(x^{1/2})\) is self-adjoint, and also Lemma 2.41, we obtain

\[
\text{tr}(w^2) = \langle P(x^{1/2})u, P(x^{1/2})u \rangle = \langle u, P(x)u \rangle \leq \lambda_{\text{max}}(u) \text{tr}(P(x)u).
\]

By using the same arguments and also \(P(x)e = x^2\) we may write

\[
\text{tr}(P(x)u) = \text{tr}(P(x)u \circ e) = \langle P(x)u, e \rangle = \langle u, P(x)e \rangle = \langle u, x^2 \rangle \leq \lambda_{\text{max}}(u) \text{tr}(x^2).
\]
Combining the above inequalities we obtain
\[ \text{tr}(w^2) \leq \lambda_{\text{max}}(P(x^{-1/2})s^{-1}) \text{tr}(x^2). \]
Due to \( P(s^{1/2})x \sim P(x^{1/2})s \sim (P(w)^{1/2}s)^2 \sim (P(w^{-1/2})x)^2 = \mu v^2 \), we have
\[ \lambda_{\text{max}}(P(x^{-1/2})s^{-1}) \text{tr}(x^2) = \frac{\text{tr}(x^2)}{\lambda_{\text{min}}(P(x^{1/2})s)} = \frac{\text{tr}(x^2)}{\mu \lambda_{\text{min}}(v)^2}. \]
Thus we obtain
\[ \text{tr}(w^2) \leq \frac{\text{tr}(x^2)}{\mu \lambda_{\text{min}}(v)^2}. \]
By noting that \( w^{-1} \) is the scaling point of \( s \) and \( x \), it follows from the above inequality, by interchanging the role of \( x \) and \( s \), that
\[ \text{tr}(w^{-2}) \leq \frac{\text{tr}(s^2)}{\mu \lambda_{\text{min}}(v)^2}. \]
By adding the last two inequalities we obtain
\[ \text{tr}(w^2 + w^{-2}) \leq \frac{\text{tr}(x^2) + \text{tr}(s^2)}{\mu \lambda_{\text{min}}(v)^2}. \]
Since \( x, s \in K \), we have \( \text{tr}(x \circ s) \geq 0 \). Hence, also using that \( \text{tr}(u^2) \leq \text{tr}(u)^2 \) for each \( u \in K \) (cf. Lemma 2.38),
\[ \text{tr}(x^2) + \text{tr}(s^2) \leq \text{tr}(x^2) + \text{tr}(s^2) + 2 \text{tr}(x \circ s) = \text{tr}((x + s)^2) \leq \text{tr}(x + s)^2. \]
Substituting yields
\[ \text{tr}(w^2 + w^{-2}) \leq \frac{\text{tr}(x + s)^2}{\mu \lambda_{\text{min}}(v)^2}, \]
which completes the proof. \( \square \)

Concluding this subsection, we have, by using that \( \mu = \nu \mu^0 = \nu \zeta^2 \),
\[ \|d_x^f\|^2_F + \|d_s^f\|^2_F \leq 2 \left[ 4(1 - \theta)^2 \delta(v)^2 + \theta^2 r \right] + 3 \theta^2 \nu \frac{\text{tr}(x + s)^2}{\mu \lambda_{\text{min}}(v)^2} \]
\[ = 2 \left[ 4(1 - \theta)^2 \delta(v)^2 + \theta^2 r \right] + 3 \theta^2 \frac{\text{tr}(x + s)^2}{\zeta^2 \lambda_{\text{min}}(v)^2}, \]  \( (4.24) \)

where \( r \) is the rank of the associated Euclidean Jordan algebra. To continue, we need an upper bound for \( \text{tr}(x + s) \), and a lower bound for \( \lambda_{\text{min}}(v) \). We handle it in the next subsection.

