Perspective-1-Ellipsoid: Formulation, Analysis and Solutions of the Ellipsoid Pose Estimation Problem in Euclidean Space

Vincent Gaudillièrè\textsuperscript{1,2}*, Gilles Simon\textsuperscript{2} and Marie-Odile Berger\textsuperscript{2}

\textsuperscript{1}SnT - Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, 29 Avenue John F. Kennedy, L-1855, Luxembourg, Luxembourg.
\textsuperscript{2}Loria - Inria - Université de Lorraine, Campus Scientifique, BP 239, F-54506, Vandoeuvre-lès-Nancy, France.

*Corresponding author(s). E-mail(s): vincent.gaudilliere@uni.lu; gilles.simon@loria.fr; marie-odile.berger@inria.fr;

Abstract

In computer vision, camera pose estimation from correspondences between 3D geometric entities and their projections into the image has been a widely investigated problem. Although most state-of-the-art methods exploit low-level primitives such as points or lines, the emergence of very effective CNN-based object detectors in the recent years has paved the way to the use of higher-level features carrying semantically meaningful information. Pioneering works in that direction have shown that modelling 3D objects by ellipsoids and 2D detections by ellipses offers a convenient manner to link 2D and 3D data. However, the mathematical formalism most often used in the related literature does not enable to easily distinguish ellipsoids and ellipses from other quadrics and conics, leading to a loss of specificity potentially detrimental in some developments. Moreover, the linearization process of the projection equation creates an over-representation of the camera parameters, also possibly causing an efficiency loss. In this paper, we therefore introduce an ellipsoid-specific theoretical framework and demonstrate its beneficial properties in the context of pose estimation. More precisely, we first show that the proposed formalism enables to reduce the ellipsoid pose estimation problem to a position or orientation-only estimation problem in which the remaining unknowns can be derived in closed-form. Then, we demonstrate that it can be further reduced to a 1 Degree-of-Freedom (1DoF) problem and provide the analytical expression of the pose as a function of that unique scalar unknown. We illustrate our theoretical considerations by visual examples. Finally, we release this work in order to contribute towards more efficient resolutions of ellipsoid-related pose estimation problems.

Keywords: Pose estimation, Object modeling, Ellipsoid, Ellipse

1 Introduction

Estimating the relative pose between a camera and a scene has been representing a core aspect of computer vision for many years. Indeed, this task is at the root of many applications, from robot navigation \cite{Bonin-Font et al., 2008} to Augmented Reality \cite{Marchand et al., 2016}.

Historically, pose estimation has been addressed by leveraging 2D-3D correspondences between low-level geometric features such as points (PnP: Perspective-n-Point) \cite{Lepetit et al., 2009, Hartley and Zisserman, 2004} or lines...
More recently, the field has been significantly impacted by the raise of deep learning and pose estimation is now widely addressed by end-to-end trainable methods (Hoque et al., 2021). However, while deep learning has proven to be indispensable in solving the problem of perception, it is still not the best choice in terms of accuracy throughout all steps of a pose estimation pipeline, as figured out in a recent object pose estimation challenge (Kisantal et al., 2020). Indeed, in this challenge, the two most accurate methods were hybrid approaches in which keypoints are located by deep regression models then used as inputs to a PnP-based solver.

Following the same proven hybrid approach, the appearance of very effective object detectors in the recent years (Redmon et al., 2016; Redmon and Farhadi, 2017; Liu et al., 2016; Girshick et al., 2014; Girshick, 2015) has been enabling the substitution of low-level primitives (e.g. points, lines), often extracted in droves and carrying limited semantic information, by object-level features providing a deeper scene understanding at a lower computational matching cost. Therefore, the choice of object representation has become crucial.

While modeling 3D objects by cuboids along with their 2D projections by bounding boxes (i.e. outputs of most object detectors) has been investigated in the context of pose computation (Li et al., 2017, 2018, 2019), it appears that the ellipse-ellipsoid modeling paradigm has the unparalleled advantage of analytically linking 2D and 3D models (Hartley and Zisserman, 2004; Eberly, 2007). In other words, ellipsoids always project onto planes in the form of ellipses, and the underlying closed-form projection equation can be leveraged to efficiently solve pose estimation problems (Crocco et al., 2016; Rubino et al., 2018; Gaudilli`ere et al., 2019a; Liao et al., 2020). As an indicator of that increasingly attractive research direction, more and more object detectors have been modeling object projections by ellipses instead of traditional bounding boxes whose sides are parallel to image axes (Li, 2019; Pan et al., 2021; Zhao et al., 2021; Zins et al., 2020; Dong et al., 2021; Dong and Isler, 2021).

However, performing pose estimation at the level of objects through ellipse/ellipsoid-modeling has mainly been formulated under the standard projective geometry formalism (Hartley and Zisserman, 2004) (Crocco et al., 2016; Gay et al., 2017; Rubino et al., 2018; Nicholson et al., 2019; Garg et al., 2021; Liao et al., 2020; Tian et al., 2021), and mostly through least square estimations where the unknowns are general quadric surfaces. This framework may present limitations since ellipses (resp. ellipsoids) are specific categories of conics (quadrics) and since the linearization of the projection equation increases the numbers of apparent unknowns (see Section 2.1).

In addition, these papers do not address the case of a small number of ellipse-ellipsoid correspondences, which is of high practical importance when a few objects are observed or when computing the solutions with minimal sampling size in RANSAC algorithms (Fischler and Bolles, 1981).

In this paper, we address the fundamental problem of camera pose estimation from one ellipse-ellipsoid correspondence, referred to as Perspective-1-Ellipsoid (PIE) in what follows. Conversely, it consists in reconstructing an ellipsoid of known size from its projected ellipse given the camera intrinsic parameters (ellipsoid pose estimation).

There are several interests in solving the PIE problem. First, on the theoretical side, and except in the case of a spherical object, we demonstrate that the solutions are a variety of dimension 1 and we provide an effective way to reconstruct the camera trajectory (i.e. solutions). This problem has been addressed in Wokes and Palmer (2010) in the particular case of a spheroid (specific ellipsoid having an axis of revolution) but was never considered for general ellipsoids. To the best of our knowledge, we are the first to propose a constructive solution to the PIE problem without resorting to any additional approximation nor prior knowledge.

The proposed formalism relies on Cartesian coordinates instead of homogeneous ones (Hartley and Zisserman, 2004), and this enables us to develop an ellipsoid-specific framework. In this study, we also consider two particular cases of important practical interest: (i) computing the ellipsoid position when the orientation is known (ii) computing the orientation when the position is known.
Besides the theoretical aspects, solving the P1E problem opens the way towards automatic positioning solutions in texture-less or low-textured environments, for instance leveraging several ellipse-ellipsoid correspondences or one with several point pairs. Industrial or other indoor scenes, in which objects would be approximated by ellipsoids, represent concrete places that could take advantage of these results.

The paper is organized as follows: In Section 2.1, we discuss the current State-of-the-Art object-based pose estimation methods and the limits of the homogeneous representation of ellipses and ellipsoids. In Section 3, we present the Euclidean formulation of the ellipsoid pose estimation problem, previously introduced in Eberly (2007). Sections 4, 5 and 6 contain our core contributions:

- In Section 4, we exhibit several mathematical properties of the P1E solutions needed in the demonstrations of sections 5 and 6.
- Section 5 is dedicated to the specific case of P1E where either the position or the orientation of the ellipsoid is known. We bring out that the problem formulation induces an inherent decoupling between orientation and position, one of which being possibly inferred in closed-form from the other one. The orientation to position solver was introduced in Gaudillièr et al. (2019a,b), then leveraged in Gaudillièr et al. (2020). In this method, nothing ensures that the reconstructed quadric is an ellipsoid, forcing the authors to add a costly post-processing non-linear optimization of the results.
- The general P1E problem is solved in Section 6. We demonstrate that the 6DoF ellipsoid pose estimation problem can be reduced to a 1DoF problem, and present an effective way to reconstruct the solutions.

2 Related Work

2.1 Quadrics-based Pose Estimation

Most methods proposing to solve pose estimation problems at the level of objects using ellipsoid modeling are based on the projective geometry formalism. Any quadric \(Q\) is thus linked to its projected conic \(C\) by the Equation

\[
C^* = PQ^*P^T
\]

where \(P = K[R|t]\) is the camera projection matrix (Hartley and Zisserman, 2004) and \(Q^*, C^*\) are the dual forms of \(Q\) and \(C\) (\(M^* = M^{-1}\) up to a scale factor for a non-singular symmetric matrix \(M\)). It is important noting that, under that formalism, ellipses and ellipsoids can be distinguished from other conics and quadrics only by certain algebraic conditions on \(C\) and \(Q\) entries, these conditions being difficult to leverage in practice.

Quadric-modeling of objects has often been implemented in the context of Semantic SLAM (Garg et al., 2020) for improving the process accuracy through multi-objective optimization (Nicholson et al., 2019; Tian et al., 2021; Liao et al., 2020). On a theoretical level, Crocco et al. (2016) addresses the object-based Structure-from-Motion (SfM) problem and introduces an analytical solution to reconstruct both quadric and affine camera poses. The problem is solved in a least square sense with a matrix \(P\) over-represented by a \(6 \times 10\) Kronecker product \(P \otimes P\). This work is extended with CAD model priors in Gay et al. (2017), while Rubino et al. (2018) present a closed-form solution to the problem of reconstructing a quadric from three calibrated pinhole camera views in which the object projections are detected.

However, in this method, nothing ensures that the reconstructed quadric is an ellipsoid, forcing the authors to add a costly post-processing non-linear optimization of the results.

Another limitation while using homogeneous quadrics and over-representation of \(P\) appears in Gaudillièr et al. (2019c). In this method, the so-called gold-standard algorithm used to retrieve the camera projection matrix from 2D-3D point correspondences (Hartley and Zisserman, 2004) is adapted to conic-quadric correspondences. To compute \(P\), 12 conic-quadric pairs are required whereas only 6 point pairs are sufficient in the same context.

We argue that these limitations may be due in part to the fact that ellipse and ellipsoid homogeneous formulations are not clearly enough distinguished from other members of their geometric families, and also to a non-minimal representation of the projection matrix. In our paper and to overcome these difficulties, we thus propose an ellipsoid-specific theoretical framework and highlight its advantages.

2.2 Perspective-1-Spheroid

In Wokes and Palmer (2010), a comprehensive study of the spheroid pose estimation problem...
is conducted. To this end, the authors introduce a spheroid-specific parameterization of the problem that enables solving it but prevent from extrapolating the method to the general case of ellipsoids.

In a nutshell, the authors demonstrates that the spheroid pose estimation problem has two distinct solutions. In 6.4, we retrieve the same result by restricting our general formulation to the case of spheroids.

3 Formulation of the Ellipsoid Pose Estimation Problem

3.1 Problem Statement

Following the notations introduced in Eberly (2007) and presented in Fig. 1, we consider an ellipsoid \((C, A)\) defined by Equation

\[
(X - C)^\top A (X - C) = 1,
\]

where \(C\) is the center of the ellipsoid, \(A\) is a real positive definite matrix characterizing its orientation and size, and \(X\) is any point on it.

