V.4 : Angle defects and related phenomena

In the previous section we showed that the angle sums of triangles in a neutral plane can behave in one of two very distinct ways. In fact, it turns out that there are essentially only two possible neutral planes, one of which is given by Euclidean geometry and the other of which does not satisfy any of the **24** properties listed in Section **2.** The purpose of this section is to study some of these properties for a non – Euclidean plane.

Definition. A neutral plane $(\mathcal{P}, \mathcal{L}, d, \mu)$ is said to be ***hyperbolic*** if Playfair's Parallel Postulate does not hold. In other words, there is ***some pair*** $(\mathbf{L}, \mathbf{X})$, where **L** is a line in $\mathcal{P}$ and **X** is a point not on **L**, for which there are ***at least two lines through*** **X** which are ***parallel to*** **L.** The study of hyperbolic planes is usually called ***hyperbolic geometry.***

The name "hyperbolic geometry" was given to the subject by F. Klein (1849 – 1925), and it refers to some relationships between the subject and other branches of geometry which cannot be easily summarized here. Detailed descriptions may be found in the references listed below:

> C. F. Adler, ***Modern Geometry: An Integrated First Course*** (2nd Ed.). McGraw – Hill, New York, 1967. ISBN: 0–070–00421–8. [see Section **8.5.3**, pp. 219 – 226]
>
> A. F. Horadam, ***Undergraduate Projective Geometry.*** Pergamon Press, New York, 1970. ISBN: 0–080–17479–5. [see pp. 271 – 272]
>
> H. Levy, ***Projective and Related Geometries.*** Macmillan, New York, 1964. ISBN: 0–000–03704–4. [see Chapter **V**, Section **7**]

A full development of hyperbolic geometry is long and ultimately highly nonelementary, and ***it requires a significant amount of differential and integral calculus.*** We shall discuss one aspect of the subject with close ties to calculus at the end of this section, but we shall only give proofs that involve "elementary" concepts and techniques.

In the previous section we showed that the angle sum of a triangle in a neutral plane is either always equal to 180° or always strictly less than 180°. We shall begin by showing that the second alternative holds in a hyperbolic plane.

Theorem 1. *In a hyperbolic plane there is a triangle* $\triangle\mathbf{ABC}$ *such that*

$$|\angle CAB| + |\angle ABC| + |\angle ACB| < 180^\circ.$$

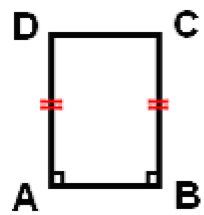
By the results of the preceding section, we immediately have numerous consequences.

Theorem 2. *In a hyperbolic plane, given an arbitrary triangle* $\triangle\mathbf{ABC}$ *we have*

$$|\angle CAB| + |\angle ABC| + |\angle ACB| < 180^\circ. \blacksquare$$

Corollary 3. *In a hyperbolic plane, suppose that we have a convex quadrilateral* $\square$**ABCD** *such that* **AB** *is perpendicular to both* **AD** *and* **BC.**

1. *If* $\square$**ABCD** *is a Saccheri quadrilateral with base* **AB** *such that* $d(\mathbf{A}, \mathbf{D}) = d(\mathbf{B}, \mathbf{C})$, *then* $|\angle ADC| = |\angle BCD| < 90^\circ$.
2. *If* $\square$**ABCD** *is a Lambert quadrilateral such that* $|\angle ABC| = |\angle BCD| = |\angle DAB| = 90^\circ$, *then* $|\angle ADC| < 90^\circ$.



***Proof of Corollary* 3.** If we split the convex quadrilateral into two triangles along the diagonal **[AC],** then by Theorem 2 we have the following:

$$|\angle \mathbf{CAB}| + |\angle \mathbf{ABC}| + |\angle \mathbf{ACB}| < 180^\circ$$

$$|\angle \mathbf{CAD}| + |\angle \mathbf{ADC}| + |\angle \mathbf{ACD}| < 180^\circ$$

Since is a convex quadrilateral we know that **C** lies in the interior or $\angle$**DAB** and **A** lies in the interior of $\angle$**BCD.** Therefore we have $|\angle \mathbf{DAB}| = |\angle \mathbf{DAC}| + |\angle \mathbf{CAB}|$ and $|\angle \mathbf{BCD}| = |\angle \mathbf{ACD}| + |\angle \mathbf{ACB}|$**;** if we combine these with the previous inequalities we obtain ***the following basic inequality, which is valid for an arbitrary convex quadrilateral in a hyperbolic plane*:**

$$|\angle \mathbf{ABC}| + |\angle \mathbf{BCD}| + |\angle \mathbf{CDA}| + |\angle \mathbf{DAB}| =$$

$$|\angle \mathbf{CAB}| + |\angle \mathbf{ABC}| + |\angle \mathbf{ACB}| + |\angle \mathbf{CAD}| + |\angle \mathbf{ADC}| + |\angle \mathbf{ACD}| < 360^\circ$$

