Quadrilateral Configurations with Invariant Centroid

Luís Machado[a,b], Knut Hüper[c], Fátima Silva Leite[b,d]

[a] *Department of Mathematics, University of Trás-os-Montes and Alto Douro (UTAD), Quinta de Prados, Vila Real, 5000-801, Portugal*
[b] *Institute of Systems and Robotics, University of Coimbra, DEEC - Polo II, 3030-290, Coimbra, Portugal*
[c] *Institute of Mathematics, Julius-Maximilians-Universität Würzburg, Campus Hubland Nord Emil-Fischer-Straße 31, 97074, Würzburg, Germany*
[d] *Department of Mathematics, University of Coimbra, Largo D. Dinis, 3000-143, Coimbra, Portugal*

Abstract

The objective of this paper is twofold. First, we characterize the geometry of the set of all possible planar configurations of four points that share the same centroid (admissible configurations). Particular configurations with this property consist of the vertices of a regular quadrilateral (square) whose diagonals intersect at the centroid. Then, we define an optimization criterion that ensures square admissible configurations are global minima of the corresponding cost function. Our goal is to characterize and classify the critical points of that cost function and understand how admissible configurations can be smoothly morphed into other configurations, in particular, into square configurations.

Keywords: Quadrilateral configurations, centroid, Riemannian manifold, cost function, critical points, Riemannian gradient and Hessian, steepest descent, Newton's method.

Contents

1. Introduction

The study of k-point configurations or arrangements on finite-dimensional Riemannian manifolds under certain geometric constraints has gained significant attention in recent years due to its many applications. Moreover, related mathematical inquiries, particularly in three-dimensional Euclidean space, have intrigued scientists for centuries. Notable examples include the comprehensive classification of Platonic solids, with more recent extensions to higher dimensions Baez (2020), and the role of crystallography in the study of five-fold symmetries and quasi-crystals. Many of these topics have a strong geometric foundation, while others are also motivated by solid-state physics, including applications in X-ray diffraction and related fields.

For over a century, the connections between these studies and algebraic concepts have been well established, particularly through group-theoretic methods. Early approaches focused on the realization and representation of discrete groups, and over time, these studies have expanded to include Lie groups, even in infinite dimensions, and aspects of invariant theory.

Another area where the study of point arrangements plays a crucial role is in computer science and engineering-related problems. Here, notable examples include packing problems, both in the classical sense of logistics Bezdek (2010, 2013) and in more abstract computational implementations related to code design Conway et al. (1996); Conway and Sloane (1999), Zong (1999), Nebe et al. (2006). The German term "Lagerungen" is often used in this context Fejes Tóth et al. (2023).

When additional geometric constraints are imposed on the desired configurations in the form of a system of equations, the problem becomes significantly more complex. In such cases, tools from differential geometry become necessary, as these equations often define a differentiable sub-Riemannian manifold embedded in Euclidean space. Numerical methods designed to find such configurations must respect these constraints, making geometric optimization techniques particularly suitable. Specifically, methods such as Riemannian gradient descent or Riemannian versions of Newton-type algorithms ensure that both the data and final results remain within the predefined constraint set Absil et al. (2008), Boumal (2023).

A closely related problem in statistics involves characterizing configurations of k distinct points on an n-dimensional Riemannian manifold that share the same centroid. Here, researchers aim to develop efficient methods for computing these configurations and their associated invariants Pennec (2006) .

In the field of robotics, specific configurations and their geometric mean play a key role in modeling how robotic networks interact with the physical environment and with each other. For planar configurations, the geometric mean (centroid) can be interpreted as a virtual anchor for movement and coordination in multi-robot formations. Individual robots rely on this balance point as a common reference to guarantee coordinated behavior (see, for instance, Bullo et al. (2009) and Mesbahi and Egerstedt (2010)). Square formations are usually preferred when planar robots work in cooperation. That choice insures more stability and prevents crashing. Reconfiguring their formation to avoid collision with obstacles and then making them return to an optimal formation is also a major objective in navigation of a team of robots, see, for instance, Alonso-Mora et al. (2017) and references therein.

Being aware of the importance of understanding the geometry and dynamics of point configurations in real applications, particularly path planning and control in cooperative robotics, was a strong motivation for initiating this research. However, following our previous work in Machado et al. (2024), the purpose here is to provide a more comprehensive understanding of these mathematical structures while complementing the theoretical results with numerical experiments and illustrations.

In the recent work Machado et al. (2024), we characterized all configurations of three points on the 2-dimensional sphere S^2 that share the same Riemannian mean and study an optimization problem that has the equilateral triangle as an optimal solution. In the meanwhile, we have noticed that the addition of a single point to the configurations increased considerably the complexity of the problem. Due to the challenges associated with the characterization of four-point configurations on S^2 with the same Riemannian mean, the present paper focus exclusively on planar quadrilateral configurations. As it will become clear, this case is already more intricate than its triangular counterparts studied in Machado et al. (2024) as a preliminary step and warm-up toward the spherical case. By combining analytical insights with computational verification, we aim to provide a more comprehensive understanding of the underlying structures, facilitating both theoretical advancements and possible practical applications.

Our main objective in the present paper is to study the geometry of the set of all possible planar configurations of four points that share the same centroid (admissible configurations), and to define an optimization criterion ensuring that square shape admissible configurations are global minima of the corresponding cost function. Furthermore, our aim is to identify and classify the critical points of that cost function and investigate how admissible configurations can be smoothly morphed into other configurations. In particular, we want to understand how a quadrilateral configuration can be continuously deformed into the vertices of a square with the same centroid.

The organization of this paper is as follows. We first introduce some of the notations that will be used throughout the paper. After defining the main objectives of this study in Section 3, we briefly present the geometry of the manifold of all planar configurations of four points that keep invariant their centroid. We then define an appropriate smooth cost function based on mutual distances between the four points and the length of the diagonals of the quadrilateral having those points as vertices. This cost function guarantees that square configurations are global minimum. Still in Section 3, we compute the Riemannian gradient of that cost and find its critical points, which turns out to be equivalent to finding the zero set of a set of multivariate third order polynomials with rational coefficients. The complete characterization of the critical points is done by an exhaustive case by case study based on elementary geometric insight combined with undergraduate linear algebra. Alternatively, we briefly describe a short cut using a Gröbner basis approach.

In section 4 we compute the Riemannian Hessian and completely classify the critical points. In particular, the square configurations arise as global minimum of the cost function. Using several routines from MATLAB toolboxes, the steepest descent and quasi-Newton algorithms on manifolds have been implemented to corroborate the theoretical outcomes. These algorithms turned out to be easy to implement, offering high accuracy and precision. To enrich the paper, meaningful plots illustrating our results are also included. The implementation of those algorithms also show how a quadrilateral configuration can be continuously deformed into the vertices of a square with the same centroid.

2. Notations

These are some of the notations used throughout the paper.

M	smooth manifold
T_pM	tangent space of M at a point $p \in M$
N_pM	normal space of M at a point $p \in M$
$P^{\perp}_{T_pM}$	orthogonal projection onto the tangent space
$P^{\perp}_{N_pM}$	orthogonal projection onto the normal space
$D\hat{F}(p)$	differential of a function $\hat{F}\colon \mathbb{R}^6 \to \mathbb{R}$ at $p \in \mathbb{R}^6$
$\nabla\hat{F}$	gradient of $\hat{F}$
∇F	Riemannian gradient of a function $F\colon M \to \mathbb{R}$
$H_{\widehat{F}}$	Hessian of $\hat{F}$
$H_{\hat{F}}$	Riemannian Hessian of F
$\lVert \cdot \rVert$	Euclidean norm

3. Quadrilateral configurations

3.1. Main objectives

Our first objective is to characterize the geometry of the set of all possible configurations of four points $\{p_0, p_1, p_2, p_3\} \subset \mathbb{R}^2$ that share the same centroid $q \in \mathbb{R}^2$ (admissible configurations), i.e,

$$q = \tfrac{1}{4}\sum_{i=0}^{3} p_i. \tag{1}$$

Clearly, particular configurations with this property consist of the vertices of a regular quadrilateral (square) whose diagonals intersect at q. Our second objective is to define an optimization criterion that ensures square admissible configurations are global minima of the corresponding cost function. Furthermore, our goal is to characterize and classify the critical points of that cost function within M, and investigate how admissible configurations can be smoothly morphed into other configurations in M. In particular, we want to understand how a quadrilateral configuration can be continuously deformed into the vertices of a square with the same centroid.

Recall that the centroid q of $\{p_0, p_1, p_2, p_3\}$ is the unique solution of the minimization problem

$$\min_{x\in\mathbb{R}^2} \sum_{i=0}^{3} \|p_i - x\|^2. \tag{2}$$

Formulas (1) and (2) can be trivially adjusted to include any number of points in $\mathbb{R}^n$, but our current focus is put on a simpler situation as the necessary preparation and insight for significantly more challenging cases. We also assume that one of the points is fixed and, without loss of generality, we fix p_0.

3.2. The configuration manifold

Under the previous assumption, the following subset of $\mathbb{R}^6$ defines the configurations having centroid q:

$$M = \big\{(p_1, p_2, p_3) \in \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \mid p_1 + p_2 + p_3 = 4q - p_0\big\}. \tag{3}$$

M is clearly a smooth manifold, since it is a 4-dimensional affine subspace of the embedding space $\mathbb{R}^6$. Moreover, we equip M with the Riemannian metric induced by the Euclidean metric of the embedding space. The geometry of this manifold is particularly important to tackle our objectives.

Lemma 1. *The tangent and the normal space of M at $p = (p_1, p_2, p_3) \in M$ are given, respectively, by*

$$T_pM = \big\{(\mathfrak{v}_1, \mathfrak{v}_2, -\mathfrak{v}_1 - \mathfrak{v}_2) \mid \mathfrak{v}_i \in \mathbb{R}^2\big\}, \quad N_pM = \big\{(\mathfrak{v}, \mathfrak{v}, \mathfrak{v}) \mid \mathfrak{v} \in \mathbb{R}^2\big\}. \tag{4}$$

Proof. M is the zero set of the function $f : \mathbb{R}^6 \to \mathbb{R}^2$, $p = (p_1, p_2, p_3) \mapsto p_1+p_2+p_3-4q+p_0$, which has maximal rank. So, the tangent space to M at p is the kernel of the linear surjection $Df(p) : \mathbb{R}^6 \to \mathbb{R}^2$, $(\mathfrak{v}_1, \mathfrak{v}_2, \mathfrak{v}_3) \mapsto \mathfrak{v}_1 + \mathfrak{v}_2 + \mathfrak{v}_3$, i.e., $T_pM = \big\{(\mathfrak{v}_1, \mathfrak{v}_2, -\mathfrak{v}_1 - \mathfrak{v}_2) \mid \mathfrak{v}_i \in \mathbb{R}^2\big\}$. Clearly $\big\{(\mathfrak{v}, \mathfrak{v}, \mathfrak{v}) \mid \mathfrak{v} \in \mathbb{R}^2\big\}$ is 2-dimensional and all vectors in this vector subspace of $\mathbb{R}^6$ are orthogonal to vectors in T_pM. So, the normal space to M at p is as given in (4). □

To simplify notations, we may represent vectors in $\mathbb{R}^6$ as row vectors or column matrices.

Lemma 2. *The orthogonal projection operators onto the tangent and normal space to M at a point $p = (p_1, p_2, p_3)$ are given, respectively, by*

$$P^{\perp}_{T_pM} \colon \mathbb{R}^6 \to T_pM, \quad (\mathfrak{v}_1, \mathfrak{v}_2, \mathfrak{v}_3) \mapsto \tfrac{1}{3} \begin{bmatrix} 2I & -I & -I \\ -I & 2I & -I \\ -I & -I & 2I \end{bmatrix} \begin{bmatrix} \mathfrak{v}_1 \\ \mathfrak{v}_2 \\ \mathfrak{v}_3 \end{bmatrix} = \tfrac{1}{3}\Big(\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \otimes I \Big) \begin{bmatrix} \mathfrak{v}_1 \\ \mathfrak{v}_2 \\ \mathfrak{v}_3 \end{bmatrix}, \tag{5}$$

and

$$P^{\perp}_{N_pM} \colon \mathbb{R}^6 \to N_pM, \quad (\mathfrak{v}_1, \mathfrak{v}_2, \mathfrak{v}_3) \mapsto \tfrac{1}{3} \begin{bmatrix} I & I & I \\ I & I & I \\ I & I & I \end{bmatrix} \begin{bmatrix} \mathfrak{v}_1 \\ \mathfrak{v}_2 \\ \mathfrak{v}_3 \end{bmatrix} = \tfrac{1}{3}\big(\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \otimes I \big) \begin{bmatrix} \mathfrak{v}_1 \\ \mathfrak{v}_2 \\ \mathfrak{v}_3 \end{bmatrix}, \tag{6}$$

where I stands for the 2×2 identity matrix.