Given a center of projection \(E\) and a projection plane of normal \(N\) which does not contain \(E\), the projection of the ellipsoid is an ellipse of center \(K\) and of semi-diameters \(a\) et \(b\). Its principal directions are represented by unit vectors \(U\) and \(V\), such that \(\{U, V, N\}\) is an orthonormal set.

3.1.1 Projection Cone

The projection cone \((E, B)\) refers to the cone of vertex \(E\) tangent to the ellipsoid. According to Eberly (2007), it is characterized by matrix

\[
B \overset{\text{def}}{=} A\Delta\Delta^\top A - (\Delta^\top A\Delta - 1)A,
\]

where \(\Delta = E - C\), so that the points \(X\) belonging to the projection cone are those who satisfy the Equation \((X - E)^\top B(X - E) = 0\). Note that \(B\) is a real, symmetric and invertible matrix which has two eigenvalues of the same sign and the third one of the opposite sign.

3.1.2 Backprojection Cone

The backprojection cone \((E, B')\) refers to the cone generated by the lines passing through \(E\) and any point on the ellipse. Eberly shows that it is characterized by matrix

\[
B' \overset{\text{def}}{=} P^\top MP - Q,
\]

where

\[
M \overset{\text{def}}{=} \frac{UU^\top}{a^2} + \frac{VV^\top}{b^2},
\]

\[
W \overset{\text{def}}{=} \frac{N}{(N \cdot (K - E))},
\]

\[
P \overset{\text{def}}{=} I - (K - E)W^\top,
\]

\[
Q \overset{\text{def}}{=} WW^\top.
\]

Here again, the points \(X\) on the backprojection cone are those who meet \((X - E)^\top B'(X - E) = 0\), while \(B'\) shares \(B\) properties (real, symmetric, invertible with signature (2,1) or (1,2)).

3.1.3 The Cone Alignment Equation

Given an ellipsoid, a central projection (center and plane) and an ellipse in the projection plane, the ellipse is the projection of the ellipsoid if and only if the projection and backprojection cones are aligned (Eberly, 2007), i.e. if and only if there is a non-zero scalar \(\sigma\) such that \(B = \sigma B'\):

\[
A\Delta\Delta^\top A + \mu A = \sigma B',
\]

where \(\mu = 1 - \Delta^\top A\Delta\).

Note that \(\mu\) encapsulates the relative configuration between the camera center \(E\) and the
ellipsoid center \( C \). It is thus negative if \( E \) is outside the ellipsoid and positive otherwise.

Equation (11), to whom we refer as the Cone Alignment Equation, encodes the ellipsoid pose estimation problem. An equivalent formulation is given by Equation (11) (see proof of equivalence in Appendix A):

\[
\Delta \Delta^\top = A^{-1} - \frac{\mu}{\sigma} B'^{-1}.
\]

### 3.2 Pose Problem Analysis

While Equation (1) has been established in the camera coordinate frame, it is also valid in the ellipsoid coordinate frame, where matrix \( A_{\text{ell}} \) is diagonal, and in the cone coordinate frame where \( P' \) is diagonal. Since \( A \) and \( B' \) are symmetric, both can be diagonalized using an orthogonal matrix \( P \) such that \( P^{-1} = P^\top \). For that reason, Equation (1) remains the same whatever the choice of the coordinate frame (camera, ellipsoid or cone) in which matrices and vectors are expressed. In the following, if there is no restriction on the coordinate frame, we will adopt notations without subscript. Otherwise, subscripts \( \text{cam, cone or ell} \) will be used.

One can theoretically distinguish between camera pose estimation, which consists in estimating the pose of the camera with respect to its environment, and its reference frame counterpart, i.e. object pose estimation, which consists in estimating the pose of an object with respect to the camera. Fundamentally, the sought transformations are the inverse of each other thus, in this paper, we may focus on estimating the pose of the ellipsoid or the pose of the camera, according to the most convenient setup for mathematical developments, without loss of generality.

The givens of the problem are the ellipse detected in the image, the camera intrinsic parameters and the ellipsoid size (i.e. lengths of its three radii). Therefore, \( B'_{\text{cone}}, B'_{\text{cam}} = \text{cam} R_{\text{cone}} B'_{\text{cone}} \text{cam} R_{\text{cone}}^\top \) and \( A_{\text{ell}} \) are known, as well as \( E_{\text{cone}}, E_{\text{cam}} \) and \( C_{\text{ell}} \) (whose entries are all zero). Since expressions of \( A \) and \( B' \) in specific coordinate frames are known, their eigenvalues are also known. In addition, matrix properties such as trace and determinant, that are fully constrained by the eigenvalues, are also known.

In the paper, we most often work in the camera frame. We thus aim at computing vector \( \Delta_{\text{cam}} = E_{\text{cam}} - C_{\text{cam}} \) and matrix

\[
A_{\text{cam}} = \text{cam} R_{\text{ell}} A_{\text{ell}} \text{cam} R_{\text{ell}}^\top = \text{cam} R_{\text{ell}} \begin{pmatrix} 1/a^2 & 0 & 0 \\ 0 & 1/b^2 & 0 \\ 0 & 0 & 1/c^2 \end{pmatrix} \text{cam} R_{\text{ell}}^\top,
\]

from which we then retrieve the ellipsoid position \( C_{\text{cam}} \) and orientation \( \text{cam} R_{\text{ell}} \).

In the other case, we aim at computing \( \Delta_{\text{ell}} \) and \( B'_{\text{ell}} \) to then derive the camera position \( E_{\text{ell}} \) and orientation \( \text{cam} R_{\text{ell}}^\top \).

Finally, vector \( \Delta \) encodes the relative position between the ellipsoid and the camera, while couple \( \{A, B'\} \) characterizes their relative orientation, denoted \( \text{cam} R_{\text{ell}} \). Indeed, its expression in the camera frame is

\[
\{ \text{cam} R_{\text{ell}} A_{\text{ell}} \text{cam} R_{\text{ell}}^\top, B'_{\text{cam}} \},
\]

while its expression in the ellipsoid frame is

\[
\{ A_{\text{ell}}, \text{cam} R_{\text{ell}}^\top B'_{\text{cam}} \text{cam} R_{\text{ell}} \},
\]

recalling that \( A_{\text{ell}} \) and \( B'_{\text{cam}} \) are known.

Solving Equation (1) therefore consists in determining the value(s) of \( \sigma \) and the expressions of \( A, B', \Delta \) in a common coordinate frame (camera or ellipsoid).

### 4 Properties of the Solutions

In this section, we demonstrate some properties of the solutions of \( \text{Equation (1)} \) and exhibit relationships between the different variables. These results are new except Result [11] which was demonstrated in Eberly (2005). It must be noted that the positive definite nature of matrix \( A \), i.e. what differentiates an ellipsoid from any other quadric, plays a fundamental role in the demonstrations of these properties.

#### 4.1 Link with a Generalized Eigenvalue Problem

Let \( \{A, B', \Delta, \sigma\} \) be a set of solutions of Equation (1). Result [1] shows that they are also solutions of a generalized eigenvalue problem (Golub and Van Loan, 1996).
Result 1. If $A, B', \Delta$ and $\sigma$ satisfy
\[ A\Delta \Delta^\top A + (1 - \Delta^\top A \Delta)A = \sigma B', \tag{1} \]
they also satisfy
\[ A \Delta = \sigma B' \Delta. \tag{2} \]

Proof Let’s right-multiply (1) by $\Delta$. Since $\Delta^\top A \Delta$ is a scalar, the right hand term can be simplified:
\[
\sigma B' \Delta = (A\Delta \Delta^\top A + (1 - \Delta^\top A \Delta)A) \Delta = (\Delta^\top A \Delta)A \Delta + A \Delta - \Delta^\top A \Delta A \Delta = A \Delta.
\]

Since $B'$ is invertible, the generalized eigenvectors and eigenvalues of pair $\{A, B'\}$ are the eigenvectors and eigenvalues of matrix $B'^{-1} A$.

4.2 Generalized Eigenvalues of $\{A, B'\}$

Result 2. The couple $\{A, B'\}$ has exactly two distinct generalized eigenvalues, that are non-zero and of opposite signs.

Proof The generalized eigenvalues of $\{A, B'\}$ are non-zero because $B'^{-1} A$ is not singular.

We can then observe that
\[ Q(x) = \mu x^2 - (\mu + 1)\sigma x + \sigma^2 \]
is an annihilator polynomial of $B'^{-1} A$ (see proof in Appendix B):
\[ \mu(B'^{-1} A)^2 - (\mu + 1)\sigma B'^{-1} A + \sigma^2 I = 0. \tag{3} \]

In linear algebra, the minimal polynomial $\pi(\cdot)$ is defined as the monic annihilator polynomial having the lowest possible degree. It can be shown [Lang 2002] that (i) $\pi(\cdot)$ divides any annihilator polynomial and (ii) the roots of $\pi(\cdot)$ are identical to the roots of the characteristic polynomial. Since $Q$ is an annihilator polynomial of degree 2, we can thus infer that $B'^{-1} A$, and thus $\{A, B'\}$, has at most two distinct eigenvalues.

We are now going to prove by contradiction that $\{A, B'\}$ has exactly two distinct eigenvalues. Let’s thus assume that the couple has only one eigenvalue with multiplicity 3 denoted $\sigma_0$.

Since $A$ is positive definite and $B'$ is symmetric, the couple $\{A, B'\}$ has the following properties (Golub and Van Loan, 1996) (Corollary 8.7.2, p. 462):

1. their generalized eigenvalues are real,
2. their reducing subspaces are of the same dimension as the multiplicity of the associated eigenvalues,
3. their generalized eigenvectors form a basis of $\mathbb{R}^3$, and those with distinct eigenvalues are $A$-orthogonal.

According to property 2. above, we have
\[ \dim(\Ker(A - \sigma_0 B')) = 3, \]
i.e.
\[ A = \sigma_0 B', \]
which is impossible because $A$ represents an ellipsoid whereas $B'$ represents a cone. So $\{A, B'\}$ has exactly two distinct generalized eigenvalues.

Let’s then denote $\sigma_1$ (multiplicity 1) and $\sigma_2$ (multiplicity 2) these two eigenvalues. Observing that $\frac{1}{\sigma_1}$ and $\frac{1}{\sigma_2}$ are the generalized eigenvalues of $\{B', A\}$, we can write, according to Avron et al. [2008] (Theorem 3)
\[ \forall X \in \mathbb{R}^3 \setminus \{0\}, \]
\[ \min \left( \frac{1}{\sigma_1}, \frac{1}{\sigma_2} \right) \leq \frac{X^\top B'X}{X^\top AX} \leq \max \left( \frac{1}{\sigma_1}, \frac{1}{\sigma_2} \right) \]
If $\sigma_1$ and $\sigma_2$ were of the same sign, then $\forall X \in \mathbb{R}^3 \setminus \{0\}, X^\top B'X$ would be of that sign (since $X^\top AX > 0$). Yet, it is impossible since $B'$ is neither positive nor negative definite (cone). We thus conclude that the two distinct eigenvalues are of opposite signs. \qed

4.3 Characterization of $\sigma$

Let’s denote $\sigma_1$ (multiplicity 1) and $\sigma_2$ (multiplicity 2) the two generalized eigenvalues of $\{A, B'\}$.

