Conjugate Direction Methods and Polarity for Quadratic Hypersurfaces

Giovanni Fasano & Raffaele Pesenti

Journal of Optimization Theory and Applications

ISSN 0022-3239

J Optim Theory Appl
DOI 10.1007/s10957-017-1180-6

Springer

J Optim Theory Appl
DOI 10.1007/s10957-017-1180-6

Conjugate Direction Methods and Polarity for Quadratic Hypersurfaces

Giovanni Fasano[1] · **Raffaele Pesenti**[1]

Received: 9 June 2016 / Accepted: 14 October 2017

Abstract We use some results from polarity theory to recast several geometric properties of Conjugate Gradient-based methods, for the solution of nonsingular symmetric linear systems. This approach allows us to pursue three main theoretical objectives. First, we can provide a novel geometric perspective on the generation of conjugate directions, in the context of positive definite systems. Second, we can extend the above geometric perspective to treat the generation of conjugate directions for handling indefinite linear systems. Third, by exploiting the geometric insight suggested by polarity theory, we can easily study the possible degeneracy (pivot breakdown) of Conjugate Gradient-based methods on indefinite linear systems. In particular, we prove that the degeneracy of the standard Conjugate Gradient on nonsingular indefinite linear systems can occur only once in the execution of the Conjugate Gradient.

Keywords Polarity in homogeneous coordinates · Quadratic hypersurfaces · Conjugate Gradient method · Indefinite linear systems

Mathematics Subject Classification 90C30 · 65K99 · 51N15

Communicated by Jérôme Bolte.

✉ Giovanni Fasano
fasano@unive.it

Raffaele Pesenti
pesenti@unive.it

1 Department of Management, University Ca' Foscari of Venice, Venice, Italy

1 Introduction

In this paper, we consider methods based on the generation of conjugate directions, namely Conjugate Gradient (CG) methods for the solution of linear systems [1–3], within the framework of polarity theory. The fact of introducing this perspective as polarity may provide a geometric framework to describe the properties of CG-based methods [4], i.e. a different viewpoint from the algorithmic usually considered in the literature. This perspective also motivates the extension of the use of CG-based methods for solving indefinite linear systems and, possibly, for the search of stationary points of polynomial functions. In addition, it allows us to recast Planar-CG methods [4–7], for solving indefinite linear systems (the relevant analysis is not included in the current paper). Finally, we conjecture that this perspective may suggest a possible geometric insight for the Quasi–Newton updates, given the relation between CG and BFGS (Broyden–Fletcher–Goldfarb–Shanno) or L-BFGS (Limited memory Broyden–Fletcher–Goldfarb–Shanno) methods [3].

To the best of the authors' knowledge, the use of polarity to detail CG-based methods was just hinted by Hestenes and Stiefel in [1], but since then little can be found in the literature. Indeed, most of the literature studies the algebraic formalization of CG-based methods, to exploit and improve their computational performance. As an example, in [8] we can read that "*At that time* [1952] *it* [the CG] *was derived from the theory of polarity. Nowadays one prefers to see it as a method which permits the transformation of a matrix in tridiagonal form*".

Although the algorithmic perspective defines a line of research of great importance for CG-based methods, it is beyond the scope of the present work. Here, we focus on the theoretic characteristics of these methods that justify their use as iterative solvers for linear systems [9], in particular within large-scale optimization problems. In this context, CG-based methods are implemented to solve either positive definite or indefinite linear systems, whose approximate solutions define *gradient-related* directions that ensure the global convergence of the optimization algorithms, along with their fast convergence to stationary points.

Truncated Newton algorithms (see [9,10]), used to solve large-scale minimization problems for a real objective function, often implement CG-based methods. Specifically, these methods are a fast choice to approximately solve Newton's equation at the h-th outer iteration, as it reduces to a linear system. As an example, in [11,12] CG-based methods are used to yield superlinear convergence to an optimal solution of large-scale unconstrained minimization problems. Within truncated Newton algorithms CG-based methods are also used to compute *negative curvature directions* for the objective function [13–15]. These directions turn out to be useful in proving the convergence of the algorithm to stationary points, along with the satisfaction of second-order optimality conditions.

Sequential Quadratic Programming procedures [9], used to solve constrained minimization problems, may also adopt CG-based methods. In this case, at the current iteration h, CG-based methods are called to solve a (possibly) indefinite linear system.

CG-based methods have been recently proposed also to deflate conjugacy loss [16], and to generate efficient preconditioners for Newton's equation in different contexts [17,18].

In this paper, initially, we show how the main characteristics of the CG in $\mathbb{R}^n$ have a counterpart in homogeneous coordinates. Then, we show that polarity justifies the application of CG-based methods to the solution of linear systems, both in the positive definite and indefinite case. Finally, we observe that fundamental theorems of polarity theory, such as the Reciprocity Theorem and the Section Theorem (see [19] or the early work [20]), apply to general nonquadratic hypersurfaces. This last observation might possibly suggest further guidelines to study the geometry underlying the nonlinear CG-based methods and BFGS. In addition, by Proposition 8.3 we are going to see the dramatic impact of CG premature stop in the indefinite case, when proving global convergence properties (see also [11]).

The rest of the paper is organized as follows. In Sect. 2, we provide an introduction to polarity theory. In Sect. 3, we detail the main concepts of polarity theory in homogeneous coordinates, for quadratic hypersurfaces. In Sects. 4–6, we analyse the relation between polarity for quadratic hypersurfaces and the solution of symmetric linear systems, both in the case where the Hessian matrix of the quadratic hypersurface is positive definite and indefinite. Then, in Sect. 7, we focus on a basic tool (namely *polar hyperplanes*) used to describe the relation between polarity and the CG. We prove that polar hyperplanes are generalizations of tangent hyperplanes to algebraic hypersurfaces. We show that the CG implicitly handles polar hyperplanes, and in particular that the latter hyperplanes coincide with tangent hyperplanes, when we consider finite points on specific quadratic hypersurfaces. Section 8 contains advances on polarity applied to study the CG. Finally, in Sect. 9 we propose some perspectives for possible further investigation, and in Sect. 10 we draw some conclusions.

We use the following notation throughout this paper. $\|\cdot\|$ is the Euclidean norm. $\mathbb{R}^n$ is the n-dimensional Cartesian space and $\mathbb{P}^n$ is the associated homogeneous coordinates projective space (in which each point is indicated as $(x_1 : x_2 : \cdots : x_n : x_0)$). Given the vector $x \in \mathbb{R}^n$ and the scalar $x_0 \in \mathbb{R}$, for brevity and with a little abuse of notation we might indicate with $(x, x_0)^T$ a vector in $\mathbb{P}^n$. $dim(S)$ is the dimension of the subspace S of $\mathbb{R}^n$. $A \succ 0$ denotes that the matrix A is symmetric positive definite and $Ker(A)$ is the null space of matrix A. Lowercase Greek letters refer to hyperplanes, either in Cartesian coordinates or homogeneous coordinates. Finally, we use terms like *hyperplane* and *linear manifold* as synonyms. When a hyperplane includes the origin we may remark the fact that it represents a *linear subspace*.

2 Introduction to Polarity Theory

This section is devoted to detail aspects of projective geometry in $\mathbb{R}^n$, with a specific attention to study algebraic hypersurfaces in homogeneous coordinates. Within this last class of geometric entities, quadratics have attractive properties which play a key role in this work, in order to exploit iterative methods for the solution of symmetric linear systems (see also [19–21] for further references).

2.1 Homogeneous Coordinates in $\mathbb{R}^n$

Cartesian coordinates $(y_1, \dots, y_n) \in \mathbb{R}^n$ define a one-to-one map between real n-tuples and finite points in $\mathbb{R}^n$. In order to handle a simple algebra, including both points with finite coordinates and points at infinity, the *homogeneous coordinates* $(x_1 : \cdots : x_n : x_0)$ can be introduced, so that

$$y_i = \frac{x_i}{x_0}, \qquad i = 1, \dots, n, \tag{1}$$

where $x_i \in \mathbb{R}$, $i = 0, \dots, n$, are $n + 1$ finite values.

Proposition 2.1 *Let us consider the relation '$\sim$' on the set $\mathbb{R}^{n+1}\backslash\{0\}$, as defined by*

$$\begin{gathered}(x_1, \dots, x_n, x_0) \sim (\hat{x}_1, \dots, \hat{x}_n, \hat{x}_0)\\ \text{iff}\\ (\hat{x}_1, \dots, \hat{x}_n, \hat{x}_0) = \rho(x_1, \dots, x_n, x_0)\end{gathered}$$

for some real value $\rho \neq 0$. Then, '$\sim$' is an equivalence relation and the point $(x_1 : \cdots : x_n : x_0) \in \mathbb{P}^n$ represents the equivalence class

$$\{\rho(x_1, \dots, x_n, x_0) \ : \ \rho \in \mathbb{R}\backslash\{0\}\}.$$

In other words, the homogeneous coordinates define the projective space $\mathbb{P}^n$ as a one-to-one map between points (possibly at infinity) in $\mathbb{R}^n$ and points $(\rho x_1 : \cdots : \rho x_n : \rho x_0)$ in $\mathbb{P}^n$, where $\rho \neq 0$, provided that $(x_1, \cdots, x_n, x_0) \neq 0$.

2.2 Basics on Polarity for Algebraic Hypersurfaces

Let $\varphi(x_1, \dots, x_n, x_0)$ be a homogeneous polynomial of degree $r \in \mathbb{N}$, then we say that the locus of points of $\mathbb{P}^n$ satisfying

$$\mathcal{F} : \left\{(x, x_0)^T \in \mathbb{P}^n \ : \ \varphi(x_1, \dots, x_n, x_0) = 0\right\} \tag{2}$$

is an *algebraic hypersurface* of order r.

Definition 2.1 Given the algebraic hypersurface (2) of order $r \geq 1$, in homogeneous coordinates, consider the point (*pole*) $P = (\bar{x}_1 : \cdots : \bar{x}_n : \bar{x}_0) \in \mathbb{P}^n$. Then, the equation

$$\sum_{i=0}^{n} \frac{\partial \varphi(x_1, \dots, x_n, x_0)}{\partial x_i} \bar{x}_i = 0 \tag{3}$$

represents an algebraic hypersurface of order $r-1$, which is said to be the *first polar* (or *polar*) of the point P with respect to the hypersurface (2), in homogeneous coordinates.

Figure 1 helps to detail the geometry behind (2) (blue continuous line) and (3) (red dashed line), being $\varphi(x_1, x_2, x_0) = x_1^3 - x_2 x_0^2 = 0$. In particular, Fig. 1 *(left)* describes

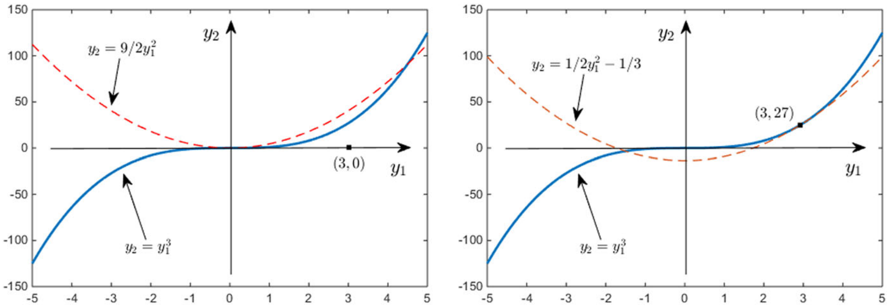


Fig. 1 The (blue) continuous line represents the hypersurface with equation $\varphi(x_1, x_2, x_0) = x_1^3 - x_2 x_0^2 = 0$, while the (red) dashed line is its polar hypersurface (3), in Cartesian coordinates. (left) The pole $(\bar{x}, 1)^T \in \mathbb{P}^2$, where $\bar{x} = (3, 0)^T \in \mathbb{R}^2$, does not belong to the hypersurface; (right) the pole $(\bar{w}, 1)^T \in \mathbb{P}^2$, where $\bar{w} = (3, 27)^T \in \mathbb{R}^2$, belongs to the hypersurface

the case in which the pole $(\bar{x}, 1)^T \in \mathbb{P}^2$, where $\bar{x} = (3, 0)^T \in \mathbb{R}^2$, is not a point of $\mathcal{F}$, while in Fig. 1 (right) the point $(\bar{w}, 1)^T \in \mathbb{P}^2$, where $\bar{w} = (3, 27)^T \in \mathbb{R}^2$, belongs to $\mathcal{F}$.

We immediately realize that in case $\varphi(x_1, \ldots, x_n, x_0) = 0$ is a quadratic hypersurface (i.e. $r = 2$ and $\partial\varphi(x_1, \ldots, x_n, x_0)/\partial x_i$ is linear) then we have from (3)

$$\sum_{i=0}^{n} \frac{\partial\varphi(x_1, \ldots, x_n, x_0)}{\partial x_i} \cdot \bar{x}_i = \sum_{i=0}^{n} \left.\frac{\partial\varphi(x_1, \ldots, x_n, x_0)}{\partial x_i}\right|_{x_i=\bar{x}_i} \cdot x_i = 0. \tag{4}$$

Using a standard taxonomy, the algebraic hypersurface (2) is often addressed as the 0-th *polar* (of any point in $\mathbb{P}^n$, with respect to itself), while the first polar of the point $P \in \mathbb{P}^n$ with respect to (2), when $r = 2$, is often called the *polar hyperplane* of P.

