Beiträge zur Algebra und Geometrie
Contributions to Algebra and Geometry
Volume 48 (2007), No. 2, 367-381.

Minimal Enclosing Hyperbolas of Line Sets

Hans-Peter Schröcker

University Innsbruck, Institute of Basic Sciences in Engineering
Unit Geometry and CAD
e-mail: hans-peter.schroecker@uibk.ac.at

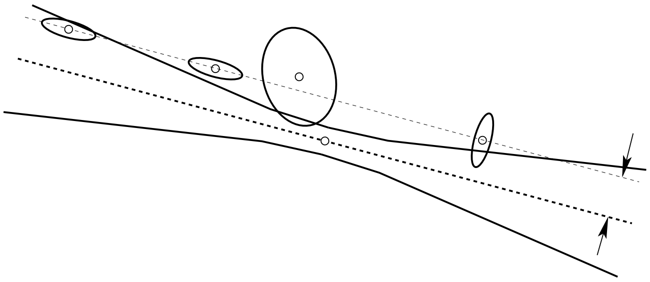
Abstract. We prove the following theorem: If H is a slim hyperbola that contains a closed set $\mathcal{S}$ of lines in the Euclidean plane, there exists exactly one hyperbola $H_{\min}$ of minimal volume that contains $\mathcal{S}$ and is contained in H. The precise concepts of "slim", the "volume of a hyperbola" and "straight lines or hyperbolas being contained in a hyperbola" are defined in the text.

1. Introduction and preliminaries

Let $\mathcal{P} \subset \mathbb{R}^d$ be a bounded, closed point set, denote the convex hull of $\mathcal{P}$ by $\mathrm{ch}(\mathcal{P})$ and assume the affine hull of $\mathcal{P}$ is $\mathbb{R}^d$. By a classical result of convex geometry there exists a unique ellipsoid $E_{\min}$ of minimal volume that contains $\mathcal{P}$ – the *Löwner ellipsoid* $E_{\min}$ of $\mathcal{P}$ (sometimes also called the Löwner-John ellipsoid or minimal ellipsoid of $\mathcal{P}$).

Similarly, there exists a unique ellipsoid $E_{\max}$ of maximal volume that is contained in $\mathrm{ch}(\mathcal{P})$ – the *John ellipsoid* or maximal ellipsoid of $\mathcal{P}$. Löwner and John ellipsoids are closely related and share many analogous properties. The uniqueness of $E_{\min}$ was first proved in [2] for $d = 2$ and in [4] for arbitrary dimension. The most important proof, however, is due to F. John [12]. John not only proved the mere uniqueness of the maximal and minimal ellipsoids, he also characterized them in terms of his famous "partition of identity". The results of John have been extended, generalized and simplified in a number of subsequent publications. To mention but a few, we cite [1], [8], [17] and, most recently, [9], [10].

Two of the attractive geometric properties of Löwner and John ellipsoids are:



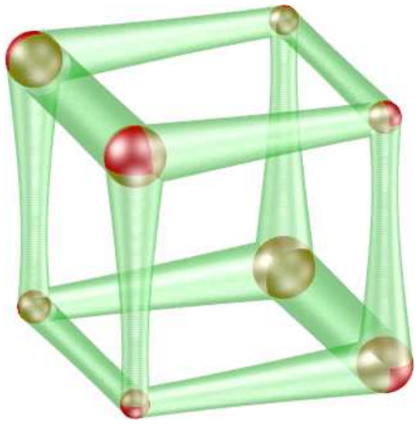


Figure 1. Examples from [5] (courtesy W. Förstner) and [20] of uncertain straight lines being represented by hyperboloids

- $E_{\min}$ and $E_{\max}$ are affinely related to $\mathcal{P}$. If $\alpha\colon \mathbb{R}^d \to \mathbb{R}^d$ is an affine transformation, then $\alpha(E_{\min})$ is the Löwner ellipsoid and $\alpha(E_{\max})$ the John ellipsoid of $\alpha(\mathcal{P})$. This affine relation can be exploited to simplify proofs of uniqueness of $E_{\min}$ and $E_{\max}$ (as for example in [4] or [13]).
- If $E_{\min}$ is scaled about its center by a factor of d^{-1}, it lies inside $\operatorname{ch}(\mathcal{P})$. If $\mathcal{P}$ is centrally symmetric, the scaling factor can be tightened to $d^{-1/2}$ (see [12] or [3, Section 8.4]). Analogously, scaling the John ellipsoid about its center by a factor of d (or by a factor of $d^{1/2}$ in case of a centrally symmetric point set) yields an ellipsoid containing $\mathcal{P}$.

Löwner and John ellipsoids have numerous applications (see [7] and the references therein), basically because they are simple shapes that can be used to reliably represent point sets. The idea of replacing point sets by enclosing or approximating ellipsoids has been used in recent publications on worst-case tolerancing in geometry [18, 21, 11]. While geometric constructions not only involve points but also straight lines or general subspaces (and many more primitives), the question of representing sets of subspaces by "simple shapes" has so far only been remotely touched. In this context, it seems natural to replace sets of lines by hyperbolas or, more generally, replace sets of subspaces by suitable hyperboloids (see Figure 1). However, theoretical fortification of this concept is not available.

In the present paper we consider a set of lines $\mathcal{S}$ and the "smallest" hyperbola $H_{\min}$ that "contains" $\mathcal{S}$. In order to do so, we need to define

1. the concept of "a straight line S being contained in a hyperbola H" and
2. the volume of a hyperbola H.

Anticipating Definition 2 we say that S is contained in H if H and S have at most one real intersection point. The volume $m(H)$ will be defined in Section 2 as the unique Euclidean measure of all straight lines contained in H.