Proof. Simple calculations show that these two linear operators are idempotent, i.e.,

$$\big(P^{\perp}_{T_pM}\big)^2 = P^{\perp}_{T_pM} \quad \text{and} \quad \big(P^{\perp}_{N_pM}\big)^2 = P^{\perp}_{N_pM}.$$

Moreover,

$$P^{\perp}_{T_pM}\Big|_{T_pM} = \mathrm{id}, \ P^{\perp}_{N_pM}\Big|_{N_pM} = \mathrm{id}, \ \ker\big(P^{\perp}_{T_pM}\big) = N_pM, \quad \text{and} \ \ker\big(P^{\perp}_{N_pM}\big) = T_pM,$$

which proves the assertion. □

3.3. The optimization problem

In order to smoothly move inside the manifold M from a polygon configuration to a square, we define the following cost function so that the square configuration emerges as a global minimum.

$$\begin{aligned} F \colon M &\to \mathbb{R}_0^+, \\ p = (p_1, p_2, p_3) &\mapsto \tfrac{1}{4}\Big(\big(d^2(p_0,p_1) - d^2(p_1,p_2)\big)^2 + \big(d^2(p_0,p_1) - d^2(p_2,p_3)\big)^2 \\ &\quad + \big(d^2(p_0,p_1) - d^2(p_3,p_0)\big)^2 + \big(d^2(p_1,p_2) - d^2(p_2,p_3)\big)^2 \\ &\quad + \big(d^2(p_1,p_2) - d^2(p_3,p_0)\big)^2 + \big(d^2(p_2,p_3) - d^2(p_3,p_0)\big)^2 \\ &\quad + \big(d^2(p_0,p_2) - d^2(p_1,p_3)\big)^2 \Big). \end{aligned} \tag{7}$$

Remark 1. *Obviously, F attains its minimum value zero when the points (p_0, p_1, p_2, p_3) form the vertices of a regular quadrilateral. The rationale behind (7) is as follows: The first six summands in (7) ensure equality between the four sides of the quadrilateral, note that $6 = \binom{4}{2}$. Whereas the last summand in (7) serves to enforce equality between its two diagonals. One of our objectives is to minimize the cost functional F and hopefully end up with this regular polygon.*

Defining

$$
\begin{aligned}
A_1 &:= d^2(p_0,p_1) - d^2(p_1,p_2) = \|p_0-p_1\|^2 - \|p_1-p_2\|^2,\\
A_2 &:= d^2(p_0,p_1) - d^2(p_2,p_3) = \|p_0-p_1\|^2 - \|p_2-p_3\|^2,\\
A_3 &:= d^2(p_0,p_1) - d^2(p_3,p_0) = \|p_0-p_1\|^2 - \|p_3-p_0\|^2,\\
C &:= d^2(p_0,p_2) - d^2(p_1,p_3) = \|p_0-p_2\|^2 - \|p_1-p_3\|^2,
\end{aligned} \tag{8}
$$

where, for simplicity of notations, we omit the dependency on p, the cost function can be rewritten as

$$
F(p) = \tfrac{1}{4}\Big(A_1^2 + A_2^2 + A_3^2 + (A_1-A_2)^2 + (A_1-A_3)^2 + (A_2-A_3)^2 + C^2\Big). \tag{9}
$$

In order to derive the Riemannian gradient of F, we first compute the differential of the function $\widehat{F}$, the latter seen as an extension of F to the embedding space.

Proposition 1. *The differential of $\widehat{F}$ at $p = (p_1,p_2,p_3) \in \mathbb{R}^6$ in the direction of the vector $\mathfrak{v} = (\mathfrak{v}_1,\mathfrak{v}_2,\mathfrak{v}_3) \in \mathbb{R}^6$ is given by*

$$
\begin{aligned}
&\mathrm{D}\widehat{F}(p)(\mathfrak{v})\\
&= \big\langle (A_1+A_2+A_3)(p_1-p_0) + (3A_1-A_2-A_3)(p_2-p_1) + C(p_3-p_1), \mathfrak{v}_1\big\rangle\\
&\quad + \big\langle (3A_1-A_2-A_3)(p_1-p_2) + (3A_2-A_1-A_3)(p_3-p_2) + C(p_2-p_0), \mathfrak{v}_2\big\rangle\\
&\quad + \big\langle (3A_2-A_1-A_3)(p_2-p_3) + (3A_3-A_1-A_2)(p_0-p_3) + C(p_1-p_3), \mathfrak{v}_3\big\rangle.
\end{aligned} \tag{10}
$$

Proof. Notice first that,

$$
F(p) = \tfrac{1}{4}\big(3A_1^2 + 3A_2^2 + 3A_3^2 - 2A_1A_2 - 2A_1A_3 - 2A_2A_3 + C^2\big), \tag{11}
$$

and

$$
\begin{aligned}
\mathrm{D}A_1(p)(\mathfrak{v}) &= 2\langle p_1-p_0, v_1\rangle - 2\langle p_1-p_2, \mathfrak{v}_1-\mathfrak{v}_2\rangle,\\
\mathrm{D}A_2(p)(\mathfrak{v}) &= 2\langle p_1-p_0, v_1\rangle - 2\langle p_2-p_3, \mathfrak{v}_2-\mathfrak{v}_3\rangle,\\
\mathrm{D}A_3(p)(\mathfrak{v}) &= 2\langle p_1-p_0, v_1\rangle - 2\langle p_3-p_0, \mathfrak{v}_3\rangle,\\
\mathrm{D}C(p)(\mathfrak{v}) &= 2\langle p_2-p_0, v_2\rangle - 2\langle p_1-p_3, \mathfrak{v}_1-\mathfrak{v}_3\rangle.
\end{aligned} \tag{12}
$$

Therefore,

$$
\begin{aligned}
\mathrm{D}\widehat{F}(p)(\mathfrak{v}) = {}& (3A_1-A_2-A_3)(\langle p_1-p_0,\mathfrak{v}_1\rangle - \langle p_1-p_2, \mathfrak{v}_1-\mathfrak{v}_2\rangle)\\
&+ (3A_2-A_1-A_3)(\langle p_1-p_0,\mathfrak{v}_1\rangle - \langle p_2-p_3, \mathfrak{v}_2-\mathfrak{v}_3\rangle)\\
&+ (3A_3-A_1-A_2)(\langle p_2-p_0,\mathfrak{v}_2\rangle - \langle p_1-p_3, \mathfrak{v}_1-\mathfrak{v}_3\rangle)\\
&+ C(\langle p_2-p_0, v_2\rangle - \langle p_1-p_3, \mathfrak{v}_1-\mathfrak{v}_3\rangle).
\end{aligned} \tag{13}
$$

After simplifying and reordering expressions, we obtain the result. □

Since $\langle \nabla\widehat{F}(p), \mathfrak{v}\rangle = \mathrm{D}\widehat{F}(p)\mathfrak{v}$, it is immediate to conclude from (10) that the Euclidean gradient of $\widehat{F}$, at the point $p \in \mathbb{R}^6$, is given by

$$
\nabla\widehat{F}(p) = \begin{bmatrix} (A_1+A_2+A_3)(p_1-p_0)+(3A_1-A_2-A_3)(p_2-p_1)+C(p_3-p_1)\\ (3A_1-A_2-A_3)(p_1-p_2)+(3A_2-A_1-A_3)(p_3-p_2)+C(p_2-p_0)\\ (3A_2-A_1-A_3)(p_2-p_3)+(3A_3-A_1-A_2)(p_0-p_3)+C(p_1-p_3) \end{bmatrix}. \tag{14}
$$

3.4. Riemannian gradient

To simplify notations, define the following variables:

$$X := 3A_1 - A_2 - A_3, \quad Y := 3A_2 - A_1 - A_3, \quad Z := 3A_3 - A_1 - A_2. \tag{15}$$

Proposition 2. *At each point $p = (p_1, p_2, p_3) \in M$, the Riemannian gradient of F, denoted ∇F, is given by*

$$\nabla F(p) = \frac{1}{3}\begin{bmatrix} 2(X+Y+Z)(p_1-p_0)+3X(p_2-p_1)+Z(p_3-p_0)+3C(p_3-p_1)+C(p_0-p_2) \\ (X+Y+Z)(p_0-p_1)+3X(p_1-p_2)+3Y(p_3-p_2)+Z(p_3-p_0)+2C(p_2-p_0) \\ (X+Y+Z)(p_0-p_1)+3Y(p_2-p_3)+2Z(p_0-p_3)+3C(p_1-p_3)+C(p_0-p_2) \end{bmatrix}. \tag{16}$$

Proof. Since $M \subset \mathbb{R}^6$ is a Riemannian submanifold, the Riemannian gradient of F at $p \in M$ is obtained by projecting the Euclidean gradient $\nabla \widehat{F}$ at p onto the tangent space of M at p. Taking into consideration that for $\mathfrak{v} = (\mathfrak{v}_1, \mathfrak{v}_2, \mathfrak{v}_3) \in T_pM$ we have $\mathfrak{v}_3 = -\mathfrak{v}_1 - \mathfrak{v}_2$ and using the notations (15), the expression (14) for the Euclidean gradient simplifies and after its projection onto T_pM, given in (4), one obtains the expression (16). □

We can now apply the steepest descent algorithm (Algorithm 1) to obtain approximate solutions to the problem.

Algorithm 1: Steepest descent with Armijo line search

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)})$ and tolerance tol
Output: Stationary point $p^* = (p_1^*, p_2^*, p_3^*)$

for $j = 0, 1, \ldots$ **do**
  Set $d^{(j)} = -\nabla F(p^{(j)})$;
  Determine the step length α_j according to Armijo rule;
  Set $p^{(j+1)} = p^{(j)} + \alpha_j d^{(j)}$;
  Stop if $F(p^{(j)}) < \mathsf{tol}$ or $\|\nabla F(p^{(j)})\| < \mathsf{tol}$
end

3.5. Characterization of the critical points

Our next goal is to characterize the critical points of the function F defined by (9). The natural identification of a point with its position vector is used in the sequel. To simplify the presentation, we further assume, without loss of generality, that q coincides with the origin in $\mathbb{R}^2$, that is, $p_1 + p_2 + p_3 = -p_0$.

Given four points in the plane, satisfying the manifold constraint, there are several ways to arrange them relative to one another. Under the assumption $q = 0$, we say that a configuration of four points is collinear if the position vector of each point with respect to the origin is a scalar multiple of a single position vector. Note that, if three of the four points are on a straight line through the origin, the manifold constraint forces the fourth point to be on the same line.

The cases listed in the tables below cover all possible arrangements of four points satisfying $p_0 + p_1 + p_2 + p_3 = 0$.

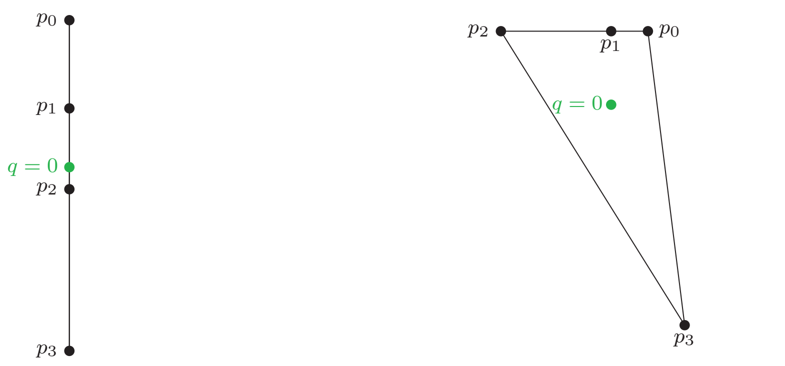
Case 1.	**Collinear configurations**
Case 1.1	all points coincide with the origin,
Case 1.2	$p_1 = p_0 = 0$, $p_2 \neq 0$,
Case 1.3	$p_1 = p_0 \neq 0$,
Case 1.4	$p_1 \neq p_0$, $p_0 \neq 0$,
Case 1.5	$p_1 \neq p_0$, $p_0 = 0$.