We report here two of the essential results for polarity, where we focus on (quadratic) hypersurfaces of order $r = 2$.

Theorem 2.1 (Reciprocity Theorem) *Consider the algebraic hypersurface* (2) *of order $r = 2$ and the points P, $Q \in \mathbb{P}^n$. If the polar hyperplane of P with respect to* (2) *passes through Q, then the polar hyperplane of Q with respect to* (2) *passes through P.*

Theorem 2.2 (Section Theorem) *Consider the algebraic hypersurface* (2) *of order $r \geq 1$. Let V_d be a linear space of dimension $d \leq n$, such that $\bar{\mathcal{F}} = \mathcal{F} \cap V_d$ and $\bar{\mathcal{F}} \neq V_d$. For every point $P \in V_d$, the section by V_d of the first polar of P [with respect to* (2)*] coincides with the first polar of P with respect to $\bar{\mathcal{F}}$.*

An example of the geometry behind the Reciprocity Theorem is detailed in Fig. 2 *(left)*, where for simplicity $n = 2$ and hyperplanes are represented by lines: the point P is the pole of ℓ_2' with respect to $\mathcal{F}$, while the dashed lines by P admit as poles their intersections with the line ℓ_2'.

The definition of polarity, along with the fact that the gradient of φ in (2) is well defined, implies that, if a finite point P satisfies (2) and $r = 2$, then the polar (i.e. the polar hyperplane) of P with respect to (2) coincides with the *tangent hyperplane* of (2) in P. On the other hand, Theorem 2.2 allows the definition of polar hyperplane also for points at infinity, where the tangent hyperplane (in Cartesian coordinates) does not exist. In this regard, the next (simplified) definition and the subsequent result (whose proof can be found in [19]) clarify the taxonomy of points.

Definition 2.2 Given the algebraic hypersurface (2) of order $r \geq 1$ and a point $P \in \mathbb{P}^n$ in homogeneous coordinates, we say that P is *self-conjugate with respect to* (2), if P belongs to its first polar with respect to (2).

In particular, it follows that, if the point P is self-conjugate and $r = 2$, then its polar hyperplane coincides with the tangent hyperplane of $\varphi(x_1, \ldots, x_n, x_0) = 0$ in P. Moreover, we have the following result.

Proposition 2.2 *Given the algebraic hypersurface* (2) *and a point $P \in \mathbb{P}^n$, P is self-conjugate if and only if P satisfies* (2)*, i.e. P belongs to the algebraic hypersurface* (2).

3 Polarity and the Geometry Behind Conjugacy

In this section, we study some basics of polarity for *quadratic hypersurfaces* [19,21], in the homogeneous coordinates projective space $\mathbb{P}^n$. Given a symmetric matrix $A = \{a_{ij}\} \in \mathbb{R}^{n\times n}$, a vector $b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$, and a scalar $c \in \mathbb{R}$, we consider the reference symmetric linear system

$$Ay = b, \tag{5}$$

and as a reference quadratic functional

$$g(y) = \frac{1}{2} y^T Ay - b^T y + c. \tag{6}$$

Initially, we associate a quadratic hypersurface $\mathcal{F}$ to g, in the corresponding $\mathbb{P}^n$ homogeneous coordinates projective space. To this end, we observe that g can be associated to the functional $f : \mathbb{P}^n \to \mathbb{R}$ in homogeneous coordinates defined, for $x_0 \neq 0$, by

$$f(x_1, \ldots, x_n, x_0) = g\left(\frac{x}{x_0}\right) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\left(\frac{x_i}{x_0}\right)\left(\frac{x_j}{x_0}\right) - \sum_{i=1}^{n} b_i\left(\frac{x_i}{x_0}\right) + c, \tag{7}$$

where $x = (x_1, \ldots, x_n)^T$. In turn, the functional $f(x_1, \ldots, x_n, x_0)$ can be associated with the quadratic hypersurface

$$\begin{aligned} \mathcal{F} &:= \{(x, x_0)^T \in \mathbb{P}^n : f(x, x_0)x_0^2 = 0\} \\ &\equiv \{(x, x_0)^T \in \mathbb{P}^n : x^T Ax - 2x_0 b^T x + 2cx_0^2 = 0\}. \end{aligned} \tag{8}$$

Hereinafter, the symbol $\mathcal{F}$ will always indicate the quadratic hypersurface in (8). Moreover, the next assumption will hold throughout the paper with the exception of a few highlighted results.

Assumption 3.1 Given the nonsingular symmetric matrix $A \in \mathbb{R}^{n\times n}$, the vector $b \in \mathbb{R}^n$ and the scalar c in (8), the next two equivalent conditions hold:

- matrix

$$\begin{pmatrix} A & -b \\ -b^T & 2c \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

  is nonsingular;
- we have $2c - b^T A^{-1} b \neq 0$.

Under Assumption 3.1, $\mathcal{F}$ in (8) has a finite centre and is nondegenerate, so that polarity identifies a *one-to-one* correspondence between points and hyperplanes in the space of homogeneous coordinates [22] (see also Fig. 4). Indeed, invoking the duality principle, we can consider the projective space $\mathbb{P}^n$ and the dual space $(\mathbb{P}^n)^*$ of hyperplanes corresponding to points in $\mathbb{P}^n$ (using the dual map in Definition 2.1, which associates the hyperplane of Eq. (3) to the point P). Then, if Assumption 3.1 holds, the dual map is nondegenerate, and the set of all the dual hyperplanes of the points of $\mathcal{F}$ in (8) coincides with its closure (i.e. the dual variety of $\mathcal{F}$—see also [23]).

We note that the transformation $y = x/x_0$ makes the *hyperplane at infinity* $x_0 = 0$ (in homogeneous coordinates) correspond to the locus of all the points at infinity of $\mathbb{R}^n$. Then, we can use the hypersurface $\mathcal{F}$ to study the properties of g at points at infinity. To this end, we introduce some definitions and propositions concerning $\mathcal{F}$ that will prove to be useful for a complete analysis.

Definition 3.1 Given $\mathcal{F}$ and the points $(x^*, x_0^*)^T, (\bar{x}, \bar{x}_0)^T \in \mathbb{P}^n$:

- $(\bar{x}, \bar{x}_0)^T$ is the pole of the hyperplane $\bar{\pi}$ if $\bar{\pi}$ is the first polar hyperplane of $(\bar{x}, \bar{x}_0)^T$ with respect to $\mathcal{F}$;
- the pole $(x^*, x_0^*)^T$ of the hyperplane $x_0 = 0$ is the centre of $\mathcal{F}$;
- the hyperplanes through the centre $(x^*, x_0^*)^T$ of $\mathcal{F}$ are the diametral hyperplanes of $\mathcal{F}$;
- the lines through the centre $(x^*, x_0^*)^T$ of $\mathcal{F}$ are the diameters of $\mathcal{F}$.

Let us stress that a finite centre for $\mathcal{F}$ can be defined only in case Assumption 3.1 holds. We also observe that, by the *Reciprocity Theorem* (Theorem 2.1), the diametral hyperplanes of $\mathcal{F}$ are polar hyperplanes of points at infinity, i.e. the pole of a diametral hyperplane is a point at infinity whose coordinates satisfy the equation $x_0 = 0$.

Definition 3.2 Two diametral hyperplanes π_1 and π_2 of $\mathcal{F}$ are conjugate, if π_i contains the pole of π_j, for $i, j \in \{1, 2\}$, $i \neq j$.

Definitions 3.1 and 3.2 imply that, if π_1 and π_2 are conjugate diametral hyperplanes, then the pole of π_1 is one of the points at infinity of π_2 and vice versa. Figure 2(left) presents the simple geometry behind polarity for quadratic hypersurfaces. The point P is the pole of the hyperplane (i.e. line, since $n = 2$) ℓ'_2. In particular, ℓ'_2 is obtained

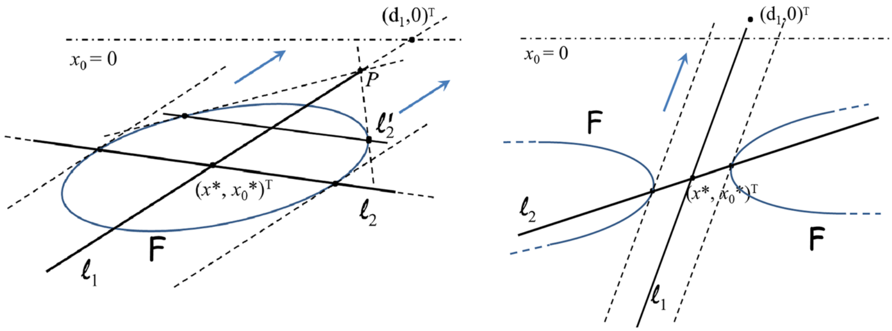


Fig. 2 Conjugate lines ℓ_1 and ℓ_2 with respect to the quadratic hypersurface $\mathcal{F}$ in (8), characterized by a positive definite matrix A (left) and an indefinite matrix A (right)

joining the tangency points of lines from P to $\mathcal{F}$. With a similar construction the point at infinity $(d_1, 0)^T$ of ℓ_1 represents also the pole of ℓ_2, being parallel the tangent lines from $(d_1, 0)^T$ to $\mathcal{F}$. As a consequence, ℓ_1 contains the pole of ℓ_2. A similar analysis also holds for Fig. 2 (right), in the indefinite case. It can be noticed that the vector d_1 is parallel to ℓ_1 and that, as $n = 2$, the line ℓ_2 may be also seen as a diametral hyperplane whose pole is $(d_1, 0)^T$.

Definition 3.3 Given $\mathcal{F}$:

- two diameters ℓ_1 and ℓ_2 of $\mathcal{F}$ are conjugate if the point at infinity $(d_i, 0)^T \in \mathbb{P}^n$ of ℓ_i, $i \in \{1, 2\}$, is the pole of a (diametral) hyperplane which contains ℓ_j, $j \in \{1, 2\}$, for $i \neq j$. Two lines ℓ_1 and ℓ_2 are conjugate, if they are, respectively, parallel to conjugate diameters.
- a diametral hyperplane π_1 is conjugate to the diameter ℓ_2 if any line contained in π_1 is conjugate to ℓ_2. A hyperplane π_1 is conjugate to the line ℓ_2 if π_1 and ℓ_2 are, respectively, parallel to a diametral hyperplane and a diameter that are conjugate.

We point out that Definition 3.3 may sound unusual. In the literature (see, e.g. [19]), lines ℓ_1 and ℓ_2 are addressed as conjugate if the polar hyperplanes of *all the points* (not just of the point at infinity) of ℓ_1 include lines parallel to ℓ_2 and vice versa. However, this latter definition and Definition 3.3 are equivalent. Indeed, let π_3 be the polar hyperplane of the point at infinity of the diameter ℓ_1 (see Fig. 3). Simple geometric considerations prove that ℓ_1 and ℓ_2 are conjugate according to Definition 3.3 (i.e. $\ell_2 \in \pi_3$). Hence, all the points on the line ℓ_1 are the poles of hyperplanes parallel to π_3. This fact is represented in Fig. 3, where a three-dimensional quartic $\mathcal{F}$ and the polar hyperplanes of three points of ℓ_1 are depicted. There, π_1 is the polar hyperplane of the point P_1 of tangency between $\mathcal{F}$ and π_1; π_2 is the polar hyperplane of a generic point $P_2 \in \ell_1$; π_3 is the polar hyperplane of $(d_1, 0)^T$, i.e. the point at infinity of ℓ_1. As the three hyperplanes π_1, π_2 and π_3 are parallel, π_1 and π_2 certainly include lines parallel to $\ell_2 \in \pi_3$. Finally, we highlight that the results in Fig. 3 (which refer to $n = 3$), can be similarly obtained when $n = 2$, as in Fig. 2 *(left)*.

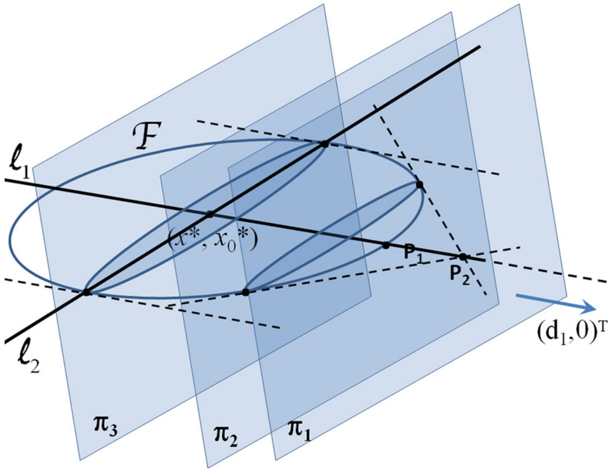


Fig. 3 Polar hyperplanes of points on the line ℓ_1, with respect to $\mathcal{F}$. π_3 is the polar hyperplane of $(d_1, 0)^T$, π_2 is the polar hyperplane of P_2 and π_1 is the polar hyperplane of P_1

4 Basic Practical Consequences of Conjugacy

In this section, we show that the geometric concept of conjugacy of lines has an immediate algebraic counterpart, involving the directional cosines of the lines. This fact will be useful to address some properties of CG-based methods in Sects. 5–7.