Our main result is a proof of uniqueness of $H_{\min}$ provided certain conditions are met (Theorem 1). It is clear that these conditions have to be more restrictive than the pre-requisites for the uniqueness of the Löwner ellipsoid. As a simple

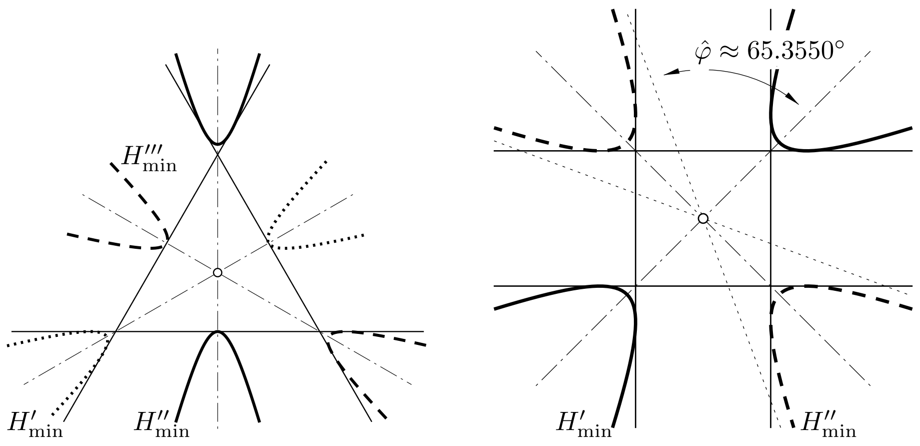


Figure 2. Minimal enclosing hyperbolas of line sets with rotational symmetry

example, consider a line set $\mathcal{S}$ with rotational symmetry; its minimal enclosing hyperbola cannot be unique (Figure 2). We will show the uniqueness of the minimal enclosing hyperbola $H_{\min}$ among all hyperbolas that enclose $\mathcal{S}$ and are "contained" in a given "slim" hyperbola H.

The main difficulty of the proof, as opposed to known proofs of uniqueness of the Löwner ellipsoids, is the lacking affine relation of $H_{\min}$ and $\mathcal{S}$. It prevents easy adaption of the geometric reasoning of [4, 16] or the optimization theoretic setup of [13]. We will instead give a computational proof of uniqueness, based on properties of the volume function $m(H)$ of hyperbolas and on the geometry of conics and dual conics.

The remaining part of this paper is organized as follows. We define the measure for sets of straight lines and derive explicit formulas for the volume of a hyperbola in Section 2. In this section we also study some properties of the volume function. Section 3 is dedicated to the proof of an auxiliary lemma concerning the volume of hyperbolas in a dual linear pencil of conics. It is used for the proof of the main theorem in Section 4. Generalizations of this paper's topic and open questions are discussed in the final Section 5.

2. The volume of a hyperbola

In this section we define a measure $m(H)$ for the set of lines $\mathcal{S}$ contained in a hyperbola H. One can think of several possible definitions for $m(H)$ but it seems sensible to use the (essentially unique) measure for sets of straight lines in the plane that is invariant with respect to Euclidean transformations.

Let $\mathcal{S}$ be a set of straight lines $S\colon x\cos\varphi + y\sin\varphi = p$. The integral

$$m(\mathcal{S}) = \int_{\mathcal{S}} dp \wedge d\varphi \tag{1}$$

is called the *density for straight lines* ([19, Chapter 3]). Up to a constant positive factor, it is the unique measure that is invariant under Euclidean transformations, i.e., $m(\mathcal{S})$ and $m(\alpha(\mathcal{S}))$ are equal for all sets of lines $\mathcal{S}$ and all proper Euclidean transformations $\alpha\colon \mathbb{R}^2 \to \mathbb{R}^2$. By convention, the measure (1) is always taken in absolute value.

Definition 1. *The interior* $\operatorname{int} C$ *of a conic section* C *is the set of intersection points of two different non-real tangents of* C. *The exterior* $\operatorname{ext} C$ *of a conic section* C *is the set of intersection points of two different real tangents of* C. *A point* $\mathbf{p}$ *is contained in* C *if* $\mathbf{p}$ *is element of* $\operatorname{int} C$ *or of* C *itself (Figure* 3*).*

Definition 2. *The line interior* l-int C *of a conic section* C *is the set of straight lines that intersect* C *in two non-real points. The line exterior* l-ext C *of a conic section* C *is the set of straight lines that intersect* C *in two real points. A straight line* S *is contained in* C *if* S *is element of* l-int C *or tangent of* C *(Figure* 3*).*

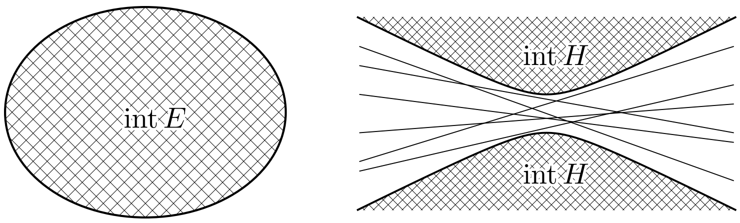


Figure 3. Interior of an ellipse E and a hyperbola H; some straight lines contained in H

The interior of an ellipse and a hyperbola are visualized in Figure 3. Note that S is contained in H if all points $\mathbf{s} \in S$ are elements of $\operatorname{ext} H$. A few straight lines contained in H are also depicted in Figure 3.

We want to compute the measure (1) for the set of lines contained in a hyperbola H. Because of the Euclidean invariance of (1) we may assume that H is given by the equation

$$H\colon \frac{y^2}{a^2} - \frac{x^2}{b^2} = 1. \tag{2}$$

The straight line $S\colon x\cos\varphi + y\sin\varphi = p$ is contained in H if and only if

$$p^2 \le a^2\sin^2\varphi - b^2\cos^2\varphi. \tag{3}$$

With $\varphi_0 = \arctan(b/a)$ and

$$p(\varphi) = \sqrt{a^2\sin^2\varphi - b^2\cos^2\varphi} \tag{4}$$

we compute

$$m(H) = \int_{\varphi_0}^{\pi-\varphi_0} \int_{-p(\varphi)}^{p(\varphi)} dp\, d\varphi = 4\int_0^{\varphi_0} p(\varphi)\, d\varphi. \tag{5}$$

The measure $m(H)$ only depends on the hyperbola's semi-axis lengths a and b. Therefore, we also write $m(H) = m(a,b)$.