**4.4.5. Bounds for \( \text{tr}(x + s) \) and \( \lambda_{\text{min}}(v) \).** Lemma 4.6. Let \( x \) and \( (y, s) \) be feasible for the perturbed problems \( (CP_\nu) \) and \( (CD_\nu) \), respectively, and \( \text{tr}(x \circ s) = \mu \text{tr}(e) \). With \( (x^0, y^0, s^0) \) as defined in (4.1) and \( \zeta \) as in (4.2), we then have
\[ \text{tr}(x + s) \leq 2 \zeta r, \]
where \( r = \text{tr}(e) \) is the rank of the associated Euclidean Jordan algebra.
This implies the lemma.

Finally, \( \text{tr}((x - \nu x^0 - (1 - \nu)x^*) \circ (s - \nu s^0 - (1 - \nu)s^*)) = 0. \)

By expanding the above equality and using the fact \( \text{tr}(x^* \circ s^*) = 0 \), we obtain

\[
\nu \left( \text{tr}(x^0 \circ s) + \text{tr}(s^0 \circ x) \right) = \text{tr}(x \circ s) + \nu^2 \text{tr}(x^0 \circ s^0) - (1 - \nu) \text{tr}(x \circ s^* + s \circ x^*) + \nu(1 - \nu) \text{tr}(x^0 \circ s^* + s^0 \circ x^*).
\]

Since \((x^0, y^0, s^0)\) is as defined in (4.1), we have

\[
\text{tr}(x^0 \circ s) + \text{tr}(s^0 \circ x) = \zeta \text{tr}(x + s),
\]

\[
\text{tr}(x^0 \circ s^0) = \zeta^2 \text{tr}(e),
\]

\[
\text{tr}(x^0 \circ s^* + s^0 \circ x^*) = \zeta \text{tr}(x^* + s^*).
\]

Due to (4.2) we have \( \text{tr}(x^* + s^*) \leq \zeta \text{tr}(e) \). Furthermore, \( \text{tr}(x \circ s) = \mu \text{tr}(e) = \nu \zeta^2 \text{tr}(e) \).

Finally, \( \text{tr}(x \circ s^* + s \circ x^*) \geq 0 \). Substitution of the relations gives

\[
\nu \zeta \text{tr}(x + s) \leq \nu \zeta^2 \text{tr}(e) + \nu^2 \zeta^2 \text{tr}(e) + \nu(1 - \nu) \zeta^2 \text{tr}(e) = 2\nu \zeta^2 \text{tr}(e).
\]

This implies the lemma. \( \square \)

**Lemma 4.7** (cf. [17, Lemma II.62]). Let \( \delta := \delta(v) \) be given by (3.13), then

\[
\frac{1}{\rho(\delta)} \leq \lambda_{\min}(v) \leq \lambda_{\max}(v) \leq \rho(\delta),
\]

where

\[
\rho(\delta) := \delta + \sqrt{1 + \delta^2}.
\]

**Proof.** By the definition of \( \delta(v) \) (cf. (3.13)), we have

\[
2\delta = \|v - v^{-1}\|_F = \sqrt{\sum_{i=1}^{\nu} \left( \lambda_i(v) - 1/\lambda_i(v) \right)^2}.
\]

Since \( \lambda_i(v) > 0 \), we derive

\[-2\delta \lambda_i(v) \leq 1 - \lambda_i(v)^2 \leq 2\delta \lambda_i(v).\]

This implies

\[
\lambda_i(v)^2 - 2\delta \lambda_i(v) - 1 \leq 0 \leq \lambda_i(v)^2 + 2\delta \lambda_i(v) - 1.
\]

Rewriting this as

\[
(\lambda_i(v) - \delta)^2 - 1 - \delta^2 \leq 0 \leq (\lambda_i(v) + \delta)^2 - 1 - \delta^2
\]
we obtain

\[(\lambda_{i}(v) - \delta)^2 \leq 1 + \delta^2 \leq (\lambda_{i}(v) + \delta)^2,\]

which implies

\[\lambda_{i}(v) - \delta \leq |\lambda_{i}(v) - \delta| \leq \sqrt{1 + \delta^2} \leq \lambda_{i}(v) + \delta.\]

