The Projection Median of a Set of Points ⋆

Stephane Durocher

Cheriton School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada [1]

David Kirkpatrick

Department of Computer Science, University of British Columbia, Vancouver, British Columbia, Canada

Abstract

Given a nonempty and finite multiset of points P in $\mathbb{R}^d$, the Euclidean median of P, denoted $M(P)$, is a point in $\mathbb{R}^d$ that minimizes the sum of the Euclidean (ℓ_2) distances from $M(P)$ to the points of P. In two or more dimensions, the Euclidean median (otherwise known as the Weber point) is unstable; small perturbations at only a few points of P can result in an arbitrarily large relative change in the position of the Euclidean median. This instability motivates us to consider alternate notions for location functions that approximate the minimum sum of distances to the points of P while maintaining a fixed degree of stability. We introduce the projection median of a multiset of points in $\mathbb{R}^2$ and compare it against the rectilinear (ℓ_1) median and the centre of mass, both in terms of approximation factor and stability. We show that a mobile facility located at the projection median of the positions of a set of mobile clients provides a good approximation of the mobile Euclidean median while ensuring both continuous motion and low relative velocity.

Key words: median, Weber point, projection, approximation, stability, continuity, facility location, mobile clients, bounded velocity

⋆ Some of these results originally appeared in the first author's doctoral thesis [Dur06] and in a preliminary version of this paper in the Proceedings of the Canadian Conference on Computational Geometry [DK05]. Funding for this research was made possible by NSERC and the MITACS project on Facility Location Optimization.

Email addresses: `sdurocher@cs.uwaterloo.ca` (Stephane Durocher), `kirk@cs.ubc.ca` (David Kirkpatrick).

[1] This work was completed while the first author was at the University of British Columbia.

1 Introduction: the Euclidean Median in $\mathbb{R}^2$

Given a multiset of points P in $\mathbb{R}$, a median of P, denoted $M(P)$, is a point that partitions the points of P such that at most $|P|/2$ points of P are greater than $M(P)$ and at most $|P|/2$ points of P are less than $M(P)$. It is straightforward to confirm that $M(P)$ is a balance point that minimizes the sum of the distances (equivalently, the average distance) from $M(P)$ to the points of P. The problem of finding a point that minimizes the sums of distances to the points of P has a natural extension to higher dimensions with applications that include geometry [Kim98, KM97], operations research [HLP$^+$87, LMW88, Web22], and robotics [CFPS03, Sch03].

Definition 1 *Given an arbitrary nonempty finite multiset P in $\mathbb{R}^2$, a* Euclidean median *of P is a point in $\mathbb{R}^2$, denoted $M(P)$, that minimizes*

$$\sum_{p \in P} ||x - p||, \tag{1}$$

when $x = M(P)$.

Def. 1 generalizes to $\mathbb{R}^d$ for any fixed dimension d; although some of our results can be generalized to higher dimensions, the primary focus our discussion concerns the case $d = 2$. We refer to the value (1) (when $x = M(P)$) as the *Euclidean median sum* of P. If the points of P are not collinear or $|P|$ is odd, then the Euclidean median is unique [KM97]. If the points of P are collinear and $|P|$ is even, then any point on the line segment joining the $(|P|/2)$nd and $(|P|/2 + 1)$st points of P is a Euclidean median of P; in this case, we assign $M(P)$ to be the midpoint of this line segment. Finally, M is invariant under similarity transformations.

The Euclidean median problem on three points in the plane was first posed by Fermat [dF91] and solved geometrically by Torricelli early in the 17th century [KV97]. Alternate geometric solution techniques were subsequently found by Cavalieri and Simpson [DKSW02]. For points on a line, a Euclidean median is easily found in $\Theta(n)$ time, where $n = |P|$, by a linear-time selection algorithm. In general, solving for the exact location of the Euclidean median in two or more dimensions is difficult. Bajaj states, "there exists no exact algorithm under models of computation where the root of an algebraic equation is obtained using arithmetic operations and the extraction of kth roots" [Baj88, p. 177]. Indeed, no polynomial-time algorithm is known, nor has the problem been shown to be NP-hard [Hak00]. The most common approximation algorithm is Weiszfeld's algorithm [Wei37], an iterative procedure that converges to the Euclidean median. Chandrasekaran and Tamir [CT90] give a polynomial-time algorithm for an ϵ-approximation of the Euclidean median. More recently, Indyk [Ind99] and Bose et al. [BMM03] both give randomized ϵ-approximations algorithms with running times linear in n and polynomial in $1/\epsilon$. Bose et al. [BMM03] also give an $O(n \log n)$-time deterministic ϵ-approximation algorithm.

The Euclidean median has been repeatedly rediscovered under a variety of names. The most common of these is *Weber point* [Baj88, BMM03, FMW05, Wes93]. Other names include Torricelli point [Kim98, Wei], Fermat point [Kim98], first Fermat point [Wei], generalized Fermat point [Wes93], first isogonic centre [Kim98, Wei], ℓ_2 median, 1-median [FMW05, Ind99], median centre [Wes93], spatial median [Wes93], minisum problem [HLP$^+$87, Wes93], Steiner problem [KM97, Wes93], bivariate median [Wes93], minimum aggregate travel point [Wes93], the point of equilibrium in a Varignon frame [Wes93], Kimberling triangle centre $X(13)$ [Kim98], or any combination of Fermat-Steiner-Torricelli-Weber point [BMM03, CT90, KM97, Wes93]. In addition, the term "median" sometimes refers to alternate generalizations of the median to higher dimensions. For example, Agarwal et al. [AdBG$^+$05], use the term in reference to a point m such that for every line l through m, at least $k|P|$ points of P lie on either side of l, where $k \in [0, \frac{1}{2}]$ is fixed. Finally, the Euclidean median is sometimes defined with a non-negative weight assigned to each point [CT90, Wes93]; when the weights are rational this reduces to Def. 1 since we allow multiplicities of points. An overview of the history and solutions to the Euclidean median problem can be found in [DKSW02, KM97, Wes93].

The remainder of the paper is organized as follows. Sec. 2 begins by observing the instability of the Euclidean median, thus motivating our search for stable approximations. We then define measures for comparing the stability and quality of different approximations. We examine two existing notions of location functions, namely, the centre of mass and the rectilinear median, in Secs. 3 and 4, respectively. Next, we introduce and analyze a new location function, the projection median, in Sec. 5. The properties of these various location functions are compared in Sec. 6, followed by a discussion of applications of location functions and, specifically, of the projection median, in defining the position of a mobile facility.