Result 3. $\sigma$ is the generalized eigenvalue of $\{A, B'\}$ with multiplicity 1:
\[ \sigma = \sigma_1. \tag{4} \]

Proof Let’s consider $(\sigma_1, \delta_1)$ and $(\sigma_2, \delta_2)$ the generalized eigenvalues and eigenvectors of $\{A, B'\}$, such that $\|\delta_i\| = 1$.

We are going to prove (1) by contradiction.

Let’s suppose that there is $k \in \mathbb{R}^n$ such that $(A, \sigma_2, k \delta_2)$ is solution of Equation (1).

By injecting these values into (1), we therefore have
\[ B' - \frac{1}{\sigma_2} A = MA, \]
where
\[ M = \frac{k^2}{\sigma^2} (A\delta_2\delta_2^\top - \delta_2^\top A\delta_2). \]

According to property 2 of the proof of Result 2,
\[ \dim(\ker(B' - \frac{1}{\sigma^2} A)) = 2, \]
whence, since \( A \) is invertible,
\[ \dim(\ker(MA)) = \dim(\ker(M)) = 2. \]

However, defining
\[ \delta_2^\perp = \{ \mathbf{X} \in \mathbb{R}^3 \mid \mathbf{X} \bot \delta_2 \} \]
the subspace of dimension 2 orthogonal to \( \delta_2 \), we observe that, \( \forall \mathbf{X} \in \delta_2^\perp \),
\[ M\mathbf{X} = \frac{k^2}{\sigma^2} A\delta_2\delta_2^\top \mathbf{X} - \frac{k^2}{\sigma^2} \delta_2^\top A\delta_2 \mathbf{X} \]
\[ = \frac{k^2}{\sigma^2} A\delta_2(\delta_2 \cdot \mathbf{X}) - \frac{k^2}{\sigma^2} \delta_2^\top A\delta_2 \mathbf{X} \]
\[ = -\frac{k^2}{\sigma^2} \delta_2^\top A\delta_2 \mathbf{X}. \]

Since \( A \) is positive definite, \( \delta_2^\top A\delta_2 > 0 \), whence it comes
\[ \forall \mathbf{X} \in \delta_2^\perp \setminus \{ \mathbf{0} \}, M\mathbf{X} \neq \mathbf{0}. \]

It means that
\[ \delta_2^\perp \cap \ker(M) = \{ \mathbf{0} \}, \]
whence the direct sum
\[ \delta_2^\perp \bigoplus \ker(M) \]
is a subspace of \( \mathbb{R}^3 \) of dimension
\[ \dim(\delta_2^\perp) + \dim(\ker(M)) = 2 + 2 = 4. \]

We end up with a contradiction since \( 4 > \dim(\mathbb{R}^3) = 3. \)

As a result, triplets \((A, \sigma_2, k\delta_2)\) cannot be solutions of (11), thus solutions are necessarily in the form \((A, \sigma_1, k\delta_1)\), where \( k \in \mathbb{R}^\ast \). \( \square \)

### 4.4 Characterizations of \( \mu \)

Result 4 demonstrates that the secondary scalar variable \( \mu \) is also closely linked to the generalized eigenvalues of \((A, B')\).

**Result 4.** \( \mu \) is equal to the ratio between the two generalized eigenvalues of \((A, B')\):
\[ \mu = \frac{\sigma_1}{\sigma_2}. \]

**Proof** Trace of \( B'^{-1}A \) is given by its eigenvalues:
\[ tr(B'^{-1}A) = \sigma_1 + 2\sigma_2. \]

Whence, by squaring the matrix,
\[ tr((B'^{-1}A)^2) = \sigma_1^2 + 2\sigma_2^2. \]

Therefore, since \( tr(I) = 3 \) and \( \sigma = \sigma_1 \), applying the operator \( tr() \) to Equation (3) leads to
\[ \mu(\sigma_1^2 + 2\sigma_2^2) - \sigma_1(\mu + 1)(\sigma_1 + 2\sigma_2) + 3\sigma_1^2 = 0, \]
which is equivalent to
\[ \mu \sigma_2(\sigma_2 - \sigma_1) = \sigma_1(\sigma_2 - \sigma_1), \]
i.e.
\[ \mu = \frac{\sigma_1}{\sigma_2}. \]

Since \( \sigma_1 \) and \( \sigma_2 \) are of opposite signs, Equation (5) shows that \( \mu < 0 \).

Result 5 now highlights the connection between \( \mu \) and \( \sigma \).

**Result 5.** \( \mu \) and \( \sigma \) are linked through Equation (6):
\[ \mu = -\sqrt{\frac{\det(B')}{\det(A)}} \sigma_3. \]

**Proof** Determinant of \( B'^{-1}A \) is given by its eigenvalues:
\[ \det(B'^{-1}A) = \sigma_1\sigma_2^2, \]
i.e.
\[ \frac{\det(A)}{\det(B')} = \sigma_1\sigma_2^2. \]

One obtains (6) by injecting (5) into (7) and using \( \mu < 0 \). \( \square \)

### 4.5 Link between \( \sigma \) and \( \| \Delta \| \)

**Result 6.** The scalar \( \sigma \) and the camera-ellipsoid distance \( \| \Delta \| \) are linked through Equation (8):
\[ (tr(B'^{-1}))^2 \sigma = \frac{\det(A)}{\det(B')} (tr(A^{-1}) - \| \Delta \|^2)^{-1}. \]

**Proof** Injecting Equations (4) and (5) into (11) leads to
\[ \Delta \Delta^\top = A^{-1} - \frac{1}{\sigma^2} B'^{-1}. \]
Applying $tr()$ then squaring:
\[
\frac{1}{\sigma_2^2} (tr(B'^{-1}))^2 = (tr(A^{-1}) - \|\Delta\|^2)^2. \tag{9}
\]

Furthermore, injecting (11) into (7) leads to the following expression for $\frac{1}{\sigma_2^2}$:
\[
\frac{1}{\sigma_2^2} = \frac{det(B')}{det(A)} \sigma. \tag{10}
\]

Equation (8) is then obtained by injecting (10) into (9).

\section{Decoupling between Orientation and Position}

In this section, we consider two sub-problems of significant practical interest: (i) computing the ellipsoid position when the orientation is known and (ii) computing the orientation when the position is known. We demonstrate that the position can be inferred in closed-form from the orientation (Section 5.1, Result 7), while the latter can be analytically derived from the former up to the ellipsoid symmetries (Section 5.2, Result 8). We assimilate these properties to a decoupling phenomenon between orientation and position.

\subsection{Position from Orientation}

In this case, $A_{\text{cam}}$ or $B'_{\text{ell}}$ is known. Since $A_{\text{ell}}$ and $B'_{\text{cam}}$ are also known, eigenvalues of \{A, B'\}, $\sigma$ (using Result 3) and $\mu$ (Result 4) can be retrieved. Result 7 then provides that $\Delta$ is unique and fully determined.

\textbf{Result 7.} Assuming that the relative camera-ellipsoid orientation is known, their relative position is given by
\[
\Delta = k \delta_1, \tag{11}
\]

where
\[
\begin{align*}
&\Delta \text{ is a unit generalized eigenvector of } \{A, B'\} \\
&\text{corresponding to } \sigma_1, \\
&\delta_1 : N < 0, \\
&k = \sqrt{tr(A^{-1}) - \frac{1}{\sigma_2} tr(B'^{-1})}.
\end{align*}
\]

\textbf{Proof} Injecting Equations (13), (15) and $\Delta = k \delta_1$ into Equation (11) leads to
\[
k^2 \delta_1 \delta_1^\top = A^{-1} - \frac{1}{\sigma_2} B'^{-1}.
\]

Whence, by applying $tr()$,
\[
k^2 \|\delta_1\|^2 = k^2 = tr(A^{-1}) - \frac{1}{\sigma_2} tr(B'^{-1}).
\]

The two vectors $\pm \sqrt{k^2} \delta_1$ define ellipsoid centers $C$ that are symmetric with respect to the camera center $E$. The only one that satisfies the chirality constraint (ellipsoid located in front of the camera) is the one whose dot product with vector $N$ is negative (see Fig. 4).

\textbf{Result 8} highlights the fact that solving Equation (11) may only determine $A_{\text{cam}}$ or $B'_{\text{ell}}$. Indeed, $\sigma$ and $\Delta$ can then be uniquely derived. In other words, the relative position is fully constrained by the orientation.

As mentioned above, this result is of high practical interest. Indeed, getting the camera orientation, e.g., from physical sensors or image analysis, is usually easier than getting the camera position, and especially indoors where the GPS is useless. In a multi-object scene, the fact that only one ellipsoid-ellipsoid association is needed to compute the pose allows using a RANSAC-like strategy with low combinatorial cost both to detect the wrong associations and to choose the correct one when a label is shared by several objects.

A re-localization algorithm based on this strategy was presented in (Gaudilli`ere et al., 2019b). The system operates in real time from YOLO detections (Redmon and Farhadi, 2018) and IMU data or vanishing points – both methods were assessed. Figure 2 shows a few qualitative results obtained with images from the standard RGB-D TUM dataset (Sturm et al., 2012). Quantitative results as well as a detailed analysis of the advantages and limitations of this algorithm can be found in (Gaudilli`ere et al., 2019b). Other applications of Result 7 are presented in (Gaudilli`ere et al., 2019a, 2020a, b).

\subsection{Orientation from Position}

In this case $\Delta_{\text{cam}}$ or $\Delta_{\text{ell}}$ is known. $A_{\text{cam}}$ and $B'_{\text{ell}}$ are unknown, but since their eigenvalues are known, $det(A)$, $tr(A^{-1})$, $det(B')$ and $tr(B'^{-1})$ are known. $\sigma$ can thus be deduced from Result 6 and $\mu$ from Result 5. Result 8 then explains how to retrieve $A_{\text{cam}}$ or $B'_{\text{ell}}$ and how to derive the orientation in closed-form, up to the cone or ellipsoid symmetries.
Result 8. Assuming that the relative camera-ellipsoid position is known, their relative orientation is given by eigenvectors of

\[ B'_{\text{ell}} = \frac{1}{\sigma} \left( A_{\text{ell}} \Delta_{\text{ell}} \Delta_{\text{ell}}^\top A_{\text{ell}} + \mu A_{\text{ell}} \right) \]

in the ellipsoid reference frame, or of

\[ A_{\text{cam}} = \frac{\sigma}{\mu} \left( B'_{\text{cam}} - \sigma B'_{\text{cam}} \Delta_{\text{cam}} \Delta_{\text{cam}}^\top B'_{\text{cam}} \right) \]

in the camera reference frame.

