TWO-DIMENSIONAL GEOMETRIES WITH ELEMENTARY AREAS

HERBERT BUSEMANN

Whereas area in spaces with a smooth Riemann metric has been widely studied, very little is known regarding area in spaces with general metrics. It is natural to ask, first, in which general spaces the most familiar types of formulas for area hold.

The present note answers this question in two cases for two-dimensional spaces in which the geodesic connection is locally unique.[1] It shows under very weak assumptions regarding the nature of area:

I. *If (and only if) locally an area exists for which triangles with equal sides have equal area, then the space is a locally isometric map of either the euclidean plane, or a hyperbolic plane, or a sphere.*

Consequently, Hero's and the corresponding non-euclidean formulas[2] are (up to constant factors) the only possible formulas for area in terms of the sides, and each formula is characteristic for its respective geometry.

II. *If (and only if) locally an area α exists such that the area of the triangle pab depends only on p, the local branch of the geodesic $\mathfrak{g}$ that contains the segment $\mathfrak{s}(a, b)$, and the distance ab* (the euclidean geometry is, of course, the special case $\alpha = p\mathfrak{g} \cdot ab/2$), *then the space is a locally isometric map of a Minkowski plane.*

The exact hypotheses regarding the space R are these: (1) *R is a metric space.* (2) *R is finitely compact.* (3) *R is two-dimensional.* (4) *R is convex.* If xy denotes the distance of x and y, let (xyz) denote the statement that the three points x, y, z are different and that $xy + yz = xz$. (5) *Every point p has a neighborhood $U(p)$ such that for any two different points x, y in $U(p)$ a point z with (xyz) exists.* (6) *If (xyz_1), (xyz_2) and $yz_1 = yz_2$, then $z_1 = z_2$.*[3]

The following facts are known to hold in R: If $S(p, \rho)$ denotes the set of points x with $px < \rho$, then a $\rho(p) > 0$ exists such that: $S(p, 3\rho(p))$ is homeomorphic to a circular disk [1, p. 29]. The segment $\mathfrak{s}(a, b)$

---

[1] The assumptions are formulated further on. They are equivalent to the conditions A, B, C, D in [1, pp. 11, 12]. Numbers in brackets refer to the references cited at the end of the paper.

[2] In spherical geometry L'Huilier's formula for the defect can be used; for a proof see [4, p. 134]. The analogous hyperbolic formula may be obtained in the same manner. A slightly different form is found in [5, p. 129].

[3] The requirements (5) and (6) are equivalent to D in [1, p. 12] or [2, p. 215], see [2, Theorem (4.1)].

is unique for any two points of $S(p, 3\rho(p))$ [2, (4.5)]. If a, b are different points of $S(p, \rho(p))$, then $\mathcal{s}(a, b) \subset S(p, 2\rho(p))$ [2, (1.15)] and $\mathcal{s}(a, b)$ is subsegment of a (unique) segment $\mathfrak{h}(a, b) = \mathcal{s}(a', b')$ with $pa' = pb' = 3\rho(p)$ and $xp < 3\rho(p)$ for $(a'xb')$. Three points c_1, c_2, c_3 of $S(p, 2\rho(p))$ will be called collinear when $c_i \in \mathfrak{h}(a, b)$ for suitable $a \neq b$.

If a, b, c are three noncollinear points of $S(p, \rho(p))$, then the union of the three segments $\mathcal{s}(a, b)$, $\mathcal{s}(b, c)$, $\mathcal{s}(c, a)$ is a closed Jordan curve in $S(p, 2\rho(p))$ and bounds exactly one convex subset $T(abc)$ of $S(p, 2\rho(p))$. (Convexity of a set E means that a, $b \in E$ implies $\mathcal{s}(a, b) \subset E$.) Therefore a segment connecting two points of $T(abc)$ is again in $S(p, 2\rho(p))$.

A symmetric function $\alpha_p(a_1a_2a_3)$ defined for any triple a_1, a_2, a_3 in $S(p, \rho(p))$ is called an area if:

A: $0 \leq \alpha_p(a_1a_2a_3) < \infty$, and $\alpha_p(a_1a_2a_3) = 0$ if and only if a_1, a_2, a_3 are collinear.