2 Approximation Measures

Point coordinates are commonly represented by discretization of real positions to nearby grid coordinates. That is, each point is approximated by the nearest grid point, resulting in a small perturbation of each point's position. Given a multiset of points P in $\mathbb{R}^2$, the Euclidean median of P is *unstable* in this sense that small perturbations at only a few points of P can result in a relatively large change (error) in the position of the Euclidean median of P. For example, let $P = \{(0,0), (0,0), (1,0), (1,\epsilon)\}$ and let $P' = \{(0,0), (0,\epsilon), (1,0), (1,0)\}$. For any $\epsilon > 0$, $M(P) = (0,0)$ and $M(P') = (1,0)$.

We refer to an arbitrary function $\Upsilon : \mathring{\mathscr{P}}(\mathbb{R}^2) \rightarrow \mathbb{R}^2$ as a *location function*, where $\mathring{\mathscr{P}}(A)$ denotes the set of all nonempty and finite multisets contained in set A. Given the instability of the Euclidean median, which may be unfit for certain applications, we are motivated to find location functions that approximate the Euclidean median while guaranteeing some degree of stability.

We formalize the notion of stability by defining κ-stability for a location function Υ as a bound on its maximum volatility. This requires a preliminary definition for an ϵ-perturbation.

Definition 2 *Given $\epsilon > 0$ and a finite nonempty multiset P in $\mathbb{R}^2$, function $f : P \rightarrow \mathbb{R}^2$ is an* ϵ-perturbation *on P if for all $p \in P$, $||p - f(p)|| \leq \epsilon$.*

Let F_ϵ^P denote the set of all ϵ-perturbations on P.

Definition 3 *A location function Υ is* κ-stable *if*

$$\forall \epsilon > 0,\ \forall f \in F_\epsilon^P,\ \kappa||\Upsilon(P) - \Upsilon(f(P))|| \leq \epsilon, \tag{2}$$

for all nonempty finite multisets P in $\mathbb{R}^2$.

The Euclidean median is not continuous even for small point sets, as demonstrated by the four-point example mentioned earlier in this section. It follows that the Euclidean median is not κ-stable for any $\kappa > 0$. In fact, this same example can be used to show that any arbitrarily-close approximation of the exact position of the Euclidean median is not κ-stable for any fixed $\kappa > 0$. Note, however, that the Euclidean median sum is stable. This gives us hope that it may be possible to approximate the Euclidean median sum while also guaranteeing some fixed degree of stability. Thus, our overall objective is to identify a location function Υ that comes close to minimizing the sum of the distances from $\Upsilon(P)$ to the points of P while maintaining a fixed degree of stability.

We formalize the notion of approximation factor by defining λ-approximation; we evaluate the approximation factor of a location function Υ as a bound on its worst-case relative approximation of (1) and not as a measure of its relative proximity to the exact position of the Euclidean median.

Definition 4 *A location function Υ is a* λ-approximation *of the Euclidean median, M, if*

$$\sum_{p \in P} ||p - \Upsilon(P)|| \leq \lambda \sum_{q \in P} ||q - M(P)||, \tag{3}$$

for all nonempty finite multisets P in $\mathbb{R}^2$.

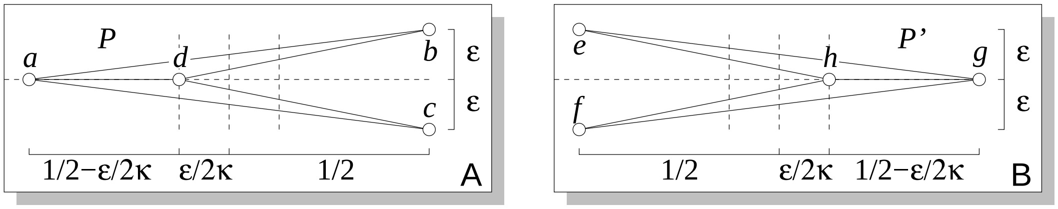

Fig. 1. illustration in support of Thm. 5

By definition, $M(P)$ is a point that minimizes (1). Consequently, for any location function, the associated approximation factor λ must be at least 1.

Stability and approximation factor are correlated. As we show formally in Thm. 5, no location function can ensure any fixed degree of stability while also guaranteeing an arbitrarily-close approximation of the Euclidean median sum.

Theorem 5 *For every $\kappa > 0$, if Υ is any κ-stable location function, then there exists some $\lambda_0 > 1$ such that Υ cannot guarantee an approximation factor less than λ_0.*

Proof. Choose any $\kappa > 0$, any $\epsilon \in (0, \kappa)$, and any κ-stable location function Υ. Let $P = \{(0,0),(0,0),(1,\epsilon),(1,-\epsilon)\}$ and let $P' = \{(0,\epsilon),(0,-\epsilon),(1,0),(1,0)\}$. Let $d = (1/2 - \epsilon/2\kappa, 0)$ and let $h = (1/2 + \epsilon/2\kappa, 0)$. The points in P and P' are labelled as $\{a, b, c\}$ and $\{e, f, g\}$, respectively, in Figs. 1A and 1B.

Since $|P| = 4$ and two points of P coincide at $(0,0)$, $M(P) = (0,0)$ [KM97]. Similarly, $M(P') = (1,0)$. The Euclidean median sum of P (and, by symmetry, P') is $2\sqrt{1+\epsilon^2}$.

Clearly there exists an ϵ-permutation of P, f, such that $f(P) = P'$. By Def. 3,

$$||\Upsilon(P) - \Upsilon(P')|| \leq \frac{\epsilon}{\kappa}. \tag{4}$$

Consequently, either $\Upsilon(P)_x \geq d_x$ or $\Upsilon(P')_x \leq h_x$, where p_x denotes the x-coordinate of a point $p \in \mathbb{R}^2$. Without loss of generality, assume $\Upsilon(P)_x \geq d_x$.

It is straightforward to show that for any point d', where $d'_x \geq d_x$, $\sum_{p\in P} ||d'_x - p|| \geq \sum_{p\in P} ||d_x - p||$. Therefore,

$$\begin{aligned}\sum_{p\in P} ||\Upsilon(P) - p|| \geq & \sum_{p\in P} ||d - p|| \\ =& 2\left(\frac{1}{2} - \frac{\epsilon}{2\kappa}\right) + 2\sqrt{\epsilon^2 + \left(\frac{1}{2} + \frac{\epsilon}{2\kappa}\right)^2}.\end{aligned} \tag{5}$$

By Def. 4, if Υ is a λ-approximation, then

$$\begin{aligned}\lambda \geq & \frac{\sum_{p\in P} ||\Upsilon(P) - p||}{\sum_{q\in P} ||M(P) - q||} \\ \geq & \frac{\frac{1}{2} - \frac{\epsilon}{2\kappa} + \sqrt{\epsilon^2 + \left(\frac{1}{2} + \frac{\epsilon}{2\kappa}\right)^2}}{\sqrt{1+\epsilon^2}}, \qquad \text{by (5)}. \end{aligned} \tag{6}$$

Let λ_1 denote the righthand expression in (6). It is straightforward to show that $\lambda_1 > 1$ for any $\epsilon \in (0, \kappa)$. Therefore, for any $\lambda_0 \in (1, \lambda_1)$, Υ is not a λ_0-approximation of the Euclidean median. □

As a consequence of Thm. 5, algorithms that achieve an arbitrarily-close approximation of the Euclidean median, such as those of Chandrasekaran and Tamir [CT90], Indyk [Ind99], and Bose et al. [BMM03] mentioned in Sec. 1, cannot be κ-stable for any $\kappa > 0$, even if they approximate the Euclidean median sum and not the exact position of the Euclidean median.