Proposition 4.1 *Two lines ℓ_1 and ℓ_2 are conjugate with respect to $\mathcal{F}$ if and only if*

$$d_1^T A d_2 = 0, \tag{9}$$

where $(d_1, 0)^T$ and $(d_2, 0)^T$ are points at infinity of, respectively, ℓ_1 and ℓ_2.

Proof The polar hyperplane of $(d_1, 0)^T \in \ell_1$, with respect to $\mathcal{F}$, is given by the diametral hyperplane whose points $(x, x_0)^T \in \mathbb{P}^n$ satisfy

$$2(Ax - bx_0)^T d_1 = 0 \tag{10}$$

(see relation (4)). Thus, by Definition 3.3, the line ℓ_2 is conjugate to ℓ_1 if and only if the point at infinity $(d_2, 0)^T$ satisfies (10), i.e. (9) holds. □

The above proposition highlights how the geometric concept of conjugacy between lines ℓ_1 and ℓ_2 simply depends on the matrix A in (8), regardless of the choice of the vector b and the scalar c in $\mathcal{F}$. The algebraic counterpart (9) of the geometric concept of conjugacy has also an additional equivalent characterization. Indeed, in case $A \succ 0$ relation (9) shows that the conjugacy between ℓ_1 and ℓ_2 is equivalent to impose the orthogonality condition $\langle d_1, d_2 \rangle_A = 0$, being $\langle \cdot, \cdot \rangle_A$ the inner product induced by A. As a consequence, in case $A = I_n$ then the conjugacy between ℓ_1 and ℓ_2 reduces to the orthogonality condition $d_1^T d_2 = 0$.

Definition 4.1 Two vectors $d_1 \in \mathbb{R}^n$ and $d_2 \in \mathbb{R}^n$ are conjugate vectors with respect to $\mathcal{F}$, if (9) holds.

We highlight that the proof of the next proposition indicates a simple but relevant property, namely, given a direction $d \in \mathbb{R}^n$, the vector Ad is orthogonal to any direction conjugate to d. We will use this fact later on in the paper (Proposition 6.2).

Proposition 4.2 *Consider a line $\hat{\ell}$ with point at infinity $(\hat{d}, 0)^T$. Then, a hyperplane π is conjugate to $\hat{\ell}$, with respect to $\mathcal{F}$, if and only if any line contained in π is conjugate to $\hat{\ell}$. Moreover, π is parallel to the polar hyperplane of $(\hat{d}, 0)^T$.*

Proof The first part of the proposition trivially follows from Definition 3.3. To prove the second part of the proposition, consider the polar hyperplane of $(\hat{d}, 0)^T$, i.e. the hyperplane whose points $(x, x_0)^T \in \mathbb{P}^n$ satisfy $(Ax - bx_0)^T \hat{d} = 0$. Let ℓ be a generic line in this hyperplane. Then ℓ can be expressed as the locus of the points $(\bar{x}, \bar{x}_0)^T + \lambda(d, 0)^T$ for all $\lambda \in \mathbb{R}$, where $(\bar{x}, \bar{x}_0)^T$ is a point of ℓ and $(d, 0)^T$ is the point at infinity of ℓ. As ℓ is in the polar hyperplane of $(\hat{d}, 0)^T$, then

$$[A(\bar{x} + \lambda d) - b\bar{x}_0]^T \hat{d} = 0 \iff (A\bar{x} - b\bar{x}_0)^T \hat{d} + \lambda(Ad)^T \hat{d} = 0, \qquad \forall \lambda \in \mathbb{R},$$

and hence $\lambda(Ad)^T \hat{d} = 0$, because $(\bar{x}, \bar{x}_0)^T \in \pi$. Consequently, ℓ is conjugate to $\hat{\ell}$ by Proposition 4.1. □

In the next proposition, we state the relation between the centre of $\mathcal{F}$ and the solution of the symmetric linear system $Ay = b$.

Proposition 4.3 (Equivalence of centre) *Consider the quadratic hypersurface $\mathcal{F}$ with centre $(x^*, x_0^*)^T$ and let Assumption 3.1 hold. Then $x_0^* = 1/(4c - 2b^T A^{-1} b)$ and the vector $v^* = (x^*/x_0^*)$ is the unique solution of the linear system $Ay = b$.*

Proof We first observe that, by definition the hyperplane $x_0 = 0$ is the polar of $(x^*, x_0^*)^T$ with respect to $\mathcal{F}$. Thus, by (4) we have that the hyperplane

$$\sum_{i=0}^{n} \left. \frac{\partial \varphi(x_1, \ldots, x_n, x_0)}{\partial x_i} \right|_{x_i = x_i^*} \cdot x_i = 0$$

must coincide with $x_0 = 0$, i.e.

$$\left. \frac{\partial \varphi(x_1, \ldots, x_n, x_0)}{\partial x_i} \right|_{x_i = x_i^*} = \begin{cases} 0 & \text{for } i = 1, \ldots, n, \\ 1 & \text{for } i = 0. \end{cases} \tag{11}$$

Recalling that by (8) we have $\varphi(x_1, \ldots, x_n, x_0) = x^T A x - 2x_0 b^T x + 2c x_0^2$, we obtain that (11) is equivalent to

$$\begin{pmatrix} 2A & -2b \\ -2b^T & 4c \end{pmatrix} \begin{pmatrix} x^* \\ x_0^* \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{12}$$

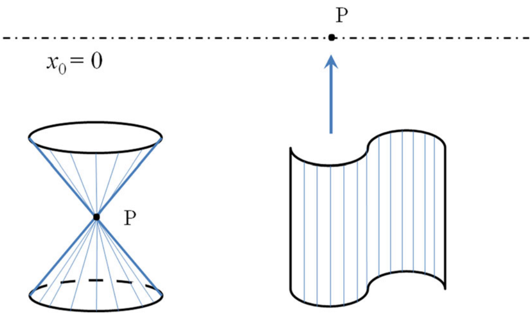


Fig. 4 Degenerate quadratic hypersurface $\mathcal{F}$ in (8) (*cone* or *cylinder*), when the coefficient matrix A in (12) has full rank n and exactly one double point P exists (i.e. $c = 1/2b^T A^{-1} b$)

By Assumption 3.1 the coefficient matrix in (12) is nonsingular. Hence, it admits a unique solution $(x^*, x_0^*)^T$ with $x_0^* = 1/(4c - 2b^T A^{-1} b)$ by Rouché-Capelli Theorem.

Finally, the first n equations of (12) imply $Ax^* = bx_0^*$, that is $\left(x^*/x_0^*\right) = A^{-1}b$. Hence, $v^* = x^*/x_0^*$ is the unique solution of $Ay = b$. □

Let us here briefly comment the hypotheses in Proposition 4.3. The non-singularity of matrix A implies that the coefficient matrix in (12) has rank greater than or equal to n. Hence, hypersurface $\mathcal{F}$ is irreducible [19], i.e. it contains at most a double point and cannot degenerate into a pair of hyperplanes. However, this hypothesis alone would not prevent $\mathcal{F}$ from being a cone or a cylinder (see Fig. 4). The additional hypothesis $c \neq 1/2b^T A^{-1} b$ in Assumption 3.1 implies $x_0^* \neq 0$, and hence ensures that the centre of $\mathcal{F}$ is not a point at infinity. This fact, along with the non-singularity of A guarantees that $\mathcal{F}$ can be only an ellipsoid or a hyperboloid or a paraboloid.

5 A Basis of Conjugate Directions: The Positive Definite Case

In this section, we report some additional results that hold as long as the matrix A in the hypersurface $\mathcal{F}$, defined in (8), is positive definite. In the spirit of this paper, the proofs of these results will rely on geometric properties suggested by polarity, instead of invoking algebraic arguments.

Proposition 5.1 (Existence of n conjugate lines, case $A \succ 0$) *Consider the quadratic hypersurface $\mathcal{F}$ and let Assumption 3.1 hold, with $A \succ 0$. Then, there exist n conjugate lines ℓ_j, $j = 1, \ldots, n$, with respect to $\mathcal{F}$. These lines are also linearly independent.*

Proof Note that by Assumption 3.1 and since $A \succ 0$, then $\mathcal{F}$ is a real hyperellipsoid which does not include points at infinity. Then, in the proof we consider a simplified expression of $\mathcal{F}$. Indeed, under the nonsingular linear transformation

$$\begin{pmatrix} x \\ x_0 \end{pmatrix} = \begin{pmatrix} I_n & A^{-1}b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{x}_0 \end{pmatrix}, \tag{13}$$

we can transform $\mathcal{F}$ into

$$\hat{\mathcal{F}} := \Big\{ (\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n \ : \\ (\hat{x} + A^{-1}b \cdot \hat{x}_0)^T A (\hat{x} + A^{-1}b \cdot \hat{x}_0) - 2\hat{x}_0 b^T (\hat{x} + A^{-1}b \cdot \hat{x}_0) + 2c\hat{x}_0^2 = 0 \Big\},$$

i.e. we obtain the simplified expression

$$\hat{\mathcal{F}} = \Big\{(\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n \ : \ \hat{f}(\hat{x}_1, \ldots, \hat{x}_n, \hat{x}_0) = 0, \\ \hat{f}(\hat{x}_1, \ldots, \hat{x}_n, \hat{x}_0) = \hat{x}^T A \hat{x} + (2c - b^T A^{-1} b)\hat{x}_0^2\Big\}, \tag{14}$$

whose centre $(\hat{x}^*, \hat{x}_0^*)^T$ is

$$\hat{x}^* = 0 \qquad \hat{x}_0^* = \frac{1}{4c - 2b^T A^{-1} b}. \tag{15}$$

Now we carry on the proof by induction and we first compute the n lines $\hat{\ell}_j$, with $j = 1, \ldots, n$. Then, by the inverse transformation of (13) we will obtain the conjugate lines ℓ_j, with $j = 1, \ldots, n$.

Base case Consider the line $\hat{\ell}_1$, with point at infinity $(\hat{d}_1, 0)^T$, and let $\hat{\pi}_1$ be the polar hyperplane of $(\hat{d}_1, 0)^T$ with respect to $\hat{\mathcal{F}}$. In case $n = 1$ the proof is over. Otherwise, we recursively define the remaining $n - 1$ conjugate lines $\hat{\ell}_j$, with $j = 2, \ldots, n$. To this end, the hyperplane $\hat{\pi}_1$ is an $(n - 1)$-dimensional hyperplane and includes the point $(0, \hat{x}_0)^T$, being $\hat{x}_0 \neq 0$.

Indeed, by (4) and since $(\hat{x}^*, \hat{x}_0^*)^T \equiv (0, 1/(4c - 2b^T A^{-1} b))^T$, we have

$$\hat{\pi}_1 \ := \ \left\{(\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : \sum_{i=0}^{n} \left.\frac{\partial \hat{f}(\hat{x}_1, \ldots, \hat{x}_n, \hat{x}_0)}{\partial \hat{x}_i}\right|_{(\hat{d}_1, 0)^T} \hat{x}_i = 0\right\},$$

which is equivalent to

$$\hat{\pi}_1 \ = \ \{(\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : \hat{d}_1^T A \hat{x} = 0\}. \tag{16}$$

Then, we show that $(\hat{d}_1, 0)^T \notin \hat{\pi}_1$. Since the (symmetric) matrix A is not singular, it easily follows from (15) that $(\hat{d}_1, 0)^T \notin \hat{\pi}_1$. Moreover, Proposition 4.2 implies that $\hat{\ell}_1$ is conjugate to any line contained in $\hat{\pi}_1$.

Induction step For $i = 1, \ldots, j - 1$, we denote by $\hat{\pi}_i$, the $(n - 1)$-dimensional polar hyperplane of $(\hat{d}_i, 0)^T$ with respect to $\hat{\mathcal{F}}$, i.e. $\hat{\pi}_i = \{(\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : \hat{d}_i^T A \hat{x} = 0\}$. Also, we denote by $\hat{\Pi}_{j-1}$ the following $[n - (j - 1)]$-dimensional hyperplane

$$\hat{\Pi}_{j-1} = \bigcap_{i=1}^{j-1} \hat{\pi}_i.$$

Then, we introduce $\hat{\ell}_j$ as an arbitrary line in $\hat{\Pi}_{j-1}$, being $(\hat{d}_j, 0)^T$ its point at infinity and $\hat{\pi}_j$ the polar hyperplane of $(\hat{d}_j, 0)^T$.

We observe that also $\hat{\pi}_j$ is a $(n-1)$-dimensional diametral hyperplane of $\hat{\mathcal{F}}$, since $(\hat{d}_j, 0)^T$ is a point at infinity. As a consequence, we can deduce that $(\hat{d}_j, 0)^T \notin \hat{\pi}_j$ by repeating for $\hat{\pi}_j$ the same reasoning which yielded (16). We also observe that the

Section Theorem 2.2 guarantees that the intersection $\hat{\pi}_j \cap \hat{\Pi}_{j-1}$ is a $(n-j)$-dimensional hyperplane which is conjugate to both $\hat{\ell}_j$ and all the lines $\hat{\ell}_i$, for $i = 1, \dots, j-1$. Finally, we note that the line $\hat{\ell}_j$ cannot be obtained as a linear combination of the lines contained in $\hat{\pi}_j$, since $(\hat{d}_j, 0)^T \notin \hat{\pi}_j$.