Proof. Injecting (2) into (1) leads to

\[ \sigma^2 B' \Delta_{\text{cam}} \Delta_{\text{cam}}^\top B' + \mu A = \sigma B' \]

whence (12) by isolating \( A \).

Once for instance \( A_{\text{cam}} \) is retrieved, one can compute its eigenvalue decomposition:

\[ A_{\text{cam}} = \frac{\text{cam}}{\text{ell}} R_{\text{ell}} \begin{pmatrix} 1/a^2 & 0 & 0 \\ 0 & 1/b^2 & 0 \\ 0 & 0 & 1/c^2 \end{pmatrix} \frac{\text{cam}}{\text{ell}} R_{\ell}^\top. \]

For a triaxial ellipsoid (\( a \neq b \neq c \)), there are 4 solutions for \( \frac{\text{cam}}{\text{ell}} R_{\ell} \). Therefore, the ellipsoid orientation can be analytically derived from its position up to the ellipsoid symmetries.

6 Closed-form Solutions

In this section, we introduce the core contribution of the paper, that is closed-form solutions to the general P1E problem. Explicit 1DoF solutions are provided based on the fact that the ellipsoid is triaxial (all \( A \) eigenvalues are different) or not, and the cone is circular (two \( B' \) eigenvalues are equal) or not.

In Section 6.1 we first present the different types of ellipsoids and cones along with their possible co-occurrences. An overview of the solutions is given in Section 6.2. In Section 6.3 we consider the case of a triaxial ellipsoid and present for the first time a Necessary and Sufficient Condition (NSC) on \( \sigma \) to be solution of Equation (1). Then we derive the analytical expressions of the other variables as functions of \( \sigma \). In Section 6.4, we address the case of the spheroid (ellipsoid with an axis of revolution). That part enables to retrieve, from another formalism, the results presented in Wokes and Palmer (2010). In Section 6.5 we finally present the solutions for the sphere.

In what follows, the problem is solved either in the ellipsoid or in the cone coordinate frame. In brief, the choice is linked to the ability to define a frame associated to the considered structure without ambiguities. The case of the triaxial ellipsoid can thus be addressed in both frames since the two structures are unambiguous. The case of the spheroid is different, and depending on the properties of the cone, solutions are derived in one or the other frame.
6.1 Preliminaries: Co-occurrences of Ellipsoid and Cone Types

In this paper, we address the full ellipsoid pose estimation problem, i.e. we cover every possible types of ellipsoids and thus cones (see Appendix D). However, a specific type of ellipsoid cannot necessarily be tangent to any type of cone, what we refer to as possible or impossible co-occurrence.

Obviously, only circular cones can be tangent to a sphere. Furthermore, we are going to prove that only a non-circular elliptic cone can be tangent to a triaxial ellipsoid.

Let’s prove it by contradiction, and assume that the projection cone has a revolution axis (circular cone).

Let’s also assume that the ellipsoid center $C$ does not belong to that axis. Since the ellipsoid is tangent to the cone, any new ellipsoid obtained by rotating the original one around the cone revolution axis shall still be tangent to the cone, thus be solution of \([1]\). Yet, in this case, the locus of ellipsoid centers would be a circle located in a plane orthogonal to that axis and whose center would belong to it. Every center would thus be at a fixed distance to the cone vertex $E$, whence there would be an infinite number of $\Delta$ solutions for the same $\sigma$, given \([3]\). However, this contradicts Equation \([19]\). Therefore, the center of the ellipsoid must belong to the revolution axis of the cone.

If the ellipsoid center belonged to the cone revolution axis, then $\Delta$ would be parallel to that axis, i.e. would be an eigenvector of $B'$, whence of $A$ given \([20]\). However, in such a case, the symmetries of the cone-ellipsoid pair would impose that the tangent ellipse (intersection between the ellipsoid and the polar plane derived from $E$ (Wylie, 2008)) belongs to a plane orthogonal to the cone revolution axis, that is also a principal axis of the ellipsoid. Therefore, that tangent ellipse should be both a circle (orthogonal section of a circular cone) and a non-circular ellipse (section of an ellipsoid by a plane parallel to one of its principal planes), which is impossible. Therefore, the cone cannot have a revolution axis.

In brief, Table 1 summarizes the possible and impossible co-occurrences between ellipsoid and cone types.

6.2 Overview of the Solutions

In the rest of this section, we determine the solutions of the Cone Alignment Equation \([1]\), and derive the camera-ellipsoid relative poses. To this end, we distinguish between the three different types of ellipsoids (triaxial, spheroid, sphere). We demonstrate, in particular, that

- there is an infinite number of triaxial fixed-size ellipsoids that are tangent to a given backprojection cone (Fig. 3),
- as already demonstrated in Wokes and Palmer (2010), there are only two fixed-size spheroids solutions (see Fig. 4).

In the first case, the infinite number of ellipsoids tangent to the cone (or, conversely, the infinite number of cones tangent to the ellipsoid) explains the infinite number of camera solutions (see Fig. 5), and provides a parameterization of them. In the second case, the infinite number of change of basis matrices between the spheroid and the camera explains the infinite number of camera solutions (see Fig. 6). The mathematical developments leading to these results are presented below.
Table 1 Possible co-occurrences of ellipsoids and projection cones according to their types. ✓ indicates that ellipsoid and projection cone of the corresponding types may occur simultaneously. × indicates that they cannot.

| Projection cone | Non-circular | Triaxial | Spheroid | Sphere |
|-----------------|-------------|----------|----------|--------|
| Circular        | ✓           | ✓        | ×        |

6.3 The Triaxial Ellipsoid

6.3.1 Solving for $\sigma$

In practice, not all $\sigma$ values give rise to a solution of the problem. When the ellipsoid has three distinct radii (triaxial), Theorem 1 provides a characterization of the scalars $\sigma$ solutions of (1).

**Theorem 1.** Let’s denote $(\lambda_{A,1}, \lambda_{A,2}, \lambda_{A,3})$ the three distinct eigenvalues of $A$. Then $\sigma = dm^2$ is solution of Equation (1) if and only if the three entries of vector $M_A^{-1}V(m)$ are all non-negative:

$$M_A = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_{A,1} & \lambda_{A,2} & \lambda_{A,3} \\ \lambda_{A,1}^2 & \lambda_{A,2}^2 & \lambda_{A,3}^2 \end{pmatrix}, \quad (13)$$

![Fig. 4](https://members.loria.fr/VGaudilliere/files/Ellipsoid.mp4) The two spheroids solutions with a non-circular backprojection cone.

![Fig. 5](https://members.loria.fr/VGaudilliere/files/Ellipsoid.mp4) Triaxial ellipsoid: locus of cone vertices i.e. camera centers.

![Fig. 6](https://members.loria.fr/VGaudilliere/files/Ellipsoid.mp4) Spheroid: locus of cone vertices i.e. camera centers.
\[ V(m) = \left( \frac{\text{tr}(A^{-1}) - \frac{\text{tr}(B't^{-1})}{d} m}{1 - m^3} \right), \tag{14} \]

with
\[ d = \sqrt[3]{\frac{\text{det}(A)}{\text{det}(B')}}. \tag{15} \]
\[ m = \sqrt[3]{\mu}. \tag{16} \]

It must be noted that \( m \) is the only unknown parameter of vector \( V(m) \) since all the others derive from \( A \) and \( B' \) eigenvalues.\[ \]

**Proof** Let’s assume that Equation (11) is satisfied:
\[ A\Delta\Delta^\top A + \mu A = \sigma B'. \tag{11} \]
Therefore, equivalent Equation (11) is also satisfied:
\[ \Delta \Delta^\top = A^{-1} - \frac{\mu}{\sigma} B'^{-1}. \tag{11} \]

Since the trace of a product of matrices does not depend on the order of the matrices, we have
\[
\text{tr}(A \Delta \Delta^\top A) = \text{tr}((\Delta^\top A)(A \Delta)) = \text{tr}(\Delta^\top A^2 \Delta) = \Delta^\top A^2 \Delta \quad \text{(scalar)},
\]
and, similarly
\[
\text{tr}(\Delta \Delta^\top) = \Delta^\top \Delta.
\]

Given, in addition, \( \mu = 1 - \Delta^\top A \Delta \), applying \( \text{tr}() \) to Equations (11) and (11) leads to the following system:
\[
\begin{aligned}
\Delta^\top \Delta &= \text{tr}(A^{-1}) - \frac{\mu}{\sigma} \text{tr}(B'^{-1}) \\
\Delta^\top A \Delta &= 1 - \mu \\
\Delta^\top A^2 \Delta &= \sigma \text{tr}(B') - \mu \text{tr}(A).
\end{aligned}
\]

Although the two scalar unknowns \( \mu \) and \( \sigma \) appear in the two scalar side, they can be expressed as functions of a third unknown. Indeed, denoting
\[ m = \sqrt[3]{\mu} \quad \text{(unknown)}, \]
and
\[ d = \sqrt[3]{\frac{\text{det}(A)}{\text{det}(B')}} \quad \text{(known)}, \]
Equation (11) can be rewritten
\[ m^3 = -\sqrt[3]{\frac{\sigma^3}{d^3}}, \]
whence, by raising it to the power of 2/3,
\[ \sigma = dm^2. \tag{17} \]

Furthermore, given
\[ \mu = m^3 \quad \text{(definition)}, \]
we have
\[ \frac{\mu}{\sigma} = \frac{m}{d}. \]

Therefore, \( \Delta \) is solution of the following system with unknown \( m \):
\[
\begin{aligned}
\Delta^\top \Delta &= \text{tr}(A^{-1}) - \frac{\mu}{\sigma} \text{tr}(B'^{-1}) \\
\Delta^\top A \Delta &= 1 - m^3 \\
\Delta^\top A^2 \Delta &= \text{tr}(B') - \sigma \text{tr}(A)m^3.
\end{aligned}
\tag{18}
\]

The above equations are independent from the considered coordinate frame. Considering the ellipsoid frame and denoting \( (\Delta_{\text{ell},x}, \Delta_{\text{ell},y}, \Delta_{\text{ell},z})^\top \) the corresponding expression of \( \Delta \), the above system can be rewritten:
\[
\begin{pmatrix}
1 \\
\lambda_{A,1} \\
\lambda_{A,2} \\
\lambda_{A,3}
\end{pmatrix}
\begin{pmatrix}
\Delta_{\text{ell},x}^2 \\
\Delta_{\text{ell},y}^2 \\
\Delta_{\text{ell},z}^2
\end{pmatrix}
= \begin{pmatrix}
\text{tr}(A^{-1}) - \frac{\mu}{\sigma} \text{tr}(B'^{-1}) \\
1 - m^3 \\
\text{tr}(B') - \sigma \text{tr}(A)m^3
\end{pmatrix},
\]

i.e.
\[ M_A \begin{pmatrix}
\Delta_{\text{ell},x}^2 \\
\Delta_{\text{ell},y}^2 \\
\Delta_{\text{ell},z}^2
\end{pmatrix} = V(m). \tag{19} \]

Since \( A \) eigenvalues are all different (triaxial ellipsoid), Vandermonde matrix \( M_A \) is not singular (cf Appendix C). Therefore, the system can be inverted:
\[
\begin{pmatrix}
\Delta_{\text{ell},x}^2 \\
\Delta_{\text{ell},y}^2 \\
\Delta_{\text{ell},z}^2
\end{pmatrix} = M_A^{-1} V(m).
\]

Left hand side elements are all non-negative, whence the result.

