The Fermat-Torricelli Problem and Weiszfeld's Algorithm in the Light of Convex Analysis

Boris S. Mordukhovich[1], Nguyen Mau Nam[2]

Dedicated to Christiane Tammer with great respect

Abstract. In the early 17th century, Pierre de Fermat proposed the following problem: given three points in the plane, find a point such that the sum of its Euclidean distances to the three given points is minimal. This problem was solved by Evangelista Torricelli and was named the *Fermat-Torricelli problem*. A more general version of the Fermat-Torricelli problem asks for a point that minimizes the sum of the distances to a finite number of given points in $\mathbb{R}^n$. This is one of the main problems in location science. In this paper we revisit the Fermat-Torricelli problem from both theoretical and numerical viewpoints using some ingredients of convex analysis and optimization.

Key words. Fermat-Torricelli problem; Weiszfeld's algorithm; Convex analysis.

AMS subject classifications. 49J52, 49J53, 90C31

1 Introduction

The Fermat-Torricelli problem asks for a point that minimizes the sum of the distances to three given points in the plane. This problem was proposed by Fermat and solved by Torricelli. Torricelli's solution states as follows: if one of the angles of the triangle formed by the three given points is greater than or equal to 120°, the corresponding vertex is the solution of the problem. Otherwise, the solution is a unique point inside of the triangle formed by the three points such that each side is seen at an angle of 120°. The first numerical algorithm for solving the general Fermat-Torricelli problem was introduced by Weiszfeld in 1937 [22]. The assumptions that guarantee the convergence along with the proof were given by Kuhn in 1972. Kuhn also pointed out an example in which Weiszfeld's algorithm fails to converge; see [11]. The Fermat-Torricelli problem has attracted great attention from many researchers not only because of its mathematical beauty, but also because of its important applications to the field of facility location. Many generalized versions of the Fermat-Torricelli and several new algorithms have been introduced to deal with generalized Fermat-Torricelli problems as well as to improve Weiszfeld's algorithm; see, e.g., [2, 4, 12, 13, 14, 15, 20, 21]. The problem has also been revisited several times from different viewpoints as in [5, 6, 17, 23] and the references therein. Among many researchers working on the field of facility location, Christiane Tammer is one of the pioneers with important contributions to the theory, algorithms, and software for solving facility location problems; see [1, 9, 10, 16] and their bibliographies.

[1]Department of Mathematics, Wayne State University, Detroit, Michigan 48202, USA (boris@math.wayne.edu). Research of this author was partly supported by the USA National Science Foundation under grants DMS-1007132 and DMS-1512846, by the USA Air Force Office of Scientific Research grant #15RT0462, and by the Australian Research Council under Discovery Project DP-190100555.

[2]Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR 97207, USA (mnn3@pdx.edu). Research of this author was partly supported by the USA National Science Foundation under grant DMS-1716057.

The main goal of this paper is to provide easy access to the problem from both theoretical and numerical aspects by using some tools of convex analysis. These tools are presented in the paper with simple proofs that are understandable for students with basic background in introduction to elementary analysis.

The paper is organized as follows. In Section 2 we establish the existence and uniqueness of the optimal solution for the Fermat-Torricelli problem generated by a finite number of points. We also present proofs of some properties of the optimal solution and derive its explicit representations for the case of three points using the *convex subdifferential*. Various advantages of employing convex analysis when solving the Fermat-Torricelli problem has been revealed in [8, 12, 13, 19]. Section 3 is devoted to revisiting Kuhn's proof of the convergence of Weiszfeld's algorithm. In this section we mainly follow the line of proving the convergence given by Kuhn [11], while with involving new ingredients from convex analysis to replace some technical tools in order to make the proof more clear.

Throughout the paper the symbol $\mathbb{B}$ denotes the closed unit ball of $\mathbb{R}^n$, and $\mathbb{B}(\bar{x}; r)$ stands for the closed ball centered at $\bar{x}$ with radius r.

2 Elements of Convex Analysis and Properties of Solutions

In this section we review several important concepts of convex analysis to study the classical Fermat-Torricelli problem as well as its general version mentioned in Section 1. Then we present elementary proofs for some properties of optimal solutions of the problem. More details of convex analysis can be found in the fundamental monograph [18].

Let $\| \cdot \|$ be the Euclidean norm in $\mathbb{R}^n$. Given a finite number of distinct points a_i for $i = 1, \ldots, m$ in $\mathbb{R}^n$, define the function

$$\varphi(x) := \sum_{i=1}^{m} \|x - a_i\|. \tag{2.1}$$

The mathematical model of the Fermat-Torricelli problem is as follows:

$$\text{minimize } \varphi(x) \text{ subject to } x \in \mathbb{R}^n. \tag{2.2}$$

The *weighted version* of this problem can be formulated and treated by a similar way.