The arguments above prove that we can iterate the recursion step $n-1$ times, i.e. for $j = 2, 3 \dots, n$, so that $n - 1 = dim(\hat{\Pi}_1) > dim(\hat{\Pi}_2) > \cdots > dim(\hat{\Pi}_{n-1}) = 0$, and the conjugate lines $\hat{\ell}_1, \dots, \hat{\ell}_n$ are linearly independent. Thus, using backwards the transformation (13), we obtain n lines ℓ_j, $j = 1, \dots, n$, with points at infinity $\{(d_j, 0)^T\}$, satisfying

$$\begin{pmatrix} d_j \\ 0 \end{pmatrix} = \begin{pmatrix} I_n & A^{-1}b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{d}_j \\ 0 \end{pmatrix}, \qquad j = 1, \dots, n.$$

The latter equalities yield $d_j \equiv \hat{d}_j$, $j = 1, \dots, n$, which implies that also $\ell_1, \dots, \ell_n$ are conjugate and linearly independent. □

As a natural consequence, we can use the above proposition to easily provide an alternative proof of the following well-known result from the literature: the solution of the positive definite linear system $Ay = b$ can be computed starting from any point $\bar{y} \in \mathbb{R}^n$, provided that up to n conjugate directions are given.

6 A Basis of Conjugate Directions: The Indefinite Case

In this section, we extend the results presented in the previous section to the case of matrix A indefinite. These contents may appear a step away from the case, where the CG can be applied and is well-posed. However, as we already pointed out, in several optimization frameworks the CG is applied to solve indefinite linear systems accepting the risk of possible failures. In this context, we explain why an extension to the projective space of homogeneous coordinates is mandatory, in order to describe possible CG failures in the indefinite case. Indeed, we show that using finite points in $\mathbb{R}^n$, as those usually generated by the CG, does not allow to grasp CG possible degeneracy.

To this end, we first denote by $\mathcal{C}_\infty$ the intersection between the hyperplane at infinity $x_0 = 0$ and the hypersurface $\mathcal{F}$, i.e.

$$\mathcal{C}_\infty := \left\{ (x, 0)^T \in \mathbb{P}^n : x^T A x = 0 \right\}. \tag{17}$$

Then, we introduce the following definitions.

Definition 6.1 The asymptotic cone of the hypersurface $\mathcal{F}$, with finite centre, is the set of all the lines connecting the centre of $\mathcal{F}$ and any point of $\mathcal{C}_\infty$.

An immediate consequence of the above definition is that, if the centre $(x^*, x_0^*)^T$ of $\mathcal{F}$ is finite, in Cartesian coordinates the equation of the asymptotic cone of $\mathcal{F}$ is given by

$$\frac{1}{2}\left(y-\frac{x^*}{x_0^*}\right)^T A\left(y-\frac{x^*}{x_0^*}\right)=0. \tag{18}$$

Indeed, we note that Eq. (18) is homogeneous in $(y - x^*/x_0^*)$, so that it is a cone, and that the line $\ell = \{y \in \mathbb{R}^n : x^*/x_0^* + \lambda d,\ \lambda \in \mathbb{R}\}$ belongs to the asymptotic cone of $\mathcal{F}$ if and only if $d^T Ad = 0$.

Definition 6.2 A direction $d \in \mathbb{R}^n \backslash \{0\}$ is auto-conjugate with respect to the hypersurface $\mathcal{F}$ with finite centre, if the point at infinity $(d, 0)^T$ belongs to the asymptotic cone of $\mathcal{F}$, i.e. $d^T Ad = 0$.

The taxonomy in Definition 6.2 aims at avoiding confusion with *self-conjugate* points, introduced in Definition 2.2. A relation exists between the two concepts as, by Definition 2.2, the polar hyperplane π of the point $(d, 0)^T$, such that $d^T Ad = 0$, is tangent to $\mathcal{F}$ in $(d, 0)^T$, inasmuch as $(d, 0)^T$ also belongs to $\mathcal{F}$. Figure 5 (left) illustrates the latter fact: the dashed lines are auto-conjugate (with respect to $\mathcal{F}$) and are tangent to $\mathcal{F}$ in points at infinity. These points are self-conjugate with respect to $\mathcal{F}$.

Relation (18) and Definition 6.2 indicate that the hypersurface $\mathcal{C}_\infty$ is what really matters in order to identify the asymptotic cone of $\mathcal{F}$. In this section, we see that the possible degeneracy of the CG in the indefinite case only deals with the points (at infinity) in $\mathcal{C}_\infty$. In Sect. 7, we will prove that when the auto-conjugate direction $d \in \mathbb{R}^n \backslash \{0\}$ is computed by the CG in Cartesian coordinates, then equivalently the point $(d, 0)^T \in \mathbb{P}^n$ is indirectly generated in homogeneous coordinates. For this point at infinity, we will prove that the polar hyperplane is defined while no tangent hyperplane exists.

Definitions 6.1 and 6.2 imply that if $A \succ 0$ the quadratic hypersurface $\mathcal{F}$ has no real (as opposed to 'complex') auto-conjugate directions, since $\mathcal{C}_\infty = \emptyset$. On the other hand, if A is indefinite, auto-conjugate directions may exist. For example, Fig. 5(left) shows the projection of the hypersurface $\mathcal{F} = \{(x, x_0)^T \in \mathbb{P}^2 : 4x_1^2 - 3x_2^2 - x_0^2 = 0\}$ in the Cartesian space $\mathbb{R}^2$. Figure 5 (right) shows the corresponding intersection between $\mathcal{F}$ and the hyperplane at infinity, i.e. $\mathcal{C}_\infty = \{(x, 0)^T \in \mathbb{P}^2 : 4x_1^2 - 3x_2^2 = 0\}$, in the homogeneous coordinate projective space $\mathbb{P}^2$. In this latter context, $d_1 = (\sqrt{3}, 2)^T$ and $d_2 = (\sqrt{3}, -2)^T$ are the two auto-conjugate directions in $\mathbb{R}^2$.

In the next proposition, conditions such that $\mathcal{F}$ admits auto-conjugate directions are given, which is a relevant result when investigating CG-based methods in the indefinite case. We also remark that Assumption 3.1 is unnecessary to prove both the next proposition and the subsequent corollary.

Proposition 6.1 (Conjugate and auto-conjugate directions) *Consider the quadratic hypersurface $\mathcal{F}$ with finite centre and A indefinite. Then:*

1. *if A is nonsingular, the asymptotic cone of $\mathcal{F}$ cannot contain pairs of directions that are both conjugate and auto-conjugate;*
2. *if A is singular, the asymptotic cone of $\mathcal{F}$ may possibly contain an infinite number of directions that are both conjugate and auto-conjugate.*

Proof We preliminarily note that if two conjugate directions d_1 and d_2 are auto-conjugate with respect to $\mathcal{F}$, then any of their linear combination $\alpha d_1 + \beta d_2$ is also

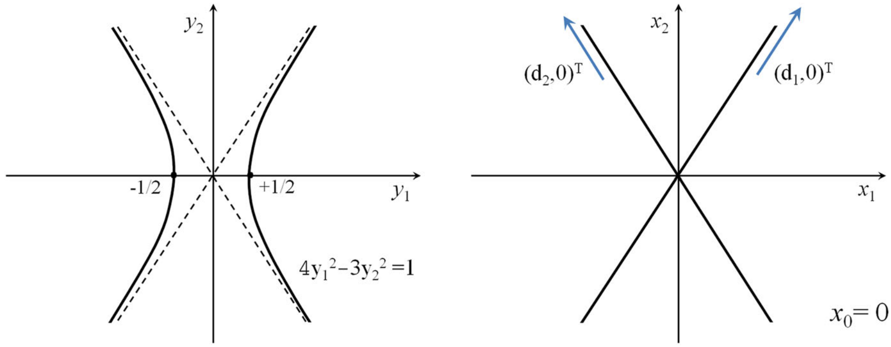


Fig. 5 (left) The quadratic hypersurface $\mathcal{F} = \{(x, x_0)^T \in \mathbb{P}^2 : 4x_1^2 - 3x_2^2 - x_0^2 = 0\}$ (in Cartesian coordinates), and (right) the intersection between $\mathcal{F}$ and the hyperplane at infinity, i.e. the hypersurface $\mathcal{C}_\infty = \{(x, 0)^T \in \mathbb{P}^2 : 4x_1^2 - 3x_2^2 = 0\}$ (in homogeneous coordinates)

auto-conjugate with respect to $\mathcal{F}$. Indeed, if $d_1^T A d_2 = 0$ and $d_1^T A d_1 = d_2^T A d_2 = 0$, then for any $\alpha, \beta \in \mathbb{R}$ we have

$$(\alpha d_1 + \beta d_2)^T A(\alpha d_1 + \beta d_2) = \alpha^2 d_1^T A d_1 + \beta^2 d_2^T A d_2 + 2\alpha\beta d_1^T A d_2 = 0. \tag{19}$$

The above fact implies that if the conjugate directions d_1 and d_2 are auto-conjugate with respect to $\mathcal{F}$, then $\mathcal{C}_\infty$ in (17) must also include the linear manifold $\text{span}\{d_1, d_2\}$. Using a similar argument, we can prove that, if *any* two conjugate directions d_1 and d_2 are auto-conjugate with respect to $\mathcal{F}$, then any pair of directions linear combinations of d_1 and d_2 are conjugate with respect to $\mathcal{F}$.

We also observe that $\mathcal{C}_\infty$ is the set of the auto-conjugate directions with respect to $\mathcal{F}$, and that $\mathcal{C}_\infty$ includes a linear manifold only if A is singular. This latter fact is trivially verifiable, as an example, by observing that if A is nonsingular, then $\mathcal{C}_\infty$ is a regular hypersurface and each point in $\mathcal{C}_\infty$ corresponds to a different gradient vector. On the contrary, when A is singular all the points $z = w + n$, where $w \in \mathcal{C}_\infty$ and $n \in Ker(A)$, belong to $\mathcal{C}_\infty$ and have the same gradient $2Aw$.

Given the above observation, we can derive those two cases considered in the problem statement: Case *1* A is nonsingular, so that $\mathcal{C}_\infty$ cannot include a hyperplane (i.e. $\mathcal{C}_\infty$ cannot degenerate into the product of two hyperplanes). Thus, if the directions d_1 and d_2 are auto-conjugate with respect to $\mathcal{F}$, then they cannot be conjugate as $\text{span}\{d_1, d_2\} \not\subseteq \mathcal{C}_\infty$.
Case *2* A is singular, $\mathcal{C}_\infty$ includes at least a linear manifold π, possibly of dimension 1. So, by definition of $\mathcal{C}_\infty$, given a pair of conjugate directions d_1 and d_2 on π, they are also auto-conjugate with respect to $\mathcal{F}$. Finally, if a pair of conjugate directions d_1 and d_2 on π exists, then any linear combination of d_1 and d_2 belongs to π by (19). □

An example of the latter result is provided by the hypersurface $\mathcal{F} = \{(x, x_0)^T \in \mathbb{P}^2 : 4x_1^2 - 3x_2^2 - x_0^2 = 0\}$ in Fig. 5, where

$$A = \begin{pmatrix} 4 & 0 \\ 0 & -3 \end{pmatrix}.$$

The two directions in $\mathbb{R}^2$ $d_1 = (\sqrt{3}, 2)^T$ and $d_2 = (\sqrt{3}, -2)^T$, though auto-conjugate, are not conjugate, being $d_1^T Ad_2 = 24 \neq 0$.

Though it is beyond the purposes of this paper, the following result can be used in order to justify that some CG-based methods from the literature (namely Planar-CG methods [4–7]) for indefinite linear systems are well-posed.

Corollary 6.1 *Consider the quadratic hypersurface $\mathcal{F}$ with finite centre and A indefinite nonsingular. Consider also two directions $d_1, d_2 \in \mathbb{R}^n$. If $(d_1, 0)^T$ and/or $(d_2, 0)^T$ belong to the asymptotic cone of $\mathcal{F}$ and d_1 is not conjugate to d_2, then*

$$(d_1^T Ad_1)(d_2^T Ad_2) - (d_1^T Ad_2)^2 \neq 0.$$

Proof The result follows immediately from Proposition 6.1 and Definitions 6.1, 6.2, as $(d_1^T Ad_1)(d_2^T Ad_2) = 0$ and $d_1^T Ad_2 \neq 0$. □

We conclude this section with a result similar to the one stated in Proposition 5.1, but referred to the case of matrix A indefinite and nonsingular. Once again, the proof is entirely based on geometric properties suggested by polarity.

Proposition 6.2 (Existence of $n - 1$ conjugate lines, case A indefinite) *Consider the quadratic hypersurface $\mathcal{F}$ with A indefinite and let Assumption 3.1 hold. Then, there exist n linearly independent lines ℓ_i, $i = 1, \dots, n$, such that at least $n - 1$ of these lines are also conjugate with respect to $\mathcal{F}$.*

Proof Preliminarily, we note that by Proposition 4.3 the centre $(x^*, x_0^*)^T$ of $\mathcal{F}$ is finite. Hence, similarly to Proposition 5.1, we can apply without loss of generality the linear transformation in (13) to $\mathcal{F}$, in order to obtain the simplified hypersurface $\hat{\mathcal{F}}$ in (14), with centre $(\hat{x}^*, \hat{x}_0^*)^T$ in (15). Then, we carry on the proof by induction, recursively defining the lines $\hat{\ell}_i$, $i = 1, \dots, n$. In the end, we compute the lines ℓ_i, $i = 1, \dots, n$, from $\hat{\ell}_i$, $i = 1, \dots, n$.