Let’s now assume that the three entries of \( M_A^{-1} V(m) \) are all non-negative.

Let \( \Delta_{\text{ell}} = (\Delta_{\text{ell},x}, \Delta_{\text{ell},y}, \Delta_{\text{ell},z})^\top \) be a vector such that
\[
\begin{pmatrix}
\Delta_{\text{ell},x}^2 \\
\Delta_{\text{ell},y}^2 \\
\Delta_{\text{ell},z}^2
\end{pmatrix} = M_A^{-1} V(m).
\]

Such a definition is possible due to the positivity of the three entries.

One can therefore demonstrate, with the help of a formal calculus software (the corresponding Maple code is provided in Appendix E as reference), that, irrespective of the sign assumptions made for \( \Delta_{\text{ell}} \) entries, the matrix
\[
\frac{\sigma}{\mu}(A^{-1} - \Delta_{\text{ell}} \Delta_{\text{ell}}^\top) = \frac{d}{m}(A^{-1} - \Delta_{\text{ell}} \Delta_{\text{ell}}^\top)
\]
has the same eigenvalues with same multiplicities as \( B'^{-1} \). Given that both matrices are diagonalizable (since symmetric), this amounts to say that they are similar, and thus that Equation (11) is satisfied. \( \square \)
6.3.2 Solving for Camera Poses

**Theorem 2.** Considering a triaxial ellipsoid, each $\sigma$ solution of (1) gives rise to eight back-projection cones $(E, B')$ tangent to the ellipsoid. These cones are symmetric with respect to the three principal planes of the ellipsoid (see Fig. 7).

In addition, each backprojection cone defines two camera solutions (see Fig. 8).

Proof Theorem 1 provides a NSC on $\sigma$ to be solution of (1). Moreover, its proof exhibits that vectors $\Delta$ solutions are expressed in the ellipsoid frame in the form

$$\Delta_{\text{ell}} = \begin{pmatrix} \pm \sqrt{\Delta_{\text{ell},x}^2} \\ \pm \sqrt{\Delta_{\text{ell},y}^2} \\ \pm \sqrt{\Delta_{\text{ell},z}^2} \end{pmatrix}$$

(20)

where

$$\begin{pmatrix} \Delta_{\text{ell},x}^2 \\ \Delta_{\text{ell},y}^2 \\ \Delta_{\text{ell},z}^2 \end{pmatrix} = M_A^{-1} V(m)$$

There are thus eight vectors $\Delta_{\text{ell}}$ solutions for a given $m$ (thus $\sigma$), and they are symmetric with respect to the three principal planes of the ellipsoid.

The cone vertices (i.e. camera positions) can then be derived:

$$E_{\text{ell}} = C_{\text{ell}} + \Delta_{\text{ell}} = \Delta_{\text{ell}}.$$
has already been addressed in [Wokes and Palmer (2010)] using a different parameterization. The authors especially show that in the general case (non-circular elliptic cone), there are only two tangent spheroids, and we retrieve this result below.

6.4.1 The Non-circular Elliptic Cone

Let us first consider a non-circular elliptic cone. Expressing the Cone Alignment Equation in the cone coordinate frame, and given that the three \( B' \) eigenvalues are different, \( \sigma \) solutions can be characterized in a similar way to Theorem 1.

**Result 9.** Let’s denote \((\lambda_{B',1}, \lambda_{B',2}, \lambda_{B',3})\) the three distinct eigenvalues of \( B' \). Then \( \sigma = dm^2 \) is solution of Equation (1) if and only if the three entries of vector \( M_{B'}^{-1}V'(m) \) are all non-negative:

\[
M_{B'} = \begin{pmatrix} 1 & 1 & 1 \\ \lambda_{B',1} & \lambda_{B',2} & \lambda_{B',3} \\ \lambda_{B',1}^2 & \lambda_{B',2}^2 & \lambda_{B',3}^2 \end{pmatrix}, \quad (22)
\]

\[
V'(m) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{dm^2} & 0 \\ 0 & 0 & \frac{1}{dm^2} \end{pmatrix} V(m). \quad (23)
\]

**Proof** The proof is based on the exact same arguments as the proof of Theorem 1. In particular, \( M_{B'}^{-1}V'(m) \) is related to the expression \( \Delta_{\text{cone}} \) of \( \Delta \) in the cone frame:

\[
\begin{pmatrix} \Delta^2_{\text{cone},x}(m) \\ \Delta^2_{\text{cone},y}(m) \\ \Delta^2_{\text{cone},z}(m) \end{pmatrix} = M_{B'}^{-1}V'(m)
\]

Let’s also consider \((\lambda_{B',1}, \lambda_{B',2}, \lambda_{B',3})\) the eigenvalues of \( B' \), where \( \lambda_{B',1} \) and \( \lambda_{B',2} \) have the same sign (opposed to the sign of \( \lambda_{B',3} \)). Finally, let’s assume, even if it means exchanging the roles, that \(|\lambda_{B',1}| > |\lambda_{B',2}|\).

**Theorem 3.** Considering a spheroid along with a non-circular backprojection cone, there is only one \( \sigma \) value solution of Equation (1):

\[
\sigma = \begin{cases} \frac{\lambda_{A,\text{single}} \lambda_{B',1}^2}{\lambda_{B',2} \lambda_{B',3}} & \text{if } \lambda_{A,\text{single}} < \lambda_{A,\text{double}} \\ \frac{\lambda_{A,\text{single}} \lambda_{B',2}^2}{\lambda_{B',1} \lambda_{B',3}} & \text{if } \lambda_{A,\text{single}} > \lambda_{A,\text{double}} \end{cases}
\]

(24)

That \( \sigma \) value gives rise to two spheroids tangent to the cone, that are symmetric with respect to one of the cone principal planes (see Fig. 4).

**Proof** According to Theorem 1 \( \sigma = dm^2 \) is solution of (1) if and only if the three entries of the following vector are all non-negative:

\[
\begin{pmatrix} \Delta^2_{\text{cone},x}(m) \\ \Delta^2_{\text{cone},y}(m) \\ \Delta^2_{\text{cone},z}(m) \end{pmatrix} = M_{B'}^{-1}V'(m).
\]

Yet, developing the right hand term leads to the following system:

\[
\begin{align*}
\Delta^2_{\text{cone},x}(m) &= \frac{1}{m^2} P_1(m) \\
\Delta^2_{\text{cone},y}(m) &= \frac{1}{m^2} P_2(m) \\
\Delta^2_{\text{cone},z}(m) &= \frac{1}{m^2} P_3(m)
\end{align*}
\]

where

\[
\begin{align*}
P_1(x) &= k_1 \left( x - \frac{\lambda_{B',1}}{\lambda_{A,\text{single}}} \right) \left( x - \frac{\lambda_{B',1}}{\lambda_{A,\text{double}}} \right)^2 \\
P_2(x) &= k_2 \left( x - \frac{\lambda_{B',2}}{\lambda_{A,\text{single}}} \right) \left( x - \frac{\lambda_{B',2}}{\lambda_{A,\text{double}}} \right)^2 \\
P_3(x) &= k_3 \left( x - \frac{\lambda_{B',3}}{\lambda_{A,\text{single}}} \right) \left( x - \frac{\lambda_{B',3}}{\lambda_{A,\text{double}}} \right)^2
\end{align*}
\]

and where

\[
\begin{align*}
k_1 &= \frac{-\lambda_{B',2} \lambda_{B',3}^2}{\lambda_{B',1}^2 (\lambda_{B',1} - \lambda_{B',2})(\lambda_{B',1} - \lambda_{B',3})} \\
k_2 &= \frac{-\lambda_{B',1} \lambda_{B',3}^2}{\lambda_{B',2}^2 (\lambda_{B',2} - \lambda_{B',1})(\lambda_{B',2} - \lambda_{B',3})} \\
k_3 &= \frac{-\lambda_{B',1} \lambda_{B',2}^2}{\lambda_{B',3}^2 (\lambda_{B',3} - \lambda_{B',1})(\lambda_{B',3} - \lambda_{B',2})}
\end{align*}
\]

The locus of scalars \( m \) solutions is the subset of \( \mathbb{R} \) on which \( P_1(x), P_2(x) \) and \( P_3(x) \) are all non-negative.
To study the variations of these three polynomials, four cases—described in Table 2 and depending on the relative order of $A$ and $B'$ eigenvalues—need to be considered. Only case #1 is addressed below, since other cases can be solved using a similar reasoning.

| $\lambda_{A,\text{single}} < \lambda_{A,\text{double}}$ | $\lambda_{B',2} < 0$ | $\lambda_{B',2} > 0$ |
|---------------------------------|------------------|------------------|
| $\lambda_{A,\text{double}}$    | #1               | #2               |
| $\lambda_{A,\text{single}} > \lambda_{A,\text{double}}$ | #3               | #4               |

Table 2 Configurations of the problem depending on $A$ and $B'$ eigenvalues.

Let’s denote $S_i$ the root of $P_i(x)$ with multiplicity 1, and $D_i$ the root with multiplicity 2, such that

$$S_i = \frac{\lambda_{B',i}}{\lambda_{A,\text{single}}} \quad \text{and} \quad D_i = \frac{\lambda_{B',3}}{\lambda_{A,\text{double}}} \cdot d.$$  

In configuration #1, studying the variations of $\sigma$ (known) one can therefore derive the expressions of $\Delta_{\text{cone}}$ expressed in the cone frame verify:

\[
\begin{align*}
\Delta_{\text{cone},x}^2 &= 0 \\
\Delta_{\text{cone},y}^2 &= \frac{1}{D_1^2} P_2(D_1) \\
\Delta_{\text{cone},z}^2 &= \frac{1}{D_1^2} P_3(D_1).
\end{align*}
\]

Then

$$\Delta_{\text{cone}} = \left(\begin{array}{c} 0 \\ \pm \sqrt{\frac{1}{D_1^2} P_2(D_1)} \\ \pm \sqrt{\frac{1}{D_1^2} P_3(D_1)} \end{array}\right).$$

Since the sign of the third entry ($\Delta_{\text{cone},z}$) is fixed under the chirality constraint (ellipsoid located in front of the camera), there remains two possible expressions for $\Delta_{\text{cone}}$. The two resulting vectors are symmetric with respect to the cone principal plane whose normal is the eigenvector corresponding to eigenvalue $\lambda_{B',2}$.