Let $f \colon \mathbb{R}^n \to \mathbb{R}$ be a real-valued function. The *epigraph* of f is a subset of $\mathbb{R}^n \times \mathbb{R}$ defined by

$$\text{epi } f := \big\{ (x, \alpha) \in \mathbb{R}^{n+1} \,\big|\, x \in \mathbb{R}^n \ \text{ and } \ \alpha \geq f(x) \big\}.$$

The function f is called *convex* if we have the inequality

$$f\big(\lambda x + (1 - \lambda)y\big) \leq \lambda f(x) + (1 - \lambda)f(y) \ \text{ for all } \ x, y \in \mathbb{R}^n \ \text{ and } \ \lambda \in (0, 1).$$

If this equality becomes strict for $x \neq y$, we say that f is *strictly convex*. It is easy to check by the definitions that f is a convex function on $\mathbb{R}^n$ if and only if its epigraph is a convex set in $\mathbb{R}^{n+1}$. We obviously have that the function φ from (2.1) is convex on $\mathbb{R}^n$.

Proposition 2.1 *Let* $f \colon \mathbb{R}^n \to \mathbb{R}$ *be a convex function. Then* f *has a local minimum at* $\bar{x}$ *if and only if* f *has an absolute/global minimum at this point.*

Proof: We only need to prove the "only if" implication since the converse is trivial. Suppose that f has a local minimum at $\bar{x}$. Then there exists a number $\delta > 0$ with

$$f(u) \geq f(\bar{x}) \ \text{ for all } \ u \in \mathbb{B}(\bar{x}; \delta).$$

For any $x \in \mathbb{R}$ it follows that $x_k := \left(1 - \dfrac{1}{k}\right)\bar{x} + \dfrac{1}{k}x \to \bar{x}$ as $k \to \infty$. Thus we get $x_k \in \mathbb{B}(\bar{x}; \delta)$ when k is sufficiently large. It follows furthermore that

$$f(\bar{x}) \leq f(x_k) \leq \left(1 - \frac{1}{k}\right)f(\bar{x}) + \frac{1}{k}f(x),$$

which implies in turn that

$$\frac{1}{k}f(\bar{x}) \leq \frac{1}{k}f(x),$$

and hence $f(\bar{x}) \leq f(x)$. It shows that f achieves an absolute minimum at $\bar{x}$. $\qquad\square$

Proposition 2.2 *The solution set of the Fermat-Torricelli problem* (2.2) *is nonempty.*

Proof: Let $m := \inf\{\varphi(x) \mid x \in \mathbb{R}^n\} \geq 0$, and let (x_k) be a sequence satisfying

$$\lim_{k \to \infty} \varphi(x_k) = m.$$

It follows from the definition that there exists $k_0 \in \mathbb{N}$ with

$$\|x_k - a_1\| \leq \varphi(x_k) \leq m + 1 \text{ for all } \ k \geq k_0,$$

which implies that $\|x_k\| \leq m + 1 + \|a_1\|$ for such k. Thus $\{x_k\}$ is a bounded sequence, and so it has a subsequence $\{x_{k_\ell}\}$ that converges to $\bar{x} \in \mathbb{R}^n$ as $\ell \to \infty$. Since φ is obviously continuous, we get

$$\varphi(\bar{x}) = \lim_{\ell \to \infty} \varphi(x_{k_\ell}) = m.$$

This shows that $\bar{x}$ is an optimal solution of the problem. $\qquad\square$

For two different points $a, b \in \mathbb{R}^n$, the line containing a and b is given by

$$\mathcal{L}(a, b) := \big\{ ta + (1 - t)b \mid t \in \mathbb{R} \big\}.$$

Proposition 2.3 *Suppose that the points* a_i *for* $i = 1, \dots, m$ *do not lie on the same line (i.e., not collinear). Then the function* φ *defined by* (2.1) *is strictly convex, and the Fermat-Torricelli problem* (2.2) *has a unique solution.*

Proof: Define $\varphi_i(x) := \|x - a_i\|$ for $i = 1, \dots, m$. Then $\varphi = \sum_{i=1}^m \varphi_i$. For any $x, y \in \mathbb{R}^n$ and $\lambda \in (0, 1)$ we have the inequality

$$\varphi_i\big(\lambda x + (1 - \lambda)y\big) \leq \lambda \varphi_i(x) + (1 - \lambda)\varphi_i(y) \text{ for } \ i = 1, \dots, m.$$

This readily implies that

$$\varphi\big(\lambda x + (1-\lambda)y\big) \le \lambda\varphi(x) + (1-\lambda)\varphi(y). \tag{2.3}$$