Base case Consider the line $\hat{\ell}_1$ with point at infinity $(\hat{d}_1, 0)^T$, and the corresponding polar hyperplane $\hat{\pi}_1$ (which includes the centre of $\hat{\mathcal{F}}$). In case $n = 1$ the proof is over. Otherwise, if $\hat{\ell}_1$ is not in the asymptotic cone of $\hat{\mathcal{F}}$ the proof proceeds by construction as the proof of Proposition 5.1, until possibly a line $\hat{\ell}_i$, $1 \le i \le n$, in the asymptotic cone of $\hat{\mathcal{F}}$ is detected. In case the latter line does not exist, then the directions $\hat{\ell}_1, \dots, \hat{\ell}_n$ are linearly independent and also conjugate. Finally, the proposition is proved using again the transformation (13) and retrieving the vectors $d_1, \dots, d_n$ as in Proposition 5.1, i.e. computing $\ell_1, \dots, \ell_n$.

Induction step Let $\hat{\ell}_i$ be a line in the asymptotic cone of $\hat{\mathcal{F}}$, being $(\hat{d}_i, 0)^T$ its point at infinity. The polar hyperplane $\hat{\pi}_i$ of $(\hat{d}_i, 0)^T$ is

$$\hat{\pi}_i := \left\{ (\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : (2A\hat{x})^T \hat{d}_i + 2(2c - b^T A^{-1} b)\hat{x}_0 \cdot 0 = 0 \right\},$$

i.e. [similarly to (16)]

$$\hat{\pi}_i := \left\{ (\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : \hat{d}_i^T A\hat{x} = 0 \right\}. \tag{20}$$

As $\hat{\ell}_i$ is in the asymptotic cone of $\hat{\mathcal{F}}$, then $\hat{d}_i^T A\hat{d}_i = 0$ and $(\hat{d}_i, 0)^T \in \hat{\pi}_i$. Let now $\hat{d}_{i+1} \in \mathbb{R}^n$ be any vector satisfying the following properties:

- $\hat{d}_i$ and $\hat{d}_{i+1}$ are linearly independent but not conjugate;
- $\hat{d}_{i+1}$ is conjugate to $\hat{d}_1, \ldots, \hat{d}_{i-1}$ (i.e. $\hat{d}_{i+1}$ is orthogonal to $A\hat{d}_1, \ldots, A\hat{d}_{i-1}$, being $\hat{d}_{i+1}^T A\hat{d}_j = 0$, $j = 1, \ldots, i-1$).

Recalling Proposition 6.1 and that $i < n$, we prove that such a vector surely exists and does not belong to $\hat{\pi}_i$ (since otherwise by (20) $\hat{d}_i$ and $\hat{d}_{i+1}$ would be conjugate). Indeed, it suffices to set

$$\hat{d}_{i+1} = \beta_{i,0} A\hat{d}_i + \sum_{j=1}^{i} \beta_{i,j}\hat{d}_j, \qquad \beta_{i,0} \neq 0, \tag{21}$$

observing that $A\hat{d}_i$ is orthogonal to $\hat{d}_i$ (see also the comment to Proposition 4.2), with $\hat{d}_j^T A\hat{d}_j \neq 0$, $j = 1, \ldots, i-1$, and compute $\{\beta_{i,j}\}$ such that

$$\hat{d}_{i+1}^T A\hat{d}_j = 0, \quad j = 1, \ldots, i-1, \qquad \text{i.e.} \quad \beta_{i,j} = -\beta_{i,0}\frac{(A\hat{d}_i)^T(A\hat{d}_j)}{\hat{d}_j^T A\hat{d}_j}. \tag{22}$$

The computation of $\beta_{i,j}$ in (22) is well-posed since $\hat{d}_1, \ldots, \hat{d}_{i-1}$ are not in the asymptotic cone of $\hat{\mathcal{F}}$. Thus, (21)–(22) yield the conjugacy of $\hat{d}_{i+1}$ with $\hat{d}_1, \ldots, \hat{d}_{i-1}$. Moreover, since $(A\hat{d}_i)^T\hat{d}_j = 0$, $j = 1, \ldots, i$, by (21) $\hat{d}_{i+1}^T A\hat{d}_i = \beta_{i,0}\|A\hat{d}_i\|^2 \neq 0$, proving that $\hat{d}_{i+1}$ and $\hat{d}_i$ are linearly independent but not conjugate.

The proof proceeds by following the guidelines of the proof of Proposition 5.1. We can thus compute the polar hyperplane $\hat{\pi}_{i+1}$ of the point $(\hat{d}_{i+1}, 0)^T$, with respect to $\hat{\mathcal{F}}$, which includes the point $(0, \hat{x}_0)^T$, being $\hat{x}_0 \neq 0$. Indeed, since $(\hat{x}^*, \hat{x}_0^*)^T \equiv (0, 1/(4c - 2b^T A^{-1}b))^T$ then $\hat{\pi}_{i+1} = \{(\hat{x}, \hat{x}_0)^T \in \mathbb{P}^n : \hat{d}_{i+1}^T A\hat{x} = 0\}$. By the *Section Theorem* 2.2 the intersection of $\hat{\pi}_{i+1}$ and $\hat{\pi}_i$ provides a $(n-2)$-dimensional hyperplane which is conjugate to $\hat{d}_1, \ldots, \hat{d}_{i-1}$. Finally, by this proposition hypotheses and Proposition 6.1, there cannot be two conjugate lines in the asymptotic cone of $\hat{\mathcal{F}}$. Thus, we can generate the remaining $\hat{d}_{i+2}, \hat{d}_{i+3}, \ldots, \hat{d}_n$ directions as in Proposition 5.1.

On the overall, the $n-1$ directions $\hat{d}_1, \ldots, \hat{d}_i, \hat{d}_{i+2}, \ldots, \hat{d}_n$ are conjugate by construction, and by (21)–(22) the n directions $\hat{d}_1, \ldots, \hat{d}_n$ are linearly independent. Finally, using again the nonsingular transformation (13), as for Proposition 5.1 we can obtain the lines ℓ_i, $i = 1, \ldots, n$, with points at infinity $\{(d_i, 0)^T\}$, such that $d_i = \hat{d}_i$, $i = 1, \ldots, n$, which completes the proof. □

Remark 6.1 Propositions 6.1 and 6.2 substantially ensure that even in case the matrix A is indefinite, though nonsingular, at least $(n-1)$ conjugate directions exist. Among these directions no more than one can be auto-conjugate. The latter simple conclusion is noteworthy, since one of the most used methods in the literature which generates conjugate directions (namely the CG method), may fail in the indefinite case, because it can iteratively yield auto-conjugate lines. The item 1. in Proposition 6.1 ensures that the latter fact can occur at most once during the execution of the CG method, even in case A is indefinite.

To better comment the last remark, the following considerations have been formulated. To the best of the authors' knowledge, the literature of the CG method based on the use of algebraic arguments has not given any results yet on the frequency of possible CG degeneracies in the nonsingular indefinite case. Such a fact is not surprising, since in the literature the main focus is often on the CG performance in practice, exploiting the current iteration rather than analysing the whole CG method within a general framework. On the contrary, Proposition 6.1 states that a geometric standpoint behind the CG indicates at most one possible degeneracy among all the iterations it can possibly perform. This also suggests another remarkable result, which will be more evident after introducing the CG in Table 1. Indeed, let the solution of a sequence of linear systems be sought (which is a typical problem from constrained numerical optimization), where the nonsingular matrix A remains unchanged and the recursion of the CG method is applied starting from a given (arbitrary) initial point $y_0 = \bar{y} \in \mathbb{R}^n$. Then, in case a degeneracy occurs when solving the h-th linear system of the sequence, we can slightly perturb the point $\bar{y}$ so that by Proposition 6.1 in none of the remaining linear systems the CG experiences a degeneracy.

The above remark also justifies the frequent use of the CG-based methods, on large indefinite linear systems, like those arising in optimization frameworks. In these cases, in order to provide gradient-related directions, a fast approximate solution of the linear system is often required. Here, CG-based methods may be preferred to more sophisticated and well-posed, but more computationally burdensome, techniques.

7 Polar and Tangent Hyperplanes: From Homogeneous to Cartesian Coordinates

In this section, we recast specific aspects of the standard CG method from the perspective of polarity theory. Specifically, we consider the case in which the CG method is used for solving the linear system $Ay = b$ in the Cartesian space $\mathbb{R}^n$, when $A \succ 0$. Our aim is to show that in this situation polarity theory gives an easy geometric interpretation of the steps of the CG. This interpretation turns out to be useful also in providing a general theoretical framework, when the matrix A is indefinite, and then the CG is not well-posed.

Throughout this section, as previously done we consider the linear system $Ay = b$ and the quadratic hypersurface $\mathcal{F}$ in $\mathbb{P}^n$ (see (8)), but also the associated quadratic functionals $g(y)$ in $\mathbb{R}^n$ and $f(x, x_0)$ in $\mathbb{R}^{n+1}$, as, respectively, defined in relations (6) and (7).

Table 1 reports the standard CG method. We can note that, when $A \succ 0$, the standard CG method solves the linear system $Ay = b$ by computing a sequence of points y_k, $k = 1, \ldots, m$, $m \le n$, along with a sequence of m linearly independent conjugate vectors $\{p_k\}$. The method iteratively generates these vectors by imposing that at each Step k the condition $p_k^T A p_j = 0$, for $j < k$, is satisfied. A similar result holds for the generalized class of CG-based methods proposed in [16].

The next three results show that the polar hyperplane of a point in homogeneous coordinates has an equivalent counterpart in Cartesian coordinates. In particular, note

Table 1 The standard CG method for solving the linear system $Ay = b$ [24]

The standard Conjugate Gradient method

Input: $y_0 \in \mathbb{R}^n$.

Step 0: Set $k = 0$, $r_0 = b - Ay_0$.

If $r_0 = 0$, then STOP. Else, set $p_0 = r_0$; $k = k + 1$.

Set $p_{-1} = 0$ and $\beta_{-1} = 0$.

Step k: Compute $\alpha_{k-1} = r_{k-1}^T p_{k-1} / p_{k-1}^T A p_{k-1}$,

$y_k = y_{k-1} + \alpha_{k-1} p_{k-1}$, $r_k = r_{k-1} - \alpha_{k-1} A p_{k-1}$.

If $r_k = 0$, then STOP. Else, set

- $\beta_{k-1} = \|r_k\|^2 / \|r_{k-1}\|^2$, $p_k = r_k + \beta_{k-1} p_{k-1}$, $k = k + 1$,
- (or equivalently set $p_k = -\alpha_{k-1} A p_{k-1} + (1 + \beta_{k-1}) p_{k-1} - \beta_{k-2} p_{k-2}$)

Go to **Step** k.

Output: $\{p_k\}$ and $\{y_k\}$.

that Lemma 7.1 and Corollary 7.1 refer to the polar hyperplane of a finite point, whereas Corollary 7.2 refers to the polar hyperplane of a point at infinity.

Lemma 7.1 (Equivalence with Polar Hyperplane 1) *Consider the quadratic hypersurface $\mathcal{F}$ in (8), the quadratic functional $g(y)$ in (6), and relation $y = x/x_0$ between Cartesian coordinates $y \in \mathbb{R}^n$ and homogeneous coordinates $(x, x_0)^T \in \mathbb{P}^n$. Let $(x^*, x_0^*)^T$ be the centre of $\mathcal{F}$. Then, the first polar π of a finite point $P = (\bar{x}, \bar{x}_0)^T \neq (x^*, x_0^*)^T$, $\bar{x}_0 \neq 0$, with respect to $\mathcal{F}$ has the following expression in Cartesian coordinates*

$$\pi := \left\{ y \in \mathbb{R}^n \ : \ 2g(\bar{y}) + \sum_{i=1}^{n} \frac{\partial g(\bar{y}_1, \cdots, \bar{y}_n)}{\partial y_i} (y_i - \bar{y}_i) = 0 \right\}. \tag{23}$$

Proof First note that $x_0 = 0$ cannot be the polar hyperplane of P, since $P \neq (x^*, x_0^*)^T$. Hence, hereinafter, together with $\bar{x}_0 \neq 0$, we can assume that $x_0 \neq 0$ and consider finite points.