2.1. The volume formula in terms of elliptic integrals

The right-hand side of equation (5) is an elliptic integral. In terms of the incomplete elliptic integral of first kind

$$E(z,k) := \int_0^z \frac{\sqrt{1-k^2t^2}}{\sqrt{1-t^2}}\,dt \tag{6}$$

it can be written as

$$m(a,b) = 4a\,E\Bigl(\frac{a}{\sqrt{a^2+b^2}}, \frac{\sqrt{a^2+b^2}}{a}\Bigr). \tag{7}$$

Numeric evaluation of this formula reveals some difficulties: Because of rounding errors, even real arguments a and b might result in complex volumes with small imaginary part. Therefore, we give an alternative expression for (7) using complete elliptic integrals

$$K(k) := \int_0^1 \frac{dt}{\sqrt{1-k^2t^2}\sqrt{1-t^2}} \quad \text{and} \quad E(k) := E(1,k) \tag{8}$$

of first and second kind. We find

$$m(a,b) = 4\sqrt{a^2+b^2}\Bigl(E\Bigl(\frac{a}{\sqrt{a^2+b^2}}\Bigr) - \frac{b^2}{a^2+b^2}K\Bigl(\frac{a}{\sqrt{a^2+b^2}}\Bigr)\Bigr). \tag{9}$$

2.2. Properties of the volume function

From equation (9) we derive the following formulas for limit cases:

- $m(0,b) = 0$, $\lim_{b\to\infty} m(a,b) = 0$: In these two cases, H degenerates and the set of lines contained in H depends on one parameter only. Its measure is zero.
- $m(a,0) = 4a$: H degenerates but contains all straight lines that meet a certain line-segment of length $2a$. The value of the volume function is measure for this set of lines. This is a special case of [19, equation (3.12)].
- $\lim_{a\to\infty} m(a,b) = \infty$: If a goes to infinity, so does the volume of H.

For any positive real s, the volume function satisfies $m(sa,sb) = s\cdot m(a,b)$. Hence, the graph

$$G := \{[a,b,m(a,b)]^T \mid a,b \geq 0\} \tag{10}$$

of the volume function is a cone in $\mathbb{R}^3$ with vertex $[0,0,0]^T$. It is depicted in Figure 4. Its intersection with the plane $a=1$ can be parameterized by

$$m(b) = m(1,b) =$$
$$4(1+b^2)^{-1/2}\Bigl((1+b^2)E\bigl((1+b^2)^{-1/2}\bigr) - b^2K\bigl((1+b^2)^{-1/2}\bigr)\Bigr), \quad b\in[0,\infty). \tag{11}$$

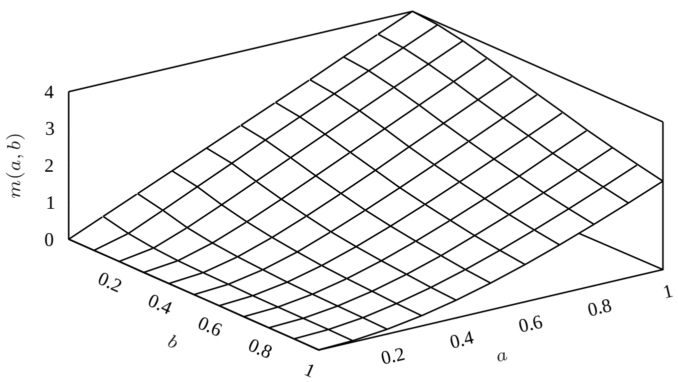


Figure 4. The graph G of the volume function $m(a,b)$

Lemma 1. *There exists exactly one value $\hat{b} > 0$ such that $m(b)$ is strictly concave on $[0,\hat{b})$ and strictly convex on $(\hat{b},\infty)$.*

Proof. The second derivative of $m(b)$ with respect to b is

$$m''(b) = 4t\big(2E(t) - K(t)\big) \quad \text{where } t = (1+b^2)^{-1/2}. \tag{12}$$

The claim of this lemma follows from Lemma 3 (Appendix, page 379). □

Corollary 1. *The volume function $m(a,b)$ is convex for $b : a > \hat{b}$.*

Proof. The corollary follows from Lemma 1 and the fact that the graph of $m(a,b)$ is a cone with vertex $[0,0,0]^T$. □

Remark 1. The numeric value of $\hat{b}$ can be computed from Lemma 3:

$$\hat{t} \approx 0.9089085575 \quad \Longrightarrow \quad \hat{b} \approx 0.4587870596. \tag{13}$$

The volume function $m(a,b)$ is convex for hyperbolas whose asymptotes enclose an angle of $\hat{\varphi} := \arctan(\hat{b}^{-1}) \approx 65.3550°$ or less with the hyperbola's minor axes.

Definition 3. *Let H be a hyperbola and denote the angle between its asymptotes and minor axis by φ. The hyperbola H is called* slim *if $\varphi < \hat{\varphi}$.*

Remark 2. By direct computation it can be shown that for the hyperbolas in the right-hand image of Figure 2 we have $\varphi = \hat{\varphi}$.

In the proof of Theorem 1 (Section 4) we will use a parametric representation of the level set

$$\tilde{m} := \{[a,b]^T \mid m(a,b) = 1\} \tag{14}$$

of the volume function. It can be computed as central projection of the curve (11) from the point $[0,0,0]^T$:

$$\tilde{\mathbf{m}}\colon \begin{cases} a(t) = \dfrac{t}{4\big(K(t)(t^2-1)+E(t)\big)} \\[2ex] b(t) = \dfrac{\sqrt{1-t^2}}{4\big(K(t)(t^2-1)+E(t)\big)} \end{cases} \qquad t \in (0,1]. \tag{15}$$

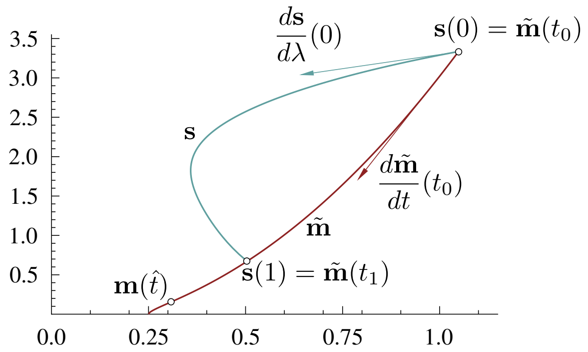


Figure 5. The level-set $\tilde{\mathbf{m}}$, the curve $[a(t), b(t)]^T$ and derivative vectors in the point $\tilde{\mathbf{m}}(t_0)$.