Equation (12) then provides the expressions of $A$ in the cone coordinate frame:

$$A_{\text{cone}} = \frac{\sigma}{\mu} B_{\text{cone}} - \frac{\sigma^2}{\mu} B'_{\text{cone}} \Delta_{\text{cone}} \Delta_{\text{cone}}^T B_{\text{cone}}.$$  

One can therefore derive the expressions of $\Delta$ and $A$ in the camera frame using $\text{cam} R_{\text{cone}}$ (known):

$$\Delta_{\text{cam}} = \text{cam} R_{\text{cone}} A_{\text{cone}} \text{cam} R_{\text{cone}}^T.$$  

Then the spheroid centers are:

$$C_{\text{cam}} = \text{E}_{\text{cam}} - \Delta_{\text{cam}} = -\Delta_{\text{cam}}.$$

Now that the spheroid solutions have been determined in the camera coordinate frame, one can deduce the poses of camera solutions.

**Result 10.** Considering a spheroid along with a non-circular backprojection cone, the axial symmetry of the spheroid leads to an infinite number of camera solutions. The solutions belong to two planes orthogonal to the revolution axis of the spheroid and located at the same distance from its center (see Fig 7).

**Proof** Since there is only one possible value for $\sigma$, vectors $\Delta$ all have the same norm (cf. Result 4), i.e. the cameras are located at a fixed distance from the spheroid center.

Furthermore, orientations $\text{cam} R_{\text{ell}}^T$ of these cameras verify:

$$A_{\text{cam}} = \text{cam} R_{\text{ell}} A_{\text{ell}} \text{cam} R_{\text{ell}}^T.$$  

Since the spheroid has a revolution axis, arbitrarily fixing $A_{\text{ell}}$ eigenvectors for instance leaves two choices for one of $A_{\text{cam}}$ eigenvectors (the one corresponding to the revolution axis) and an infinite number of choices for the other two.

### 6.4.2 The Circular Cone

Considering the case of a circular cone (elliptic cone with a revolution axis), we are going to demonstrate that there is only one spheroid tangent to it (Result 11).

In this case, $B'$ has two distinct eigenvalues. Let’s denote them $\lambda_{B',\text{single}}$ (multiplicity 1) and $\lambda_{B',\text{double}}$ (multiplicity 2).

**Result 11.** Considering a spheroid along with a circular backprojection cone, there is only one $\sigma$ value solution of Equation (1):

$$\sigma = \frac{\lambda_{A,\text{single}}}{\lambda_{B',\text{single}}}.\quad (25)$$

That $\sigma$ value gives rise to one spheroid tangent to the cone, and both revolution axes coincide (See
Perspective-Ellipsoid gives and that observing that \( \top \) minimal polynomial is \( A \)

\[
A \text{double} = \begin{bmatrix}
\lambda_A,\text{single} \\
\lambda_A,\text{double}
\end{bmatrix}
\]

Left-multiplying by \( \Delta \) and right-multiplying by \( \Delta \) gives

\[
\Delta^\top A^2 \Delta = \alpha \Delta^\top A \Delta + \beta \Delta^\top \Delta.
\]

 Injecting the expressions of \( \Delta^\top A^2 \Delta, \Delta^\top A \Delta \) and \( \Delta^\top \Delta \) as functions of \( m \) from [13] into [27] and observing that

\[
\tr(A) = \lambda_A,\text{single} + 2\lambda_A,\text{double}
\]

and that \( \tr(A^{-1}) = \frac{1}{\lambda_A,\text{single}} + \frac{2}{\lambda_A,\text{double}} \), we obtain, after simplification

\[
m^3 = \frac{\tr(B')d}{\lambda_A,\text{double}} m^2 + \frac{\lambda_A,\text{single} \tr(B'^{-1})}{d} m \frac{\lambda_A,\text{single}}{\lambda_A,\text{double}}
\]

Let’s call \( P_{\text{spheroid}}(x) \) the above polynomial whose \( m \) is root:

\[
P_{\text{spheroid}}(x) = x^3 - \frac{\tr(B')d}{\lambda_A,\text{double}} x^2 - \frac{\lambda_A,\text{single} \tr(B'^{-1})}{d} x + \frac{\lambda_A,\text{single}}{\lambda_A,\text{double}}.
\]

Developing \( \tr(B') \) and \( \tr(B'^{-1}) \), one can observe that it can be rewritten

\[
P_{\text{spheroid}}(x) = \left(x - \frac{\lambda_B,\text{single} d}{\lambda_A,\text{double}} \right) \left(x - \frac{\lambda_B,\text{double} d}{\lambda_A,\text{double}} \right)^2.
\]

At this stage, one can note that the sign of \( \lambda_B,\text{single} \) is root:

\[
d = \frac{1}{\lambda_A,\text{double}} \frac{\det(A)}{\det(B')}
\]

Thus the signs of the roots of \( P_{\text{spheroid}}(x) \) are:

\[
\frac{\lambda_B,\text{single}}{\lambda_A,\text{double}} d > 0 \quad \text{and} \quad \frac{\lambda_B,\text{double}}{\lambda_A,\text{double}} d < 0.
\]

Since \( m < 0 \), the only possible value for it is the second one. Therefore, \( \sigma \) is given by:

\[
\sigma = \frac{1}{\lambda_B,\text{single}} \frac{\det(A)}{\det(B')}
\]

Let’s now focus on \( \Delta \), and consider the first two equations of System [13]

\[
\Delta^\top A \Delta = 1 - m^3
\]

Considering the ellipsoid frame, its left hand side can be rewritten in a matrix form:

\[
\begin{pmatrix}
\frac{1}{\lambda_A,\text{single}} & \lambda_A,\text{double} \\
\lambda_A,\text{single} & \lambda_A,\text{double}
\end{pmatrix}
\begin{pmatrix}
\Delta_{\text{ell},x}^2 \\
\Delta_{\text{ell},y}^2 + \Delta_{\text{ell},z}^2
\end{pmatrix}
\]

That Vandermonde matrix is not singular given that \( \lambda_A,\text{single} \neq \lambda_A,\text{double} \), thus the system can be inverted.

After developing the right hand side, we finally obtain

\[
\Delta_{\text{ell},x}^2 = \frac{1}{\lambda_A,\text{single}} \left(1 - \frac{\lambda_A,\text{single} \lambda_B,\text{double}}{\lambda_A,\text{double} \lambda_B,\text{single}}\right),
\]

and

\[
\Delta_{\text{ell},y}^2 + \Delta_{\text{ell},z}^2 = 0.
\]

Therefore,

\[
\Delta_{\text{ell}} = \left(\pm \frac{1}{\lambda_A,\text{single}} \left(1 - \frac{\lambda_A,\text{single} \lambda_B,\text{double}}{\lambda_A,\text{double} \lambda_B,\text{single}}\right), 0, 0\right)
\]
\[ \Delta \] is thus an eigenvector of \( A \) corresponding to eigenvalue \( \lambda_{A,\text{single}} \), i.e. coincides with the revolution axis of the spheroid.

Equation (2) then requires that \( \Delta \) is also an eigenvector of \( B' \). It must be the one corresponding to the revolution axis of the cone since, if not, the ellipsoid center would be located outside of the cone. In that respect, both axes of revolutions (cone and spheroid) coincide, and \( \| \Delta \| \) is given by:

\[
\| \Delta \| = \sqrt{\frac{1}{\lambda_{A,\text{single}}} \left( 1 - \frac{\lambda_{A,\text{double}} \lambda_{B',\text{double}}}{\lambda_{B',\text{single}}} \right)}.
\]

\[ A_{\text{cone}} \] can then be obtained from (12), then \( A_{\text{cam}} \) and \( C_{\text{cam}} \) using \( \text{cam}R_{\text{cone}} \) (known).

Now the spheroid solution has been determined in the camera frame, we can infer the corresponding camera poses.

**Result 12.** When considering a spheroid along with a circular backprojection cone, there are two solutions for camera position, that are located on the revolution axis of the spheroid and at the same distance from its center, and an infinite number of solutions for camera orientation.

**Proof** Equation (28) gives the two possible solutions for \( \Delta \). For reasons of symmetry, both of them are actual solutions.

Furthermore, orientations \( \text{cam}R_{\text{ell}}^T \) of these cameras verify:

\[ A_{\text{cam}} = \text{cam}R_{\text{ell}}A_{\text{ell}}\text{cam}R_{\text{ell}}^T. \]

Just as when considering a non-circular elliptic cone (Result 10), we conclude on the infinite number of camera orientations. □

### 6.5 The Sphere

When the ellipsoid is a sphere, the matrix \( A \) has the same expression in every basis:

\[ A = \lambda_{A,\text{triple}} I, \tag{29} \]

where \( R = \frac{1}{\sqrt{\lambda_{A,\text{triple}}}} \) is the sphere radius.

Given its observation as an circle in the image, there is obviously an infinite number of camera solutions, that are located at the same distance from the sphere center.

Using the formalism of our study, we can precise these properties with the following result:

**Result 13.** When considering a sphere, there is only one \( \sigma \) solution of Equation (1):

\[ \sigma = \frac{\lambda_{A,\text{triple}}}{\lambda_{B',\text{single}}}. \tag{30} \]

That \( \sigma \) value defines a unique sphere tangent to the cone, whose center belongs the revolution axis of the cone. The distance from the cone vertex to sphere center is given by:

\[
\| \Delta \| = \sqrt{\frac{1}{\lambda_{A,\text{triple}}} \left( 1 - \frac{\lambda_{B',\text{double}}}{\lambda_{B',\text{single}}} \right)}.
\]

The locus of camera positions is then a sphere with radius \( \| \Delta \| \) around the ellipsoid center.

Proof of Result 13 is provided in Appendix H.

### 6.6 Examples of Retrieved Poses

We provide in Fig. 10 a few examples of retrieved camera trajectories from one ellipse-ellipsoid correspondence on a real scene from the T-LESS dataset (Sturm et al., 2012). Ellipsoidal models of objects (all triaxial) were reconstructed using Rubino et al. (2018). Ellipsoids have then been reprojected into one image of the sequence using the groundtruth projection matrix. Finally, the trajectories of the camera solutions were retrieved using our method (Section 6.3). In the figure, ellipse and ellipsoid colors coincide with those of the trajectories. Naturally, all the trajectories intersect at the ground-truth camera position \( E \). This example illustrate the practical interest of our method. However, a comprehensive study of numerical aspects, e.g. noise robustness, are outside the scope of this paper.

### 7 Conclusion

We propose in this paper a complete characterization of the P1E problem and noticeably extend previous works that only partially addressed this problem. Besides its theoretical interest, this paper also proposes a constructive solution of the camera trajectories. The closed-form solution provided for the position-from-orientation case has proven its practical interest. The orientation-from-position solution, even if not yet exploited, may
also represent a convenient manner to simplify the pose estimation problem.

Future investigations concern numerical aspects and especially a sensitivity analysis of the method to image noise. Another important concern will be the joint use of the method with a minimal number of other image features, such as points or other ellipse-ellipsoid correspondences, to ensure the computation of a unique solution.

**Appendix A  Equivalent Problem Formulations**

To prove that Equations (1) and (1') are equivalent, we demonstrate below that (1) implies (1') then (1') implies (1).