Table 1: Possible arrangements of four collinear points in the plane

Case 2.	**Non-collinear configurations**
Case 2.1	$\{p_0, p_1\}$ linearly independent,
Case 2.2	$\{p_0, p_1\}$ linearly dependent and $p_0 \neq 0$,
Case 2.3	$\{p_0, p_1\}$ linearly dependent and $p_1 \neq 0$.

Table 2: Possible arrangements of four non-collinear points in the plane

The following figure shows one configuration from Case 1., another from Case 2..



(a) Collinear configuration (Case 1.4)

(b) Non-collinear configuration (Case 2.1)

Figure 1: A collinear and a non-collinear configuration of four points in $\mathbb{R}^2$

The cases described in Table 1 and Table 2 above will be used to prove the following theorem which characterizes the critical points of F.

Theorem 1. *Any critical point $p = (p_1, p_2, p_3) \in M \subset \mathbb{R}^6$ of the function F defined by (9) fulfills one of the following conditions:*

1. $$p_0 = p_1 = p_2 = p_3 = 0; \tag{17}$$

2. $$\begin{aligned} &p_2 = p_0, \\ &p_1 = p_3 = -p_0; \end{aligned} \tag{18}$$

3. $$\begin{aligned} &p_1 = p_3 = \lambda p_0, \quad \textit{for} \ \ \lambda = -\tfrac{5 \pm 2\sqrt{2}}{17}, \\ &p_2 = -p_0 + 2p_1; \end{aligned} \tag{19}$$

4. $$\begin{aligned} &p_2 = 0, \\ &p_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}^\top, \\ &p_1 = \tfrac{1}{2}\begin{bmatrix} -x_0 - y_0 & x_0 - y_0 \end{bmatrix}^\top \ \ \textit{or} \ \ p_1 = \tfrac{1}{2}\begin{bmatrix} y_0 - x_0 & -x_0 - y_0 \end{bmatrix}^\top, \\ &p_3 = -p_0 - p_1 - p_2; \end{aligned} \tag{20}$$

5. $$\begin{aligned} &p_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}^\top, \\ &p_1 = \tfrac{1}{5}\begin{bmatrix} -x_0 \mp y_0\sqrt{4 \mp \sqrt{15}} & -y_0 \pm x_0\sqrt{4 \mp \sqrt{15}} \end{bmatrix}^\top, \\ &p_2 = \alpha\, p_0 + \tfrac{2\alpha}{1-\alpha}\, p_1, \quad \textit{for} \ \ \alpha = \pm\sqrt{\tfrac{3}{5}}, \\ &p_3 = -p_0 - p_1 - p_2; \end{aligned} \tag{21}$$

6. p_0, p_1, p_2, p_3 *form the vertices of a regular quadrilateral (square).* (22)

Proof. The critical points are triples $(p_1, p_2, p_3) \in M \subset \mathbb{R}^6$ with directional derivative $\mathrm{D}F(p_1, p_2, p_3)(\mathfrak{v}_1, \mathfrak{v}_2, -\mathfrak{v}_1 - \mathfrak{v}_2) = 0$, for all $\mathfrak{v}_1, \mathfrak{v}_2 \in \mathbb{R}^2$.

Using (10), the critical points (p_1, p_2, p_3) are the solutions of the two coupled equations

$$\begin{aligned} (X + Y + Z)(p_1 - p_0) + X(p_2 - p_1) + Y(p_3 - p_2) + Z(p_3 - p_0) + 2C(p_3 - p_1) &= 0, \\ X(p_1 - p_2) + 2Y(p_3 - p_2) + Z(p_3 - p_0) + C(p_3 - p_1) + C(p_2 - p_0) &= 0. \end{aligned} \tag{23}$$

Subtracting both equations and replacing the first by that, one gets

$$\begin{aligned} (X + Y + Z)(p_1 - p_0) + 2X(p_2 - p_1) + Y(p_2 - p_3) + C(p_3 - p_1) + C(p_0 - p_2) &= 0, \\ X(p_1 - p_2) + 2Y(p_3 - p_2) + Z(p_3 - p_0) + C(p_3 - p_1) + C(p_2 - p_0) &= 0, \end{aligned} \tag{24}$$

or, equivalently,

$$\begin{aligned}(-X-Y-Z+C)p_0+(-X+Y+Z-C)p_1+(2X+Y-C)p_2+(C-Y)p_3=0,\\ (-Z-C)p_0+(X-C)p_1+(-X-2Y+C)p_2+(2Y+Z+C)p_3=0.\end{aligned} \tag{25}$$

The proof consists in finding solutions of (25), for all arrangements in Table 1 and Table 2.

Case 1. Collinear configurations.

Case 1.1 All points coincide with the origin.

If $p_0 = p_1 = p_2 = p_3 = 0$, system (25) is trivially satisfied. So, this configuration is a critical point.

Case 1.2 $p_0 = p_1 = 0$, $p_2 \neq 0$.

In this case $p_3 = -p_2$. Simple calculations, using (8) and (15), give $X = Z = 2\|p_2\|^2$, $Y = -10\|p_2\|^2$, $C = 0$, and system (25) reduces to the unsolvable

$$X + Y = 0, \quad X - 2Y = 0.$$

Case 1.3 $p_0 = p_1 \neq 0$.

In this case $p_2 = \lambda p_0$, for some real scalar λ, and $p_3 = -(2+\lambda)p_0$. So, replacing in (8) and (15), we obtain

$$\begin{cases} A_1 &= -(\lambda-1)^2\|p_0\|^2 \\ A_2 &= -4(1+\lambda)^2\|p_0\|^2 \\ A_3 &= -(3+\lambda)^2\|p_0\|^2 \\ C &= -8(\lambda+1)\|p_0\|^2 \end{cases} \implies \begin{cases} X &= 2(5+10\lambda+\lambda^2)\|p_0\|^2 \\ Y &= -2(1+10\lambda+5\lambda^2)\|p_0\|^2 \\ Z &= 2(-11-6\lambda+\lambda^2)\|p_0\|^2 \end{cases} \tag{26}$$

and, after some simplifications, system (25) reduces to

$$2\lambda^3 + \lambda^2 + 4\lambda + 1 = 0, \quad 9\lambda^3 + 27\lambda^2 + 47\lambda + 29 = 0, \tag{27}$$

which has no real solutions.

Case 1.4 $p_0 \neq p_1$, $p_0 \neq 0$.

In this case, there exist real scalars λ and μ such that

$$p_1 = \lambda p_0, \quad p_2 = \mu p_0, \quad p_3 = -(1+\lambda+\mu)p_0. \tag{28}$$

Replacing in (8) we obtain

$$\begin{aligned} A_1 &= (\mu-1)(2\lambda-\mu-1)\|p_0\|^2, \quad & A_2 &= -4(\mu+1)(\lambda+\mu)\|p_0\|^2, \\ A_3 &= -(\mu+3)(2\lambda+\mu+1)\|p_0\|^2, \quad & C &= -4(\lambda+1)(\lambda+\mu)\|p_0\|^2. \end{aligned} \tag{29}$$

After some tedious calculations we obtain

$$\begin{aligned} X &= 2\left(\lambda(6\mu+2)+\mu^2+4\mu+3\right)\|p_0\|^2, \\ Y &= -2\left(\lambda(6\mu+2)+5\mu^2+4\mu-1\right)\|p_0\|^2, \\ Z &= -2\left(2\lambda(\mu+3)-\mu^2+4\mu+5\right)\|p_0\|^2, \end{aligned} \tag{30}$$

and the system of equations (25) reduces to

$$\begin{aligned} 2\lambda^3 - 2\lambda^2(3\mu+2) - \lambda\left(3\mu^2+6\mu+1\right) - 4\mu^3+\mu^2+4\mu+1 &= 0, \\ 2\lambda^3 + 2\lambda^2(6\mu+5) + \lambda\left(15\mu^2+26\mu+13\right) + 9\mu^3+12\mu^2+9\mu+4 &= 0. \end{aligned} \tag{31}$$

Let $f(\lambda,\mu)$ and $g(\lambda,\mu)$ denote the polynomials on the left hand side of the previous equations. To solve the system, we look for the common zeros of the bivariate polynomials f and g using Sylvester's resultant method. Sylvester's resultant is the determinant of the Sylvester matrix of the two polynomials. We then use the fact that a necessary and sufficient condition for f and g to have a common root is that its resultant vanishes, see Cox et al. (2007).

The Sylvester matrix for the polynomials in (31) is

$$S = \begin{bmatrix} 2 & -6\mu-4 & -3\mu^2-6\mu-1 & -4\mu^3+\mu^2+4\mu+1 & 0 & 0 \\ 0 & 2 & -6\mu-4 & -3\mu^2-6\mu-1 & -4\mu^3+\mu^2+4\mu+1 & 0 \\ 0 & 0 & 2 & -6\mu-4 & -3\mu^2-6\mu-1 & -4\mu^3+\mu^2+4\mu+1 \\ 2 & 12\mu+10 & 15\mu^2+26\mu+13 & 9\mu^3+12\mu^2+9\mu+4 & 0 & 0 \\ 0 & 2 & 12\mu+10 & 15\mu^2+26\mu+13 & 9\mu^3+12\mu^2+9\mu+4 & 0 \\ 0 & 0 & 2 & 12\mu+10 & 15\mu^2+26\mu+13 & 9\mu^3+12\mu^2+9\mu+4 \end{bmatrix} \tag{32}$$

Note that

$$\det(S) = 4096(\mu-1)(\mu+1)^6\left(17\mu^2+14\mu+1\right). \tag{33}$$

When $\mu = 1$, the system of equations (31) reduces to

$$\lambda^3 - 5\lambda^2 - 5\lambda + 1 = 0, \quad \lambda^3 + 11\lambda^2 + 27\lambda + 17 = 0, \tag{34}$$

whose solution is $\lambda = -1$. This yields the critical points described in item 2. of the theorem.

When $\mu = -1$, the system of equations (31) reduces to

$$\lambda^3 + \lambda^2 + \lambda + 1 = 0, \quad \lambda^3 - \lambda^2 + \lambda - 1 = 0, \tag{35}$$

which is easily seen to have only $\lambda = \pm\imath$ as solutions.

Finally, when $\mu = \frac{-7\pm4\sqrt{2}}{17}$, one gets $\lambda = \frac{-5\pm2\sqrt{2}}{17}$, which gives the description in item 3. of the theorem.

Case 1.5 $p_1 \neq p_0$, $p_0 = 0$

In this case, $p_1 \neq 0$ and instead of (28) and (29) we have

$$p_0 = 0, \quad p_2 = \mu p_1, \quad p_3 = -(1+\mu)p_1, \tag{36}$$

$$\begin{aligned} A_1 &= -\mu(\mu-2)\|p_1\|^2, \\ A_2 &= -4\mu(\mu+1)\|p_1\|^2, \\ A_3 &= -\mu(\mu+2)\|p_1\|^2, \\ C &= -4(\mu+1)\|p_1\|^2, \end{aligned} \tag{37}$$

from what follows

$$X = 2\mu(\mu+6)\|p_1\|^2, \quad Y = -2\mu(5\mu+6)\|p_1\|^2, \quad Z = 2\mu(\mu-2)\|p_1\|^2. \tag{38}$$

Then, system (25) reduces to

$$4\mu^3 + 3\mu^2 + 6\mu - 2 = 0, \quad 9\mu^3 + 15\mu^2 + 12\mu + 2 = 0, \tag{39}$$

which has no real solutions.

Concluding, the only critical points with collinear configurations are the ones stated in the theorem.

Case 2. Non-collinear configurations

Case 2.1 p_0 and p_1 are linearly independent.

In this case, we will get critical points that fulfill the descriptions of items 4., 5. and 6. of the theorem.

Let α and β be real scalars such that $p_2 = \alpha p_0 + \beta p_1$. Since $p_0 + p_1 + p_2 + p_3 = 0$, then $p_3 = -(1+\alpha)p_0 - (1+\beta)p_1$. Replacing these values of p_2 and p_3 in (25), we obtain the following system of linear equations

$$\begin{aligned} 0 &= (2\alpha-1)X + 2\alpha Y - Z - 2\alpha C, \\ 0 &= (2\beta-1)X + 2(1+\beta)Y + Z - 2(1+\beta)C, \\ 0 &= -\alpha X - 2(1+2\alpha)Y - (2+\alpha)Z - 2C, \\ 0 &= -(\beta-1)X - 2(1+2\beta)Y - (1+\beta)Z - 2C, \end{aligned} \tag{40}$$

that can be written as the matrix equation

$$\underbrace{\begin{bmatrix} 2\alpha-1 & 2\alpha & -1 & -2\alpha \\ 2\beta-1 & 2\beta+2 & 1 & -2(\beta+1) \\ -\alpha & -2(2\alpha+1) & -\alpha-2 & -2 \\ 1-\beta & -2(2\beta+1) & -\beta-1 & -2 \end{bmatrix}}_{N} \begin{bmatrix} X \\ Y \\ Z \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \det N = 32(2\alpha - \beta + \alpha\beta). \tag{41}$$

If $\det N \neq 0$, then $X = Y = Z = C = 0$, which implies that p_0, p_1, p_2, p_3 form the vertices of a regular quadrilateral, thus obtaining the critical point described by item 6. in the theorem.