In summary, we seek to identify functions $\Upsilon : \mathring{\mathscr{P}}(\mathbb{R}^d) \to \mathbb{R}^d$ with stability κ and approximation factor λ under the dual objective of maximizing stability (maximize κ) while minimizing the approximation of the Euclidean median sum (minimize λ). Before introducing the projection median (Sec. 5), we first examine the stabilities and approximation factors of the centre of mass and the rectilinear median (Secs. 3 and 4).

3 The Centre of Mass

Definition 6 *Given an arbitrary nonempty finite multiset P in $\mathbb{R}^2$, the* centre of mass *of P is the function whose value, denoted $C(P)$, is the point in $\mathbb{R}^2$ given by*

$$C(P) = \frac{1}{|P|} \sum_{p\in P} p. \tag{7}$$

Function C is invariant under affine transformations. The position of the centre of mass is is easily constructed in $\Theta(n)$ time. Furthermore, $C(P)$ is the unique point that minimizes the sum of the squares of the distances to the points of P [Sch73, Wes93], suggesting C as a candidate for approximating (1).

The centre of mass is also known as geometric centroid [Wei], centroid [Wes93], mean, 1-mean, centre of gravity [Sch73, Wes93], and Kimberling triangle centre $X(2)$ [Kim98].

We now establish a tight bound on the approximation factor of the centre of mass in Thm. 9. Necessary to the proof of Thm. 9 is Lem. 8 which shows that

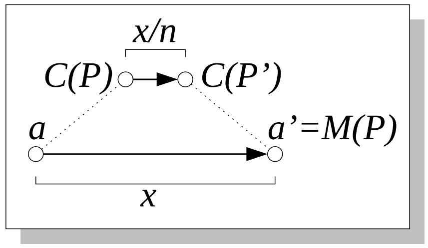

Fig. 2. illustration in support of Lem. 8

for any finite multiset P, if some point $a \neq M(P)$ is moved to coincide with $M(P)$, then the Euclidean median of the new multiset P' remains unchanged. Lems. 7 and 8 and Thm. 9 refer to the following definitions for P, a, x, and n. Let P denote a finite multiset in $\mathbb{R}^2$ such that $a \neq M(P)$ for some $a \in P$. Let $a' = M(P)$, let $P' = (P \setminus \{a\}) \cup \{a'\}$, let $x = ||a - a'||$, and let $n = |P|$. See Fig. 2.

Lemma 7 *Point $M(P)$ is a Euclidean median of P'.*

Proof. Suppose $M(P)$ is not a Euclidean median of P'. Thus,

$$\sum_{p \in P'} ||p - M(P')|| < \sum_{p \in P'} ||p - M(P)||. \tag{8}$$

Therefore,

$$\begin{aligned}
\sum_{p \in P} ||p - M(P')|| &= ||a - M(P')|| + \sum_{p \in P \setminus \{a\}} ||p - M(P')|| \\
&\leq x + ||a' - M(P')|| + \sum_{p \in P \setminus \{a\}} ||p - M(P')|| \\
&= x + ||a' - M(P')|| + \sum_{p \in P' \setminus \{a\}} ||p - M(P')|| \\
&= x + \sum_{p \in P'} ||p - M(P')|| \\
&< x + \sum_{p \in P'} ||p - M(P)||, \text{ by our assumption,} \\
&= \sum_{p \in P} ||p - M(P)||. \qquad (9)
\end{aligned}$$

Thus, $M(P)$ did not minimize $\sum_{p \in P} ||p - M(P)||$. Consequently, $M(P)$ cannot be a median of P, deriving a contradiction. Therefore $M(P') = M(P)$. □

Also necessary to the proof of Thm. 9 is Lem. 8 which relates the sum of the distances from $C(P)$ to the points of P to the corresponding value for P'.

Lemma 8

$$\sum_{p \in P} ||p - C(P)|| - \sum_{p \in P'} ||p - C(P')|| \leq 2x \left(1 - \frac{1}{n}\right). \tag{10}$$

Proof. Since all points remain static except for point a, $C(P) - C(P') = \frac{1}{n}(a - a')$. See Fig. 2. Consequently, the distance from a to the centre of mass changes by at most $\pm(x - x/n)$. For each of the $n-1$ points in $P \setminus \{a\}$, the corresponding distance changes by at most $\pm x/n$. The result follows. □

Theorem 9 *The centre of mass provides a* $(2 - 2/n)$*-approximation of the Euclidean median.*

Proof. Let a, a', x, and P' be as defined in Lem. 7. Let $m = \sum_{p \in P} ||p - M(P)||$ and let $c = \sum_{p \in P} ||p - C(P)||$. Let m' and c' denote the corresponding values for P'. Assume P is a multiset that maximizes the approximation factor of C such that $c > m(2 - 2/n)$. Observe that a point $a \neq M(P)$ must exist under this assumption, otherwise all points of P would be collocated with $M(P)$ and $C(P)$. Thus,

$$c > m\left(2 - \frac{2}{n}\right),$$

$$\Rightarrow \qquad cx - cm > 2mx\left(1 - \frac{1}{n}\right) - cm,$$

since $a \neq a'$ and, consequently, $x = ||a - a'|| > 0$,

$$\Rightarrow \qquad c(x - m) > m\left[2x\left(1 - \frac{1}{n}\right) - c\right],$$

$$\Rightarrow \qquad c(m - x) < m\left[c - 2x\left(1 - \frac{1}{n}\right)\right],$$

$$\Rightarrow \qquad c(m - x) < mc',$$

by Lem. 8,

$$\Rightarrow \qquad cm' < mc',$$

since $M(P) = M(P')$ by Lem. 7 and, consequently, $m = m' + x$,

$$\Rightarrow \qquad \frac{c}{m} < \frac{c'}{m'}, \tag{11}$$

since m and m' are sums of non-negative terms.