On the contrary, suppose that φ is not strictly convex. It means that there exist $\bar{x}, \bar{y} \in \mathbb{R}^n$ with $\bar{x} \ne \bar{y}$ and $\lambda \in (0,1)$ for which (2.3) holds as equality. Then

$$\varphi_i\big(\lambda\bar{x} + (1-\lambda)\bar{y}\big) = \lambda\varphi_i(\bar{x}) + (1-\lambda)\varphi_i(\bar{y}) \text{ for } i = 1,\ldots,m,$$

which can be rewritten as

$$\|\lambda(\bar{x} - a_i) + (1-\lambda)(\bar{y} - a_i)\| = \|\lambda(\bar{x} - a_i)\| + \|(1-\lambda)(\bar{y} - a_i)\| \text{ for } i = 1,\ldots,m.$$

If $\bar{x} \ne a_i$ and $\bar{y} \ne a_i$, then there exists $t_i > 0$ such that

$$t_i\lambda(\bar{x} - a_i) = (1-\lambda)(\bar{y} - a_i).$$

This tells us that $\bar{x} - a_i = \gamma_i(\bar{y} - a_i)$, where $\gamma_i := \dfrac{1-\lambda}{t_i\lambda}$. Since $\bar{x} \ne \bar{y}$, we obtain $\gamma_i \ne 1$ and

$$a_i = \frac{1}{1-\gamma_i}\bar{x} - \frac{\gamma_i}{1-\gamma_i}\bar{y} \in \mathcal{L}(\bar{x}, \bar{y}).$$

In the case where either $\bar{x} = a_i$ or $\bar{y} = a_i$, it is obvious that $a_i \in \mathcal{L}(\bar{x}, \bar{y})$. Thus we arrive at $a_i \in \mathcal{L}(\bar{x}, \bar{y})$ for $i = 1,\ldots,m$, which is a contradiction that completes the proof. $\square$

Recall that a vector $v \in \mathbb{R}^n$ is a *subgradient* of a convex function $f\colon \mathbb{R}^n \to \mathbb{R}$ at the given point $\bar{x} \in \mathbb{R}^n$ if it satisfies the inequality

$$f(x) \ge f(\bar{x}) + \langle v, x - \bar{x}\rangle \text{ for all } x \in \mathbb{R}^n, \tag{2.4}$$

where $\langle \cdot, \cdot \rangle$ stands for the usual scalar product in $\mathbb{R}^n$. The set of all subgradients of f at $\bar{x}$ is called the *subdifferential* of this function at $\bar{x}$ and is denoted by $\partial f(\bar{x})$.

Directly from the definition, we deduce the following *subdifferential Fermat rule*:

$$f \text{ achieves an absolute minimum at } \bar{x} \text{ if and only if } 0 \in \partial f(\bar{x}). \tag{2.5}$$

The proposition below shows that the subdifferential of a convex function at a given point reduces to the gradient at that point when the function is differentiable.

Proposition 2.4 *Suppose that $f\colon \mathbb{R}^n \to \mathbb{R}$ is convex and (Fréchet) differentiable at $\bar{x}$. Then*

$$\langle \nabla f(\bar{x}), x - \bar{x}\rangle \le f(x) - f(\bar{x}) \text{ for all } x \in \mathbb{R}^n. \tag{2.6}$$

Furthermore, we have $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$.

Proof: Since f is differentiable at $\bar{x}$, it follows from the definition that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$-\varepsilon\|x - \bar{x}\| \le f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x}\rangle \le \varepsilon\|x - \bar{x}\| \text{ whenever } \|x - \bar{x}\| < \delta.$$

Consider further the function

$$\psi(x) := f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle + \varepsilon \|x - \bar{x}\|,$$

which satisfies $\psi(x) \geq \psi(\bar{x}) = 0$ for all $x \in \mathbb{B}(\bar{x}; \delta)$. By the convexity of ψ we have $\psi(x) \geq \psi(\bar{x})$ whenever $x \in \mathbb{R}^n$. Hence

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq f(x) - f(\bar{x}) + \varepsilon \|x - \bar{x}\| \quad \text{for all} \ \ x \in \mathbb{R}^n.$$

Letting now $\varepsilon \downarrow 0$ justifies the estimate in (2.6).

Equality (2.6) yields $\nabla f(\bar{x}) \in \partial f(\bar{x})$. Taking any $v \in \partial f(\bar{x})$, we get by definition that

$$\langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all} \ \ x \in \mathbb{R}^n.$$

The differentiability of f at $\bar{x}$ also implies that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\langle v - \nabla f(\bar{x}), x - \bar{x} \rangle \leq \varepsilon \|x - \bar{x}\| \quad \text{whenever} \ \ \|x - \bar{x}\| < \delta.$$

Thus we have $\|v - \nabla f(\bar{x})\| \leq \varepsilon$, which shows that $v = \nabla f(\bar{x})$ since $\varepsilon > 0$ was chosen arbitrarily. It verifies that $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$ and thus completes the proof of the proposition. $\square$

The subdifferential formula for the norm function derived in the next example plays a crucial role in our subsequent analysis to solve the Fermat-Torricelli problem.