B: If (a_2ba_3) then $\alpha_p(a_1a_2b) + \alpha_p(a_1ba_3) = \alpha_p(a_1a_2a_3)$.

Statement I will be proved by showing:

(1) If in $S(p, 2\delta(p))$, $0 < \delta(p) \leq \rho(p)/2$, an area $\alpha(a_1a_2a_3)$ exists such that $a_ia_k = a_i' a_k'$ implies $\alpha(a_1a_2a_3) = \alpha(a_1' a_2' a_3')$, then $S(p, \delta(p))$ is isometric to the interior of a circle of the euclidean plane, a hyperbolic plane or a sphere.

Consider two different points a, b in $S(p, \delta(p))$ and two points c_1, c_2 (not on $\mathcal{s}(a, b)$) in $S(p, \delta(p))$ with $c_ia = c_ib$ such that $\mathcal{s}(a, b)$ and $\mathcal{s}(c_1, c_2)$ intersect at a point q. Such c_i exist, compare [1, Theorem III, 2.5']. Then

(2) $$\alpha(c_1c_2a) = \alpha(c_1c_2b).$$

If $am = mb = ab/2$, then

(3) $$\alpha(c_iam) = \alpha(c_ibm), \qquad i = 1, 2.$$

These relations and properties A, B of α imply that q coincides with m. For if this were not the case it may be assumed that (amq), because the triangle inequality implies that q cannot coincide with a or b. Then $\alpha(c_imq) > 0$, but $\sum_i [\alpha(c_iam) + \alpha(c_imq)] = \sum_i \alpha(c_iaq) = \alpha(c_1c_2a)$ $= \alpha(c_1c_2b) = \sum_i \alpha(c_iqb) = \sum_i [\alpha(c_imb) - \alpha(c_imq)]$, which contradicts (3). Hence the points c_1, c_2, m are collinear.

Let c be any other point in $S(p, \delta(p))$ which is on the bisector $B(a, b)$ of a and b, that is the locus $xa = xb$. Then either $\mathcal{s}(c_1, c)$ or $\mathcal{s}(c_2, c)$ intersects $\mathcal{s}(a, b)$, compare [1, Theorem III, 2.5']. Therefore c is collinear either with c_1 and m or with c_2 and m. Hence $c \in \mathfrak{h}(c_1, c_2)$ so that $B(a, b)$ is linear. (1) follows then from [2, (15.2)].

Theorem I would become meaningless if the passage: "an area

exists" were replaced by "every area that satisfies A and B." For instance, if $\phi(\xi, \eta)$ is any nonconstant positive continuous function in the euclidean plane, then $\alpha(abc) = \int_{T(abc)}\phi(\xi, \eta)d\xi d\eta$ will satisfy A and B but not the relation $\alpha(a_1' a_2' a_3') = \alpha(a_1 a_2 a_3)$ for any two triangles with equal (euclidean) sides. This remark applies also to Theorem II.

This theorem will be proved in the following form:

(4) If every point p of R has a neighborhood $S(p, 2\delta(p))$, $0 < \delta(p) \leq \rho(p)/2$, such that in $S(p, 2\delta(p))$ an area $\alpha_p(xab)$ exists which depends only on x, $\mathfrak{h}(a, b)$,[4] and ab, then the universal covering space of R is a Minkowski plane.

The *proof* for (4) is not as easy as for (1). If (abc), then $\mathfrak{h}(a, b) = \mathfrak{h}(b, c)$. Property B of area and standard arguments[5] show that for any point x not on $\mathfrak{h}(a, b)$

$$(5) \qquad\qquad \alpha(xab) : \alpha(xbc) = ab : bc,$$

where α stands for α_p.