This contradicts our assumption that P maximizes the approximation factor of C. Therefore, $c \leq m(2 - 2/n)$. That is, for all nonempty finite multisets P,

$$\sum_{p \in P} ||p - C(P)|| \leq \left(2 - \frac{2}{n}\right) \sum_{p \in P} ||p - M(P)||,$$

where $n = |P|$. □

The approximation bound $2 - 2/n$ is realized by $n - 1$ points located at the origin and a single point located at $(1, 0)$.

As shown by Bespamyatnikh et al. [BBKS00], any function defined by a convex combination of a set of mobile points moves with maximum relative velocity at most one. Since the centre of mass is a convex combination of the points of P, this result implies that the centre of mass is 1-stable. The bound is trivially tight, as demonstrated by any translation of the points of P.

4 The Rectilinear Median

The rectilinear median is defined analogously to the Euclidean median with respect to the ℓ_1 norm instead of the ℓ_2 norm.

Definition 10 *Given an arbitrary nonempty finite multiset P in $\mathbb{R}^2$, a* rectilinear median *of P is a point in $\mathbb{R}^2$, denoted $R(P)$, that minimizes*

$$\sum_{p \in P} ||x - p||_1, \qquad (12)$$

when $x = R(P)$ and where $|| \cdot ||_1$ denotes the ℓ_1 norm.

Function R is invariant under translation and uniform scaling, but not under rotation or reflection. If $|P|$ is even, then R may not be unique; in this case, we assign $R(P)$ to be the midpoint of the rectangular region of points that define rectilinear medians of P.

The rectilinear median is found in $\Theta(n)$ time by solving two independent one-dimensional median problems on the x- and y-coordinates of the points of P. The rectilinear median is also known as coordinate median [Wes93] and ℓ_1 median.

Bespamyatnikh et al. [BBKS00] show that the relative velocity of the rectilinear median of a set of mobile points in $\mathbb{R}^2$ is at most $\sqrt{2}$. Furthermore, this bound is tight. It is straightforward to show that maximum relative velocity is inversely related to stability, implying that R is $(1/\sqrt{2})$-stable.

Bespamyatnikh et al. [BBKS00] also show that the rectilinear median provides a $\sqrt{2}$-approximation of the Euclidean median. We show this bound is tight in the following example. Let $2k$ points lie at $(1, 0)$, let $k + 1$ points lie at $(0, 1)$, and let $k + 1$ points lie at $(0, -1)$. See Fig. 3. The unique rectilinear median of P lies at $(0, 0)$. Since the points of P are not collinear, the position of the Euclidean median of P is also unique. Consequently, by the symmetry of P and the invariance of $M(P)$ under reflection, $M(P)$ must lie on the x-axis.

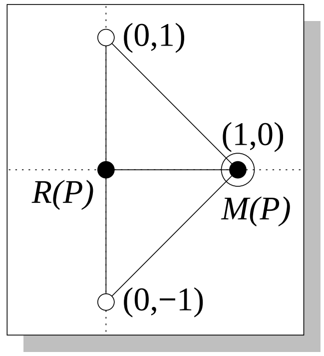

Fig. 3. example realizing the approximation factor of the rectilinear median

The Euclidean median sum of P is

$$2k|1 - M_x(P)| + 2(k+1)\sqrt{M_x(P)^2 + 1}. \tag{13}$$

For any $k \geq 3$, it is straightforward to show that (13) is minimized at $M_x(P) = 1$. Therefore, the Euclidean median of P lies at $(1, 0)$. We obtain the following lower bound on the approximation factor of R:

$$\begin{aligned} \lambda \geq & \lim_{k\to\infty} \frac{\sum_{p\in P} ||p - R(P)||}{\sum_{q\in P} ||q - M(P)||} \\ = & \lim_{k\to\infty} \frac{2k+1}{(k+1)\sqrt{2}} \\ = & \sqrt{2}. \end{aligned} \tag{14}$$

5 The Projection Median

The definition of the Euclidean median is the most natural generalization of the one-dimensional median to higher dimensions. Eq. (1), however, suggests other possible generalizations. One possibility is to project points onto a line through the origin, to find the one-dimensional median of the projection, and to integrate these one-dimensional medians for all lines through the origin.

Let l_θ denote the line through the origin parallel to the unit vector $u_\theta = (\cos\theta, \sin\theta)$. Expressed in slope-intercept form, l_θ is the line $y = \tan\theta\ x$. Given a multiset of points P in $\mathbb{R}^2$ and an angle $\theta \in [0, \pi)$, let P_θ denote the multiset defined by the projection of P onto line l_θ. See Fig. 4A. That is,

$$P_\theta = \{u_\theta\langle p, u_\theta\rangle \mid p \in P\}. \tag{15}$$

Let $p \in \mathbb{R}^2$ be any fixed point. The average over all projections of p onto lines

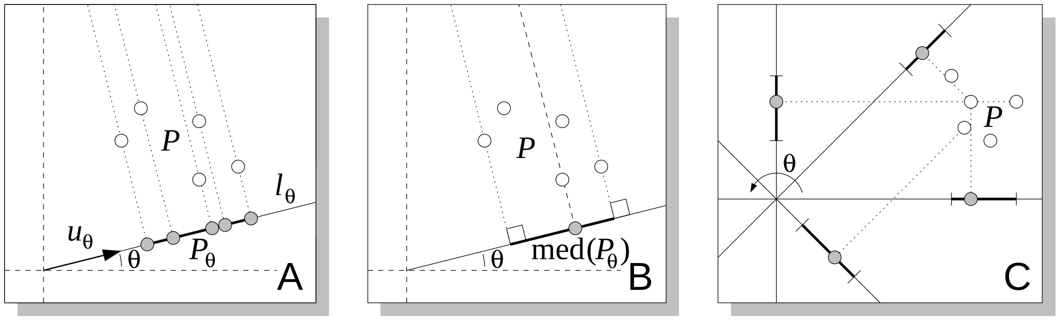

Fig. 4. defining the projection median

l_θ is

$$\frac{1}{\pi}\int_0^\pi u_\theta\langle p, u_\theta\rangle \; d\theta = \frac{p}{2},$$

suggesting an additional factor of 2 is necessary in the following definition for a location function:

Definition 11 *The* projection median *of a nonempty finite multiset P in $\mathbb{R}^2$ is*

$$\Pi(P) = \frac{2}{\pi}\int_0^\pi \mathrm{med}(P_\theta)\; d\theta, \tag{16}$$

where $\mathrm{med}(P_\theta)$ *is the median of the projection of P onto the line $y = \tan\theta\ x$.*

If $|P|$ is even, then P_θ may not have a unique median. In this case, let $\mathrm{med}(P_\theta)$ denote the midpoint of the interval of points on l_θ that define medians of P_θ. It is straightforward to show that Π is invariant under similarity transformations.