Example 2.5 Let $p(x) := \|x\|$, the Euclidean norm function on $\mathbb{R}^n$. Then we have

$$\partial p(x) = \begin{cases} \mathbb{B} & \text{if} \ \ x = 0, \\ \left\{ \dfrac{x}{\|x\|} \right\} & \text{otherwise.} \end{cases}$$

Since the function $p(\cdot)$ is differentiable with $\nabla p(x) = x/\|x\|$ for $x \neq 0$, it suffices to verify the claimed formula for $x = 0$. It follows from (2.4) that $v \in \partial p(0)$ if and only if

$$\langle v, x \rangle = \langle v, x - 0 \rangle \leq p(x) - p(0) = \|x\| \text{ for all} \ \ x \in \mathbb{R}^n.$$

Letting $x := v$, we get $\langle v, v \rangle \leq \|v\|$, which implies that $\|v\| \leq 1$, i.e., $v \in \mathbb{B}$. Conversely, for $v \in \mathbb{B}$ the Cauchy-Schwarz inequality tells us that

$$\langle v, x - 0 \rangle = \langle v, x \rangle \leq \|v\| \cdot \|x\| \leq \|x\| = p(x) - p(0) \text{ whenever} \ \ x \in \mathbb{R}^n,$$

and hence $v \in \partial p(0)$. Thus we arrive at $\partial p(0) = \mathbb{B}$.

Solving the Fermat-Torricelli problem involves the usage of the following subdifferential rule for sums of two convex functions one of which is differentiable while the other may be not.

Proposition 2.6 *Let $f_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2$, be two convex functions such that f_2 is differentiable at $\bar{x}$. Then we have the equality*

$$\partial(f_1 + f_2)(\bar{x}) = \partial f_1(\bar{x}) + \nabla f_2(\bar{x}). \tag{2.7}$$

5

Proof: Fix any $v \in \partial(f_1 + f_2)(\bar{x})$ and get for each $x \in \mathbb{R}^n$ that

$$\langle v, x - \bar{x} \rangle \leq f_1(x) - f_1(\bar{x}) + f_2(x) - f_2(\bar{x}) = f_1(x) - f_1(\bar{x}) + \langle \nabla f_2(\bar{x}), x - \bar{x} \rangle + o(\|x - \bar{x}\|).$$

For any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$0 \leq \langle \nabla f_2(\bar{x}) - v, x - \bar{x} \rangle + f_1(x) - f_1(\bar{x}) + \varepsilon \|x - \bar{x}\| \text{ whenever } x \in \mathbb{B}(\bar{x}; \delta).$$

The convexity of f_1 ensures that the latter holds for all $x \in \mathbb{R}^n$. Letting $\varepsilon \downarrow 0$ yields

$$0 \leq \langle \nabla f_2(\bar{x}) - v, x - \bar{x} \rangle + f_1(x) - f_1(\bar{x}) \text{ whenever } x \in \mathbb{R}^n.$$

By (2.4) it tells us that $v - \nabla f_2(\bar{x}) \in \partial f_1(\bar{x})$ and hence $v \in \partial f_1(\bar{x}) + \nabla f_2(\bar{x})$, which justifies the inclusion "$\subset$" in (2.7). Representing $f_1 = f_1 + f_2 + (-f_2)$ and applying the obtained inclusion, we have the relationships

$$\partial f_1(\bar{x}) \subset \partial(f_1 + f_2)(\bar{x}) + \nabla(-f_2)(\bar{x}) = \partial(f_1 + f_2)(\bar{x}) - \nabla f_2(\bar{x}),$$

which verify the opposite inclusion "$\supset$" in (2.7) and thus complete the proof. $\qquad\square$

Let us now use subgradients of the norm function to obtain the classical solution of the Fermat-Torricelli problem. Given two nonzero vectors $u, v \in \mathbb{R}^n$, denote

$$\cos(u, v) := \frac{\langle u, v \rangle}{\|u\| \cdot \|v\|}.$$

Fix $\bar{x} \neq a_i$ and define the unit vectors

$$v_i := \frac{\bar{x} - a_i}{\|\bar{x} - a_i\|}, \quad i = 1, 2, 3.$$

Each v_i is the unit vector pointing in the direction from the vertex a_i to $\bar{x}$. Observe that the Fermat-Torricelli problem formulated above always has a unique solution even if the three given points are on the same line. It is easy to see in the latter case that the middle point is the solution of the problem. The next proposition completely characterizes the solution of the Fermat-Torricelli problem in the general three-point setting of the n-dimensional space, not just on the plane as in the original framework.