Thus we arrive at

\[-\delta + \sqrt{1 + \delta^2} \leq \lambda_{i}(v) \leq \delta + \sqrt{1 + \delta^2} = \rho(\delta).\]

For the left-hand side expression we write

\[-\delta + \sqrt{1 + \delta^2} = \frac{1}{\delta + \sqrt{1 + \delta^2}} = \frac{1}{\rho(\delta)}.\]

This proves the lemma. \(\square\)

Substitute the above two results in (4.24), with \(\delta := \delta(v)\), we derive an upper bound for \(\|d^f\|^2_F + \|d^l\|^2_F\) as follows.

\[
\|d^f\|^2_F + \|d^l\|^2_F \leq 2 \left[4(1 - \theta)^2\delta^2 + \theta^2 r\right] + 3\theta^2 \frac{2\zeta r^2 \rho(\delta)^2}{\zeta^2} = 2 \left[4(1 - \theta)^2\delta^2 + \theta^2 r\right] + 12\theta^2 r^2 \rho(\delta)^2, \tag{4.25}\]

where \(r = \text{tr}(c)\) is the rank of the associated Euclidean Jordan algebra.

4.5. Analysis of the centering step. Assume that after the feasibility step \((x^f, y^f, s^f)\) lies in the quadratic convergence neighborhood with respect to the \(\mu\)+-center of the perturbed problems \((\text{CP}_{\mu^+})\) and \((\text{CD}_{\mu^+})\). From Section 3 we precisely know how to analyze these steps. If \(\delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2}\), then by Corollary 3.6, after \(k\) centering steps we will have iterates \((x^+, y^+, s^+)\) that are still feasible for \((\text{CP}_{\mu^+})\) and \((\text{CD}_{\mu^+})\) and such that

\[\delta(x^+, s^+; \mu^+) \leq (1/\sqrt{2})^{2^k}.\]

From this one easily deduces that \(\delta(x^+, s^+; \mu^+) \leq \tau\) will hold after at most

\[
\left\lfloor \log_2 \left( \log_2 \frac{1}{\tau} \right) \right\rfloor = 1 + \left\lfloor \log_2 \left( \log_2 \frac{1}{\tau} \right) \right\rfloor \tag{4.26}\]

centering steps.

4.6. Updating the barrier parameter \(\mu\). We want to choose \(\theta\), with \(0 < \theta < 1\) (as large as possible), and such that \((x^f, y^f, s^f)\) lies in the quadratic convergence neighborhood with respect to the \(\mu\)+-center of the perturbed problems \((\text{CP}_{\mu^+})\) and \((\text{CD}_{\mu^+})\), i.e., \(\delta(v^f) \leq 1/\sqrt{2}\). By (4.16), we derive this is the case when

\[
\frac{1}{4} \left( \frac{\|d^f\|^2_F + \|d^l\|^2_F}{1 - \theta} \right)^2 \leq 2.
\]

Considering \(\frac{\|d^f\|^2_F + \|d^l\|^2_F}{1 - \theta}\) as a single term, and by some elementary calculation, we obtain that

\[
\frac{\|d^f\|^2_F + \|d^l\|^2_F}{1 - \theta} \leq 2\sqrt{3} - 2 \approx 1.4641.
\]
By (4.25), the above equation holds if
\[ 2 \left[ 4(1 - \theta)^2\delta^2 + \theta^2 r \right] + 12 \theta^2 r^2 \rho(\delta)^2 \leq (2\sqrt{3} - 2)(1 - \theta). \]
Choosing \( \tau = 1/16 \), one may easily verify that if \( \theta = 1/4r \), (4.27) then the above inequality is satisfied.