(1) \implies (1') Multiply (1) on the right by \(\Delta\) to obtain \(A\Delta = \sigma B'\Delta\) (see Appendix B). Whence

\[A\Delta \Delta^\top A + \mu A = \sigma B' = \sigma B'\Delta \Delta^\top A = \mu A\]

Left-multiplying by \(B'^{-1}\)

\[\sigma I - \sigma \Delta \Delta^\top A = \mu B'^{-1} A\]

Then right-multiplying by \(A^{-1}\)

\[\sigma A^{-1} - \sigma \Delta \Delta^\top = \mu B'^{-1}\]

Finally (\(\sigma \neq 0\))

\[A^{-1} = \frac{\mu}{\sigma} B'^{-1}\]

(1') \implies (1) Multiply (1') on the right by \(A\) then on the left by \(B'\) to obtain

\[B' - B'\Delta \Delta^\top A = \frac{\mu}{\sigma} A\]

Then right-multiply by \(\Delta\) to obtain \(\mu B'\Delta = \frac{\mu}{\sigma} A\Delta\), whence \((\mu \neq 0)\) \(A\Delta = \sigma B'\Delta\).

Injecting that result into he previous equation leads to

\[\sigma B' - A\Delta \Delta^\top A = \mu A\]

**Appendix B  Proving that**

\[Q(B'^{-1}A) = 0\]

Replacing (2) into (1), we obtain:

\[\sigma^2 B'\Delta \Delta^\top B' - (\sigma \Delta^\top B'\Delta - 1) A = \sigma B'\]

We can then deduce the following expression for \(A\):

\[A = \frac{\sigma}{1 - \sigma \Delta^\top B'\Delta}(B' - \sigma B'\Delta \Delta^\top B')\]
Whence, denoting $I$ the identity matrix and defining $f = \frac{1}{\sigma \mathbf{A}}$, then left-multiplying by $B'^{-1}$, we obtain
\[
B'^{-1} A = f(I - \sigma \mathbf{A} \mathbf{A}^\top B')
\]
Squaring that expression leads to
\[
(B'^{-1} A)^2 = f^2(I - \sigma \mathbf{A} \mathbf{A}^\top B')^2 = f^2(2\sigma \mathbf{A} \mathbf{A}^\top B' + \sigma^2 (\mathbf{A}^\top B' \mathbf{A}) \mathbf{A}^\top B') = f^2(2\sigma \mathbf{A} \mathbf{A}^\top B' + \sigma^2 (\mathbf{A}^\top B' \mathbf{A}) \mathbf{A}^\top B') = f^2(I - \sigma(2 - \sigma \mathbf{A}^\top B') \mathbf{A} \mathbf{A}^\top B')
\]
Defining $\mu = 1 - \sigma \mathbf{A}^\top B' \mathbf{A} = 1 - \mathbf{A} \mathbf{A}^\top \mathbf{A}$:
\[
(B'^{-1} A)^2 = f^2(I - \sigma(\mu + 1) \mathbf{A} \mathbf{A}^\top B') = f^2((\mu + 1)(I - \sigma \mathbf{A} \mathbf{A}^\top B') - \mu I) = f(\mu + 1)B'^{-1} A - f^2 \mu I = \frac{\sigma}{\mu}(\mu + 1)B'^{-1} A - \frac{\sigma^2}{\mu} I
\]
Finally, we have
\[
\mu(B'^{-1} A)^2 = \sigma(\mu + 1)B'^{-1} A - \sigma^2 I
\]
Whence, denoting $Q(x) = \mu x^2 - (\mu + 1)\sigma x + \sigma^2$,
\[
Q(B'^{-1} A) = 0
\]

Appendix C  Vandermonde Matrix

A Vandermonde matrix is a matrix with the terms of a geometric progression in each row or column:

\[
V = \begin{pmatrix}
1 & 1 & 1 & \ldots & 1 \\
x_1 & x_2 & x_3 & \ldots & x_n \\
x_1^2 & x_2^2 & x_3^2 & \ldots & x_n^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
x_1^{m-1} & x_2^{m-1} & x_3^{m-1} & \ldots & x_n^{m-1}
\end{pmatrix}
\]

The determinant of a square Vandermonde matrix (when $m = n$) is given by
\[
det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i).
\]
Therefore, $V$ is not singular (i.e. $det(V) \neq 0$) if and only if all $x_i$ are distinct.

Appendix D  Ellipsoid and Cone types

See Tables D1 and D2.

Appendix E  Theorem 9: Maple code

\[
> \text{with(linalg):}
> A:= matrix([[1A1, 0, 0], [0, 1A2, 0], [0, 0, 1A3]]);
> B:= matrix([[1B1, 0, 0], [0, 1B2, 0], [0, 0, 1B3]]);
> M_B:= transpose(vandermonde([[1B1, 1B2, 1B3]]));
> d:= (det(A)/det(B))^1/3;
> V:= map(i -> inverse(A)-trace(inverse(A))+m^2*trace(B)*d^3*m^3), M_B);
> eigenvalues(inverse(V));
\]

Appendix G  Solving the Polynomial Equation in Theorem 3

In case #1, the signs of $B'$ eigenvalues ensure that
\[
d = \sqrt[3]{\frac{\lambda_{A,\text{single}} \lambda_{A,\text{double}}}{\lambda_{B',1} \lambda_{B',2} \lambda_{B',3}}} > 0,
\]
|                  | Triaxial ellipsoid | Spheroid | Sphere |
|------------------|-------------------|----------|--------|
| **Illustration** | ![Triaxial ellipsoid](image1) | ![Spheroid](image2) | ![Sphere](image3) |
| **Lengths of principal axes** | $a, b, c$ | $a, b$ | $r$ |
| **Eigenvalues of $A$** | $\frac{1}{a^2}, \frac{1}{b^2}, \frac{1}{c^2}$ | $\frac{1}{a^2}, \frac{1}{b^2}, \frac{1}{r^2}$ | $\frac{1}{r^2}, \frac{1}{r^2}, \frac{1}{r^2}$ |
| **Signs of eigenvalues** | $+, +, +$ | $+, +, +$ | $+, +, +$ |
| **Characteristic polynomial of $A$** | $P_A(x) = (x-a)(x-b)(x-c)$ | $P_A(x) = (x-a)(x-b)^2$ | $P_A(x) = (x-r)^3$ |
| **Minimal polynomial of $A$** | $\pi_A(x) = (x-a)(x-b)(x-c)$ | $\pi_A(x) = (x-a)(x-b)$ | $\pi_A(x) = (x-r)$ |

**Table D1** The different types of ellipsoids.

|                  | Non-circular elliptic cone | Circular cone |
|------------------|---------------------------|--------------|
| **Illustration** | ![Non-circular elliptic cone](image4) | ![Circular cone](image5) |
| **Eigenvalues of $B$** | $\lambda_1, \lambda_2, \lambda_3$ | $\lambda_1, \lambda_2, \lambda_2$ |
| **Signs of eigenvalues** | $-, +, +$ | $-, +, +$ |
| **Charateristic polynomial of $B$** | $P_B(x) = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$ | $P_B(x) = (x-\lambda_1)(x-\lambda_2)^2$ |
| **Minimal polynomial of $B$** | $\pi_B(x) = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$ | $\pi_B(x) = (x-\lambda_1)(x-\lambda_2)$ |

**Table D2** The different types of elliptic cones.
then that

\[ k_1 < 0, \, k_2 > 0 \text{ and } k_3 < 0. \]

Let’s denote \( S_i \) the root of \( P_i(x) \) with multiplicity 1, and \( D_i \) the root with multiplicity 2, such that

\[ S_i = \frac{\lambda_{B',i}}{\lambda_{A,\text{single}}} d \quad \text{ and } \quad D_i = \frac{\lambda_{B',i}}{\lambda_{A,\text{double}}} d. \]

Considering the signs of \( \lambda_{B',i} \) and the fact that \( \lambda_{A,\text{single}} < \lambda_{A,\text{double}} \), it comes

\[ S_1 < D_1 < 0, \]
\[ S_2 < D_2 < 0, \]
\[ 0 < D_3 < S_3. \]

One can therefore note that the roots of \( P_1(x) \) and \( P_2(x) \) are negative, while those of \( P_3(x) \) are positive. Since, in addition, \( k_3 < 0 \), we have

\[ \forall x \leq 0, \, P_3(x) > 0. \]

Let’s now focus on the signs of \( S \) and \( B \), to determine the locus of possible \( m \) values. Since \( \lambda_{B',1} < \lambda_{B',2} \), their roots verify

\[ S_1 < S_2, \]
\[ D_1 < D_2. \]

Therefore, one can distinguish between two configurations regarding the roots order:

\[ S_1 < D_1 < S_2 < D_2, \]

or

\[ S_1 < S_2 < D_1 < D_2. \]

For the first case, the variations of the three polynomials are presented in Table G3.

One can observe that the three polynomials are never all non-negative, thus this configuration is impossible. For the second case, however, there is one value for which all three polynomials are non-negative: \( D_1 \) (see Table G3).

**Appendix H  Proof of Result**

**Proof** Left-multiplying \( A = \lambda_{A,\text{triple}} I \) by \( \Delta^T \) right-multiplying it by \( \Delta \), we obtain

\[ \Delta^T A \Delta = \lambda_{A,\text{triple}} \Delta^T \Delta. \]

Injecting the first two equations of System (13), we then have

\[ 1 - m^3 = \lambda_{A,\text{triple}} \left( tr(A^{-1}) - \frac{tr(B'^{-1})}{d} m \right). \]

Developing \( tr(A^{-1}) \) and \( tr(B'^{-1}) \) leads to

\[ 1 - m^3 = \frac{\lambda_{A,\text{triple}}}{d} \lambda_{B',\text{double}} \lambda_{B',\text{simple}} m \]

Furthermore, one can observe that

\[ d = \frac{\sqrt{\det(A)}}{\det(B')} = \frac{\lambda_{A,\text{triple}}}{\lambda_{B',\text{simple}}^{1/3} \lambda_{B',\text{double}}^{2/3}}. \]

Whence, injecting this into the former equation:

\[ 0 = m^3 - \frac{\lambda_{B',\text{double}} + 2 \lambda_{B',\text{simple}}}{\lambda_{B',\text{double}} \lambda_{B',\text{simple}}^{1/3}} m + 2 \]

Denoting

\[ R = \frac{\sqrt{\lambda_{B',\text{double}}}}{\lambda_{B',\text{simple}}}, \]

the last equation means that \( m \) is root of the polynomial

\[ P_{\text{sphere}}(x) = x^3 - \left( R^2 + \frac{2}{R} \right) x + 2. \]

Yet, \( R \) is an obvious root of this polynomial:

\[ P_{\text{sphere}}(x) = (x - R) \left( x^2 + R x - \frac{2}{R} \right) \]

Even if obtaining a formal expression of the two other roots is not straightforward, Vieta’s formulas provide the following constraints:

\[ \begin{cases} x_1 + x_2 + R = 0, \\ x_1 x_2 R = -2. \end{cases} \]

If roots \( x_1 \) are \( x_2 \) complex, then \( R \) is the only possible value for \( m \). If they are real, then, since \( R < 0 \) (\( B' \) eigenvalues are of opposite signs), the second formula requires that \( x_1 \) and \( x_2 \) are of the same sign, and the first formula requires that they are positive. Finally, \( m = R \).