The formulation of the projection median displays some resemblance to the Steiner centre, which can be expressed similarly to (16) in $\mathbb{R}^2$ by replacing $\mathrm{med}(P_\theta)$ with $\frac{u_\theta}{2}(\min_{p\in P}\langle p, u_\theta\rangle + \max_{q\in P}\langle q, u_\theta\rangle)$, the centre of P_θ [DK06, Dur06].

Although this paper examines location functions defined over finite multisets, these can also be defined over bounded regions in $\mathbb{R}^d$ with an associated density function. In this case, the sums in Defs. 1, 6, 10, and 11 are replaced by integrals. This family of problems is referred to as *continuous facility location.* See Fekete et al. [FMW05] for an examination of the continuous rectilinear median.

The projection median can be found using techniques similar to those developed by Bespamyatnikh et al. [BKS00]. In brief, as θ varies from 0 to π, the point(s) in P that induce $\mathrm{med}(P_\theta)$ are identified by maintaining a line (perpendicular to l_θ) that partitions P into two sets of at most $\lfloor n/2\rfloor$ points each. The convex hull of each partition is maintained as the line rotates, requiring $O(\log^2 n)$ time per update [OvL81]. See Fig. 5 Since the dual problem to maintaining these partitions corresponds to an $(n/2)$-level, we require at most $O(n^{4/3})$ such updates [Dey98]. Between updates, the contribution to $\Pi(P)$ of

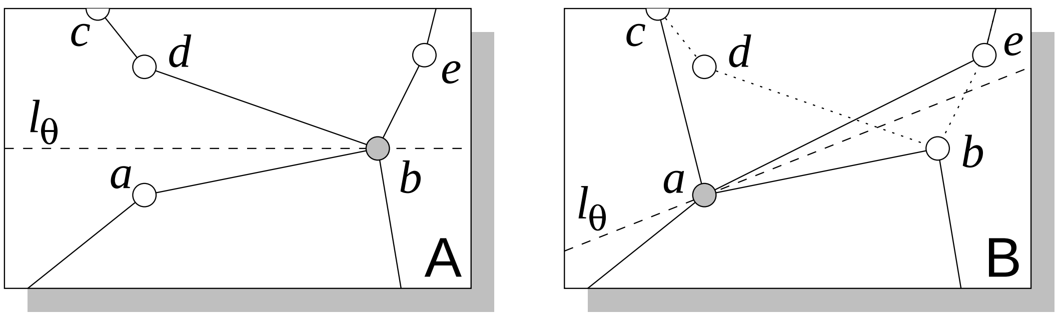

Fig. 5. maintaining the convex hulls of both partitions as l_θ rotates

the point(s) that induce $\mathrm{med}(P_\theta)$ is calculated in $O(1)$ time. Together, these give an $O(n^{4/3}\log^2 n)$-time algorithm. This can be improved to $O(n^{4/3}\log^{1+\epsilon} n)$ amortized time using the dynamic convex hull data structure of Chan [Cha01] or $O(n^{4/3})$ expected time [Cha99]. Providing details of this algorithm is not the goal of this paper; rather, we focus on the properties of approximation factor and stability.

Let d_ϕ denote the ℓ_1 norm relative to a rotation by ϕ of the reference axis. That is, $d_\phi(x) = ||f_\phi(x)||_1$, where f_ϕ is a clockwise rotation about the origin by ϕ. Let $R_\phi(P) = f_\phi^{-1}(R(f_\phi(P)))$ denote the rectilinear median with respect to norm d_ϕ. Observe that $R_\phi(P) = \mathrm{med}(P_\phi) + \mathrm{med}(P_{\phi+\pi/2})$. Consequently,

$$\begin{aligned}\Pi(P) =& \frac{2}{\pi}\int_0^\pi \mathrm{med}(P_\theta)\ d\theta\\ =& \frac{2}{\pi}\left[\int_0^{\pi/2} \mathrm{med}(P_\theta)\ d\theta + \int_{\pi/2}^{\pi} \mathrm{med}(P_\theta)\ d\theta\right]\\ =& \frac{2}{\pi}\int_0^{\pi/2} \mathrm{med}(P_\theta) + \mathrm{med}(P_{\theta+\pi/2})\ d\theta\\ =& \frac{2}{\pi}\int_0^{\pi/2} R_\theta(P)\ d\theta. \end{aligned} \tag{17}$$

Theorem 12 *The projection median provides a $(4/\pi)$-approximation of the Euclidean median.*

Proof. Let P denote any nonempty finite multiset of points in $\mathbb{R}^2$. We bound the approximation factor of $\Pi(P)$:

$$\begin{aligned}
&\frac{\sum_{p\in P}||\Pi(P)-p||}{\sum_{q\in P}||M(P)-q||} \\
&= \frac{\sum_{p\in P}\left|\left|\frac{2}{\pi}\int_0^{\pi/2}R_\theta(P)\ d\theta - p\right|\right|}{\sum_{q\in P}||M(P)-q||}, && \text{by (17)}, \\
&= \frac{\sum_{p\in P}\left|\left|\frac{2}{\pi}\int_0^{\pi/2}R_\theta(P)\ d\theta - \frac{2}{\pi}\int_0^{\pi/2}p\ d\theta\right|\right|}{\sum_{q\in P}||M(P)-q||} \\
&= \frac{2}{\pi}\frac{\sum_{p\in P}\left|\left|\int_0^{\pi/2}R_\theta(P) - p\ d\theta\right|\right|}{\sum_{q\in P}||M(P)-q||} \\
&\le \frac{2}{\pi}\frac{\sum_{p\in P}\int_0^{\pi/2}||R_\theta(P) - p||\ d\theta}{\sum_{q\in P}||M(P)-q||}, && \text{by } \triangle \text{ ineq.}, \\
&\le \frac{2}{\pi}\frac{\sum_{p\in P}\int_0^{\pi/2}d_\theta(R_\theta(P) - p)\ d\theta}{\sum_{q\in P}||M(P)-q||}, && \text{(18a)}
\end{aligned}$$

since $\forall x\ ||x||_1 \ge ||x||$ and, similarly, $\forall x \forall \phi\ d_\phi(x) \ge ||x||$,

$$\begin{aligned}
&= \frac{2}{\pi}\frac{\int_0^{\pi/2}\sum_{p\in P}d_\theta(R_\theta(P) - p)\ d\theta}{\sum_{q\in P}||M(P)-q||} \\
&\le \frac{2}{\pi}\frac{\int_0^{\pi/2}\sum_{p\in P}d_\theta(M(P) - p)\ d\theta}{\sum_{q\in P}||M(P)-q||}, && \text{(18b)}
\end{aligned}$$