Proposition 2.7 *Let the points $a_1, a_2, a_3 \in \mathbb{R}^n$ generate the Fermat-Torricelli problem in $\mathbb{R}^n$. Then we have the following descriptions of the optimal solution to (2.2) with φ taken from (2.1):*
(i) *In the case where $\bar{x} \notin \{a_1, a_2, a_3\}$, $\bar{x}$ is the solution of the problem if and only if*

$$\cos(v_1, v_2) = \cos(v_2, v_3) = \cos(v_3, v_1) = -1/2.$$

(ii) *Consider the case where $\bar{x} \in \{a_1, a_2, a_3\}$ and suppose for definiteness that $\bar{x} = a_1$. Then $\bar{x}$ is the solution of the problem if and only if*

$$\cos\langle v_2, v_3 \rangle \leq -1/2.$$

Proof: In case (i) we have that the function φ from (2.1) is differentiable at $\bar{x}$. Since φ is convex, $\bar{x}$ is the solution to the Fermat-Torricelli problem if and only if

$$\nabla\varphi(\bar{x}) = v_1 + v_2 + v_3 = 0.$$

Remembering that $\|v_i\| = 1$ for $i = 1, 2, 3$, we get

$$\langle v_1, v_2 \rangle + \langle v_1, v_3 \rangle = -1$$
$$\langle v_2, v_1 \rangle + \langle v_2, v_3 \rangle = -1$$
$$\langle v_3, v_1 \rangle + \langle v_3, v_2 \rangle = -1.$$

Solving this system of equations yields

$$\langle v_i, v_j \rangle = \cos(v_i, v_j) = -1/2 \text{ for } i \neq j, \ i, j \in \{1, 2, 3\}.$$

If furthermore $\langle v_i, v_j \rangle = -1/2$ for $i \neq j$, $i, j \in \{1, 2, 3\}$, then

$$\|v_1 + v_2 + v_3\|^2 = \sum_{i=1}^{3} \|v_i\| + \sum_{i,j=1,i\neq j}^{3} \langle v_i, v_j \rangle = 0,$$

which gives us $v_1 + v_2 + v_3 = 0$ and thus completes the proof in the case.

In case (ii) we deduce from the subdifferential Fermat rule (2.5) and the subdifferential sum rule (2.7) that $\bar{x} = a_1$ is the solution to the Fermat-Torricelli problem if and only if

$$0 \in \partial\varphi(a_1) = \mathbb{B} + v_2 + v_3.$$

This is equivalent to $\|v_2 + v_3\|^2 \leq 1$ or, equivalently, to $\|v_2\|^2 + \|v_3\|^2 + 2\langle v_2, v_3 \rangle \leq 1$. Since v_2 and v_3 are unit vectors, we obtain

$$\langle v_2, v_3 \rangle = \cos(v_2, v_3) \leq -1/2$$

and complete the proof of the proposition. $\square$

Next we present an example with a figure illustrating the obtained solution in the classical case of three points on the plane.

Example 2.8 Consider the Fermat-Torricelli problem given by three points A, B, and C on the plane as shown in the figure. If one of the angles of the triangle ABC is greater than or equal to $120°$, then the corresponding vertex is the solution to the problem by Proposition 2.7(ii). Let us examine the case where none of the angles of the triangle is greater than or equal to $120°$. Construct two equilateral triangles ABD and ACE and let S be the intersection of DC and BE as in the figure. Two quadrilaterals $ADBC$ and $ABCE$ are convex, and hence S lies inside the triangle ABC. It is clear that two triangles DAC and BAE are congruent. A rotation of $60°$ about A maps the triangle DAC to the triangle BAE. The rotation maps CD to BE, so $\angle DSB = 60°$. Let T be the image of S through this rotation. Then T belongs to BE. It follows that $\angle AST = \angle ASE = 60°$. Moreover, $\angle DSA = 60°$, and hence $\angle BSA = 120°$. It is now clear that $\angle ASC = 120°$ and $\angle BSC = 120°$. Proposition 2.7(i) tells us that the point S is the solution to this classical Fermat-Torricelli problem.
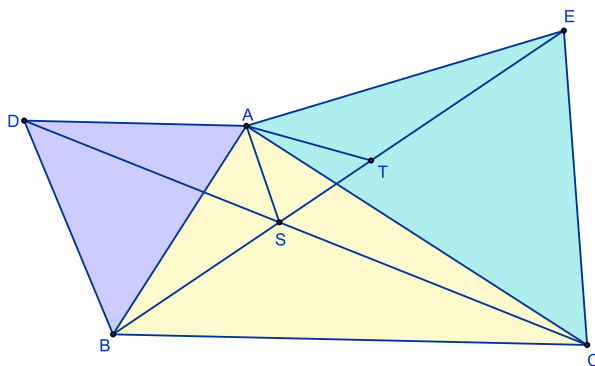
3 Weiszfeld's Algorithm

In this section we revisit Kuhn's proof [11] of the convergence of Weiszfeld's algorithm [22] for solving the Fermat-Torricelli problem (2.2). With some additional ingredients of convex analysis we are able to provide a more clear picture of Kuhn's proof. Throughout this section, assume that the points a_i for $i = 1, \ldots, m$ are not collinear.