If $\det N = 0$, then $\beta = \frac{2\alpha}{1-\alpha}$ (note that this is well defined, since for $\alpha = 1$, $\det N = 64$). Using the definition of X, Y, Z and C in (15) and (8), and assuming that $\alpha \neq 0$ (the case $\alpha = 0$ is considered later), the above system of equations is equivalent to

$$\begin{aligned} 0 &= 3A_1 - A_2 - A_3 - \tfrac{1+\alpha(3+\alpha-\alpha^2)}{4\alpha^2}C, \\ 0 &= -A_1 + 3A_2 - A_3 - \tfrac{1+\alpha(-3+\alpha+\alpha^2)}{4\alpha^2}C, \\ 0 &= -A_1 - A_2 + 3A_3 + \tfrac{1+\alpha(-1+\alpha+3\alpha^2)}{4\alpha^2}C, \end{aligned} \tag{42}$$

and the solution is thus given by

$$A_1 = \tfrac{1+\alpha(2+\alpha-2\alpha^2)}{8\alpha^2}C, \quad A_2 = \tfrac{(1-\alpha)(1+\alpha^2)}{8\alpha^2}C, \quad A_3 = \tfrac{\alpha(1-3\alpha^2)}{8\alpha^2}C. \tag{43}$$

Using the definition of A_1, A_2, A_3 and C in (8), the above is still equivalent to

$$\begin{aligned}
0 &= \|p_1\|^2(1+2\alpha+4\alpha^3-17\alpha^4+2\alpha^5) - p_0^\top p_1(\alpha-1)(2\alpha^5-17\alpha^4+6\alpha^3+4\alpha+1)\\
&\quad + \|p_0\|^2\alpha(\alpha-1)^2(-4\alpha^3+\alpha^2+4\alpha+1),\\
0 &= \|p_1\|^2(\alpha+1)(\alpha^4-18\alpha^3+2\alpha^2-2\alpha+1) - p_0^\top p_1\alpha(\alpha-1)^2(9\alpha^3+7\alpha^2+\alpha-1)\\
&\quad + \|p_0\|^2(\alpha-1)(\alpha^5-27\alpha^4-14\alpha^3-10\alpha^2+\alpha+1),\\
0 &= \|p_1\|^2(3\alpha^4-12\alpha^2+1) - p_0^\top p_1(\alpha-1)(3\alpha^4-10\alpha^3-12\alpha^2-10\alpha+1)\\
&\quad - \|p_0\|^2\alpha(\alpha-1)^2(5\alpha^2+8\alpha+5).
\end{aligned} \tag{44}$$

By computing the determinant of the coefficient matrix N_1 of system (44), it yields

$$\det N_1 = 32(\alpha-1)^3\alpha^4\left(\alpha^2+1\right)^3\left(5\alpha^2-3\right).$$

If this determinant does not vanish, the only solution of (44) is $\|p_1\| = \|p_0\| = p_0^\top p_1 = 0$, which contradicts the assumption that p_0 and p_1 are linearly independent. Therefore, it suffices to study the cases where $\alpha = \pm\sqrt{3/5}$. Next, we only consider the case $\alpha = \sqrt{3/5}$, since the study when $\alpha = -\sqrt{3/5}$ is analogous.

For $\alpha = \sqrt{3/5}$, the above system of equations reduces to

$$\|p_0\|^2 = -5p_0^\top p_1, \quad \|p_1\|^2 = \tfrac{\sqrt{15}-5}{5}p_0^\top p_1, \tag{45}$$

or, equivalently, to

$$\|p_0\|^2 = -5p_0^\top p_1, \quad \|p_1\|^2 = \tfrac{5-\sqrt{15}}{25}\|p_0\|^2. \tag{46}$$

The second equation tells us that p_1 is on the circle centered at the origin with radius $\frac{\sqrt{5-\sqrt{15}}}{5}\|p_0\|$, and since $p_0^\top p_1 = \|p_0\|\|p_1\|\cos\theta$, where θ is the angle between p_0 and p_1, we also get from (46) that $\cos\theta = \frac{-1}{\sqrt{5-\sqrt{15}}}$, which implies $\sin\theta = \sqrt{\frac{5-\sqrt{15}}{10}}$. Therefore, there are two possibilities for the localization of p_1, corresponding to applying a clockwise or an anticlockwise rotation of angle θ to the point $\frac{\sqrt{5-\sqrt{15}}}{5}p_0$. More precisely, writing $p_0 = \begin{bmatrix}x_0 & y_0\end{bmatrix}^\top$ and $p_1 = \begin{bmatrix}x_1 & y_1\end{bmatrix}^\top$,

$$\left[\begin{smallmatrix}x_1\\ y_1\end{smallmatrix}\right] = \frac{\sqrt{5-\sqrt{15}}}{5}\begin{bmatrix}\frac{-1}{\sqrt{5-\sqrt{15}}} & -\sqrt{\frac{5-\sqrt{15}}{10}}\\ \sqrt{\frac{5-\sqrt{15}}{10}} & \frac{-1}{\sqrt{5-\sqrt{15}}}\end{bmatrix}\left[\begin{smallmatrix}x_0\\ y_0\end{smallmatrix}\right] = \frac{1}{5}\begin{bmatrix}-x_0-\frac{5-\sqrt{15}}{\sqrt{10}}y_0\\ \frac{5-\sqrt{15}}{\sqrt{10}}x_0-y_0\end{bmatrix}, \tag{47}$$

and

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \frac{\sqrt{5-\sqrt{15}}}{5} \begin{bmatrix} \frac{-1}{\sqrt{5-\sqrt{15}}} & \sqrt{\frac{5-\sqrt{15}}{10}} \\ -\sqrt{\frac{5-\sqrt{15}}{10}} & \frac{-1}{\sqrt{5-\sqrt{15}}} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -x_0 + \frac{5-\sqrt{15}}{\sqrt{10}} y_0 \\ -\frac{5-\sqrt{15}}{\sqrt{10}} x_0 - y_0 \end{bmatrix}. \tag{48}$$

Since $\frac{5-\sqrt{15}}{\sqrt{10}} = \sqrt{4-\sqrt{15}}$, the above gives two solutions for the descriptions in item 5. of the theorem. The other two are obtained with $\alpha = -\sqrt{3/5}$.

Finally, we deal with the situation when $\alpha = 0$, which implies $\beta = 0$, and consequently $p_2 = 0$. In this case, the system of equations (41) reduces to

$$\underbrace{\begin{bmatrix} -1 & 0 & -1 & 0 \\ -1 & 2 & 1 & -2 \\ 0 & -2 & -2 & -2 \\ 1 & -2 & -1 & -2 \end{bmatrix}}_{N} \begin{bmatrix} X \\ Y \\ Z \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{49}$$

from where we get $Y = X = -Z$ and $C = 0$. Using identities (15) and (8), the former is equivalent to $A_1 = A_2$, $A_3 = 0$ and $C = 0$. And since in this case $p_3 = -p_0 - p_1$, we easily get the following conditions on p_1, in terms of p_0.

$$\|p_1\|^2 = \tfrac{1}{2}\|p_0\|^2, \quad \|p_1\|^2 = -p_0^\top p_1. \tag{50}$$

Using this and the fact that $p_0^\top p_1 = \|p_0\|\|p_1\|\cos\theta$, where θ is the angle between p_0 and p_1, we obtain $\cos\theta = -\sqrt{2}/2$, from where the two solutions in the description of item 4. will emerge. Indeed, writing $p_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}^\top$ and $p_1 = \begin{bmatrix} x_1 & y_1 \end{bmatrix}^\top$, the first equation in (50) tells us that the point p_1 must live on the circle centered at the origin with radius $\frac{1}{\sqrt{2}}\|p_0\|$, and so there are 2 possibilities for the localization of p_1, corresponding to applying a clockwise or an anticlockwise rotation of angle $\theta = \frac{3\pi}{4}$ to the point $\frac{1}{\sqrt{2}} p_0$. More precisely,

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\frac{3\pi}{4} & -\sin\frac{3\pi}{4} \\ \sin\frac{3\pi}{4} & \cos\frac{3\pi}{4} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -x_0 - y_0 \\ x_0 - y_0 \end{bmatrix}, \tag{51}$$

and

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos\frac{3\pi}{4} & \sin\frac{3\pi}{4} \\ -\sin\frac{3\pi}{4} & \cos\frac{3\pi}{4} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -x_0 + y_0 \\ -x_0 - y_0 \end{bmatrix}. \tag{52}$$

This completes the proof of case 1.1.

Case 2.2 p_0 and p_1 are linearly dependent, $p_0 \neq 0$.

Under these conditions, the manifold constraint also forces p_0 and p_2 to be linearly independent. We will show that there are no critical points in this case.

Let $p_1 = \lambda p_0$, for some $\lambda \in \mathbb{R}$, and $p_3 = -(1+\lambda)p_0 - p_2$. Replacing these values of p_1 and p_3 in (25), and using the fact that p_0 and p_2 are linearly independent, we obtain the following linear matrix equation.

$$\underbrace{\begin{bmatrix} -(\lambda+1) & 2\lambda & \lambda-1 & -2\lambda \\ 2 & 2 & 0 & -2 \\ \lambda & -2(\lambda+1) & -(\lambda+2) & -2(\lambda+1) \\ -1 & -4 & -1 & 0 \end{bmatrix}}_{N} \begin{bmatrix} X \\ Y \\ Z \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{53}$$

It happens that $\det N = -32\lambda$. If $\lambda \neq 0$, this equation only has the trivial solution $X = Y = Z = C = 0$, which implies $A_1 = A_2 = A_3 = C = 0$, meaning that the points p_0, p_1, p_2, and p_3 are the vertices of a regular quadrilateral. However, due to the assumption that p_0 and p_1 are linearly dependent, this configuration cannot happen. So, $\lambda = 0$, $p_1 = 0$, and the equation (53) is indeterminate. Simple calculations, show that in this case $X = C$, $Y = 0$ and $Z = -C$, or, equivalently, using identities (15), $A_1 = C/4$, $A_2 = 0$, $A_3 = -C/4$. Then, written in terms of the points, the solutions of (53) must satisfy

$$\begin{cases} \|p_0\|^2 - \|p_2\|^2 = -p_0^\top p_2 \\ p_0^\top p_2 + \|p_2\|^2 = 0 \\ -3\|p_0\|^2 - \|p_2\|^2 = 5p_0^\top p_2 \end{cases} \Leftrightarrow \begin{cases} \|p_0\|^2 = -2p_0^\top p_2 \\ \|p_2\|^2 = -p_0^\top p_2 \\ p_0^\top p_2 = 0 \end{cases} \Leftrightarrow p_0 = p_2 = 0, \tag{54}$$

which contradicts the assumption $p_0 \neq 0$. This case doesn't produce any critical points.

Case 2.3 p_0 and p_1 are linearly dependent, $p_1 \neq 0$.

Under these conditions, the manifold constraint also forces p_1 and p_2 to be linearly independent. We will show that in this case there are no critical points.