since $R_\theta(P)$ minimizes the sum of the d_θ distances to points of P,

$$= \frac{2}{\pi}\frac{\int_0^{\pi/2}\sum_{p\in P}\Big[|\sin(\theta-\alpha_p)| + |\cos(\theta-\alpha_p)|\Big]\cdot||M(P)-p||\ d\theta}{\sum_{q\in P}||M(P)-q||}, \qquad \text{(18c)}$$

where $\alpha_p = \arctan[(M_y(P)-p_y)/(M_x(P)-p_x)] \bmod \frac{\pi}{2}$ (see Fig. 6),

$$\begin{aligned}
&= \frac{2}{\pi}\frac{\int_0^{\pi}\sum_{p\in P}|\sin(\theta-\alpha_p)|\cdot||M(P)-p||\ d\theta}{\sum_{q\in P}||M(P)-q||} \\
&= \frac{2}{\pi}\frac{\int_0^{\pi}\sum_{p\in P}|\sin\theta|\cdot||M(P)-p||\ d\theta}{\sum_{q\in P}||M(P)-q||} \\
&= \frac{2}{\pi}\frac{\sum_{p\in P}||M(P)-p||}{\sum_{q\in P}||M(P)-q||}\int_0^{\pi}|\sin\theta|\ d\theta \\
&= \frac{2}{\pi}\int_0^{\pi}|\sin\theta|\ d\theta \\
&= \frac{4}{\pi} \\
&\approx 1.2732. && \text{(18d)}
\end{aligned}$$

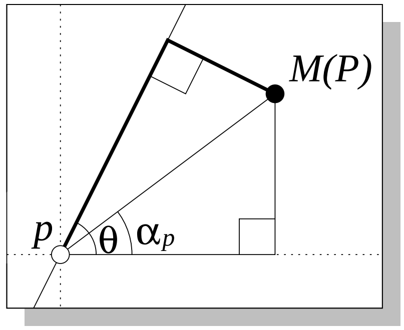

Fig. 6. $d_\theta(M(P)-p) = \big[|\sin(\theta-\alpha_p)| + |\cos(\theta-\alpha_p)|\big] \cdot ||M(P)-p||$

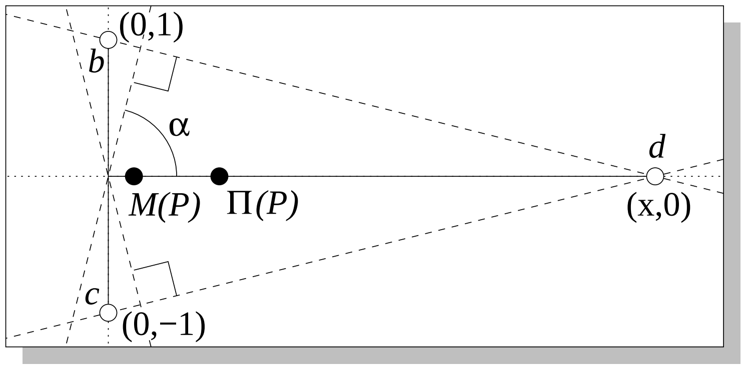

Fig. 7. example realizing the lower bound in Thm. 13

Therefore, for any nonempty finite multiset of points P in $\mathbb{R}^2$,

$$\sum_{p\in P} ||\Pi(P)-p|| \le \frac{4}{\pi} \sum_{q\in P} ||M(P)-q||. \ \square \tag{19}$$

Although we have not been able to prove that the bound in (19) is tight, we provide the following lower bound:

Theorem 13 *The projection median cannot guarantee an approximation factor less than* $\sqrt{4/\pi^2+1}$ *in the worst case.*

Proof. Let multiset P be defined by k points located at $b=(0,1)$, k points located at $c=(0,-1)$, and a single point located at $d=(x,0)$, for some $k\in\mathbb{N}$ and $x\in\mathbb{R}^+$. Let $\alpha = \pi/2 - \arctan(1/x) = \arctan x$. See Fig. 7.

We first derive the position of $M(P)$. Since the points of P are not collinear, the position of the Euclidean median of P is unique. Consequently, by the symmetry of P and the invariance of $M(P)$ under reflection, $M(P)$ must lie on the x-axis. The Euclidean median sum of P is

$$2k\sqrt{1+M_x(P)^2} + |x - M_x(P)|. \tag{20}$$

Simple calculus shows that (20) is minimized at $M_x(P) = 1/\sqrt{4k^2-1}$. Consequently, $M(P) = (1/\sqrt{4k^2-1}, 0)$.

By (16), the projection median of P is located at

$$\begin{aligned}\Pi(P) &= \frac{2}{\pi}\left[\int_0^{\alpha} u_\theta\langle b, u_\theta\rangle \; d\theta + \int_{\alpha}^{\pi-\alpha} u_\theta\langle d, u_\theta\rangle \; d\theta + \int_{\pi-\alpha}^{\pi} u_\theta\langle c, u_\theta\rangle \; d\theta\right] \\ &= \frac{2}{\pi}\left[\int_0^{\alpha} u_\theta \sin\theta \; d\theta + \int_{\alpha}^{\pi-\alpha} x u_\theta \cos\theta \; d\theta - \int_{\pi-\alpha}^{\pi} u_\theta \sin\theta \; d\theta\right] \\ &= \left(\frac{2x}{\pi}\arctan\left(\frac{1}{x}\right), 0\right). \end{aligned} \tag{21}$$

The approximation factor λ is at least

$$\begin{aligned}\lambda &\geq \lim_{\substack{x\to\infty \\ k\to\infty}} \frac{\sum_{p\in P} ||\Pi(P) - p||}{\sum_{q\in P} ||M(P) - q||} \\ &= \lim_{\substack{x\to\infty \\ k\to\infty}} \frac{2k\sqrt{\frac{4x^2}{\pi^2}\arctan^2\left(\frac{1}{x}\right) + 1} + x - \frac{2x}{\pi}\arctan\left(\frac{1}{x}\right)}{2k\sqrt{\frac{1}{4k^2-1} + 1} + x - \frac{1}{\sqrt{4k^2-1}}} \\ &= \lim_{x\to\infty} \sqrt{\frac{4x^2}{\pi^2}\arctan^2\left(\frac{1}{x}\right) + 1} \\ &= \sqrt{\frac{4}{\pi^2} + 1} \\ &> 1.1854. \; \square \end{aligned} \tag{22}$$

We now derive a tight bound on the stability of Π.