The following relations hold given $y = x/x_0$ and the chain rule:

$$\begin{aligned} \frac{\partial f(x_1, \cdots, x_n, x_0)}{\partial x_i} &= \frac{\partial g(y_1, \cdots, y_n)}{\partial y_i} \cdot \frac{\partial y_i}{\partial x_i} \\ &= \frac{\partial g(y_1, \cdots, y_n)}{\partial y_i} \cdot \left(\frac{1}{x_0} \right), \quad i = 1, \ldots, n. \end{aligned} \tag{24}$$

Then, using Eq. (4) and relation

$$\frac{\partial f(x, x_0)}{\partial x_0} = -\nabla_x f(x, x_0)^T \left(\frac{x}{x_0}\right),$$

we can write the first polar of the point $(\bar{x}, \bar{x}_0)^T$ as:

$$\begin{aligned}
&\left\{(x, x_0)^T \in \mathbb{P}^n : \sum_{i=0}^{n} \frac{\partial \left[2 f(\bar{x}, \bar{x}_0)\bar{x}_0^2\right]}{\partial x_i} x_i = 0\right\} \\
&\equiv \Bigg\{(x, x_0)^T \in \mathbb{P}^n : \\
&\quad 2\left[\frac{\partial f(\bar{x}, \bar{x}_0)}{\partial x_0}\bar{x}_0^2 + f(\bar{x}, \bar{x}_0)\cdot 2\bar{x}_0\right] x_0 + 2\sum_{i=1}^{n}\frac{\partial f(\bar{x}, \bar{x}_0)}{\partial x_i}\bar{x}_0^2 x_i = 0\Bigg\} \\
&\equiv \Big\{(x, x_0)^T \in \mathbb{P}^n : \\
&\quad 2\left[-\nabla_x f(\bar{x}, \bar{x}_0)^T \bar{x}\bar{x}_0 + g(\bar{y})\cdot 2\bar{x}_0\right] x_0 + 2\bar{x}_0^2 \nabla_x f(\bar{x}, \bar{x}_0)^T x = 0\Big\},
\end{aligned}$$

so that dividing by $2\bar{x}_0 x_0 \neq 0$ and using (24) we obtain for π the expression

$$\left\{(x, x_0)^T \in \mathbb{P}^n : -\nabla g(\bar{y})^T \bar{y} + 2g(\bar{y}) + \nabla g(\bar{y})^T y = 0\right\}. \tag{25}$$

Finally, we note that (25) (in homogeneous coordinates) becomes (23) in Cartesian coordinates as $y = x/x_0$. □

The polar hyperplane (23) is a generalization of the tangent hyperplane to an algebraic hypersurface. The result is formally detailed in the next corollary.

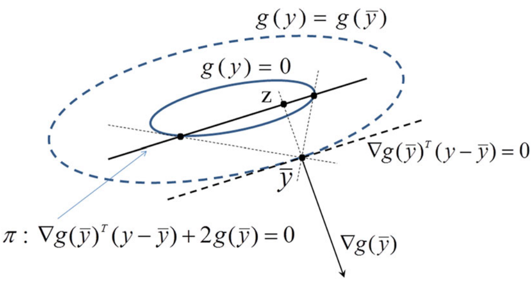
Corollary 7.1 *Consider the quadratic hypersurface $\mathcal{F}$ in (8), the quadratic functional $g(y)$ in (6), and relation $y = x/x_0$ between Cartesian coordinates $y \in \mathbb{R}^n$ and homogeneous coordinates $(x, x_0)^T \in \mathbb{P}^n$.*

1. *Let $(x^*, x_0^*)^T \in \mathbb{P}^n$ be the centre of $\mathcal{F}$. The first polar π in (23) of a finite point $P \equiv (\bar{x}, \bar{x}_0)^T \neq (x^*, x_0^*)^T$, when expressed in Cartesian coordinates, coincides with the tangent hyperplane to $g(y) = 0$ at $\bar{y} = \bar{x}/\bar{x}_0$, if and only if $g(\bar{y}) = 0$;*
2. *Let A be nonsingular and $c \neq 1/2b^T A^{-1}b$, then the set π in (23) is nonempty if and only if $A\bar{y} \neq b$;*
3. *Let A be nonsingular and $c = 1/2b^T A^{-1}b$, then the set π in (23) is nonempty.*

Proof As regards point *1.*, observe that if $g(\bar{y}) = 0$ then the tangent hyperplane to $g(y) = 0$ at $\bar{y}$ is unique, having the expression $\nabla g(\bar{y})^T(y - \bar{y}) = 0$. The latter equation coincides with (23) if and only if $g(\bar{y}) = 0$. As regards point *2.*, for the sufficient condition, by (6) and (23) we have

$$\pi := \left\{y \in \mathbb{R}^n : \bar{y}^T A\bar{y} - 2b^T\bar{y} + 2c + (A\bar{y} - b)^T(y - \bar{y}) = 0\right\}$$

Fig. 6 The geometry, in Cartesian coordinates, behind the polar hyperplane π in (23). $\bar{y}$ is the pole of π, while the point z is given by $z = \bar{y} - 2g(\bar{y})/\|\nabla g(\bar{y})\|^2 \nabla g(\bar{y})$



$$\equiv \left\{ y \in \mathbb{R}^n \ : \ (2c - b^T \bar{y}) + (A\bar{y} - b)^T y = 0 \right\}. \tag{26}$$

Thus, if $A\bar{y} \neq b$ then the point

$$y = \frac{b^T \bar{y} - 2c}{\|A\bar{y} - b\|^2}(A\bar{y} - b) \tag{27}$$

belongs to π, so that π is nonempty. Conversely, for the necessary condition, let π be nonempty and assume by contradiction $A\bar{y} = b$. Then, (26) would imply that the relation $0 = 2c - b^T \bar{y} = 2c - b^T A^{-1} b$ holds for any $y \in \mathbb{R}^n \setminus \{0\}$, in contradiction with the assumption $c \neq 1/2 b^T A^{-1} b$.

As regards point *3.*, we distinguish between the two cases: $A\bar{y} = b$ and $A\bar{y} \neq b$. In the first case $\pi \equiv \mathbb{R}^n$. In the second case, again the vector in (27) belongs to π. □

Figure 6 depicts (in Cartesian coordinates) the geometry behind the first polar π in (23). Here consider the hypersurface $g(y) = 0$, along with the tangent hyperplane $\nabla g(\bar{y})^T (y - \bar{y}) = 0$ to $g(y) = g(\bar{y})$ at $\bar{y}$. Then, observe that π is parallel to the latter hyperplane and contains the point $z = \bar{y} - 2g(\bar{y})/\|\nabla g(\bar{y})\|^2 \nabla g(\bar{y})$.

The next corollary defines the first polar of points at infinity, and will represent an essential tool to study the possible CG degeneracy in the indefinite case.

Corollary 7.2 (Equivalence with Polar Hyperplane 2) *Consider the quadratic hypersurface $\mathcal{F}$ in (8), the quadratic functional $g(y)$ in (6), and relation $y = x/x_0$ between Cartesian coordinates $y \in \mathbb{R}^n$ and homogeneous coordinates $(x, x_0)^T \in \mathbb{P}^n$. Let $(x^*, x_0^*)^T$ be the centre of $\mathcal{F}$. Then, the first polar π of the point at infinity $(\bar{p}, 0)^T$, $\bar{p} \in \mathbb{R}^n$, with respect to $\mathcal{F}$, has the following expression in Cartesian coordinates*

$$\pi \ := \left\{ y \in \mathbb{R}^n : \nabla g(y)^T \bar{p} = 0 \right\}. \tag{28}$$

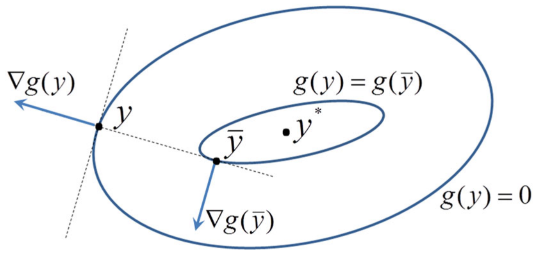
If A is nonsingular, then the hyperplane π is nonempty.

Proof The proof follows immediately from Definition 2.1 and (24). □

We now observe that, for $\bar{y} \in \mathbb{R}^n$ such that $g(\bar{y}) = 0$, the equation that describes the polar hyperplane π both in (23) and in (28) is equivalent to the Ritz–Galerkin condition

$$\nabla g(\bar{y})^T (y - \bar{y}) = 0, \tag{29}$$

Fig. 7 A graphical description in $\mathbb{R}^n$ of the general iteration of the CG, for $A \succ 0$ and $n = 2$



imposed by the standard CG-based methods at point $\bar{y} \in \mathbb{R}^n$ [2]. Indeed, in Table 1, at Step k, the coefficient α_{k-1} is chosen by imposing the Ritz–Galerkin condition $0 = r_k^T r_{k-1} = r_k^T p_{k-1}$, i.e. $\nabla g(y_k)^T(y_k - y_{k-1}) = 0$. Then, given β_k, the new conjugate direction p_k is determined using r_k and p_{k-1}. As an example with $n = 2$, consider Fig. 7. The ellipsoids in $\mathbb{R}^n$ represent level sets of the function $g(y)$ in (6). Starting from the point y the standard CG method finds the new point $\bar{y}$, by setting $\bar{\alpha}$ along the direction $-\nabla g(y)$ so that (29) is satisfied. Finally, note that (28) may easily reduce to (23) after setting $\bar{y} = \bar{x}/\bar{x}_0$ in (28), with $\bar{x} = \bar{p}$ and $\bar{x}_0 \to 0$.

8 CG-Based Methods and Polarity Theory: Advances

In this section, we further analyse the relation between polarity theory for the quadratic functional $g(y)$ in (6) and the CG method, in order to get some advances.

For the ease of notation and without loss of generality, in the current and in the following section we assume to have first performed the change of variables

$$y = \hat{y} - \tilde{y}, \qquad \tilde{y} = -A^{-1}b, \tag{30}$$

and hence (6) becomes

$$\hat{g}(\hat{y}) = \frac{1}{2}\hat{y}^T A\hat{y} + \left(c - \frac{1}{2}b^T A^{-1}b\right). \tag{31}$$

Then, in (30) we re-denominate $\hat{y}$ and $\hat{g}$ as y and g, respectively. This allows us to address the linear system $Ay = 0$, in place of (5).

As long as A is nonsingular, the linear transformation (30) leaves unchanged the Hessian matrix, and the quadratic functional $g(y)$ in (6) becomes

$$g(y) = \frac{1}{2}y^T Ay + \xi, \qquad \xi \in \mathbb{R}. \tag{32}$$

Then, if $A \succ 0$, $y^* = 0$ is trivially the optimal solution of the linear system $Ay = 0$ and coincides also with the ratio x^*/x_0^*, being $(x^*, x_0^*)^T \equiv (0, -1/(4\gamma))^T$ the centre of the family (see Proposition 4.3, with $c = -\gamma$ and $b = 0$) of quadratic hypersurfaces (ellipsoids)

$$\frac{1}{2}y^T Ay = \gamma, \qquad \gamma \geq 0. \tag{33}$$

In homogeneous coordinates, the hypersurfaces (33) are described by the following quadratic homogeneous functions

$$\mathcal{F}_\gamma := \left\{ (x, x_0)^T \in \mathbb{P}^n : \frac{1}{2} x^T A x - \gamma x_0^2 = 0 \right\}. \tag{34}$$

In this section:

- we prove that the standard CG method generates at any iteration k a hyperplane in $\mathbb{R}^n$ that is equivalent to a diametral hyperplane of $\mathcal{F}_\gamma$, for some $\gamma > 0$, in $\mathbb{P}^n$. In particular, we show that the resulting diametral hyperplane has the point at infinity $(p_k, 0)^T$ as a pole, being p_k the conjugate direction generated at the iteration k (see Table 1);
- we show that all the directions $\{p_k\}$, generated by the standard CG method in $\mathbb{R}^n$, are parallel to lines contained in polar hyperplanes with respect to different quadratic hypersurfaces $\mathcal{F}_\gamma$, for some $\gamma > 0$, in $\mathbb{P}^n$;
- we give evidence that the standard CG method generates at each iteration a pair hyperplane-point in $\mathbb{R}^n$. This pair hyperplane-point corresponds to a pair polar hyperplane-pole in $\mathbb{P}^n$, and the pole is a finite self-conjugate point as in Definition 2.2;
- we provide a geometric motivation for the fact that, if A is indefinite nonsingular in (34), the standard CG method may fail to provide at current iteration a diametral hyperplane of $\mathcal{F}_\gamma$, for some $\gamma > 0$, in $\mathbb{P}^n$.