Corresponding \( \sigma \) value is:

\[ \sigma = dR^2 \]
Table G3 Signs of $P_i(x)$ when $S_1 < D_1 < S_2 < D_2$.

| $x$   | $-\infty$ | $S_1$ | $D_1$ | $S_2$ | $D_2$ | 0 |
|-------|-----------|-------|-------|-------|-------|---|
| $P_1(x)$ | + | 0 | - | 0 | - | 0 |
| $P_2(x)$ | - | 0 | + | 0 | + | 0 |
| $P_3(x)$ | + |   |   |   |   |   |

Table G4 Signs of $P_i(x)$ when $S_1 < S_2 < D_1 < D_2$.

| $x$   | $-\infty$ | $S_1$ | $S_2$ | $D_1$ | $D_2$ | 0 |
|-------|-----------|-------|-------|-------|-------|---|
| $P_1(x)$ | + | 0 | - | 0 | - | 0 |
| $P_2(x)$ | - | 0 | + | 0 | + | 0 |
| $P_3(x)$ | + |   |   |   |   |   |

$\lambda_{A,\text{triple}} = \lambda_{B',\text{double}}^{2/3} \lambda_{B',\text{single}}^{1/3}$

$\lambda_{B',\text{single}} = \lambda_{A,\text{triple}}$.

Applying $tr()$ to Equation (1) gives the value of $\|\Delta\|^2$:

$$\|\Delta\|^2 = \frac{3}{\lambda_{A,\text{triple}}} - \frac{\mu}{\sigma} \left( \frac{1}{\lambda_{B',\text{single}}} + \frac{2}{\lambda_{B',\text{double}}} \right).$$

Given

$$\sigma = \frac{\lambda_{A,\text{triple}}}{\lambda_{B',\text{single}}},$$

and

$$\mu = m^3 = \frac{\lambda_{B',\text{double}}}{\lambda_{B',\text{single}}},$$

the right hand side can be rewritten

$$\frac{3}{\lambda_{A,\text{triple}}} - \frac{\lambda_{B',\text{double}}}{\lambda_{A,\text{triple}}} \left( \frac{1}{\lambda_{B',\text{single}}} + \frac{2}{\lambda_{B',\text{double}}} \right),$$

i.e.

$$\|\Delta\|^2 = \frac{3}{\lambda_{A,\text{triple}}} - \frac{\lambda_{B',\text{double}}}{\lambda_{A,\text{triple}}} \frac{1}{\lambda_{B',\text{single}}} = \frac{2}{\lambda_{A,\text{triple}}},$$

$$= \left( 1 - \frac{\lambda_{B',\text{double}}}{\lambda_{B',\text{single}}} \right) \frac{1}{\lambda_{A,\text{triple}}}. \quad \Box$$

References

Avron, H., E. Ng, and S. Toledo. 2008, Aug. A generalized courant-fischer minimax theorem. [https://escholarship.org/uc/item/4gb4t762](https://escholarship.org/uc/item/4gb4t762).

Bonin-Font, F., A. Ortiz, and G. Oliver. 2008. Visual navigation for mobile robots: A survey. *J. Intell. Robotic Syst.* 53(3): 263–296.

Crocco, M., C. Rubino, and A. Del Bue 2016, June. Structure from motion with objects. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Dong, W. and V. Isler. 2021. Ellipse regression with predicted uncertainties for accurate multi-view 3d object estimation. CoRR abs/2101.05212.

Dong, W., P. Roy, C. Peng, and V. Isler. 2021. Ellipse R-CNN: learning to infer elliptical object from clustering and occlusion. *IEEE Trans. Image Process.* 30: 2193–2206.

Eberly, D. 2007, May. Reconstructing an ellipsoid from its perspective projection onto a plane. [https://www.geometrictools.com/](https://www.geometrictools.com/) Updated version: March 1, 2008.

Fischler, M.A. and R.C. Bolles. 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24(6): 381–395.

Garg, S., N. Sünderhauf, F. Dayoub, D. Morrison, A. Cosgun, G. Carneiro, Q. Wu, T.J. Chin, I. Reid, S. Gould, P. Corke, and M. Milford. 2020. Semantics for robotic mapping, perception and interaction: A survey. *Foundations and Trends in Robotics* 8(1–2): 1–224.

Gaudilliére, V., G. Simon, and M.O. Berger. 2019a, November. Camera Pose Estimation with Semantic 3D Model. In *IROS 2019 - 2019*
IEEE/RSJ International Conference on Intelligent Robots and Systems, Macau, Macau SAR China.

Gaudilli`ere, V., G. Simon, and M.O. Berger. 2019b, October. Camera Relocalization with Ellipsoidal Abstraction of Objects. In ISMAR 2019 - 18th IEEE International Symposium on Mixed and Augmented Reality, Beijing, China.

Gaudilli`ere, V., G. Simon, and M.O. Berger. 2019c, Mar. Perspective-12-Quadric: An analytical solution to the camera pose estimation problem from conic - quadric correspondences.

Gaudilli`ere, V., G. Simon, and M.O. Berger. 2020a, October. Perspective-2-Ellipsoid: Bridging the Gap Between Object Detections and 6-DoF Camera Pose. In IROS 2020 – 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, Las Vegas, United States.

Gaudilli`ere, V., G. Simon, and M.O. Berger. 2020b. Perspective-2-Ellipsoid: Bridging the Gap Between Object Detections and 6-DoF Camera Pose. IEEE Robotics and Automation Letters: 5189–5196.

Gay, P., C. Rubino, V. Bansal, and A. Del Bue 2017, Oct. Probabilistic structure from motion with objects (psfmo). In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Girshick, R. 2015, December. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

Girshick, R., J. Donahue, T. Darrell, and J. Malik 2014, June. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Golub, G.H. and C.F. Van Loan. 1996. Matrix Computations (Third ed.). The Johns Hopkins University Press.

Hartley, R.I. and A. Zisserman. 2004. Multiple View Geometry in Computer Vision (Second ed.). Cambridge University Press.

Hoque, S., M.Y. Arafat, S. Xu, A. Maiti, and Y. Wei. 2021. A comprehensive review on 3d object detection and 6d pose estimation with deep learning. IEEE Access 9: 143746–143770.

Kisantal, M., S. Sharma, T.H. Park, D. Izzo, M. Mårtens, and S. D’Amico. 2020. Satellite pose estimation challenge: Dataset, competition design, and results. IEEE Trans. Aerosp. Electron. Syst. 56(5): 4083–4098.

Lang, S. 2002. Algebra. Graduate Texts in Mathematics, vol. 211 (Revised third ed.).

Lepetit, V., F. Moreno-Noguer, and P. Fua. 2009. Epnp: An accurate O(n) solution to the pnp problem. Int. J. Comput. Vis. 81(2): 155–166.

Li, J., D. Meger, and G. Dudek 2017. Context-coherent scenes of objects for camera pose estimation. In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2017, Vancouver, BC, Canada, September 24-28, 2017, pp. 655–660. IEEE.

Li, J., Z. Xu, D. Meger, and G. Dudek 2018, May. Semantic scene models for visual localization under large viewpoint changes. In 15th Conference on Computer and Robot Vision (CRV), Toronto, Canada.

Li, Y. 2019. Detecting lesion bounding ellipses with gaussian proposal networks. In Machine Learning in Medical Imaging - 10th International Workshop, MLMI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13, 2019, Proceedings, Volume 11861 of Lecture Notes in Computer Science, pp. 337–344. Springer.

Liao, Z., W. Wang, X. Qi, X. Zhang, L. Xue, J. Jiao, and R. Wei. 2020. Object-oriented SLAM using quadrics and symmetry properties for indoor environments. CoRR abs/2004.05303.
Liu, W., D. Anguelov, D. Erhan, C. Szegedy, S.E. Reed, C. Fu, and A.C. Berg. 2016. SSD: single shot multibox detector. In Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I, Volume 9905 of Lecture Notes in Computer Science, pp. 21–37. Springer.

Marchand, É., H. Uchiyama, and F. Spindler. 2016. Pose estimation for augmented reality: A hands-on survey. IEEE Trans. Vis. Comput. Graph. 22(12): 2633–2651.

Nicholson, L., M. Milford, and N. Sünderhauf. 2019, Jan. Quadricslam: Dual quadrics from object detections as landmarks in object-oriented slam. IEEE Robotics and Automation Letters 4(1): 1–8.

Pan, S., S. Fan, S.W.K. Wong, J.V. Zidek, and H. Rhodin 2021. Ellipse detection and localization with applications to knots in sawn lumber images. In IEEE Winter Conference on Applications of Computer Vision, WACV 2021, Waikoloa, HI, USA, January 3-8, 2021, pp. 3891–3900. IEEE.

Redmon, J., S. Divvala, R. Girshick, and A. Farhadi 2016, June. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Redmon, J. and A. Farhadi 2017, July. Yolo9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Redmon, J. and A. Farhadi. 2018. Yolov3: An incremental improvement. http://arxiv.org/abs/1804.02767

Rubino, C., M. Crocco, and A.D. Bue. 2018, Jun. 3d object localisation from multi-view image detections. IEEE Transactions on Pattern Analysis and Machine Intelligence 40(6): 1281–1294.

Sturm, J., N. Engelhard, F. Endres, W. Burgard, and D. Cremers 2012. A benchmark for the evaluation of RGB-D SLAM systems. In 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2012, Vilamoura, Algarve, Portugal, October 7-12, 2012, pp. 573–580. IEEE.

Tian, R., Y. Zhang, Y. Feng, L. Yang, Z. Cao, S. Coleman, and D. Kerr. 2021. Accurate and robust object-oriented SLAM with 3d quadric landmark construction in outdoor environment. CoRR abs/2110.08977.

Wokes, D.S. and P.L. Palmer. 2010. Perspective reconstruction of a spheroid from an image plane ellipse. International Journal of Computer Vision 90(3): 369–379.

Wylie, C.R. 2008. Introduction to projective geometry. New York, NY: Dover.

Xu, C., L. Zhang, L. Cheng, and R. Koch. 2017. Pose estimation from line correspondences: A complete analysis and a series of solutions. IEEE Trans. Pattern Anal. Mach. Intell. 39(6): 1209–1222.

Zhao, M., X. Jia, L. Fan, Y. Liang, and D. Yan. 2021. Robust ellipse fitting using hierarchical gaussian mixture models. IEEE Trans. Image Process. 30: 3828–3843.

Zins, M., G. Simon, and M.O. Berger 2020. 3d-aware ellipse prediction for object-based camera pose estimation. In 8th International Conference on 3D Vision, 3DV 2020, Virtual Event, Japan, November 25-28, 2020, pp. 281–290. IEEE.