Let $p_0 = \lambda p_1$, for some $\lambda \in \mathbb{R}$, and $p_3 = -(1+\lambda)p_1 - p_2$. Although the procedure is similar to the previous case, we still need to do computations since the formulas involved are not symmetrical with respect to p_0 and p_1. Replacing the values of p_0 and p_3 in (25), and using the fact that p_1 and p_2 are linearly independent, we now obtain

$$\underbrace{\begin{bmatrix} -(\lambda+1) & 2 & 1-\lambda & -2 \\ 2 & 2 & 0 & -2 \\ 1 & -2(\lambda+1) & -(2\lambda+1) & -2(\lambda+1) \\ -1 & -4 & -1 & 0 \end{bmatrix}}_{N} \begin{bmatrix} X \\ Y \\ Z \\ C \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{55}$$

It also happens that $\det N = -32\lambda$. If $\lambda \neq 0$, this equation only has the trivial solution $X = Y = Z = C = 0$, which implies $A_1 = A_2 = A_3 = C = 0$, meaning that the points p_0, p_1, p_2, and p_3 are the vertices of a regular quadrilateral. However, as before, this contradicts the assumption that p_0 and p_1 are linearly dependent. So, consider that $\lambda = 0$ and consequently $p_0 = 0$. In this case, solutions of (55) satisfy $Y = -X$, $Z = 3X$, and $C = 0$, or, equivalently, using identities (15), $A_1 = 2A_2$, $A_3 = 3A_2$, $C = 0$. Taking into consideration that, in this case, $p_2 - p_3 = 2p_2 + p_1$, $p_1 - p_3 = 2p_1 + p_2$, $A_1 = -\|p_2\|^2 + 2p_1^\top p_2$, $A_2 = -4(\|p_2\|^2 + p_1^\top p_2)$, $A_3 = -\|p_2\|^2 - 2p_1^\top p_2$, and $C = -4(\|p_1\|^2 + p_1^\top p_2)$, the solutions written in terms of the points simplify to

$$\begin{cases} 7\|p_2\|^2 = -10p_1^\top p_2 \\ 11\|p_2\|^2 = -10p_1^\top p_2 \\ \|p_1\|^2 = -p_1^\top p_2 \end{cases} \Leftrightarrow p_1 = p_2 = 0, \tag{56}$$

contradicting the assumption $p_1 \neq 0$. So, this case does not produce further critical points.

□

Remark 2. *Note that the second configuration in the description 4. of Theorem 1 results from a reflection of the first configuration across the line generated by* p_0. *The same happens to the two configurations in the description 5. corresponding to the same value of the parameter* α. *Obviously, there are also two square configurations in the description 6. that are related by the same reflection that interchanges the points* p_1 *and* p_3. *In all cases, the reflection responsible for these linear transformation has the matrix representation*

$$R = \frac{1}{x_0^2+y_0^2}\begin{bmatrix} x_0^2-y_0^2 & 2x_0y_0 \\ 2x_0y_0 & -x_0^2+y_0^2 \end{bmatrix}. \tag{57}$$

The configurations in the description 5. corresponding to a particular value of the parameter α, *for instance* $\alpha = \sqrt{3/5}$, *are also related to those corresponding to* $\alpha = -\sqrt{3/5}$, *but in this case one needs a rotation matrix* $\Theta \in SO(6)$, *namely*

$$\Theta = \begin{bmatrix} 0 & 0 & I_2 \\ 0 & I_2 & 0 \\ I_2 & 0 & 0 \end{bmatrix}, \tag{58}$$

that keeps p_2 *invariant and interchanges* p_1 *and* p_3.

Remark 3. *There is an alternative proof of Theorem 1, we want to sketch here. It is, however, heavily supported by computations by means of the computer algebra system Mathematica, Version 14.2, cf. Wolfram Research, Inc..*

The set of critical points is defined by the zero set of the third order polynomials, with rational coefficients, in six real variables, see (7). *We have been successful in solving this system by straight forward geometric insight, as seen above. A more modern approach to tackle such a problem could be by using Gröbner bases. Due to well-known issues concerning complexity, see, for instance, Bose (1995) or Cox et al. (2007), usually one does not know in advance if a derived Gröbner basis, i.e. a new set of polymomials, often greater in number and possibly of much higher degree, even though generating the same ideal, is more suited for finding the solution set explicitly. For the current case, however, we have been successful in confirming the closed form solution and its geometric interpretation.*

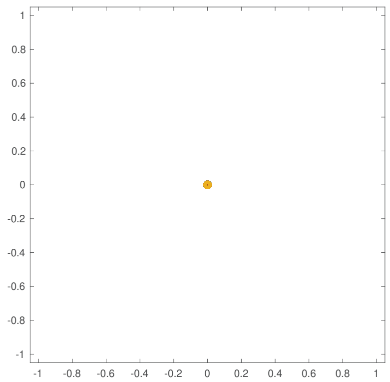
The setting is as follows:

- *Define six equations by the zero set of* (7) *and add the two affine equations, defining the constraint set* M.
- *We might first assume that the centroid of the four points is equal to zero, and also ignore the trivial case when all points coincide.*
- *Without loss of generality, we might secondly assume that one of the four points is nonzero and in particular equal to a standard basis vector in* $\mathbb{R}^2$. *The latter is correct, as the critical point condition is invariant under nonzero scaling, and also invariant under a collective rotation of all four points in* $\mathbb{R}^2$.

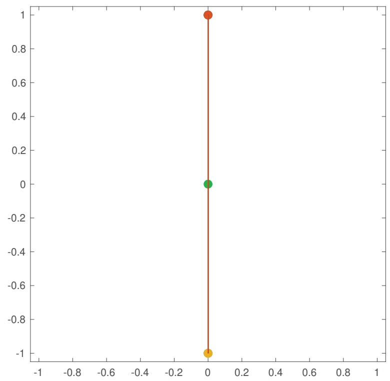
- *For the corresponding new polynomial system, i.e., eight equations, which generate the same ideal as* (7) *does on M, we compute a Gröbner basis. By our geometric insight, we know already that the critical point set is zero dimensional (i.e. a discrete set), ensuring by the so called Finite Theorem (Thm 2.1.2 in Dickenstein and Emiris (2010)), that such an approach is feasible.*

Many details have to be omitted here. However, we mention that the Gröbner basis we computed consists of ten equations, i.e. two more than the earlier eight, and the polynomials showing up are of 14-th degree. At first glance, this seems a potential disaster, but looking carefully to this result, one was even able to solve that system at least in principle using paper and pencil (including some minor Mathematica support), essentially by some sort of back substitution. Here, one also needs to use to full capacity, the ability of Mathematica to deal with radicals and root reducing methods from ordinary algebra. The amazing point we want to mention here is that all the solutions already found by geometric insight above, are recovered, actually the solutions are all given by algebraic numbers or, in other words, by explicit symbolic representations in the sense of computer algebra.

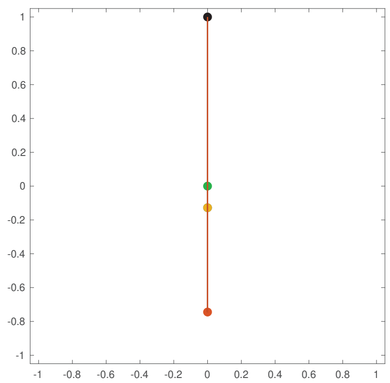
The figure below contains all the critical points (p_1, p_2, p_3) when p_0 is chosen to be $\begin{bmatrix} 0 & 1 \end{bmatrix}^\top$, up to a reflection across the y-axis or, for some configurations in 5., up to the rotation that interchanges p_1 and p_3.



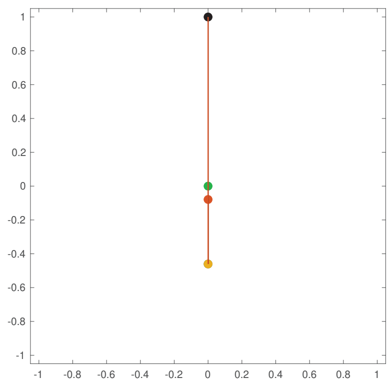
(a) $p_0 = p_1 = p_2 = p_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



(b) $p_0 = p_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $p_1 = p_3 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

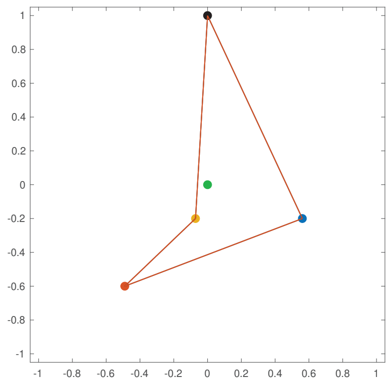


(c) $p_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $p_1 = p_3 = \frac{1}{17} \begin{bmatrix} 0 \\ 2\sqrt{2}-5 \end{bmatrix}$,
$p_2 = \frac{1}{17} \begin{bmatrix} 0 \\ -4\sqrt{2}-7 \end{bmatrix}$

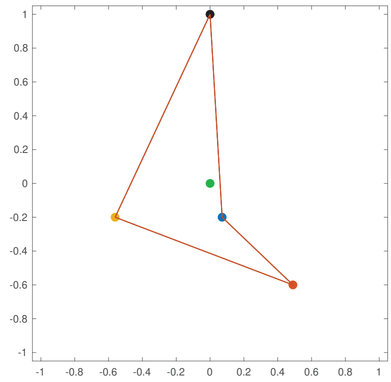


(d) $p_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $p_1 = p_3 = \frac{1}{17} \begin{bmatrix} 0 \\ -5-2\sqrt{2} \end{bmatrix}$,
$p_2 = \frac{1}{17} \begin{bmatrix} 0 \\ 4\sqrt{2}-7 \end{bmatrix}$

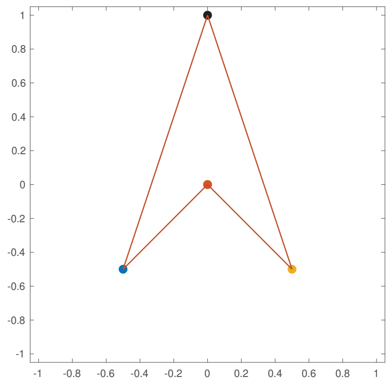
Figure 2: Collinear configurations of critical points described in Theorem 1



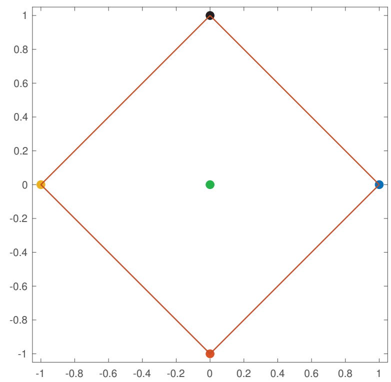
(a) $p_1 = \frac{1}{5}\begin{bmatrix}\sqrt{4+\sqrt{15}}\\ -1\end{bmatrix}$, $p_2 = -\frac{1}{5}\begin{bmatrix}\sqrt{6}\\ 3\end{bmatrix}$, $p_3 = -\frac{1}{5}\begin{bmatrix}\sqrt{4-\sqrt{15}}\\ 1\end{bmatrix}$



(b) $p_1 = \frac{1}{5}\begin{bmatrix}\sqrt{4-\sqrt{15}}\\ -1\end{bmatrix}$, $p_2 = \frac{1}{5}\begin{bmatrix}\sqrt{6}\\ -3\end{bmatrix}$, $p_3 = -\frac{1}{5}\begin{bmatrix}\sqrt{4+\sqrt{15}}\\ 1\end{bmatrix}$



(c) $p_1 = -\frac{1}{2}\begin{bmatrix}1\\ 1\end{bmatrix}$, $p_2 = \begin{bmatrix}0\\ 0\end{bmatrix}$, $p_3 = \frac{1}{2}\begin{bmatrix}1\\ -1\end{bmatrix}$



(d) $p_1 = \begin{bmatrix}1\\ 0\end{bmatrix}$, $p_2 = \begin{bmatrix}0\\ -1\end{bmatrix}$, $p_3 = \begin{bmatrix}-1\\ 0\end{bmatrix}$

Figure 3: Non-collinear configurations of critical points from Theorem 1 for $p_0 = \begin{bmatrix}0\\ 1\end{bmatrix}$