The curve $\tilde{\mathbf{m}}(t)$ is depicted in Figure 5. Parameter values t in the vicinity of zero belong to large values of $a(t)$ and $b(t)$ while $\tilde{\mathbf{m}}(1) = [1/4, 0]^T$ yields the smallest possible values of a and b such that $m(a, b) = 1$. The point $\mathbf{m}(\hat{t})$ is an inflection point.

For later reference we state that the derivative vector of $\tilde{\mathbf{m}}(t)$ has the direction and orientation of

$$d\tilde{\mathbf{m}}(t) := \begin{bmatrix} E(t) - K(t) \\ -tE(t)/\sqrt{1-t^2} \end{bmatrix}. \tag{16}$$

Regardless of the value of t, both entries of this vector are negative (Figure 5).

3. Dual linear pencils of conics

In this section we prove an auxiliary lemma on the volume of hyperbolas in a dual linear pencil of conics. It will be used in the proof of Theorem 1.

Let H_0 and H_1 be two hyperbolas with equations

$$H_i\colon \mathbf{x}^T \cdot \mathbf{H}_i \cdot \mathbf{x} = 0, \quad i = 0, 1, \tag{17}$$

where $\mathbf{x} = [1, x, y]^T$ and $\mathbf{H}_i$ is a regular symmetric matrix of dimension three. The *dual conic* $\overline{H}_i$ is the set of tangents of H_i. The equation of a tangent of H_i reads $u_0 + u_1 x + u_2 y = 0$ such that the vector $\mathbf{u} = [u_0, u_1, u_2]^T$ satisfies

$$\overline{H}_i\colon \mathbf{u}^T \cdot \overline{\mathbf{H}}_i \cdot \mathbf{u} = 0, \tag{18}$$

where $\overline{\mathbf{H}}_i = \varrho \mathbf{H}_i^{-1}$ and $\varrho \in \mathbb{R} \setminus \{0\}$. The *linear pencil of dual conics* spanned by $\overline{H}_0$ and $\overline{H}_1$ is the set of conics

$$\overline{H}(\lambda)\colon \mathbf{x}^T \cdot \overline{\mathbf{H}}(\lambda) \cdot \mathbf{x} = 0 \tag{19}$$

where $\overline{\mathbf{H}}(\lambda) = (1-\lambda)\overline{\mathbf{H}}_0 + \lambda \overline{\mathbf{H}}_1$ and $\lambda \in \mathbb{R} \cup \{\infty\}$. The matrix $\overline{H}(\infty)$ is defined as

$$\overline{H}(\infty) := \overline{\mathbf{H}}_1 - \overline{\mathbf{H}}_0 = \lim_{\lambda \to \infty} \lambda^{-1} \overline{\mathbf{H}}(\lambda). \tag{20}$$

We call the set

$$H(\lambda) = \overline{\overline{H}}(\lambda), \quad \lambda \in \mathbb{R} \cup \{\infty\} \tag{21}$$

a *dual linear pencil of conics.* As opposed to a linear pencil of dual conics (which consists of dual conics), its elements are ordinary conics that share four (possible complex or coinciding) tangents.

The dual linear pencil of conics $H(\lambda)$ contains exactly one parabola P. Its parameter value λ_p is the solution of a linear equation obtained by setting the entry $\overline{h}_{00}(\lambda)$ in the first row and first column of $\overline{\mathbf{H}}(\lambda)$ to zero. By suitably normalizing the matrices $\overline{\mathbf{H}}_0$ and $\overline{\mathbf{H}}_1$ (they are only unique up to a constant factor) it is no loss of generality to assume $\lambda_p = \infty$, i.e., $\overline{h}_{00}$ is actually independent of λ.

Lemma 2. *Let H_0 and H_1 be two regular slim hyperbolas of equal volume $m_0 = m(H_0) = m(H_1)$. Let $H(\lambda)$ be a parameterization of the dual linear pencil of conics spanned by $\overline{H}_0$ and $\overline{H}_1$ such that $H_0 = H(0)$, $H_1 = H(1)$ and assume further that $P = H(\infty)$ is the unique parabola in the dual linear pencil $H(\lambda)$. Then there exists a parameter value $\lambda \in (0, 1)$ such that $H_{\min} := H(\lambda)$ is a hyperbola of volume $m(H_{\min}) < m_0$.*

Proof. The major and minor semi-axis lengths of H_i be denoted by a_i and b_i. Because H_0 and H_1 are of equal volume it is no loss of generality to assume

$$a_0 \geq a_1 \quad \text{and} \quad b_0 \geq b_1. \tag{22}$$

Furthermore, we can scale the hyperbolas H_0 and H_1 appropriately, so that $m_0 = 1$. Therefore, there exist values $0 < t_0 \leq t_1 < 1$ such that

$$\tilde{\mathbf{m}}(t_0) = [a_0, b_0]^T \quad \text{and} \quad \tilde{\mathbf{m}}(t_1) = [a_1, b_1]^T. \tag{23}$$

In a suitable coordinate frame, the algebraic equations of $\overline{H}_1$ and $\overline{H}_2$ read

$$\overline{H}_0 \colon \mathbf{u} \cdot \overline{\mathbf{H}}_0 \cdot \mathbf{u}^T = 0 \quad \text{and} \quad \overline{H}_1 \colon \mathbf{u} \cdot \overline{\mathbf{H}}_1 \cdot \mathbf{u}^T = 0, \tag{24}$$

and where

$$\overline{\mathbf{H}}_0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -a_0^2 & 0 \\ 0 & 0 & b_0^2 \end{bmatrix}, \tag{25}$$

$$\overline{\mathbf{H}}_1 = \begin{bmatrix} 1 & t_x & t_y \\ t_x & t_x^2 + b_1^2 - (a_1^2 + b_1^2)\cos^2\varphi & t_x t_y - (a_1^2 + b_1^2)\sin\varphi\cos\varphi \\ t_y & t_x t_y - (a_1^2 + b_1^2)\sin\varphi\cos\varphi & t_y^2 - a_1^2 + (a_1^2 + b_1^2)\cos(\varphi)^2 \end{bmatrix}. \tag{26}$$

The point $[t_x, t_y]^T$ is the center of H_1. The angle between the minor axis of H_1 and the x-axis of the coordinate frame is φ.