Theorem 14 *The projection median is $(\pi/4)$-stable.*

Proof. Choose any nonempty and finite P in $\mathbb{R}^2$. Let $f : P \to \mathbb{R}^2$ be any ϵ-perturbation of P. Let multiset $Q = f(P)$. Since Π is invariant under rotation and translation, without loss of generality assume $\Pi(P)$ and $\Pi(Q)$ lie on the x-axis. The one-dimensional median is 1-stable. Consequently, for any ϵ-perturbation of P, f,

$$||\text{med}(P_\theta) - \text{med}(Q_\theta)|| \leq \max_{p\in P} ||p - f(p)||.$$

Thus, for any θ,

$$\begin{aligned}|\text{med}(P_\theta)_x - \text{med}(Q_\theta)_x| &= |\cos\theta| \cdot ||\text{med}(P_\theta) - \text{med}(Q_\theta)|| \\ &\leq |\cos\theta| \cdot \max_{p\in P} ||p - f(p)|| \\ &\leq |\cos\theta| \cdot \epsilon. \end{aligned} \tag{23}$$

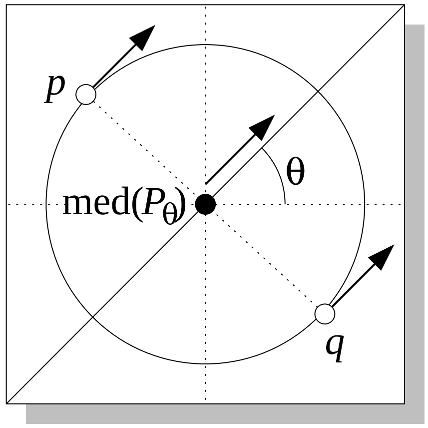

Fig. 8. example realizing the bound in Thm. 14

We bound the stability of Π from below by

$$\begin{aligned} ||\Pi(P) - \Pi(f(P))|| &= |\Pi(P)_x - \Pi(Q)_x| \\ &= \left|\frac{2}{\pi}\int_0^\pi \mathrm{med}(P_\theta)_x \, d\theta - \frac{2}{\pi}\int_0^\pi \mathrm{med}(Q_\theta)_x \, d\theta\right| \\ &= \frac{2}{\pi}\left|\int_0^\pi \mathrm{med}(P_\theta)_x - \mathrm{med}(Q_\theta)_x \, d\theta\right| \\ &\le \frac{2}{\pi}\int_0^\pi |\mathrm{med}(P_\theta)_x - \mathrm{med}(Q_\theta)_x| \, d\theta \\ &\le \frac{2}{\pi}\int_0^\pi |\cos\theta| \cdot \epsilon \, d\theta \\ &= \frac{4\epsilon}{\pi}. \end{aligned} \tag{24}$$

Hence, $\kappa \ge \frac{\pi}{4}$. Therefore, for all nonempty finite multisets P in $\mathbb{R}^2$,

$$\forall \epsilon > 0, \ \forall f \in F_\epsilon^P, \ \frac{\pi}{4}||\Pi(P) - \Pi(f(P))|| \le \epsilon. \ \square \tag{25}$$

The bound in (25) is shown to be tight by the following example. Let P be an even number of points uniformly distributed on the unit circle centred at the origin. Choose any $\epsilon \in (0, 1)$ and define an ϵ-perturbation such that points above the x-axis move right (clockwise) in a direction tangent to the circle while points below the x-axis move right (counter-clockwise) in the opposite direction. Every point p in P has a corresponding point in P, $q = -p$, opposite the origin from p. The midpoint of each such pair of points p and q defines $\mathrm{med}(P_\theta)$ for some P_θ (corresponding to the projection onto the line perpendicular to $p - q$). The resulting change in the position of $\mathrm{med}(P_\theta)$ is identical to the change at p and q. See Fig. 8. The resulting stability corresponds exactly to that derived in (25).

location function	notation	approximation	stability
Euclidean median	M	1	0
rectilinear median	R	$\sqrt{2} \approx 1.4142$	$1/\sqrt{2} \approx 0.7071$
centre of mass	C	2	1
projection median	Π	$[\sqrt{4/\pi^2+1}, 4/\pi]$	$\pi/4$
		$\approx [1.1854, 1.2732]$	≈ 0.7854

Table 1
comparing location functions in $\mathbb{R}^2$

6 Conclusion

6.1 Evaluation

As shown in Sec. 2, the Euclidean median, M, is arbitrarily unstable. Guaranteeing any degree of stability in a location function implies an increase in the Euclidean median sum (1) and necessitates approximation by a location function. In this paper we introduced the projection median, Π, as a stable approximation of the Euclidean median. We now compare the stability and approximation factor of Π against those of two common location functions: the rectilinear median, R, and the centre of mass, C. These results are summarized in Tab. 1.

Observe that Π is more stable and guarantees a better approximation factor than R. Similarly, Π guarantees a better approximation than C, but one that is not as stable. Depending on the degree of stability required and approximation factor necessary for a particular application, either the centre of mass or the projection median may be preferred.

6.2 Applications to Mobile Facility Location

The projection median's benefits extend beyond its definition as a median of a set of static points. Recently, several questions of facility location have been posed within the setting of mobile facility location (e.g., [AGG02, AH01, BBKS00, DK06, DK07, Dur06, Her05]). Given a set of mobile clients moving continuously and with bounded velocity in $\mathbb{R}^2$, the fitness of a mobile facility is determined both by its approximation factor and also by its maximum velocity and continuity of its motion. The stability of a location function is inversely related to the maximum velocity of a mobile facility, providing further motivation for the need of stability in a location function. Thus, the projection median defines the position of a mobile facility that approximates the mobile

Euclidean median with a factor of $4/\pi$ while maintaining a maximum velocity of at most $4/\pi$ relative to the velocity of the clients.

6.3 Directions for Future Research

Def. 11 has a natural generalization to $\mathbb{R}^d$, suggesting that the properties that make the projection median a good location function might not be limited to $\mathbb{R}^2$, but may extend to three or higher dimensions. The projection median is (2/3)-stable in three dimensions [Dur06]. Obtaining a good bound on its approximation factor in three dimensions, however, remains open.

Sec. 2 begins with an example consisting of a set of four points and a perturbation of those points that illustrate the instability of the Euclidean median in $\mathbb{R}^2$. The four points in the example are nearly collinear. To what extent is the instability of the Euclidean median attributable to this degeneracy (collinearity)? Expanding on this idea, can the stability of the Euclidean median be bounded by a function of the ratio of a point set's width to its diameter? If so, then conceivably some kind of hybrid approach combining the Euclidean median and the projection median might provide a better stable approximation by adapting to this ratio. Alternatively, are there unstable configurations of point sets for which this ratio is bounded? Similarly, to what extend does instability arise from parity? Are there unstable configurations of odd-sized point sets? These questions remain open.