Proposition 8.1 (CG-Polar Hyperplane 1) *Let the standard CG method perform m steps to solve the linear system $Ay = 0$, with $A \succ 0$. Then, for every $k < m$ the linear manifold*

$$y_{k+1} + span\{p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_m\}$$

represents in Cartesian coordinates a diametral hyperplane of the homogeneous hypersurface $\mathcal{F}_\gamma$ in (34), for any $\gamma > 0$. This diametral hyperplane is the polar hyperplane of the pole $(p_k, 0)^T$, with respect to $\mathcal{F}_\gamma$, and can be written as

$$\pi_{k+1} = \{y \in \mathbb{R}^n : (Ap_k)^T y = 0\}.$$

Proof We observe that, at the $(k+1)$-th iteration, the standard CG method detects the minimum point y_{k+1} of $g(y)$ along the line $y = y_k + \alpha p_k$, where p_k is a suitable search direction and $\alpha \in \mathbb{R}$ (see Table 1 or, for $n = 3$, Fig. 8). In particular, the method determines the steplength α_k by solving the unconstrained problem

$$\min_\alpha \; g(y_k + \alpha p_k).$$

Hence, we have $\nabla g(y_k + \alpha_k p_k)^T p_k = 0$, or equivalently

$$0 = \nabla g(y_{k+1})^T p_k = (Ay_{k+1})^T p_k. \tag{35}$$

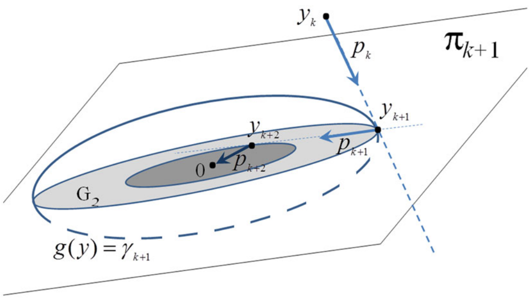


Fig. 8 A detail of the geometry of the $(k+1)$-th CG iteration: the hyperplane π_{k+1} (which includes the span of the directions p_{k+1} and p_{k+2}) is conjugate to direction p_k, with respect to $g(y) = \gamma_{k+1}$, $\gamma_{k+1} = 1/2 y_{k+1}^T A y_{k+1}$

On the other hand, for any $\gamma > 0$, the polar hyperplane π_{k+1} of $(p_k, 0)^T$ with respect to the hypersurface $\mathcal{F}_\gamma$ can be rewritten in Cartesian coordinates as

$$\pi_{k+1} := \{y \in \mathbb{R}^n \ : \ p_k^T A y = 0\}, \tag{36}$$

by Corollary 7.2. In fact, explicitly computing the polar (diametral) hyperplane of the point $(p_k, 0)^T$, with respect to $\mathcal{F}_\gamma$, we obtain

$$\begin{aligned} &\left\{(x, x_0)^T \in \mathbb{P}^n \ : \ (Ap_k)^T x + (-2\gamma \cdot 0)x_0 = 0\right\} \\ &\equiv \left\{(x, x_0)^T \in \mathbb{P}^n \ : \ (Ap_k)^T x = 0\right\}, \end{aligned}$$

which indeed yields (36) in Cartesian coordinates. Thus, by the *Reciprocity Theorem* the latter hyperplane contains the centre $(x^*, x_0^*)^T \equiv (0, -1/(4\gamma))^T$ of $\mathcal{F}_\gamma$. Hence, π_{k+1} in (36) contains $x^*/x_0^* = 0$ and is equivalently a diametral hyperplane (in Cartesian coordinates) of $\mathcal{F}_\gamma$. Furthermore, the ellipsoid $g(y) = \gamma$, for any $\gamma > 0$, intersects the hyperplane π_{k+1} into an $(n-2)$-dimensional ellipsoid (see also the contours of the shaded areas in Fig. 8, where G_2 represents an ellipse).

It remains to show that the manifold $y_{k+1} + \text{span}\{p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_m\}$ satisfies equation (36). Indeed, by simple substitution in (36) and recalling (35) we have

$$\begin{aligned} &(Ap_k)^T \left[y_{k+1} + \text{span}\{p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_m\}\right] \\ &= \ (Ap_k)^T \left[\text{span}\{p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_m\}\right] = 0, \end{aligned}$$

where the last equality follows from the conjugacy among the m vectors $\{p_1, \ldots, p_m\}$, generated by the CG. □

Proposition 8.2 (CG-Polar Hyperplane 2) *Let the standard CG method perform m steps to solve the linear system $Ay = 0$, with $A \succ 0$. Then, at Step k, $k < m$, the standard CG method generates a hyperplane in Cartesian coordinates equivalent to the polar hyperplane of the point y_k, with respect to the quadratic hypersurface $\mathcal{F}_{\gamma_k}$, $\gamma_k = 1/2 y_k^T A y_k$. This hyperplane has equation*

$$\tilde{\pi}_k := \left\{y \in \mathbb{R}^n \ : \ (Ay_k)^T (y - y_k) = 0\right\} \tag{37}$$

and contains the line $y_{k-1} + \alpha p_{k-1}$, $\alpha \in \mathbb{R}$.

Proof Let us consider Step k in Table 1 (for $n = 3$, see Fig. 9). From point y_{k-1}, the standard CG method moves along p_{k-1} and determines the new point y_k on a hyperplane $\tilde{\pi}_k$, which is tangent to $g(y) = 1/2 y_k^T A y_k$ at y_k. Thus, the hyperplane $\tilde{\pi}_k$ has equation in Cartesian coordinates

$$(A y_k)^T (y - y_k) = 0. \tag{38}$$

Since $y_k \in \tilde{\pi}_k$ is finite, Definition 2.2 guarantees that $\tilde{\pi}_k$ represents (in Cartesian coordinates) the polar hyperplane of the pole $(x_k, x_{0k})^T$, where $y_k = x_k / x_{0k}$, $x_{0k} \neq 0$, with respect to $\mathcal{F}_{\gamma_k}$, and $\gamma_k = 1/2 y_k^T A y_k$. Indeed, as

$$\mathcal{F}_{\gamma_k} := \left\{ (x, x_0)^T \in \mathbb{P}^n \ : \ \frac{1}{2} x^T A x - \gamma_k x_0^2 = 0, \quad \gamma_k = \frac{1}{2} y_k^T A y_k \right\}, \tag{39}$$

the polar hyperplane of the pole $(x_k, x_{0k})^T$ is given by

$$(A x_k)^T x - 2\gamma_k x_{k0} x_0 = 0. \tag{40}$$

As for $x_{0k} \neq 0$, $x_0 \neq 0$, $y = x/x_0$, then the equivalence between (38) and (40) holds. When $y = y_{k-1} + \alpha p_{k-1}$, $\alpha \in \mathbb{R}$, the equation (38) yields

$$(A y_k)^T p_{k-1} = 0, \tag{41}$$

showing that $\tilde{\pi}_k$ contains the line $y_{k-1} + \alpha p_{k-1}$, $\alpha \in \mathbb{R}$, which completes the proof. □

8.1 CG Iterations and CG Failure: A Geometric Viewpoint

Here we preliminarly show how the CG iterations can be rewritten in terms of polarity; then, we present a geometric motivation of the CG failure when A is indefinite. To this end, we recall that $y^* = 0$ is a solution of $Ay = 0$, by (30)–(31) and therein comments. Hence, unless the standard CG method has reached the very last iteration, the direction p_k computed at Step k has the extremes which are both nonzero.

In the following, we consider the polar hyperplane π_k (see also Fig. 9) of $(p_{k-1}, 0)^T$, with respect to $\mathcal{F}_{\gamma_k}$, both written in homogeneous coordinates

$$\begin{aligned} \pi_k &:= \left\{ (x, x_0)^T \in \mathbb{P}^n \ : \ (A p_{k-1})^T x - 2\gamma_k \cdot 0 \cdot x_0 = 0 \right\} \\ &\equiv \left\{ (x, x_0)^T \in \mathbb{P}^n \ : \ (A p_{k-1})^T x = 0 \right\}, \end{aligned}$$

and in Cartesian coordinates (see also Corollary 7.2)

$$\pi_k := \{ y \in \mathbb{R}^n \ : \ (A p_{k-1})^T y = 0 \}. \tag{42}$$

The representation of π_k in Cartesian coordinates allows one to conclude that π_k includes the origin. The representation of π_k in homogeneous coordinates allows to use Proposition 4.2, showing that π_k has the pole $(p_{k-1}, 0)^T$, with respect to $\mathcal{F}_{\gamma_k}$, and contains all the directions conjugate to the vector p_{k-1}. We also consider in homogeneous coordinates the manifold $V_d \subset \mathbb{P}^n$

$$V_d \;:=\; \left\{(x, x_0)^T \in \mathbb{P}^n \;:\; a^T x = 0, \;\; a \in \mathbb{R}^n \backslash \{0\}\right\},$$

and the quadratic hypersurface

$$\bar{\mathcal{F}} \;:= \mathcal{F}_{\gamma_k} \cap V_d = \begin{cases} x^T A x - 2\gamma_k x_0^2 = 0 \\ a^T x = 0. \end{cases}$$

We note that the polar hyperplane of $(x_k, x_{k0})^T \in \bar{\mathcal{F}}$ with respect to $\mathcal{F}_{\gamma_k}$ has equation $(Ax_k)^T x - 2\gamma_k x_{k0} x_0 = 0$, and the intersection ℓ between the latter hyperplane and the manifold V_d is given by

$$\ell \;:= \begin{cases} (Ax_k)^T x - 2\gamma_k x_{k0} x_0 = 0 \\ a^T x = 0. \end{cases}$$

The *Section Theorem* guarantees that ℓ also coincides with the polar hyperplane of $(x_k, x_{k0})^T$ with respect to $\bar{\mathcal{F}}$. Thus, the point $(x_k, x_{k0})^T \in \bar{\mathcal{F}}$ satisfies

$$\begin{array}{l} x_k^T A x_k - 2\gamma_k x_{k0}^2 = 0 \\ (Ax_k)^T x - 2\gamma_k x_{k0} x_0 = 0 \,\big|_{(x,x_0)^T = (x_k, x_{k0})^T} \\ a^T x_k = 0. \end{array} \tag{43}$$

In order to understand how the above concepts are of interest for the standard CG method, let us consider again Fig. 9. The hyperplane π_k contains the centre of $g(y) = \gamma_k$ and is therefore both a diametral hyperplane in $\mathbb{R}^n$ and a subspace. Then consider the linear manifold ℓ, i.e. the intersection between π_k (polar hyperplane of $(p_{k-1}, 0)^T$) and the tangent hyperplane $\tilde{\pi}_k$ to $g(y) = 1/2 y_k^T A y_k$ at $y_k = x_k / x_{k0}$. Now, ℓ is equivalent in $\mathbb{R}^n$ to the polar hyperplane of the point y_k with respect to the hypersurface Γ_k (contour of the shaded area in Fig. 9), which is the intersection between $g(y) = \gamma_k$ and π_k, i.e.

$$\Gamma_k \;:\; \begin{cases} g(y) = \gamma_k, & \text{where} \quad \gamma_k = \frac{1}{2} y_k^T A y_k, \\ y \in \pi_k. \end{cases} \tag{44}$$

The above observation implies that the k-th iteration of the standard CG method can be simply analysed in terms of the linear manifold ℓ and the $(n-1)$-dimensional hyperplane π_k (which is sketched also in Fig. 10). It also implies that

$$\pi_k \;:=\; y_k + \text{span}\{p_1, \ldots, p_{k-2}, p_k, \ldots, p_m\},$$

being $p_1, \ldots, p_{k-2}, p_k, \ldots, p_m$ conjugate to p_{k-1}. In particular, note that at Step k the standard CG method computes the residual r_k orthogonal to $\tilde{\pi}_k$ (and hence also

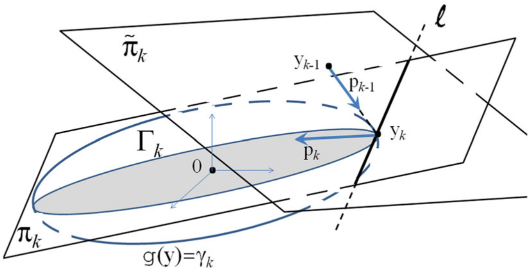


Fig. 9 An application (in Cartesian coordinates) of *Section Theorem*, within the k-th CG iteration: we have $y_{k-1} \in \tilde{\pi}_k$. The direction p_{k-1} is conjugate to the hyperplane π_k; $\tilde{\pi}_k$ represents in Cartesian coordinates the polar hyperplane of y_k with respect to $g(y) = \gamma_k$

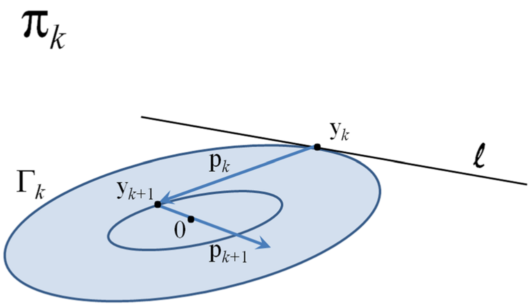


Fig. 10 The $(k + 1)$-th iteration of the CG reduces to analyse the $(n - 1)$-dimensional hyperplane π_k of Fig. 9

orthogonal to ℓ), and defines a vector p_k linearly independent from $p_1, \dots, p_{k-1}$ and belonging to π_k.

These facts allow to focus, at Step k of the standard CG method, only on π_k, Γ_k and ℓ. i.e. Step k of the standard CG method can be rewritten only in terms of π_k, Γ_k and ℓ, as in Fig. 10, without any explicit reference to directions $p_1, \dots, p_{k-1}$. In Fig. 10 one can see the conjugate directions p_k and p_{k+1} generated at both the k-th and the $(k + 1)$-th iteration.

As a general achievement, we can equivalently conclude that the $(k+1)$-th iteration of the standard CG method can be described by limiting the analysis to the hyperplane π_k in (42) and the hypersurface Γ_k in (44). This is partially summarized in the next corollary to Propositions 8.1 and 8.2.

Corollary 8.1 *Let the standard CG method perform m steps to solve the linear system $Ay = 0$, with $A \succ 0$. Then, at Step k, $k \leq m$, the polar hyperplane of $(p_{k-1}, 0)^T$, with respect to $g(y) = \gamma_k$, includes the point y_k, and is conjugate to p_{k-1}.*

Proof First observe that Proposition 8.1 implies that the polar hyperplane π_k of $(p_{k-1}, 0)^T$, with respect to $g(y) = \gamma_k$, satisfies $(Ap_{k-1})^T y = 0$. Recall that this last equation defines the locus of all the conjugate directions to p_{k-1}. Then, π_k coincides with $z + \mathrm{span}\{p_1, \dots, p_{k-2}, p_k, \dots, p_m\}$, where z is such that $(Ap_{k-1})^T z = 0$. Consequently, the proof follows immediately from the proof of Propositions 8.1 and 8.2, since (41) imposes

$$(Ap_{k-1})^T y_k = (Ay_k)^T p_{k-1} = 0,$$

which implies that $y_k \in \pi_k$. $\square$

The next proposition proposes a geometric insight of the standard CG method failure (also addressed as *pivot breakdown* in the literature [2]), in case the matrix A is indefinite. In particular, it turns out to be useful in explaining why the method may fail to generate a further search direction, when a line in the asymptotic cone (see Definition 6.1) of $\mathcal{F}_\gamma$ is detected.