Any dual conic $\overline{H}(\lambda)$ in the linear pencil of dual conics spanned by $\overline{H}_0$ and $\overline{H}_1$ can be described by an equation of the form

$$\overline{H}(\lambda) \colon \mathbf{u} \cdot \overline{\mathbf{H}}(\lambda) \cdot \mathbf{u}^T = 0, \text{ where } \overline{\mathbf{H}}(\lambda) = (1-\lambda)\overline{\mathbf{H}}_0 + \lambda\overline{\mathbf{H}}_1 \text{ and } \lambda \in \mathbb{R} \cup \{\infty\}. \tag{27}$$

The matrices $\overline{\mathbf{H}}_i$ are already normalized such that the parabola in this dual linear pencil belongs to $\lambda_p = \infty$, as required by the lemma's assumptions.

Because H_0 is a regular hyperbola, there exists a certain neighbourhood N of 0 in $(0, 1)$ such that all conics $H(\lambda)$ with $\lambda \in N$ are hyperbolas. We denote by $a(\lambda)$ and $b(\lambda)$ the major and minor semi-axis length of $H(\lambda)$ and consider the curve

$$\mathbf{s}(\lambda) = [a(\lambda), b(\lambda)]^T, \quad \lambda \in N \tag{28}$$

in the $[a, b]$-plane. A plot of $\mathbf{s}(\lambda)$ is depicted in Figure 5. Note that the shape of the curve $\mathbf{s}(\lambda)$ depends on the hyperbola's semi-axis lengths (i.e., indirectly on t_0 and t_1), on the center $[t_x, t_y]^T$ of H_1 and on the angle φ.

We will show that the derivative vector

$$d\mathbf{s}(0) := \left[\frac{da}{d\lambda}(0), \frac{db}{d\lambda}(0)\right]^T \tag{29}$$

points "to the left" of the level set curve $\tilde{\mathbf{m}}(t)$ of the volume function graph (25) (compare Figure 5). This already implies the existence of $\lambda_{\min} \in N$ such that $H_{\min} = H(\lambda_{\min})$ is a hyperbola of volume less than m_0. "Pointing to the left" means that the determinant $d := \det[d\mathbf{s}(0), d\tilde{\mathbf{m}}(t_0)]$ is positive.

From equations (25), (26) and (27) we can compute expressions for the semi-axis lengths $a(\lambda)$ and $b(\lambda)$ of $H(\lambda)$ in the following way: We invert $\overline{\mathbf{H}}(\lambda)$ to obtain $\mathbf{H}(\lambda)$. Then we translate the conic $H(\lambda)$ so that $[0, 0]^T$ becomes its new center. The equation of the translated conic $\tilde{H}(\lambda)$ is $\mathbf{x}^T \cdot \tilde{\mathbf{H}}(\lambda) \cdot \mathbf{x} = 0$ where $\tilde{\mathbf{H}} = (\tilde{h}_{ij})$, $\tilde{h}_{00} = -1$ and $\tilde{h}_{0i} = \tilde{h}_{i0} = 0$ for $i > 0$. The eigenvalues of the matrix $\tilde{\mathbf{H}}$ are -1, ν_0 and ν_1 where

$$\nu_0 = 1/2\Big(\tilde{h}_{22} + \tilde{h}_{11} + \sqrt{(\tilde{h}_{11} - \tilde{h}_{22})^2 + 4\tilde{h}_{12}^2}\Big), \quad \text{and} \tag{30}$$

$$\nu_1 = 1/2\Big(\tilde{h}_{22} + \tilde{h}_{11} - \sqrt{(\tilde{h}_{11} - \tilde{h}_{22})^2 + 4\tilde{h}_{12}^2}\Big). \tag{31}$$

By assumption $\tilde{H}(\lambda)$ is a hyperbola; therefore ν_0 is positive and ν_1 is negative. The major semi-axis length of $H(\lambda)$ is $a(\lambda) = (\nu_0)^{-1/2}$, its minor semi-axis length is $b(\lambda) = (-\nu_1)^{-1/2}$. The explicit formulas for $a(\lambda)$ and $b(\lambda)$ in terms of a_0, b_0, a_1, b_1, t_x, t_y and φ are lengthy. However, the derivatives at $\lambda = 0$ are of simple shape:

$$\begin{aligned} \frac{da}{d\lambda}(0) &= \frac{(a_1^2 + b_1^2)\cos^2\varphi - a_0^2 - b_1^2 - t_x^2}{2a_0}, \\ \frac{db}{d\lambda}(0) &= \frac{(a_1^2 + b_1^2)\cos^2\varphi - a_1^2 - b_0^2 + t_y^2}{2b_0}. \end{aligned} \tag{32}$$

The derivative $da/d\lambda(0)$ equals zero if $\cos^2\varphi = 1$, $t_x = 0$ and $a_1 = a_0$ ($\implies b_1 = b_0$). The derivative $db/d\lambda(0)$ equals zero if additionally $t_y = 0$. Hence, the vector (29) vanishes if and only if H_0 and H_1 are equal. This case has been excluded.