Note that by (4.15) and the above bound for \( \|d^f_x\|_F + \|d^f_s\|_F \), we have
\[ \|\lambda(d^f_x \circ d^f_s)\|_\infty \leq \frac{1}{2}(\|d^f_x\|_F^2 + \|d^f_s\|_F^2) \leq (\sqrt{3} - 1)(1 - \theta), \]
which by Lemma 4.2, means that \((x^f, y^f, s^f)\) are strictly feasible.

**4.7. Iteration bound.** In the previous sections we have found that if at the start of an iteration the iterates satisfy \( \delta(x, s; \mu) \leq \tau \), with \( \tau = 1/16 \), then after the feasibility step, with \( \theta \) as defined in (4.27), the iterates are strictly feasible and satisfy \( \delta(x^f, s^f; \mu^+) \leq 1/\sqrt{2} \).

According to (4.26), at most three centering steps then suffice to get iterates \((x^+, y^+, s^+)\) that satisfy \( \delta(x^+, s^+; \mu^+) \leq \tau \) again. So each main iteration consists of at most four so-called inner iterations, in each of which we need to compute a search direction (for either a feasibility step or a centering step).

It has become a custom to measure the complexity of an IPM by the required number of inner iterations. In each main iteration both the duality gap and the norms of the residual vectors are reduced by the factor \( 1 - \theta \). Hence, using \( \text{tr}(x^0 \circ s^0) = r\zeta^2 \), the total number of main iterations is bounded above by
\[ \frac{1}{\theta} \log \max \left\{ \frac{r\zeta^2, \|x^0\|_F, \|s^0\|_F}{\varepsilon} \right\}. \]

Due to (4.27) we may take
\[ \theta = \frac{1}{4r} \]

Hence the total number of inner iterations is bounded above by
\[ 16r \log \max \left\{ \frac{r\zeta^2, \|x^0\|_F, \|s^0\|_F}{\varepsilon} \right\}. \]

Thus we may state without further proof the main result of the paper.

**Theorem 4.8.** If (CP) has an optimal solution \( x^* \) and (CD) an optimal solution \((y^*, s^*)\), which satisfy \( \text{tr}(x^* \circ s^*) = 0 \) and \( x^* + s^* \preceq K \zeta e \) for some \( \zeta > 0 \), then after at most
\[ 16r \log \max \left\{ \frac{r\zeta^2, \|x^0\|_F, \|s^0\|_F}{\varepsilon} \right\} \]
inner iterations the algorithm finds an \( \varepsilon \)-solution of (CP) and (CD). Here \( r = \text{tr}(e) \) is the rank of the associated Euclidean Jordan algebra.

Note that this bound is slightly better than that in [16, Theorem 4.8]. As in LO, it is derived under the assumption that there exists some optimal solutions of (CP)
and (CD) with vanishing duality gap and $x^* + s^* \preceq_{\mathcal{K}} \zeta \epsilon$. One might ask what happens if this condition is not satisfied. In that case, during the course of the algorithm it may happen that after some main steps the proximity measure $\delta$ (after the feasibility step) exceeds $1/\sqrt{2}$, because otherwise there is no reason why the algorithm would not generate an $\epsilon$-solution. So if this happens it tells us that either the problem (CP) and (CD) do not have optimal solutions (with zero duality gap) or the value of $\zeta$ has been chosen too small. In the latter case one might run the algorithm once more with a larger $\zeta$.

5. Concluding remarks. Using Jordan algebras, we generalize the IIPM for LO of Roos [16] to symmetric optimization. This unifies the analysis for LO, SOCO and SDO problems. The order of the iterations bounds coincide with the bounds derived for LO, which is the currently best known iteration bounds for symmetric optimization.

In this paper, we use NT-direction. A natural generalization is to use directions in the commutative class, or more generally the directions in MZ-family (maybe with different proximity measures). Another topic for further research is to consider large-update variants of the algorithm, since such methods are much more efficient in practice. Finally, a question that might be considered is whether full step methods can be made efficient by using dynamic updates of the barrier parameter. This will not improve the theoretical complexity, but it will enhance the practical performance of the algorithm significantly.

REFERENCES