Proposition 8.3 (CG Failure) *Let the standard CG method solve the linear system $Ay = 0$, with A indefinite nonsingular. Suppose that at Step k the CG computes the vector p_k satisfying $p_k^T A p_k = 0$, i.e. the point at infinity $(p_k, 0)^T$ belongs to the asymptotic cone of $\mathcal{F}_\gamma$ in (34), for some $\gamma > 0$. Then,*

(i) the point $(p_k, 0)^T$ belongs to its polar hyperplane with respect to $\mathcal{F}_\gamma$ and is self-conjugate with respect to $\mathcal{F}_\gamma$;
(ii) p_k belongs to the span of all the directions conjugate to p_k.

Proof We prove *(i)* by observing that the intersection between the asymptotic cone of $\mathcal{F}_\gamma$ in (34) and the hyperplane at infinity $x_0 = 0$ is given by $\{(x, 0)^T \in \mathbb{P}^n : x^T A x = 0\}$. Then, by the hypotheses, $(p_k, 0)^T$ is also a point of $\mathcal{F}_\gamma$. Proposition 2.2 and Definition 2.2 guarantee that $(p_k, 0)^T$ is self-conjugate with respect to $\mathcal{F}_\gamma$, and it satisfies the equation of its polar hyperplane $\pi_{k+1} = \{(x, x_0)^T \in \mathbb{P}^n : (Ap_k)^T x - 2\gamma \cdot 0 \cdot x_0 = 0\}$ since

$$(Ap_k)^T p_k = p_k^T A p_k = 0.$$

As regards *(ii)*, we observe that π_{k+1} is the locus of all the points $(d, 0)^T \in \mathbb{P}^n$ such that d and p_k are conjugate as in (4). Hence p_k is in the span of its conjugate directions. □

Let us remark that, if the point $(p_k, 0)^T$ satisfies the equation of the asymptotic cone of the quadratic hypersurface $\mathcal{F}_\gamma$, then nevertheless p_k might be appealing as a direction along which to search the centre of the hypersurface. Unfortunately, the standard CG method gets stuck if this direction is generated, since it is unable to compute a suitable finite steplength along it. In addition, we highlight that item *(i)* in the previous proposition does not imply also that the point $(x_k, x_{0k})^T$, with $y_k = x_k / x_{0k}$, is in the asymptotic cone of $\mathcal{F}_\gamma$. Indeed *(i)* in general implies the situation depicted in Fig. 11 (left), where $(p_k, 0)^T \in \mathcal{C}_\infty$, but the point y_k does not satisfy Eq. (18). Figure 11 (right) gives a graphical representation of item *(ii)* of Proposition 8.3. In the case that Step m provides a direction p_m in the asymptotic cone of $\mathcal{F}_\gamma$ (i.e. $p_m^T A p_m = 0$), then a direction p_{m+1} can not be generated since it would be parallel to p_m, unlike p_{h+1} or p_{k+1} at points (respectively) P_k and P_h.

Finally, Proposition 8.3 may have a dramatic impact in optimization frameworks. Indeed, we recall that whenever $A \succ 0$, the standard CG method iteratively computes *gradient-related* directions with respect to the functional $g(y)$. Specifically, at Step k, the following condition holds $\nabla g(y_k)^T p_k \le -\varepsilon_1 \|\nabla g(y_k)\|^2$ with $\varepsilon_1 > 0$, and $\|p_k\|$ is bounded as long as $\nabla g(y_k) \neq 0$. As a consequence, p_k can be used within linesearch procedures of Armijo-type or Wolfe-type (see for instance [9]), in order to guarantee global convergence properties. Differently, if A is indefinite and at Step k we have

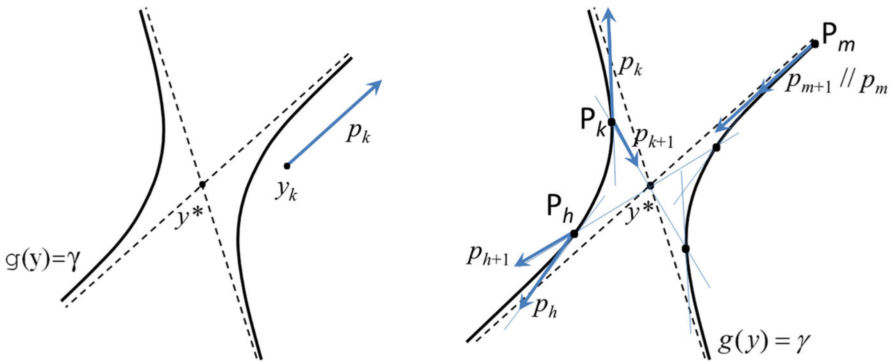


Fig. 11 A failure of the standard CG method when the matrix A is indefinite, in Cartesian coordinates. (left) The point $(p_k, 0)^T$ in homogeneous coordinates belongs to the asymptotic cone of $\mathcal{F}_\gamma$, for any γ. (right) When the current direction p_m generated at Step m of the CG approaches the asymptotic cone of $\mathcal{F}_\gamma$ (i.e. $p_m^T A p_m \approx 0$), for some γ, then the angle between p_m and p_{m+1} tends to zero

$p_k^T A p_k \approx 0$, then the standard CG method stops prematurely and p_{k+1} is no more generated, which claims for some adjustments in the optimization framework, in order to preserve convergence properties (see e.g. [11]).

9 Perspectives

In this brief section, we suggest items which can represent issues of interest for further research, where polarity may possibly play a key role. Since our analysis in the previous sections follows a different perspective, with respect to the current literature on CG-based methods, we conjecture that our hints here can give the opportunity for novel discussions and investigations on well-known topics.

- It may be worth exploring the possibility that the properties of polarity in Sects. 3-7 can lead to possible extensions of the Nonlinear Conjugate Gradient (NCG) method (see [9]). In this regard there is the chance that in the nonlinear case, not only the pairs polar hyperplane-point may play a keynote role, but also the so-called h-th polars [19] of points (with $h > 1$) may be helpful. The h-th polar of an algebraic hypersurface can capture in fact h-order information, which is unavailable in case the analysis includes only polar hyperplanes.
- Polarity perspective may also provide some contribution in the definition of possibly inexact linesearch procedures for NCG methods. Indeed, the results of Lemma 7.1 suggest that in Cartesian coordinates the polar hyperplane of a point not only includes first order information (i.e. the gradient of the function at the current iterate), but also information on the function value at that point. More specific linesearch procedures may rely on this latter information.
- The proof of Proposition 8.3 guarantees that if in (8) the matrix A is indefinite, then the polar hyperplane of a point at infinity P in the asymptotic cone is well defined. On the contrary, we cannot formally define at the latter point the tangent hyperplane in Cartesian coordinates. This suggests that a possible failure of the

CG, in case A is indefinite, might be possibly recovered by suitably alternating homogeneous coordinates and Cartesian coordinates in CG iterations. Indeed, by Proposition 8.3 a CG failure occurs in case at Step k we have $p_k^T A p_k = 0$, i.e. p_k is auto-conjugate with respect to the quadratic hypersurface. Thus, from the definition of asymptotic cone, p_k is in principle still a possible search direction, in order to detect the centre of the hypersurface, by computing a suitable steplength. Of course, since possibly y_k in Table 1 does not satisfy equation (18), in general in Cartesian coordinates the line $y_k + \lambda p_k$, $\lambda \in \mathbb{R}$, does not include the centre y^* of $\mathcal{F}$. Unfortunately, the CG is unable to compute a finite steplength along p_k (see also Fig. 11), so that it stops prematurely. Nevertheless, an *ad hoc* inexact linesearch procedure along p_k could be conceived.
- The case when the CG detects a *nearly* auto-conjugate direction p_k (i.e. such that $p_k^T A p_k \approx 0$ but $p_k^T A p_k \neq 0$) represents another intriguing scenario to theoretically investigate from a geometric standpoint. In fact, the latter case makes the CG well-posed but possibly numerically unstable (see e.g. [25]).
- Finally, it is worth considering the role of polarity theory also to study stationary points of cubic functions. They have recently gained the attention of the optimization community, due to their role within regularization problems (see the seminal paper [26]).

10 Conclusions

In this paper, we have investigated the role of polarity in homogeneous coordinates for quadratic hypersurfaces, in order to provide general tools for CG-based methods, from a geometric perspective. We have presented both the analytical properties and the geometric insight revealed by polarity, justifying the fact that in CG-based methods n-dimensional vectors are typically introduced, without recurring to homogeneous coordinates. Our use of polarity in homogeneous coordinates has not required the quadratic hypersurface in (8) to be an hyperellipsoid (i.e. $A \succ 0$). This allowed us to prove additional general results, to treat indefinite linear systems. Moreover, we showed to what extent the premature stop of the CG, in the indefinite case, is likely to be a very rare event.

Acknowledgements The authors are indebted with the Editorial Board and the reviewers, for their constructive and valuable comments. The work of G. Fasano is partially supported by the Italian Flagship Project RITMARE, coordinated by the Italian National Research Council (CNR) and funded by the Italian Ministry of Education, within the National Research Program 2012–2016.

References

1. Hestenes, M., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**, 409–435 (1952)
2. Stoer, J.: Solution of large linear systems of equations by conjugate gradient type methods. In: Bachem, A., Grötschel, M., Korte, B. (eds.) Mathematical Programming. The State of the Art, pp. 540–565. Springer, Berlin (1983)
3. Nazareth, L.: A relationship between the BFGS and conjugate gradient algorithms and its implications for new algorithms. SIAM J. Numer. Anal. **16**, 794–800 (1979)

4. Hestenes, M.: Conjugate direction methods in optimization. Springer, New York (1980)
5. Fasano, G.: Conjugate gradient (CG)-type method for the solution of Newton's equation within optimization frameworks. Optim. Methods Softw. **19**, 267–290 (2004)
6. Fasano, G.: Planar-conjugate gradient algorithm for large scale unconstrained optimization, part 2: application. J. Optim. Theory Appl. **125**, 543–558 (2005)
7. Luenberger, D.: Hyperbolic pairs in the method of conjugate gradients. SIAM J. Appl. Math. **17**, 1263–1267 (1969)
8. Lakshmikantham, V., Trigiante, D.: Theory of Difference Equations: Numerical Methods and Applications. Marcel Dekker Inc, New York (2002)
9. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (2000)
10. Nash, S.: A survey of truncated-Newton methods. J. Comput. Appl. Math. **124**, 45–59 (2000)
11. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. SIAM J. Numer. Anal. **19**, 400–408 (1982)
12. Grippo, L., Lampariello, F., Lucidi, S.: A truncated Newton method with nonmonotone line search for unconstrained optimization. J. Optim. Theory Appl. **60**, 401–419 (1989)
13. Fasano, G., Roma, M.: Iterative computation of negative curvature directions in large scale optimization. Comput. Optim. Appl. **38**, 81–104 (2007)
14. Gould, N.I.M., Lucidi, S., Roma, M., Toint, PhL: Exploiting negative curvature directions in linesearch methods for unconstrained optimization. Optim. Methods Softw. **14**, 75–98 (2000)
15. Fasano, G., Lucidi, S.: A nonmonotone truncated Newton–Krylov method exploiting negative curvature directions, for large scale unconstrained optimization. Optim. Lett. **3**, 521–535 (2009)
16. Fasano, G.: A framework of conjugate direction methods for symmetric linear systems in optimization. J. Optim. Theory Appl. **164**, 883–914 (2015)
17. Gratton, S., Sartenaer, A., Tshimanga, J.: On a class of limited memory preconditioners for large scale linear systems with multiple right-hand sides. SIAM J. Optim. **21**, 912–935 (2011)
18. Fasano, G., Roma, M.: Preconditioning Newton–Krylov methods in non-convex large scale optimization. Comput. Optim. Appl. **56**, 253–290 (2013)
19. Beltrametti, M.C., Carletti, E., Gallarati, D., Monti Bragadin, G.: Lectures on Curves. European Mathematical Society, Surfaces and Projective Varieties: A Classical View of Algebraic Geometry (2009)
20. Casey, J.: A Treatise on the Analytical Geometry of the Point, Line, Circle and Conic Sections. Dublin University Press, Dublin (1885)
21. Schreirer, O., Sperner, E.: Projective Geometry of n Dimensions. Chelsea Publishing Company, New York (1961)
22. Seidenberg, A.: Lectures in Projective Geometry. Dover Publications Inc, Mineola (2005)
23. Akivis, M.A., Goldberg, V.V.: Differential Geometry of Varieties with Degenerate Gauss Maps. Springer, New York (2004)
24. Higham, N.J.: Accuracy and Stability of Numerical Algorithms. SIAM, Philadelphia (1996)
25. Greenbaum, A.: Iterative Methods for Solving Linear Systems. SIAM, Philadelphia (1997)
26. Nesterov, Y., Polyak, B.T.: Cubic regularization of Newton method and its performance. Math. Program. Ser. A **108**, 177–205 (2006)