Consider now three noncollinear points a, b_1, b_2 in $S(p, \delta(p))$. Let m be the center of $\mathfrak{s}(b_1, b_2)$, (mqa) and put $c_1 = \mathfrak{h}(b_2, q) \cap \mathfrak{s}(b_1, a)$, $c_2 = \mathfrak{h}(b_1, q) \cap \mathfrak{s}(b_2, a)$, $n = \mathfrak{s}(a, m) \cap \mathfrak{s}(c_1, c_2)$. Then (anq).
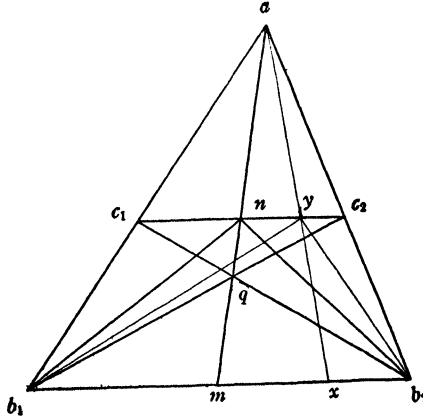
The relation (5) yields $\alpha(ab_1m) = \alpha(ab_2m)$ and $\alpha(qb_1m) = \alpha(qb_2m)$, hence

$$(6) \qquad\qquad \alpha(qb_1a) = \alpha(qb_2a)$$

---

[4] $\mathfrak{h}(a, b)$ is the exact form of the term "local branch of the geodesic that contains $\mathfrak{s}(a, b)$" employed in II. It is necessary to use that branch instead of the geodesic itself because the example of a torus with a euclidean metric shows that a geodesic may cross an arbitrary neighborhood of a given point infinitely often.

[5] See for instance [6, pp. 1, 2].

so that by (5)

$$\frac{\alpha(qb_1c_1)}{\alpha(qc_1a)} = \frac{b_1c_1}{c_1a} = \frac{\alpha(b_2b_1c_1)}{\alpha(b_2c_1a)} = \frac{\alpha(qb_1b_2) + \alpha(qb_1c_1)}{\alpha(qb_2a) + \alpha(qc_1a)} = \frac{\alpha(qb_1b_2)}{\alpha(qb_2a)}$$

and in the same way

$$\frac{\alpha(qb_2c_2)}{\alpha(qc_2a)} = \frac{b_2c_2}{c_2a} = \frac{\alpha(qb_1b_2)}{\alpha(qb_1a)}\,.$$

Therefore it follows from (6) that

(7) $b_1c_1:c_1a = b_2c_2:c_2a,$

(8) $\alpha(qb_1c_1) = \alpha(qb_2c_2),$ and $\alpha(qc_1a) = \alpha(qc_2a).$

Now (5) yields

$$\frac{\alpha(ac_1n)}{\alpha(qc_1n)} = \frac{an}{qn} = \frac{\alpha(ac_2n)}{\alpha(qc_2n)},$$

hence, by (8), $\alpha(ac_1n) = \alpha(ac_2n)$ and then by (5)

(9) $c_1n = c_2n.$

Applying (9) and (5)

$$\frac{\alpha(ac_1b_2)}{\alpha(anb_2)} = \frac{2\alpha(anc_2) + 2\alpha(b_2nc_2)}{\alpha(anc_2) + \alpha(b_2nc_2)} = 2$$

so that

(10) $$\frac{an}{am} = \frac{2\alpha(anb_2)}{2\alpha(amb_2)} = \frac{\alpha(ac_1b_2)}{\alpha(ab_1b_2)} = \frac{ac_1}{ab_1}\,.$$

Successive dyadic subdivisions of $\mathcal{S}(b_1, b_2)$ and applications of (9) and (10) yield:

(11) If (b_1xb_2) and $\mathcal{S}(a, x) \cap \mathcal{S}(c_1, c_2) = y$ then $ay:ax = ac_i:ab_i$ and $c_1y:c_2y = b_1x:b_2x.$

This leads to:

(12) If z_1, z_2 are points of $\mathfrak{h}(a, b) \cap S(p, \delta(p))$ then $\alpha(z_1z_2y)$ is constant for fixed z_i and $y \in \mathfrak{h}(c_1, c_2) \cap S(p, 2\delta(p)).$

For

$$\frac{\alpha(ab_1b_2)}{\alpha(c_1b_1b_2)} = \frac{ab_1}{c_1b_1} = \frac{ax}{yx} = \frac{\alpha(ab_ix)}{\alpha(yb_ix)} = \frac{\alpha(ab_1b_2)}{\alpha(yb_1b_2)}$$

so that $\alpha(b_1b_2y)$ is constant. The general statement follows easily from (5).