4. Classification of the critical points

From the Euclidean gradient given in (14), we can now proceed with the Hessian. Similarly to the gradient, we first compute the Euclidean Hessian $H_{\widehat{F}}$ (recall that $\widehat{F}$ is an extension of F to the embedding space), and then project it onto the tangent space to M to obtain the Riemannian Hessian H_F. To simplify notations, we also use the same letters for the matrix representation of the Hessians. To compute the Hessian of $\widehat{F}$, we proceed as follows, with $p = (p_1, p_2, p_3)$, and A_1, A_2, A_3 and C as defined in (37):

$$
\begin{aligned}
\mathrm{D}^2_{11}\widehat{F}(p) &= 6(p_1 - p_0)(p_1 - p_0)^\top + 2(p_1 - p_0)(p_2 - p_1)^\top \\
&\quad + 2(p_2 - p_1)(p_1 - p_0)^\top + 6(p_1 - p_2)(p_1 - p_2)^\top \\
&\quad + 2(p_1 - p_3)(p_1 - p_3)^\top - (2A_1 - 2A_2 - 2A_3 + C)I, \\
\mathrm{D}^2_{12}\widehat{F}(p) &= 2(p_1 - p_0)(p_1 - p_2)^\top - 2(p_1 - p_0)(p_2 - p_3)^\top \\
&\quad - 6(p_1 - p_2)(p_1 - p_2)^\top + 2(p_2 - p_1)(p_2 - p_3)^\top \\
&\quad + 2(p_3 - p_1)(p_2 - p_0)^\top + (3A_1 - A_2 - A_3)I, \\
\mathrm{D}^2_{13}\widehat{F}(p) &= 2(p_1 - p_0)(p_2 - p_3)^\top - 2(p_1 - p_0)(p_3 - p_0)^\top \\
&\quad - 2(p_2 - p_1)(p_2 - p_3)^\top + 2(p_2 - p_1)(p_3 - p_0)^\top \\
&\quad + 2(p_3 - p_1)(p_1 - p_3)^\top + CI, \\
\mathrm{D}^2_{22}\widehat{F}(p) &= 6(p_1 - p_2)(p_1 - p_2)^\top + 2(p_1 - p_2)(p_2 - p_3)^\top \\
&\quad - 6(p_3 - p_2)(p_2 - p_3)^\top - 2(p_3 - p_2)(p_1 - p_2)^\top \\
&\quad + 2(p_2 - p_0)(p_2 - p_0)^\top - (2A_1 + 2A_2 - 2A_3 - C)I, \\
\mathrm{D}^2_{23}\widehat{F}(p) &= -2(p_1 - p_2)(p_2 - p_3)^\top + 2(p_1 - p_2)(p_3 - p_0)^\top \\
&\quad + 6(p_3 - p_2)(p_2 - p_3)^\top + 2(p_3 - p_2)(p_3 - p_0)^\top \\
&\quad + 2(p_2 - p_0)(p_1 - p_3)^\top + (3A_2 - A_1 - A_3)I, \\
\mathrm{D}^2_{33}\widehat{F}(p) &= 6(p_2 - p_3)(p_2 - p_3)^\top + 2(p_2 - p_3)(p_3 - p_0)^\top \\
&\quad + 6(p_3 - p_0)(p_3 - p_0)^\top - 2(p_0 - p_3)(p_2 - p_3)^\top \\
&\quad + 2(p_1 - p_3)(p_1 - p_3)^\top + (2A_1 - 2A_2 - 2A_3 - C)I.
\end{aligned}
\tag{59}
$$

So, the matrix representation of the Euclidean Hessian is, at each $p \in \mathbb{R}^6$,

$$
H_{\widehat{F}}(p) = \begin{bmatrix} \mathrm{D}^2_{11}\widehat{F}(p) & \mathrm{D}^2_{12}\widehat{F}(p) & \mathrm{D}^2_{13}\widehat{F}(p) \\ (\mathrm{D}^2_{12}\widehat{F})^\top(p) & \mathrm{D}^2_{22}\widehat{F}(p) & \mathrm{D}^2_{23}\widehat{F}(p) \\ (\mathrm{D}^2_{13}\widehat{F})^\top(p) & (\mathrm{D}^2_{23}\widehat{F})^\top(p) & \mathrm{D}^2_{33}\widehat{F}(p) \end{bmatrix}. \tag{60}
$$

4.1. The Riemannian Hessian

Proposition 3. *The matrix representation of the Riemannian Hessian, in Euclidean coordinates of the embedding space $\mathbb{R}^6$, at every $p \in M$, is the 6×6 symmetric matrix with block structure,*

$$
H_F(p) = \frac{1}{9}\begin{bmatrix} H_{11} & H_{12} & -H_{11} - H_{12} \\ H_{12}^\top & -H_{12}^\top - H_{23} & H_{23} \\ -H_{11} - H_{12}^\top & H_{23}^\top & H_{11} + H_{12}^\top - H_{23}^\top \end{bmatrix}, \tag{61}
$$

where

$$
\begin{aligned}
H_{11} &= 40v_1v_1^\top + 56v_2v_2^\top + 24v_3v_3^\top - 8v_1v_2^\top - 8v_2v_1^\top - 4v_1v_3^\top - 4v_3v_1^\top - 12v_2v_3^\top - 12v_3v_2^\top \\
&\quad + (-22A_1 + 14A_2 + 10A_3 - 8C)I, \\
H_{12} &= -8v_1v_1^\top - 40v_2v_2^\top + 16v_1v_2^\top + 16v_2v_1^\top - 16v_1v_3^\top - 16v_3v_1^\top - 24v_2v_3^\top + 24v_3v_2^\top \\
&\quad + (26A_1 - 10A_2 - 14A_3 - 2C)I, \\
H_{23} &= -8v_1v_1^\top - 40v_2v_2^\top - 48v_3v_3^\top + 16v_1v_2^\top + 16v_2v_1^\top + 48v_3v_2^\top + 8v_1v_3^\top + 8v_3v_1^\top \\
&\quad + (-10A_1 + 26A_2 - 2A_3 - 2C)I,
\end{aligned}
\tag{62}
$$

A_1, A_2, A_3 *and* C *are defined in* (37)*, and* v_1, v_2, v_3 *are defined by*

$$
v_1 := p_1 - p_0, \quad v_2 := p_1 - p_2, \quad v_3 := p_1 - p_3. \tag{63}
$$

Proof. The Riemannian Hessian H_F is the restriction of the Euclidean Hessian to the tangent space. Consequently,

$$
H_F(p) = P^\perp_{T_pM}\ H_{\widehat{F}}(p)\ P^\perp_{T_pM}, \tag{64}
$$

where $P^\perp_{T_pM}$ is the matrix that defines the orthogonal projection onto T_pM. Taking into consideration that, according to (63),

$$
p_2 - p_0 = v_1 - v_2, \quad p_2 - p_3 = v_3 - v_2, \quad p_3 - p_0 = v_1 - v_3, \tag{65}
$$

one easily obtains the matrix representation of the Riemannian Hessian after some computations. □

4.2. Damped Newton method

Using the Riemannian Hessian, we can apply a damped Newton method (Algorithm 2) to numerically compute solutions for our minimization problem. In general, this method locally converges faster than steepest descent methods as it incorporates second-order information, allowing for more precise and efficient steps toward the minimum.

Algorithm 2: Damped Newton's method

Input : Initial point $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)})$, $\lambda^{(0)} > 0$, $\alpha^{(0)} > 0$, and tolerance tol
Output: Stationary point $p^* = (p_1^*, p_2^*, p_3^*)$

for $k = 0, 1, \ldots$ **do**
  Set $B^{(k)} = H_F(p^{(k)}) + \lambda^{(k)} I_n$;
  while $\mathrm{rcond}(B^{(k)}) < \mathsf{tol}$ **do**
    Increase $\lambda^{(k)} = 2\lambda^{(k)}$;
    Recompute $B^{(k)} = H_F(p^{(k)}) + \lambda^{(k)} I_n$;
  end
  Solve $d^{(k)}$ from $B^{(k)} d^{(k)} = -\nabla F(p^{(k)})$;
  Perform line search to update $\alpha^{(k)}$;
  Update $p^{(k+1)} = p^{(k)} + \alpha^{(k)} d^{(k)}$;
  Stop if $F(p^{(k)}) < \mathsf{tol}$ or $\|\nabla F(p^{(k)})\| < \mathsf{tol}$
end

Figure 4 illustrates the comparative performance of the steepest-descent method and the damped Newton method when applied to the same initial configuration of points, represented in each case as the vertices of a polygon with black edges. The points in the final configuration are joined by red edges. The results clearly indicate that the damped Newton method converges significantly faster, requiring fewer iterations to reach a solution than the steepest-descent approach. These illustrations corroborate the increased performance of damped Newton methods, reinforcing their benefits in optimization problems.

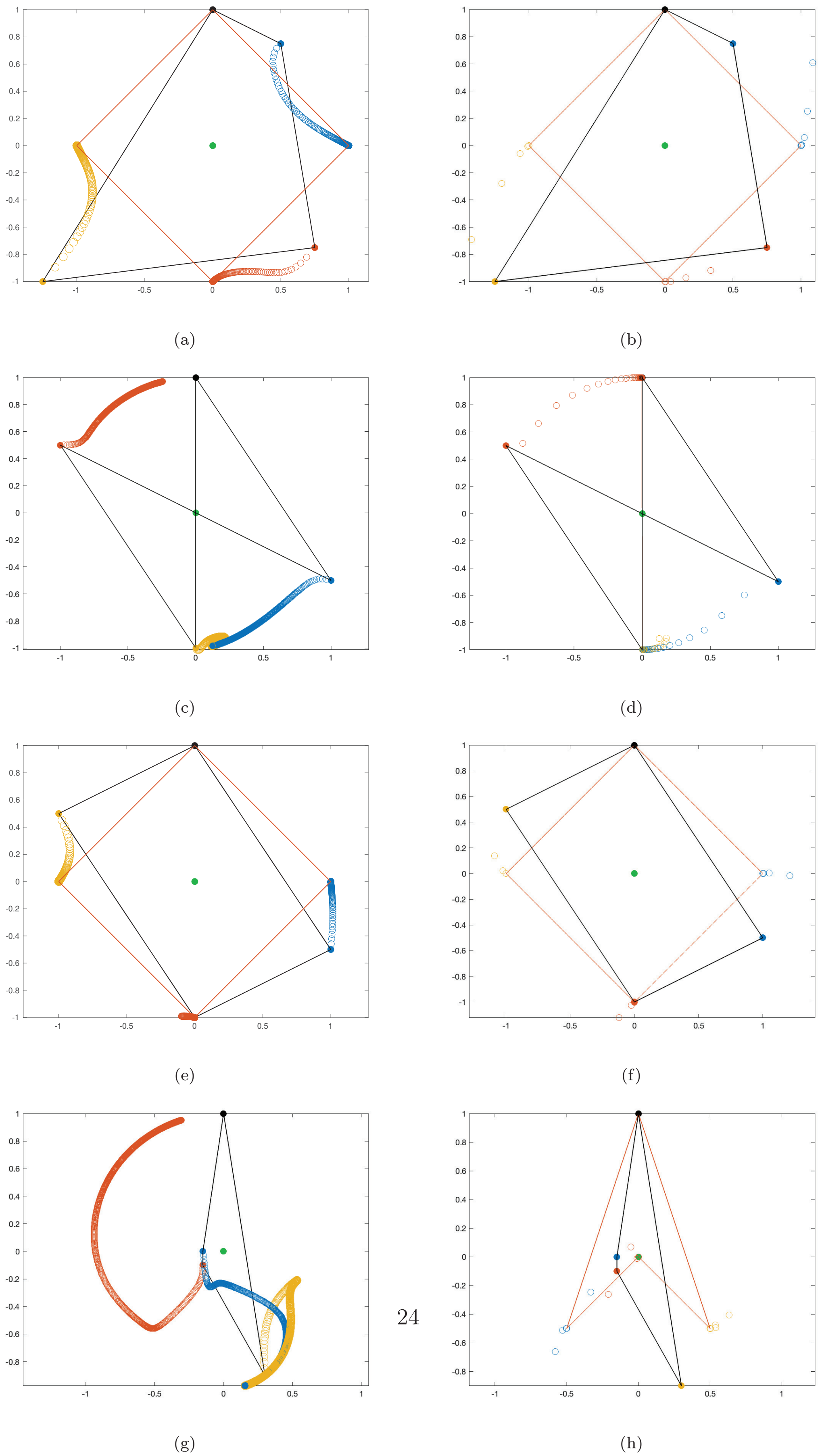


Figure 4: Configurations given by the steepest-descent algorithm (on the left) and by the damped Newton algorithm (on the right) for the same initial configuration

4.3. Classification of the critical points

We start with some remarks that will be used to substantially simplify the proof of the next theorem.