From (22) we conclude $da/d\lambda(0) \le 0$. If $db/d\lambda(0)$ is not negative, Lemma 2 is proved (because both entries of (14) are negative). Otherwise, the extreme case $t_x = t_y = 0$ can be assumed. We substitute (23) into (29) and compute the determinant $d = \det[d\mathbf{s}(0), d\tilde{\mathbf{m}}(t_0)]$. It can be written in the form

$$d = \frac{P - \cos^2\varphi(2E_0 - K_0)(K_0(t_0^2 - 1) + E_0)}{32t_0\sqrt{1-t_0^2}(K_0(t_0^2-1)+E_0)^2(K_1(t_1^2-1)+E_1)^2} \tag{33}$$

where $E_i = E(t_i)$, $K_i = K(t_i)$ and P is a polynomial expression in t_0, t_1, E_0, E_1, K_0 and K_1. In order to show that (33) is positive we distinguish two cases:

Case 1 ($t_0 = t_1$): Equation 33 simplifies to

$$d = (2E_0 - K_0)(K_0(t_0^2 - 1) + E_0)\sin^2\varphi. \tag{34}$$

Because H_0 is slim, Lemma 3 implies that the first factor in this expression is positive. The last factor is positive because H_0 and H_1 are different. The middle factor is positive because of Lemma 4 on page 380. Hence we conclude $d > 0$ and the lemma is proved.

Case 2 ($t_0 < t_1$): By Lemma 3 and Lemma 4 the coefficient of $\cos^2\varphi$ in (33) is negative. Hence, we may restrict ourselves to the extreme case $\cos^2\varphi = 1$. We compute

$$d = (K_1(t_1^2 - 1) + E_1)^2 - (K_0(t_0^2 - 1) + E_0)(K_0(t_1^2 - 1) + E_0). \tag{35}$$

The positivity of d follows from Lemma 3 together with the fact that K is strictly monotone increasing and E is strictly monotone decreasing on $(0, 1)$. □

4. Uniqueness of the minimal enclosing hyperboloid

In this section we prove the main theorem of this article. It is a counterpart of the theorem on the uniqueness of the Löwner ellipse to a bounded, closed point set $\mathcal{P} \subset \mathbb{E}^2$. To begin with, we clarify the notion of a "hyperbola being contained in a hyperbola".

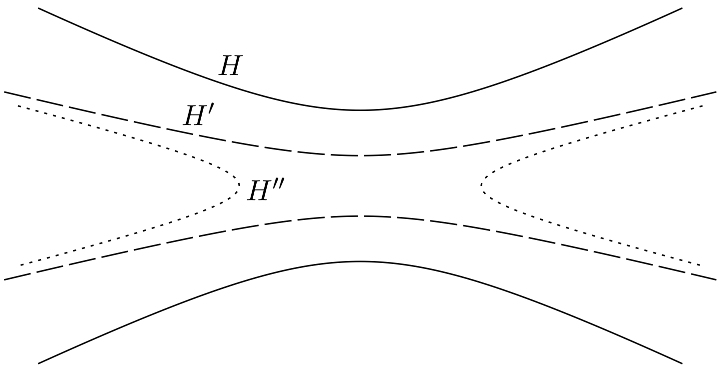


Figure 6. The hyperbola H' is contained in the hyperbola H while H'' is not contained in H

Definition 4. *A hyperbola H' is said to be* contained *in a hyperbola H if all points of H' are contained in the closure of* $\operatorname{ext} H$ *and all points of H are contained in the closure of* $\operatorname{int} H'$.

The hyperbola H' of Figure 6 is contained in the hyperbola H while H'' is not. Note that Definition 4 is tailored for the use in Theorem 1 and differs from usual concepts of "a conic being contained in a conic". In Theorem 1 we use the notion of a *closed* set $\mathcal{S}$ of lines. This means that $\mathcal{S}$ can be mapped continuously and bijectively onto a closed set of points, for example via the polarity at a conic section.

Theorem 1. *Let $\mathcal{S}$ be a closed set of lines such that*

- *not all elements of $\mathcal{S}$ have a common point or are parallel and*
- *all elements of $\mathcal{S}$ are contained in a regular, slim hyperbola H.*

Then there exists a unique hyperbola $H_{\min}$ of minimal volume that is contained in H and contains $\mathcal{S}$ (see Figure 7).

Proof. a) Existence: Consider an ellipse E whose center $\mathbf{e}$ is element of $\operatorname{int} H$. The polarity ε at E is the mapping that associates to every point $\mathbf{p} \in \mathbb{P}^2$ its polar $\varepsilon(\mathbf{p})$ and to every line $S \in \overline{\mathbb{P}}^2$ its pole $\varepsilon(S)$ with respect to E. Denote the set of all hyperbolas contained in H and containing $\mathcal{S}$ by $\mathcal{H}$ and let $\overline{\mathcal{H}} := \{\overline{H} \mid H \in \mathcal{H}\}$ be the corresponding set of dual hyperbolas. The polar image of $\overline{\mathcal{H}}$ is a set $\mathcal{E}$ of ellipses contained in E and containing $\varepsilon(\mathcal{S})$. As shown in [4], the set $\mathcal{E}$ (and hence also $\mathcal{H}$) can be mapped continuously to a bounded subset of $\mathbb{R}^5$ (the coefficients of all ellipse equations in a certain normal form are bounded). The set $\mathcal{E}$ is closed because $\mathcal{S}$ is closed. Hence, the existence of a minimal volume hyperbola $H_{\min}$ follows from Weierstrass' theorem on the existence of extremal values of continuous functions.

b) Uniqueness: In order to show uniqueness, we assume there exist two minimal hyperbolas H_0, $H_1 \in \mathcal{H}$. We let $m_0 = m(H_0) = m(H_1)$. Because H_0 and H_1 are contained in H, they are slim. Hence we can apply Lemma 2 and we find that the dual linear pencil of conics $H(\lambda)$ spanned by $\overline{H}_0$ and $\overline{H}_1$ contains a hyperbola $H_{\min}$ of volume $m(H_{\min}) < m_0$. If $H(\lambda)$ is parameterized so that $H(0) = H_0$, $H(1) = H_1$ and $H(\infty) = P$ is the unique parabola in $H(\lambda)$, $H_{\min}$ belongs to a parameter value $\lambda_{\min} \in (0, 1)$. We will show that $H_{\min}$ is contained in H and contains $\mathcal{S}$.