We call $\mathfrak{h}_2 = \mathfrak{h}(c_1, c_2)$ parallel to $\mathfrak{h}_1 = \mathfrak{h}(b_1, b_2)$. (8) and (9) show $\alpha(c_1 c_2 b_1) = \alpha(c_1 c_2 b_2)$ so that also $\mathfrak{h}_1$ is parallel to $\mathfrak{h}_2$. The relation (5) yields:

(13) Let $\mathfrak{h}_3$ be a parallel to $\mathfrak{h}_1$ and between $\mathfrak{h}_1$ and $\mathfrak{h}_2$, and $s_i, t_i \in \mathfrak{h}_i$, $(t_1 t_3 t_2)$, $(s_1 s_3 s_2)$, then $s_1 s_2 : s_3 s_1 = t_1 t_2 : t_3 t_1$. Hence $\mathfrak{h}_3$ is also parallel to $\mathfrak{h}_2$.

Every point x of $S(p, \delta(p))$, in particular p itself, is an interior point of a parallelogram with vertices in $S(p, \delta(p))$, where parallelogram is defined as a quadrangle in which opposite sides lie on parallels in the present sense. For, if $\mathfrak{s}(a, c)$ and $\mathfrak{s}(b, d)$ are noncollinear segments with the same center, then (13) shows that $abcd$ is a parallelogram.

Call P the closed convex set in $S(p, 2\delta(p))$ bounded by the parallelogram $ouwv$, where the vertices are in $S(p, \delta(p))$ and w is opposite to o. Coordinates ξ, η can be introduced in P as follows: The parallels to $\mathfrak{h}(o, v)$ and $\mathfrak{h}(o, u)$ through a given point $q \in P$ may intersect $s(o, u)$ and $s(o, v)$ in q_1 and q_2 respectively. Then $oq_1 = \xi$ and $oq_2 = \eta$ are the coordinates of q. A line $L = \mathfrak{h}(a, b) \cap P \neq 0$ intersects the boundary of P. Let, for instance, $\mathfrak{s}(o, u) \cap L = a = (\beta, 0)$. Choose any two different points $q_0 = (\xi_0, \eta_0)$ and $q = (\xi, \eta)$ on L which are different from a. If $\xi_0 = \beta$, then $\xi = \beta$ is the equation of L. If $\xi_0 \neq \beta$ let the parallel to $\mathfrak{h}(o, v)$ through a intersect the parallels to $\mathfrak{h}(o, u)$ through q_0 and q_1 at b_0 and b_1 respectively. Then by (11) and (13)

$$\frac{\xi - \beta}{\xi_0 - \beta} = \frac{aq}{aq_0} = \frac{ab}{ab_0} = \frac{\eta}{\eta_0}$$

so that ξ and η satisfy a linear equation.

Besides the given distance $q_1 q_2$ introduce in P a euclidean metric $e(q_1, q_2) = [(\xi_1 - \xi_2)^2 + (\eta_1 - \eta_2)^2]^{1/2}$ where $q_i = (\xi_i, \eta_i)$. Then the lines L coincide with the euclidean straight line pieces in P; moreover because of (13), the distances $e(q_1, q_2)$ and $q_1 q_2$ are proportional when q_1, q_2 vary on a fixed line L.

This and (13) imply that the universal covering space $\overline{R}$ of R is the whole euclidean plane with the euclidean straight lines as geodesics, and that on each straight line the distance $\bar{q}_1 \bar{q}_2$ in $\overline{R}$ is proportional to the euclidean distance $e(\bar{q}_1, \bar{q}_2)$. Therefore (13) holds for R, the parallels now meaning ordinary euclidean lines. It also follows that the circles $\bar{p}\bar{x} = \rho$ about a fixed point p are homothetic. To see that the metric in $\overline{R}$ is Minkowskian it must be shown that circles with different centers $\bar{p}_1, \bar{p}_2$ are also homothetic.