Remark 4. *Since a plane rotation can bring any point in the plane to the y-axis, to simplify notations we may assume that, for the configurations in Theorem 1,* $p_0 = \begin{bmatrix} 0 & y_0 \end{bmatrix}^\top$*, for some* $y_0 \in \mathbb{R}\backslash\{0\}$*. The nature (minimum, maximum, or saddle) of the critical points will not be affected by this coordinate change. Indeed, if* R *denotes the rotation matrix in* $SO(2)$ *that brings* p_0 *to the y-axis, then the rotation matrix*

$$\Theta = \begin{bmatrix} R & 0 & 0 \\ 0 & R & 0 \\ 0 & 0 & R \end{bmatrix} \in SO(6),$$

rotates the other three points (identified with a point in $\mathbb{R}^6$*) accordingly, and the Hessian matrix for the new configuration is* $\Theta\, H_F\, \Theta^\top$*, thus keeping the set of eigenvalues invariant, as well as its signature. A similar argument can be used to reduce the classification of critical points whenever two configurations are related by a rotation or a reflection, an orthogonal matrix) as well. Remark 2 already identified such configurations.*

Remark 5. *The nature of a critical point* p *depends on the signature of the Riemannian Hessian. By considering the relationship* (64) *between Euclidean and Riemannian Hessians, it becomes clear that to analyse the sign of the quadratic form* $u^\top H_F u$ *at* p *we can either consider a generic direction* $u \in \mathbb{R}^6$ *or a restricted direction* $u \in T_pM$*. In the proof of the Theorem 2, our choice for the direction* u *will be based on how it simplifies calculations.*

Since the Hessian matrix has a block structure, to compute $u^\top H_F u$ it is convenient to use the notation $u = \begin{bmatrix} u_1^\top & u_2^\top & u_3^\top \end{bmatrix}^\top$ for a vector in $u \in \mathbb{R}^6$, where $u_i \in \mathbb{R}^2$, $i = 1, 2, 3$.

Theorem 2. *The critical points of the function* F *on* M*, characterized in Theorem 1, are classified as:*

(a) The critical points (p_1, p_2, p_3) *corresponding to the descriptions 1., 2. and 6. of Theorem 1 are global minima;*

(b) The critical points (p_1, p_2, p_3) *corresponding to the descriptions 3., 4. and 5. of Theorem 1 are saddle points.*

Proof. The statement is obvious for $p_0 = p_1 = p_2 = p_3 = 0$.

When $p_2 = p_0$, $p_1 = p_3 = -p_0$, we have $v_1 = v_2$, $v_3 = 0$, $A_1 = A_2 = A_3 = C = 0$, and the Hessian matrix reduces to

$$H_F = \frac{16}{9} \begin{bmatrix} 5v_1v_1^\top & -v_1v_1^\top & -4v_1v_1^\top \\ -v_1v_1^\top & 2v_1v_1^\top & -v_1v_1^\top \\ -4v_1v_1^\top & -v_1v_1^\top & 5v_1v_1^\top \end{bmatrix} = \frac{16}{9} \underbrace{\begin{bmatrix} 5 & -1 & -4 \\ -1 & 2 & -4 \\ -4 & -1 & 5 \end{bmatrix}}_{S} \otimes (v_1v_1^\top). \tag{66}$$

We now consider the eigenvalues of certain matrices. The spectrum of S is $\sigma(S) = \{9, 3, 0\}$, $\sigma(v_1 v_1^\top) = \{\|v_1\|^2, 0\}$, and $\sigma(S \otimes v_1 v_1^\top) = \{9\|v_1\|^2, 3\|v_1\|^2, 0, 0, 0, 0\}$, consequently H_F has 2 positive eigenvalues and 4 eigenvalues equal to zero. Together with the fact that in this case the function F attains its minimum value (zero), it tells us that the critical point for the description 2. in Theorem 1 is a global minimum.

Now, when the four points are at the vertices of a regular quadrilateral, we have $v_3 = v_1 + v_2$, $A_1 = A_2 = A_3 = C = 0$, and the main blocks of the Hessian matrix simplify to

$$\begin{aligned} H_{11} &= 56(v_1 v_1^\top + v_2 v_2^\top) \\ H_{12} &= H_{23} = -8(5v_1 v_1^\top + 5v_2 v_2^\top - 3v_1 v_2^\top + 3v_2 v_1^\top) \end{aligned} \tag{67}$$

For the critical point $p = (p_1, p_2, p_3)$ of this configuration, we take $u \in T_p M$, so that $u_3 = -u_1 - u_2$. After simplification, the quadratic form becomes

$$\begin{aligned} u^\top H_F u = \tfrac{1}{9}\big(& u_1^\top 2(2H_{11} + H_{12}^\top)u_1 + u_2^\top (H_{11} - 3H_{12}^\top - H_{12})u_2 \\ &+ 2u_1^\top (2H_{11} + 2H_{12} - H_{12}^\top)u_2\big). \end{aligned} \tag{68}$$

After further simplification, one gets

$$\begin{aligned} u^\top H_F u &= \tfrac{72}{9}\big(2(u_1^\top v_1)^2 + 2(u_1^\top v_2)^2 + 2(u_1^\top v_1)(u_2^\top v_2) + 2(u_1^\top v_2)(u_2^\top v_2) \\ &\quad + 2(u_1^\top v_1)(u_2^\top v_2) - 2(u_1^\top v_2)(u_2^\top v_1) + 3(u_2^\top v_1)^2 + 3(u_2^\top v_2)^2\big) \\ &= 72\big((u_1^\top v_1 + u_2^\top v_1)^2 + (u_1^\top v_1 + u_2^\top v_2)^2 + (u_1^\top v_2 + u_2^\top v_2)^2 \\ &\quad + (u_1^\top v_2 - u_2^\top v_1)^2 + (u_1^\top v_2 + u_2^\top v_2)^2 + (u_2^\top v_2)^2 + (u_2^\top v_1)^2\big). \end{aligned} \tag{69}$$

Since all summands are pure squares, the quadratic form is clearly non-negative, thus proving that the critical point for the regular quadrilateral described in Theorem 1 is a local minimum. Actually, since F takes the value zero at this configuration, it is a global minimum.

Next we prove that the remaining critical points are saddle points. The strategy to prove that a critical point is a saddle will be to show that there exist two directions $u, w \in \mathbb{R}^6$ such that the quadratic forms $u^\top H_F u$ and $w^\top H_F w$ have opposite signs at that point.

For the critical points described by item 3. in Theorem 1, we have $p_1 = p_3 = \lambda p_0$, $p_2 = -(1 + 2\lambda)p_0$, where $\lambda = -\frac{5 \pm 2\sqrt{2}}{17}$. In this case,

$$A_1 = A_2 = -8\lambda(\lambda + 1)\|p_0\|^2, \quad A_3 = 0, \quad C = 4(\lambda + 1)^2\|p_0\|^2. \tag{70}$$

Consider the direction $u \in \mathbb{R}^6$ with components $u_1 = v_1$, and $u_2 = u_3 = 0$. Then,

$$\begin{aligned} u^\top H_F u &= \tfrac{1}{9}\big((-8A_1 - 8C)v_1^\top v_1 + 40(v_1^\top v_1)^2 + 56(v_1^\top v_2)^2 - 16(v_1^\top v_1)(v_1^\top v_2)\big) \\ &= \tfrac{16}{9}(\lambda - 1)^2(33\lambda^2 + 18\lambda + 5))\|p_0\|^4 > 0. \end{aligned} \tag{71}$$

Now consider the direction $w \in \mathbb{R}^6$, with $w_2 = w_3 = 0$ and $w_1 \in \mathbb{R}^2$ such that $w_1^\top v_1 = 0$. Using these conditions, we obtain

$$w^\top H_F w = \tfrac{1}{9}(-8A_1 - 8C)w_1^\top w_1 = -\tfrac{32}{9}(1-\lambda^2)\|p_0\|^2\|w_1\|^2 < 0. \tag{72}$$

Therefore, the critical points described by item 3. in Theorem 1 are saddle points.

For the critical points described by item 4. in Theorem 1, we consider $p_2 = 0$, $p_0 = \begin{bmatrix} x_0 & y_0 \end{bmatrix}^\top$, $p_1 = \frac{1}{2}\begin{bmatrix} y_0 - x_0 & x_0 - y_0 \end{bmatrix}^\top$, $p_3 = -p_0 - p_1 - p_2$. To simplify notations, we consider the vector $r = \begin{bmatrix} -y_0 & x_0 \end{bmatrix}^\top$ and rewrite the points p_i also in terms of this vector, which is orthogonal to p_0 and satisfies $\|r\| = \|p_0\|$. So, $p_1 = -\frac{1}{2}(p_0 + r)$, $p_2 = 0$ and $p_3 = -\frac{1}{2}(p_0 - r)$. Moreover,

$$v_1 = -\tfrac{1}{2}(3p_0 + r), \quad v_2 = -\tfrac{1}{2}(p_0 + r), \quad v_3 = -r; \quad A_1 = A_2 = 2\|p_0\|^2, \quad A_3 = C = 0,$$

and the main blocks of the Hessian matrix reduce to

$$\begin{aligned} H_{11} &= 92p_0p_0^\top + 28rr^\top + 24p_0r^\top + 24rp_0^\top - 16\|p_0\|^2 I, \\ H_{12} &= -4p_0p_0^\top - 20rr^\top - 36p_0r^\top - 12rp_0^\top + 32\|p_0\|^2 I, \\ H_{23} &= -4p_0p_0^\top - 20rr^\top + 12p_0r^\top + 36rp_0^\top + 32\|p_0\|^2 I. \end{aligned}$$

Now consider those directions $u, w \in \mathbb{R}^6$, where $u_1 = p_0$, $u_2 = u_3 = 0$ and $w_2 = p_0$, $w_1 = w_3 = 0$. A simple calculation, taking into account that r is orthogonal to p_0, gives

$$u^\top H_F u = p_0^\top H_{11} p_0 = \tfrac{76}{9}\|p_0\|^4, \quad w^\top H_F w = p_0^\top(-H_{12}^\top - H_{23})p_0 = -\tfrac{56}{9}\|p_0\|^4.$$

So, the critical points described by item 4. in Theorem 1 are saddle points.

Finally, we consider the configurations given at item 5. of Theorem 1. According to Remark 4, it is enough to choose only one configuration and also take $x_0 = 0$. Defining $\mu := \sqrt{(4 - \sqrt{15})}$ and $\nu := \sqrt{6}$ (to simplify notations), we consider

$$p_0 = \tfrac{y_0}{5}\begin{bmatrix} 0 \\ 5 \end{bmatrix}, \quad p_1 = \tfrac{y_0}{5}\begin{bmatrix} \mu \\ -1 \end{bmatrix}, \quad p_2 = \tfrac{y_0}{5}\begin{bmatrix} \nu \\ -3 \end{bmatrix}, \quad p_3 = \tfrac{y_0}{5}\begin{bmatrix} -\mu-\nu \\ -1 \end{bmatrix}, \tag{73}$$

which correspond to $\alpha = \sqrt{\frac{3}{5}}$. In this case,

$$\begin{aligned} A_1 &= \tfrac{y_0^2}{25}(2\mu\nu - \nu^2 + 32), \\ A_2 &= \tfrac{-4y_0^2}{25}(\mu\nu + \nu^2 - 8), \\ A_3 &= -\tfrac{y_0^2}{25}\nu(2\mu + \nu), \\ C &= \tfrac{-4y_0^2}{25}(\mu^2 + \mu\nu - 16) = \tfrac{12y_0^2}{5}. \end{aligned} \tag{74}$$

Let us consider a direction $u \in \mathbb{R}^6$, with $u_1 = p_0$, $u_2 = 0$ and $u_3 = -u_1$. After simplifications, this yields

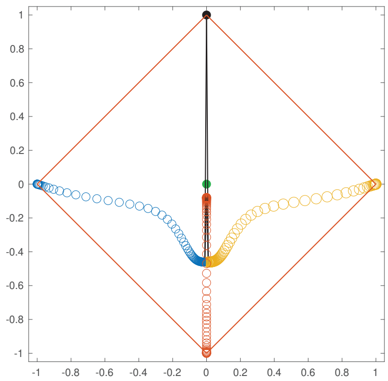
$$u^\top H_F u = \tfrac{1}{9} p_0^\top (4H_{11} + 3H_{12}^\top - H_{23}^\top) p_0 = \tfrac{144}{9} \big((p_0^\top v_1)^2 + (p_0^\top v_2)^2 - \tfrac{3}{5} y_0^2 (p_0^\top p_0) \big) = 16 y_0^4 > 0. \quad (75)$$

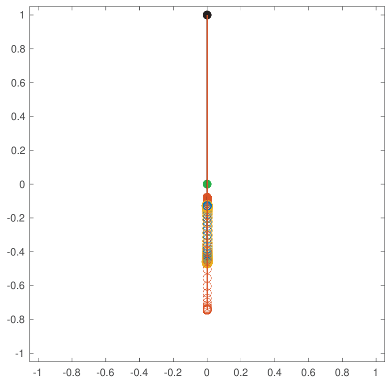
Choosing a direction w, with $w_1 = p_3 - p_0$, $w_2 = -u_1$, and $w_3 = 0$, and performing some lengthy computations, one obtains

$$w^\top H_F w = \tfrac{1}{9} w_1^\top (H_{11} - 3H_{12}^\top - H_{23}) w_1 = \tfrac{8}{125} (245 - 76\sqrt{15}) y_0^4 < 0. \quad (76)$$

So, the configurations described in 5. are saddle points. The proof is now complete. □

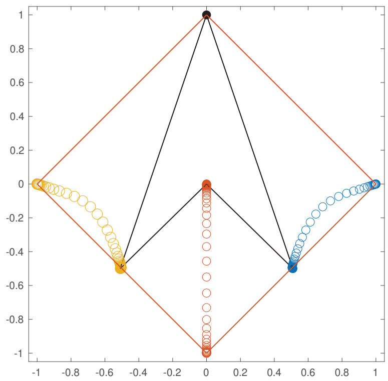
Figure 5 illustrates the behavior of the points p_1, p_2, p_3 in a neighborhood of the saddle points described in Theorem 2. The illustrations were generated using the steepest-descent algorithm. Depending on the deviation of the initial configuration from the critical points in figures 2 and 3, we observe either convergence to a regular quadrilateral configuration or to a collinear arrangement where $p_2 = p_0$ and $p_1 = p_3 = -p_0$, both of which are local minima of the functional F.