Because $\operatorname{int} H \setminus \operatorname{int} P$ has inner points, it is no loss of generality to assume that the ellipse center $\mathbf{e}$ is contained in $\operatorname{int} H \setminus \operatorname{int} P$. We study the polar image of the conics $\overline{H}(\lambda)$:

- The set $\{E(\lambda) = \varepsilon(\overline{H}(\lambda)) \mid \lambda \in \mathbb{R} \cup \{\infty\}\}$ is a linear pencil of conics.
- The conic $E_i := \varepsilon(\overline{H}_i)$ is an ellipse contained in E and containing $\varepsilon(\mathcal{S})$.
- The conic $\varepsilon(P)$ is a hyperbola.

By Lemma 5, $E(\lambda_{\min}) = \varepsilon(\overline{H}_{\min})$ is an ellipse containing $\varepsilon(\mathcal{S})$ and contained in $\operatorname{int} E$. Therefore $\overline{H}_{\min}$ contains $\mathcal{S}$ and is contained in H. This contradicts the assumed minimality of H_0 and H_1. □

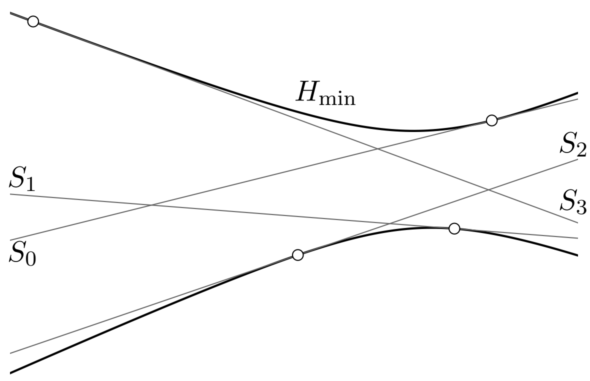


Figure 7. The minimal volume hyperbola $H_{\min}$ to four straight lines S_0, S_1, S_2 and S_3

5. Conclusion and future research

The main contribution of this text is the proof of uniqueness of a minimal hyperbola among all hyperbolas that enclose a given line set and are contained in a slim hyperbola H. We are currently unable to prove or refute the claim of Theorem 1 when H is not slim. It would be desirable to clarify this question. Further ideas for future research are collected in the remaining part of this section.

5.1. Different volume formulas

We have defined the volume of a hyperbola via the density for straight lines (1) and obtained a volume formula in terms of elliptic integrals (9). This definition seems reasonable but it is not the only way one can think of. An axiomatic characterization of a volume function v for hyperbolas might look as follows:

- $v(H) = v(a, b)$, i.e., it depends only on the hyperbola's semi-axis lengths.
- $v(a, b)$ is strictly monotone increasing in a and strictly monotone decreasing in b.
- $v(a, b)$ has the limit behavior of $m(a, b)$ as presented at the beginning of Section 2.2 but with exception of $v(a, 0) = 4a$. We only require that $a > 0$ implies $v(a, 0) > 0$.

Which of the functions characterized in this way allow unique minimal volume hyperboloids? Is it possible to find a volume for hyperboloids such that $H_{\min}$ is affinely related to the line set $\mathcal{S}$? Further suggestions and questions raised below can also be discussed in connection with axiomatically defined volume functions.

5.2. Maximal volume hyperboloids contained in convex line sets

Consider a set of lines $\mathcal{S}$ in $\mathbb{R}^2$ and a point $\mathbf{p}$ such that $\mathbf{p} \notin S$ for all $S \in \mathcal{S}$. We call the line set $\mathcal{S}$ $\mathbf{p}$-*convex* if the polar image $\mathcal{P}$ of $\mathcal{S}$ with respect to an ellipse E centered at $\mathbf{p}$ is convex. The $\mathbf{p}$-*convex hull* of $\mathcal{S}$ is the polar image of the convex hull of $\mathcal{P}$. It is easy to see that these concepts are well-defined, i.e., they do not depend on the particular choice of E. A hyperbola H is said to be contained in the **p**-convex line set $\mathcal{S}$ if every point of H is contained in a line $S \in \mathcal{S}$.

Given the **p**-convex line set $\mathcal{S}$, is there a unique hyperbola $H_{\max}$ of maximal volume contained in $\mathcal{S}$? What can be said about the relation between the minimal and maximal hyperboloids? Are their centers identical as are the centers of the Löwner and John ellipses (see [2])? Are there results on the approximation quality of $H_{\min}$ and $H_{\max}$ similar to those mentioned in the introduction? If yes, by what factor do we have to scale $H_{\min}$ in order to ensure it is contained in the **p**-convex hull of $\mathcal{S}$?

5.3. Minimal volume enclosing quadrics of sets of subspaces

A generalization of Theorem 1 to sets of lines or planes in $\mathbb{R}^3$ is obvious. Straight lines or planes in $\mathbb{R}^3$ can be enclosed by hyperboloids of one or two sheets, respectively. The volume of these hyperboloids can be defined via the density of lines or planes in $\mathbb{R}^3$ (see [19, Section 12.2]). Hence we can ask for existence and uniqueness of minimal enclosing hyperboloids of lines and planes in $\mathbb{R}^3$ or, more general, for existence and uniqueness of minimal enclosing hyperboloids of k-spaces in $\mathbb{R}^d$.

5.4. Computational issues

The actual computation of $H_{\min}$ is an optimization problem. For producing the images in this article we found the routines in standard software packages to be sufficient. However, theoretical results on the reliability or efficiency are not available.

In the past years, some attention has been paid to the computation of Löwner and John ellipsoids. It is a typical instance of convex programming (see [3, 13, 14, 15]). A randomized algorithm for the computation of $E_{\min}$ has been proposed by [22]; its primitive steps are the topic of [6]. While it seems possible to adapt at least some of the methods of [22] and [6], we are currently unable to formulate the computation of minimal hyperbolas as instance of a convex optimization problem.

A. Auxiliary results

In the appendix we prove a few technical results. They are needed at certain points in the preceding text but are of minor importance otherwise.