The gradient of the function φ from (2.1) is computed by

$$\nabla \varphi(x) = \sum_{i=1}^{m} \frac{x - a_i}{\|x - a_i\|}, \ x \notin \{a_1, a_2, \ldots, a_m\}.$$

Solving the gradient equation $\nabla \varphi(x) = 0$ gives us the formula

$$x = \frac{\displaystyle\sum_{i=1}^{m} \frac{a_i}{\|x - a_i\|}}{\displaystyle\sum_{i=1}^{m} \frac{1}{\|x - a_i\|}} := F(x). \tag{3.8}$$

To keep the continuity, define $F(x) := x$ for $x \in \{a_1, a_2, \ldots, a_m\}$.

Weiszfeld introduced the following algorithm: choose a starting point $x_0 \in \mathbb{R}^n$ and define

$$x_k = F(x_{k-1}) \text{ for } k \in \mathbb{N}.$$

He also claimed that if $x_0 \notin \{a_1, a_2, \ldots, a_m\}$ where a_i for $i = 1, \ldots, m$ are not collinear, then $\{x_k\}$ converges to the unique optimal solution of the problem. A correct statement and the proof of the convergence were given by Kuhn in [11].

The next proposition guarantees that the cost function value in (2.2) decreases after each iteration of the Weiszfeld algorithm.

Proposition 3.1 *If $F(x) \neq x$, then $\varphi(F(x)) < \varphi(x)$.*

Proof: It is clear that x is not a vertex, since otherwise we get $F(x) = x$. Moreover, the point $F(x)$ is the unique minimizer of the following strictly convex function:

$$g(z) := \sum_{i=1}^{m} \frac{\|z - a_i\|^2}{\|x - a_i\|}.$$

Indeed, we can easily check that $F(x)$ is the unique solution of the equation $\nabla g(z) = 0$. Using $F(x) \neq x$ tells us that $g(F(x)) < g(x) = \varphi(x)$. Furthermore, it clearly follows from the constructions above that

$$
\begin{aligned}
g(F(x)) &= \sum_{i=1}^{m} \frac{\|F(x) - a_i\|^2}{\|x - a_i\|} \\
&= \sum_{i=1}^{m} \frac{(\|x - a_i\| + \|F(x) - a_i\| - \|x - a_i\|)^2}{\|x - a_i\|} \\
&= \varphi(x) + 2\big(\varphi(F(x)) - \varphi(x)\big) + \sum_{i=1}^{m} \frac{(\|F(x) - a_i\| - \|x - a_i\|)^2}{\|x - a_i\|},
\end{aligned}
$$

which verifies the strict inequality

$$2\varphi(F(x)) + \sum_{i=1}^{m} \frac{(\|F(x) - a_i\| - \|x - a_i\|)^2}{\|x - a_i\|} < 2\varphi(x)$$

and hence yields the claimed decreasing property $\varphi(F(x)) < \varphi(x)$. $\qquad\square$

Now we investigate behavior of the algorithm mapping F near a vertex and deal with the case where a vertex is not the solution of the Fermat-Torricelli problem (2.2). Let us first present a necessary and sufficient condition for a vertex to be the optimal solution of the problem. Define

$$R_j := \sum_{i=1, i \neq j}^{m} \frac{a_i - a_j}{\|a_i - a_j\|}, \quad j = 1, \ldots, m,$$

Proposition 3.2 *The vertex a_j is the optimal solution to* (2.2) *if and only if* $\|R_j\| \leq 1$.

Proof: Employing the subdifferential Fermat rule (2.5) and the subdifferential sum rule from Proposition 2.6 ensures that the vertex a_j is the optimal solution of the problem if and only if

$$0 \in \partial\varphi(a_j) = -R_j + \mathbb{B},$$

which can be equivalently rewritten as $\|R_j\| \leq 1$. $\qquad\square$

The obtained result allows us to significantly simplify the proof of the next proposition taken from [11, Subsection 3.2].

Proposition 3.3 *Suppose that the vertex a_j is not the optimal solution to* (2.2). *Then there is a number $\delta > 0$ such that the condition $0 < \|x - a_j\| \leq \delta$ yields the existence of a positive integer q for which we have the estimates*

$$\|F^q(x) - a_j\| > \delta \text{ and } \|F^{q-1}(x) - a_j\| \leq \delta \tag{3.9}$$

with using the notation

$$F^q(x) := F\left(F^{q-1}(x)\right) \quad \text{whenever} \quad q = 1, 2, \ldots \quad \text{and} \quad F^0(x) := x. \tag{3.10}$$