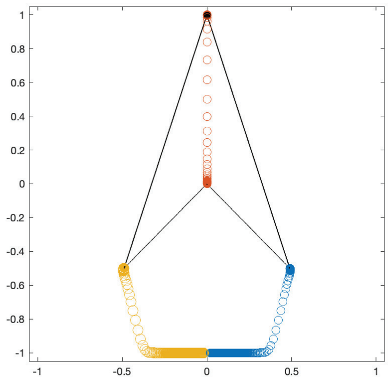




(a) $p_1 = \begin{bmatrix} -0.01 \\ -\frac{2\sqrt{2}+5}{17} \end{bmatrix}$, $p_2 = \begin{bmatrix} 0 \\ \frac{4\sqrt{2}-7}{17} \end{bmatrix}$, $p_3 = \begin{bmatrix} 0.01 \\ -\frac{2\sqrt{2}+5}{17} \end{bmatrix}$

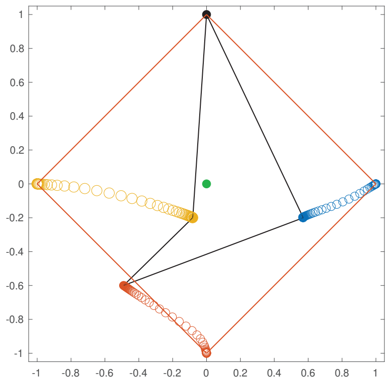
(b) $p_1 = \begin{bmatrix} 0 \\ -\frac{2\sqrt{2}+5}{17}+0.01 \end{bmatrix}$, $p_2 = \begin{bmatrix} 0 \\ \frac{4\sqrt{2}-7}{17} \end{bmatrix}$, $p_3 = \begin{bmatrix} 0 \\ -\frac{2\sqrt{2}+5}{17}-0.01 \end{bmatrix}$

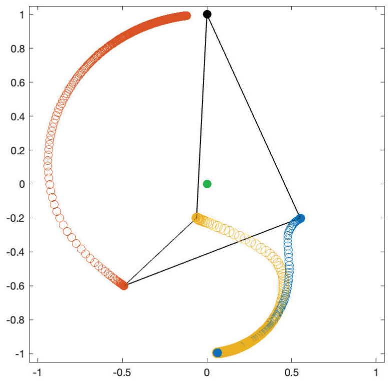




(c) $p_1 = \begin{bmatrix} \frac{1}{2}+0.01 \\ -\frac{1}{2} \end{bmatrix}$, $p_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $p_3 = \begin{bmatrix} -\frac{1}{2}-0.01 \\ -\frac{1}{2} \end{bmatrix}$

(d) $p_1 = \begin{bmatrix} \frac{1}{2}-0.01 \\ -\frac{1}{2} \end{bmatrix}$, $p_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $p_3 = \begin{bmatrix} -\frac{1}{2}+0.01 \\ -\frac{1}{2} \end{bmatrix}$



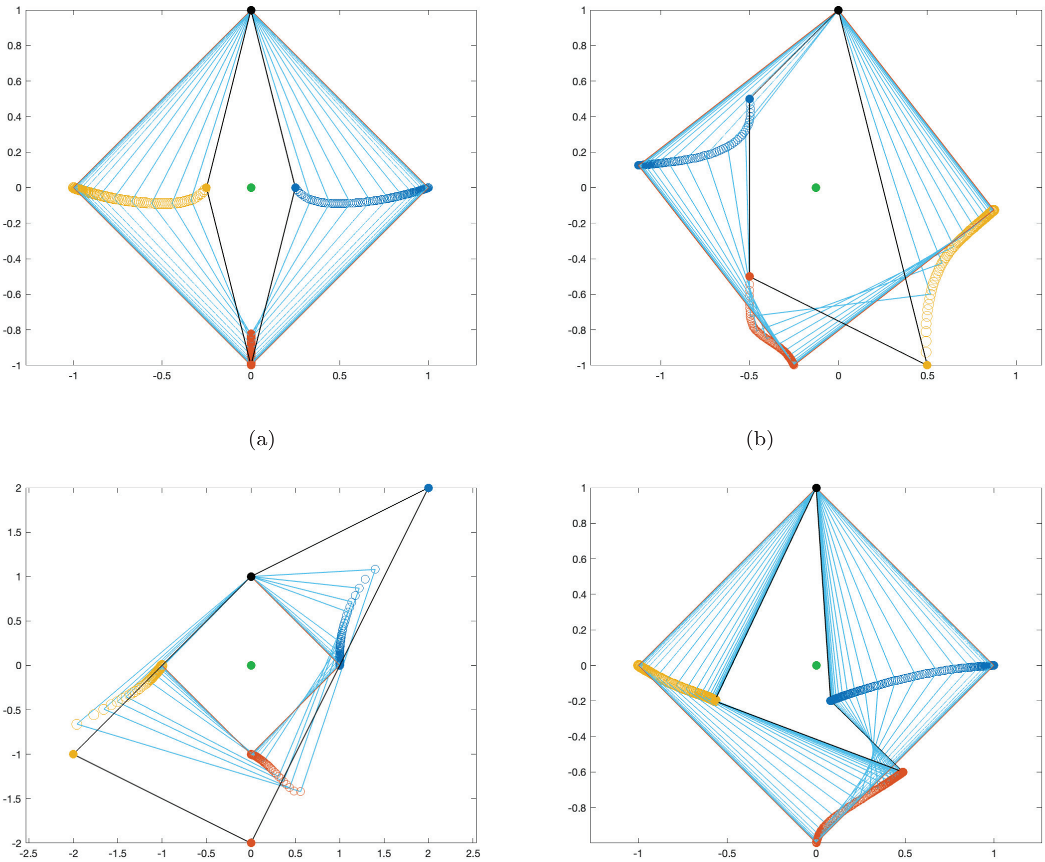


(e) $p_1 = \frac{1}{5}\begin{bmatrix} \sqrt{4+\sqrt{15}}+0.05 \\ -1 \end{bmatrix}$, $p_2 = \frac{1}{5}\begin{bmatrix} -\sqrt{6} \\ -3 \end{bmatrix}$, $p_3 = \frac{1}{5}\begin{bmatrix} -\sqrt{4-\sqrt{15}}-0.05 \\ -1 \end{bmatrix}$

(f) $p_1 = \frac{1}{5}\begin{bmatrix} \sqrt{4+\sqrt{15}}-0.05 \\ -1 \end{bmatrix}$, $p_2 = -\frac{1}{5}\begin{bmatrix} \sqrt{6} \\ 3 \end{bmatrix}$, $p_3 = \frac{1}{5}\begin{bmatrix} -\sqrt{4-\sqrt{15}}+0.05 \\ -1 \end{bmatrix}$

Figure 5: Behavior in a neighborhood of a saddle point, when $p_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

We finalize this paper with some pictures that illustrate how quadrilateral configurations are continuously deformed into the vertices of a square with the same centroid. These configurations were generated using the steepest descent algorithm. The quadrilaterals observed in the pictures correspond to specific iterations of the process, highlighting its progressive deformation towards the final square shape.



(a) (b)

(c) (d)

Figure 6: Morphing a quadrilateral configuration into a square with the same centroid, using the steepest descent algorithm

Acknowledgements

The work of L. Machado and F. Silva Leite has been supported by Fundação para a Ciência e Tecnologia (FCT) under the project UIDB/00048/2020 (`https://doi.org/10.`

54499/UIDB/00048/2020). K. Hüper has been supported by the German Federal Ministry of Education and Research (BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA).

References

Absil, P.A., Mahony, R., Sepulchre, R., 2008. Optimization algorithms on matrix manifolds. Princeton University Press.

Alonso-Mora, J., Baker, S., Rus, D., 2017. Multi-robot formation control and object transport in dynamic environments via constrained optimization. The International Journal of Robotics Research 36, 1000–1021. URL: https://doi.org/10.1177/0278364917719333, doi:10.1177/0278364917719333.

Baez, J., 2020. Platonic solids in all dimensions. URL: https://math.ucr.edu/home/baez/platonic.html.

Bezdek, K., 2010. Classical topics in discrete geometry. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York. URL: https://doi.org/10.1007/978-1-4419-0600-7, doi:10.1007/978-1-4419-0600-7.

Bezdek, K., 2013. Lectures on sphere arrangements—the discrete geometric side. volume 32 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON. URL: https://doi.org/10.1007/978-1-4614-8118-8, doi:10.1007/978-1-4614-8118-8.

Bose, N.K., 1995. Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory. Springer Netherlands, Dordrecht. pp. 89–127. URL: https://doi.org/10.1007/978-94-017-0275-1_4, doi:10.1007/978-94-017-0275-1_4.

Boumal, N., 2023. An introduction to optimization on smooth manifolds. Cambridge University Press. URL: https://www.nicolasboumal.net/book, doi:10.1017/9781009166164.

Bullo, F., Cortés, J., Martínez, S., 2009. Distributed Control of Robotic Networks. Applied Mathematics Series, Princeton University Press. Electronically available at http://coordinationbook.info.

Conway, J., Sloane, N., 1999. Sphere packings, lattices and groups. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov. 3rd ed. Grundlehren der Mathematischen Wissenschaften. 290. New York, NY: Springer. lxxiv, 703 p. .

Conway, J.H., Hardin, R.H., Sloane, N.J.A., 1996. Packing lines, planes, etc.: packings in Grassmannian spaces. Experiment. Math. 5, 139–159.

Cox, D.A., Little, J.B., O'Shea, D., 2007. Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra. 3rd ed., Springer, New York.

Dickenstein, A., Emiris, I.Z., 2010. Solving Polynomial Equations: Foundations, Algorithms, and Applications. 1st ed., Springer Publishing Company, Incorporated.

Fejes Tóth, L., Fejes Tóth, G., Kuperberg, W., 2023. Lagerungen. Arrangements in the plane, on the sphere, and in space. Translated from the German. Cham: Springer. doi:10.1007/978-3-031-21800-2.

Machado, L., Hüper, K., Krakowski, K., Silva Leite, F., 2024. Spherical triangular configurations with invariant geometric mean. Linear Algebra and its Applications URL: https://www.sciencedirect.com/science/article/pii/S0024379524002714, doi:https://doi.org/10.1016/j.laa.2024.06.017.

Mesbahi, M., Egerstedt, M., 2010. Graph Theoretic Methods in Multiagent Networks. Princeton University Press, Princeton. URL: https://doi.org/10.1515/9781400835355, doi:doi:10.1515/9781400835355.

Nebe, G., Rains, E.M., Sloane, N.J.A., 2006. Self-dual codes and invariant theory. Berlin: Springer. doi:10.1007/3-540-30731-1.

Pennec, X., 2006. Intrinsic statistics on riemannian manifolds: Basic tools for geometric measurements. Journal of Mathematical Imaging and Vision 36, 127–154.

Wolfram Research, Inc., . Mathematica, Version 14.2. URL: https://www.wolfram.com/mathematica. champaign, IL, 2024.

Zong, C., 1999. Sphere packings. Universitext. New York, NY: Springer.