Lemma 3. *The function $f_1(t) = 2E(t) - K(t)$, $t \in [0, 1)$ has exactly one zero $\hat{t}$. It is positive for $t < \hat{t}$ and negative for $t > \hat{t}$. The numeric value of $\hat{t}$ is approximately $\hat{t} \approx 0.9089085575$.*

Proof. The first and second complete elliptic integrals have the following well-known properties:

- E is strictly monotone decreasing,
- K is strictly monotone increasing on $(0, 1)$,
- $E(0) = K(0) = \pi/2$, $E(1) = 1$, $\lim_{x\to 1} K(x) = \infty$.

These properties imply the existence of a unique zero $\hat{t}$ of $f_1(t)$. □

Lemma 4. *The function* $f_2(t) = (t^2 - 1)K(t) + E(t)$ *is strictly monotone increasing and positive on* $(0, 1)$.

Proof. The first derivative of $f_2(t)$ is $f_2'(t) = tK(t)$. It is positive for $t > 0$; hence f_2 is strictly monotone increasing on $(0, 1)$. Because of $f_2(0) = 0$ the function $f_2(t)$ is positive on $(0, 1)$. □

Lemma 5. *Let* $C(\lambda)$ *be a linear pencil of conics such that* $C(0)$ *and* $C(1)$ *are ellipses with* $\operatorname{int} C_0 \cap \operatorname{int} C_1 \neq \varnothing$ *and assume there exists a value* $\lambda_h \notin [0, 1]$ *such that* $C(\lambda_h)$ *is a hyperbola. Then for every* $\lambda_0 \in (0, 1)$ *the conic* $C(\lambda_0)$ *is an ellipse and*

$$\operatorname{int} C(0) \cap \operatorname{int} C(1) \subset \operatorname{int} C(\lambda_0). \tag{36}$$

Proof. A proof of this lemma (in slightly different formulation) is already given at other places, for example in [6] or in the proof of Theorem 1 in [4]. □

References

[1] Ball, K. M.: *Ellipsoids of maximal volume in convex bodies.* Geom. Dedicata **41** (1992), 241–250. Zbl 0747.52007

[2] Behrend, F.: *Über die kleinste umbeschriebene und die größte einbeschriebene Ellipse eines konvexen Bereiches.* Math. Ann. **115** (1938), 379–411. Zbl 0018.17502 and JFM 64.0731.05

[3] Boyd, S.; Vandenberghe, L.: *Convex Optimization.* Cambridge University Press, 2004. Zbl 1058.90049

[4] Danzer, L.; Laugwitz, D.; Lenz, H.: *Über das Löwnersche Ellipsoid und sein Analogon unter den einem Eikörper einbeschriebenen Ellipsoiden.* Arch. Math. **8** (1957), 214–219. Zbl 0078.35803

[5] Förstner, W.: *Handbook of Geometric Computing.* Chapter Uncertainty and Projective Geometry, 493–535. Springer, 2005.

[6] Gärtner, B.; Schönherr, S.: *Exact primitives for smallest enclosing ellipses.* In: Proc. 13th Annual ACM Symposium on Computational Geometry, pages 430–432, New York 1997. ACM press. Zbl 0914.00075

[7] Gärtner, B.; Schönherr, S.: *Smallest enclosing ellipses – fast and exact.* Technical report, Free University Berlin 1997.

[8] Giannopoulos, A.; Perissinaki, I.; Tsolomitits, A.: *John's theorem for an arbitrary pair of convex bodies.* Geom. Dedicata **84**(1–3) (2001), 63–79. Zbl 0989.52004

[9] Gordon, Y.; Litvak, A. E.; Meyer, M.; Pajor, A.: *John's decomposition in the general case and applications.* J. Differential Geom. **68**(1) (2004), 99–119. Zbl pre05033781

[10] Gruber, P. M.; Schuster, F. E.: *An arithmetic proof of John's ellipsoid theorem.* Arch. Math. **85** (2005), 82–88. Zbl 1086.52002

[11] Hu, S.-M.; Wallner, J.: *Error propagation through geometric transformations.* J. Geom. Graph. **8**(2) (2004), 171–183. Zbl 1076.51009

[12] John, F.: *Extremum problems with inequalities as subsidiary conditions.* Studies Essays, pres. to R. Courant, 187–204, Interscience Publ. Inc., New York 1948. Zbl 0034.10503

[13] Juhnke, F.: *Volumenminimale Ellipsoidüberdeckungen.* Beitr. Algebra Geom. **30** (1990), 143–153. Zbl 0746.52026

[14] Juhnke, F.: *Embedded maximal ellipsoids and semi-infinite optimization.* Beitr. Algebra Geom. **35**(2) (1994), 163–171. Zbl 0819.52006

[15] Juhnke, F.: *Polarity of embedded and circumscribed ellipsoids.* Beitr. Algebra Geom. **36**(1) (1995), 17–24. Zbl 0819.52007

[16] Laugwitz, D.: *Differentialgeometrie.* Teubner, 3. edition, 1977. Zbl 0411.53001

[17] Pełczyński, A.: *Remarks on John's theorem on the ellipsoid of maximal volume inscribed into a convex symmetric body in* $\mathbb{R}^n$. Note Mat. **10**(Suppl. 2) (1990), 395–410. Zbl 0785.46011

[18] Pottmann, H.; Odehnal, B.; Peternell, M.; Wallner, J.; Haddou, R. A.: *On optimal tolerancing in Computer-Aided Design.* In: R. Martin and W. Wang, editors: Geometric Modeling and Processing 2000, 347–363. IEEE Computer Society, Los Alamitos, Calif., 2000.

[19] Santaló, L.: *Integral Geometry and Geometric Probability.* Cambridge Mathematical Library. Cambridge University Press, 2nd edition, 2004. Zbl pre01849873

[20] Schröcker, H.-P.; Wallner, J.: *Curvatures and tolerances in the Euclidean motion group.* Result. Math. **47** (2005), 132–146. Zbl 1075.53007

[21] Wallner, J.; Krasauskas, R.; Pottmann, H.: *Error propagation in geometric constructions.* Computer Aided Design **32** (2000), 631–641.

[22] Welzl, E.: *New Results and New Trends in Computer Science.* Lecture Notes in Computer Science **555**, Chapter: Smallest enclosing disks (balls and ellipsoids), 359–370. Springer-Verlag (1991). Zbl 0825.00080

Received January 24, 2006