Acknowledgements

The authors would like to thank Mark Keil with whom preliminary ideas on the projection median were discussed. In addition, the authors acknowledge an anonymous reviewer for his/her helpful suggestions, some of which inspired the discussion in Sec. 6.3.

References

[AdBG+05] Pankaj K. Agarwal, Mark de Berg, Jie Gao, Leonidas J. Guibas, and Sariel Har-Peled. Staying in the middle: Exact and approximate medians in $\mathbb{R}^1$ and $\mathbb{R}^2$ for moving points. In *Proceedings of the Canadian Conference on Computational Geometry*, volume 17, pages 42–45, 2005.

[AGG02] Pankaj K. Agarwal, Jie Gao, and Leonidas J. Guibas. Kinetic medians and *kd*-trees. In *Proceedings of the Tenth European Sym-*

posium on Algorithms, volume 2461 of *Lecture Notes in Computer Science*, pages 5–16, 2002.

[AH01] Pankaj K. Agarwal and Sariel Har-Peled. Maintaining approximate extent measures of moving points. In *Proceedings of the Symposium on Discrete Algorithms*, pages 148–157. ACM Press, 2001.

[Baj88] Chanderjit Bajaj. The algebraic degree of geometric optimization problems. *Discrete and Computational Geometry*, 3:177–191, 1988.

[BBKS00] Sergei Bespamyatnikh, Binay Bhattacharya, David Kirkpatrick, and Michael Segal. Mobile facility location. In *Proceedings of the International ACM Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, volume 4, pages 46–53, 2000.

[BKS00] Sergei Bespamyatnikh, David Kirkpatrick, and Jack Snoeyink. Generalizing ham sandwich cuts to equitable subdivisions. *Discrete and Computational Geometry*, 24(4):605–622, 2000.

[BMM03] Prosenjit Bose, Anil Maheshwari, and Pat Morin. Fast approximations for sums of distances, clustering and the Fermat-Weber problem. *Computational Geometry: Theory and Applications*, 24(3):135–146, 2003.

[CFPS03] Mark Cieliebak, Paola Flocchini, Giuseppe Prencipe, and Nicoal Santoro. Solving the robots gathering problem. In *Proceedings of the Thirtieth International Colloquium on Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 1181–1196, 2003.

[Cha99] Timothy M. Chan. Remarks on k-level algorithms in the plane. Manuscript, 1999.

[Cha01] Timothy M. Chan. Dynamic planar convex hull operations in near-logarithmic amortized time. *Journal of the ACM*, 48(1):1–12, 2001.

[CT90] R. Chandrasekaran and A. Tamir. Algebraic optimization: The Fermat-Weber location problem. *Mathematical Programming*, 46:219–224, 1990.

[Dey98] Tamal K. Dey. Improved bounds for planar k-sets and related problems. *Discrete and Computational Geometry*, 19:373–382, 1998.

[dF91] Pierre de Fermat. Tome I. In M. M. P. Tannery and C. Henry, editors, *Oeuvres*. Gauthier-Villars et Fils, Paris, 1891.

[DK05] Stephane Durocher and David Kirkpatrick. The projection median of a set of points in $\mathbb{R}^2$. In *Proceedings of the Canadian Conference on Computational Geometry*, volume 17, pages 46–50, 2005.

[DK06] Stephane Durocher and David Kirkpatrick. The Steiner centre: Stability, eccentricity, and applications to mobile facility location.

International Journal of Computational Geometry and Applications, 16(4):345–371, 2006.

[DK07] Stephane Durocher and David Kirkpatrick. Bounded-velocity approximations of mobile Euclidean 2-centres. *International Journal of Computational Geometry and Applications*, 2007. To appear.

[DKSW02] Zvi Drezner, Kathrin Klamroth, Anita Schöbel, and George O. Wesolosky. The Weber problem. In Zvi Drezner and Horst W. Hamacher, editors, *Facility Location: Applications and Theory*, pages 1–36. Springer, New York, 2002.

[Dur06] Stephane Durocher. *Geometric Facility Location under Continuous Motion*. PhD thesis, University of British Columbia, 2006.

[FMW05] Sándor P. Fekete, Joseph S. B. Mitchell, and Karin Weinbrecht. On the continuous Fermat-Weber problem. *Operations Research*, 53:61–76, 2005.

[Hak00] S. Louis Hakimi. Location theory. In Rosen, Michaels, Gross, Grossman, and Shier, editors, *Handbook of Discrete and Combinatorial Mathematics*. CRC Press, 2000.

[Her05] John Hershberger. Smooth kinetic maintenance of clusters. *Computational Geometry: Theory and Applications*, 31:3–30, 2005.

[HLP+87] Pierre Hansen, Martine Labbé, Dominique Peeters, Jacques-François Thisse, and Vernon J. Henderson. *Systems of Cities and Facility Location*. Harwood Academic Publishers, New York, 1987.

[Ind99] Piotr Indyk. Sublinear time algorithms for metric space problems. In *Proceedings of the Symposium on the Theory of Computing*, volume 31, pages 428–434, 1999.

[Kim98] Clark Kimberling. Triangle centers and central triangles. *Congressus Numerantium*, 129:1–295, 1998.

[KM97] Y. S. Kupitz and H. Martini. Geometric aspects of the generalized Fermat-Torricelli problem. In *Intuitive Geometry*, volume 6, pages 55–127. Bolyai Society Mathematical Studies, Budapest, 1997.

[KV97] J. Krarup and S. Vajda. On Torricelli's geometrical solution to a problem of Fermat. *IMA Journal of Management Mathematics*, 8(3):215–224, 1997.

[LMW88] Robert F. Love, James G. Morris, and George O. Wesolowsky. *Facilities Location*. North-Holland, New York, 1988.

[OvL81] Mark H. Overmars and Jan van Leeuwen. Maintenance of configurations in the plane. *Journal of Computer and System Sciences*, 23(2):166–204, 1981.

[Sch73] Alain Schärlig. About the confusion between the center of gravity and Weber's optimum. *Regional and Urban Economics*, 3(4):371–382, 1973.

[Sch03] Konrad Schlude. From robotics to facility location: Contraction functions, Weber point, convex core. Technical Report 403, Eidgenössische Technische Hochschule Zürich, 2003.

[Web22] Alfred Weber. *Über den Standort der Industrie*. J. C. B. Mohr (Paul Siebeck), Tübingen, 1922.

[Wei] Eric W. Weisstein. Mathworld – a Wolfram web resource. http://mathworld.wolfram.com.

[Wei37] Endre Weiszfeld. Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Mathematical Journal*, 43:355–386, 1937.

[Wes93] George O. Wesolowsky. The Weber problem: History and perspectives. *Location Science*, 1(1):5–23, 1993.