Proof: If x is not a vertex, then we get from (3.8) that

$$F(x) = \frac{\displaystyle\sum_{i=1}^{m} \frac{a_i}{\|x - a_i\|}}{\displaystyle\sum_{i=1}^{m} \frac{1}{\|x - a_i\|}}.$$

which implies in turn that

$$F(x) - a_j = \frac{\displaystyle\sum_{i=1, i\neq j}^{m} \frac{a_i - a_j}{\|x - a_i\|}}{\displaystyle\sum_{i=1}^{m} \frac{1}{\|x - a_i\|}}.$$

Taking now the limit as $x \to a_j$ leads us to

$$\lim_{x \to a_j} \frac{F(x) - a_j}{\|x - a_j\|} = \lim_{x \to a_j} \frac{\displaystyle\sum_{i=1, i\neq j}^{m} \frac{a_i - a_j}{\|x - a_i\|}}{1 + \displaystyle\sum_{i=1, i\neq j}^{m} \frac{\|x - a_j\|}{\|x - a_i\|}} = R_j.$$

This implies by Proposition 3.2 that

$$\lim_{x \to a_j} \frac{\|F(x) - a_j\|}{\|x - a_j\|} = \|R_j\| > 1 \tag{3.11}$$

and thus allows us to find positive numbers ε and δ with

$$\frac{\|F(x) - a_j\|}{\|x - a_j\|} \geq (1 + \varepsilon) \text{ whenever } 0 < \|x - a_j\| \leq \delta. \tag{3.12}$$

Remembering the notation in (3.10), if

$$0 < \|F^{p-1}(x) - a_j\| \leq \delta \text{ for all } p = 1, \ldots, q,$$

then by (3.12) we have

$$\|F^q(x) - a_j\| \geq (1 + \varepsilon)\|F^{q-1}(x) - a_j\| \geq \ldots \geq (1 + \varepsilon)^q\|x - a_j\|.$$

Taking into account that $(1 + \varepsilon)^q\|x - a_j\| \to \infty$ as $q \to \infty$ verifies the estimates in (3.9) and thus completes the proof of the proposition. $\qquad\square$

We finally present the following simplified and improved proof (with taking into account Propositions 3.2 and (3.3) above) of Kuhn's convergence result [11] for Weiszfeld's algorithm to solve the Fermat-Torricelli problem.

Theorem 3.4 *Let $\{x_k\}$ be the sequence of iterates generated by Weiszfeld's algorithm, and let $x_k \notin \{a_1, a_2, \ldots, a_m\}$ for all $k = 0, 1, \ldots$. Then $\{x_k\}$ converges to the optimal solution $\bar{x}$ of the Fermat-Torricelli problem* (2.2).

Proof: Observe first that if $x_k = x_{k+1}$ for some $k = k_0$, then x_k is a constant sequence for all $k \geq k_0$, which therefore converges to x_{k_0}. Since $F(x_{k_0}) = x_{k_0}$ and x_{k_0} is not a vertex, the point x_{k_0} is the solution of the problem. Hence we can proceed by assuming that $x_{k+1} \neq x_k$ for every k. Proposition 3.1 tells us that the sequence $\{\varphi(x_k)\}$ is nonnegative and decreasing, and thus it converges, which means that

$$\lim_{k \to \infty} \big(\varphi(x_k) - \varphi(x_{k+1})\big) = 0. \tag{3.13}$$

It follows from the algorithm that $x_k \in \mathrm{co}\,\{a_1, a_2, \ldots, a_m\}$ (the convex hull) for all $k \geq 1$. Since the latter set is compact in $\mathbb{R}^n$, we have the convergence of some subsequence of $\{x_k\}$. Take a subsequence $\{x_{k_\ell}\}$ of $\{x_k\}$ that converges to a point $\bar{y}$. It suffices to prove that $\bar{y} = \bar{x}$. To proceed, deduce from (3.13) that

$$\lim_{\ell \to \infty} \big(\varphi(x_{k_\ell}) - \varphi(F(x_{k_\ell}))\big) = 0$$

and conclude by the continuity of φ that $\varphi(\bar{y}) = \varphi(F(\bar{y}))$. This clearly yields $F(\bar{y}) = \bar{y}$.

If $\bar{y}$ is not a vertex, then it is the solution of the problem, so $\bar{y} = \bar{x}$. Let us consider the case where $\bar{y}$ is a vertex, say a_1. Arguing by contradiction, suppose that $\bar{y} \neq \bar{x}$. Choose $\delta > 0$ sufficiently small such that the properties in Proposition 3.3 hold, and that the ball $\mathbb{B}(a_1; \delta)$ does not contain $\bar{x}$ and a_i for $i = 2, \ldots, m$. Since $x_{k_\ell} \to a_1 = \bar{y}$, we assume without loss of generality that the sequence is contained in $\mathbb{B}(a_1; \delta)$.