Consider two parallel lines $\overline{L}_1$ and $\overline{L}_2$ and a point $\bar{p}$ between them. Let $\bar{f}_1$ be a foot of $\bar{p}$ on $\overline{L}_1$ and let the line through $\bar{p}$ and $\bar{f}_1$ intersect $\overline{L}_2$ at $\bar{f}_2$. Let another line through $\bar{p}$ intersect $\overline{L}_i$ at $\bar{a}_i$. Then by (13)

$1 \geq \bar{p}\bar{f}_1/\bar{p}\bar{a}_1 = \bar{p}\bar{f}_2/\bar{p}\bar{a}_2$, hence $\bar{f}_2$ is a foot of $\bar{p}$ on $\bar{L}_2$. If $(\bar{p}\bar{q}\bar{f}_2)$, then $\bar{f}_2$ is the only foot of $\bar{p}$ on $\bar{L}_2$ (see [2, (7.9)]), and by the preceding argument $\bar{f}_1$ is a foot of $\bar{q}$ on $\bar{L}_1$, hence $\bar{f}_1$ is the only foot of $\bar{p}$ on $\bar{L}_1$. This means that the circles in R are convex [1, Theorems IV, 3.1, 3]. If the lines through $\bar{p}_1$ and $\bar{f}_1$ and through $\bar{p}_2$ and $\bar{f}_2$ are parallel, and $\bar{L}_1$ is a supporting line to the circle $\bar{p}_1\bar{x} = \bar{p}_1\bar{f}_1$ at $\bar{f}_1$, then the parallel to $\bar{L}_1$ through $\bar{f}_2$ is a supporting line of $\bar{p}_2\bar{x} = \bar{p}_2\bar{f}_2$ [1, Theorem IV, 7.1]. Therefore the two circles are homothetic and the metric of R is Minkowskian.

The "only if" part in II follows from the fact that in a Minkowskian plane the ordinary area of a euclidean plane with the same geodesics will satisfy A and B.

The plane, cylinder, torus, Moebius strip, and the one-sided torus are the only two-dimensional manifolds that can carry locally euclidean metrics (compare [3, Chapter III, §VII]). The arguments used to prove this fact show also that these *are the only manifolds that can carry locally Minkowskian metrics.* If a Minkowski metric is given, then cylinders and tori exist which have locally this metric. But a *Moebius strip and a one-sided torus with this metric exist, if and only if the Minkowski metric admits a reflection in some straight line,* see [3, pp. 81–85].

The preceding discussion suggests the question: in which geometries is the locus of the vertex x of a triangle xab with the fixed base $\mathfrak{s}(a, b)$, $a \neq b$, a fixed area, and on a given side of $\mathfrak{h}(a, b)$ a straight line? However, as long as area only satisfies A and B, this question is not reasonable, as the following example shows: Represent the hyperbolic plane by a euclidean model, in which the straight line pieces in the interior K of the euclidean unit circle are the hyperbolic geodesics. For three points a, b, c in K define $\alpha(abc)$ as the content of the triangle in the sense of the imbedding euclidean metric. Then α satisfies A and B and the above condition.

REFERENCES

1. H. Busemann, *Metric methods in Finsler spaces and in the foundations of geometry*, Annals of Mathematics Studies, no. 8, Princeton, 1942.
2. ———, *Local metric geometry*, Trans. Amer. Math. Soc. vol. 56 (1944) pp. 200–274.
3. E. Cartan, *Leçons sur la géometrie des espaces de Riemann*, Paris, 1928.
4. G. Hessenberg, *Ebene und sphärische Trigonometrie*, Berlin, 1934.
5. H. Liebmann, *Nichteuklidische Geometrie*, 1st ed., Leipzig, 1905.
6. E. Picard, *Leçons sur quelques équations fonctionnelles*, Paris, 1928.

SMITH COLLEGE