For $x = x_{k_1}$, choose q_1 such that $x_{q_1} \in \mathbb{B}(a_1; \delta)$ and $F(x_{q_1}) \notin \mathbb{B}(a_1; \delta)$. Selecting further an index $k_\ell > q_1$ and applying Proposition 3.3, we find $q_2 > q_1$ such that $x_{q_2} \in \mathbb{B}(a_1; \delta)$ and $F(x_{q_2}) \notin \mathbb{B}(a_1; \delta)$. Repeating this procedure gives us a sequence $\{x_{q_\ell}\}$ with $x_{q_\ell} \in \mathbb{B}(a_1; \delta)$ and $F(x_{q_\ell})$ not belonging to the ball. Extracting yet another subsequence, suppose that $x_{q_\ell} \to \bar{z}$. It follows from the above that $F(\bar{z}) = \bar{z}$. If $\bar{z}$ is not a vertex, then it must be the solution, which is a contradiction because the solution $\bar{x}$ is not in $\mathbb{B}(a_1; \delta)$. Thus $\bar{z}$ is a vertex that should be a_1, since the other vertices do not belong to the ball as well. It tells us that

$$\lim_{\ell \to \infty} \frac{\|F(x_{q_\ell}) - a_1\|}{\|x_{q_\ell} - a_1\|} = \infty,$$

which contradicts Proposition 3.2 via (3.11) and thus completes the proof of the theorem. $\square$

References

[1] S. Alzorba, C. Gunther, N. Popovici and C. Tammer, A new algorithm for solving planar multi-objective location problems involving the Manhattan norm. European J. Oper. Res. 258 (2017), 35–46.

[2] A. Beck and S. Sabach, Weiszfeld's method: old and new results. J. Optim. Theory Appl. 164 (2015), 1–40.

[3] J. Brimberg. The Fermat Weber location problem revisited. Math. Program. 71 (1995), 71–76.

[4] J. Brimberg, R. Chen and D. Chen, Accelerating convergence in the Fermat-Weber location problem. Oper. Res. Lett. 22 (1998), 151–157.

[5] L. Cooper and I. Katz, The Weber problem revisited. Comput. Math. Appl. 7 (1981), 225–234.

[6] Z. Drezner, On the convergence of the generalized Weiszfeld algorithm. Ann. Oper. Res. 167 (2009), 327–336.

[7] U. Eckhardt, Weber's problem and Weiszfeld's algorithm in general spaces. Math. Program. 18 (1980) 186–196.

[8] F. Giannessi, Constrained Optimization and Image Space Analysis, Vol. 1. Separation of Sets and Optimality Conditions. Math. Concepts Methods Sci. Engrg. 49, Springer, New York (2005).

[9] C. Gunther and C. Tammer, Relationsship between constrained and unconstrained multi-objective optimization and application in location theory. Math. Methods Oper. Res. 84 (2016), 359–387.

[10] C. Gunther, C. Tammer, M. Hillmann and B. Winkler, Project Facility Location Optimizer. The Martin Luther University of Halle-Wittenberg (2018).

[11] H.W. Kuhn, A note on Fermat-Torricelli problem. Math. Program. 4 (1973), 98–107.

[12] B.S. Mordukhovich and N.M. Nam, Applications of variational analysis to a generalized Fermat-Torricelli problem. J. Optim. Theory Appl. 148 (2011), 431–454.

[13] B.S. Mordukhovich and N. M. Nam, An Easy Path to Convex Analysis and Applications. Morgan & Claypool Publishers, San Rafael, CA (2014).

[14] B.S. Mordukhovich, N.M. Nam and J. Salinas, Applications of variational analysis to a generalized Heron problem. Applic Anal. 91 (2012), 1915–1942.

[15] J.G. Morris and W.A. Verdini, Minisum ℓ_p distance location problems solved via a perturbed problem and Weiszfeld's algorithm. Oper. Res. 27 (1979), 1180–1188.

[16] R. Patz, J. Spitzner and C. Tammer, Decision support for location problems in town planning. Inter. Trans. Oper. Res. 9 (2002), 261–278.

[17] F. Plastria, The Weiszfeld algorithm: proof, amendments and extensions. H.A. Eiselt and V. Marianov (Eds.). Foundations of Location Analysis. International Series in Operations Research and Management Science, Vol. 155, pp. 357–389. Springer, New York (2011).

[18] R.T. Rockafellar, Convex Analysis. Princeton University Press, Princeton, NJ (1970).

[19] A. Ruszczyński, Nonlinear Optimization. Princeton University Press, Princeton, NJ, 2006.

[20] H. Üster and R.F. Love, The convergence of the Weiszfeld algorithm. Comput. Math. Appl. 40 (2000), 443–451.

[21] Y. Vardi and C.-H. Zhang, A modified Weiszfeld algorithm for the Fermat-Weber location problem. Math. Program. 90 (2001), 559–566.

[22] E. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnés est minimum, Tôhoku Mathematics Journal 43 (1937), 355–386.

[23] E. Weiszfeld and F. Plastria, On the point for which the sum of the distances to n given points is minimum. Ann Oper Res 167 (2009), 7–41.