To prove the first statement, suppose that ☐**ABCD** is a Saccheri quadrilateral, so that $|\angle \mathbf{ADC}| = |\angle \mathbf{BCD}|$ by the results of the previous section. Since $|\angle \mathbf{DAB}| = |\angle \mathbf{ABC}| = 90^\circ$, the preceding inequality reduces to

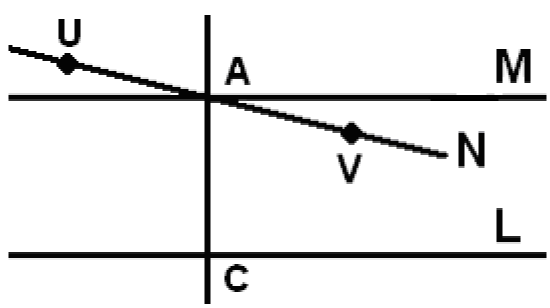
$$180 + |\angle \mathbf{BCD}| + |\angle \mathbf{CDA}| = 180^\circ + 2|\angle \mathbf{BCD}| = 180^\circ + 2|\angle \mathbf{CDA}| < 360^\circ$$

which implies $|\angle \mathbf{ADC}| = |\angle \mathbf{BCD}| < 90^\circ$.

To prove the first statement, suppose that ☐**ABCD** is a Lambert quadrilateral, so that $|\angle \mathbf{BCD}| = 90^\circ$. Since $|\angle \mathbf{DAB}| = |\angle \mathbf{ABC}| = 90^\circ$, the general inequality specializes in this case to $270^\circ + |\angle \mathbf{CDA}| < 360^\circ$, which implies $|\angle \mathbf{ADC}| < 90^\circ$.■

***Proof of Theorem* 1.** In a hyperbolic plane, we know that there is some line **L** and some point **A** not on **L** such that there are at least two parallel lines to **L** which contain **A.**

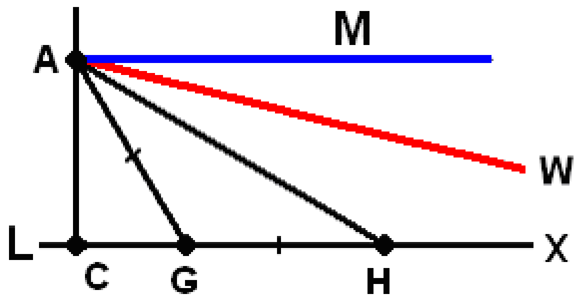
Let **C** be the foot of the unique perpendicular from **A** to **L,** and let **M** be the unique line through **A** which is perpendicular to **AC** in the plane of **L** and **A.** Then we know that **L** and **M** have no points in common (otherwise there would be two perpendiculars to **AC** through some external point). By the choice of **A** and **L** we know that there is a second line **N** through **A** which is disjoint from **L.**



The line **N** contains points **U** and **V** on each side of **AC,** and they must satisfy **U*A*V.** Since **N** is not perpendicular to **AC** and $|\angle CAU| + |\angle CAV| = 180^\circ$, it follows that one of $|\angle CAU|$, $|\angle CAV|$ must be less than 90°. Choose **W** to be either **U** or **V** so that we have $\theta = |\angle CAW| < 90^\circ$.

The line **L** also contains points on both sides of **AC,** so let **X** be a point of **L** which is on the same side of **AC** as **W.**

<u>CLAIM:</u> *If* **G** *is a point of* **(CX,** *then there is a point* **H** *on* **(CX** *such that* **C*G*H** *and* $|\angle CHA| \leq \frac{1}{2}|\angle CGA|$.



To prove the claim, let **H** be the point on **(CX** such that $d(C, H) = d(C, G) + d(G, A)$; it follows that **C*G*H** holds and also that $d(G, H) = d(A, G)$. The Isosceles Triangle Theorem then implies that $|\angle GHA| = |\angle GAH|$, and by a corollary to the Saccheri – Legendre Theorem we also have $|\angle CGA| \geq |\angle GHA| + |\angle GAH| = 2|\angle GHA| = 2|\angle CHA|$, where the final equation holds because $\angle GHA = \angle CHA$. ***<u>This proves the claim.</u>***

Proceeding inductively, we obtain a sequence of points $B_0, B_1, B_2, \ldots$ of points on **(CH** such that $|\angle CB_{k+1}A| \leq \frac{1}{2}|\angle CB_kA|$, and it follows that for each n we have

$$|\angle CB_nA| \leq 2^{-n}|\angle CB_0A|.$$

If we choose n large enough, we can make the right hand side (hence the left hand side) of this inequality less than $\frac{1}{2}(90 - \theta)$. Furthermore, we can also choose n so that

$$|\angle CB_nA| < \theta = |\angle CAW|$$

and it follows that the angle sum for $\triangle AB_nC$ will be

$$|\angle CAB_n| + |\angle AB_nC| + |\angle ACB_n| < \tfrac{1}{2}(90^\circ - \theta) + \theta + 90^\circ <$$

$$(90^\circ - \theta) + \theta + 90^\circ = 180^\circ.$$

Therefore we have constructed a triangle whose angle sum is less than 180°, as required.■

<u>Definition.</u> Given $\triangle ABC$ in a hyperbolic plane, its ***angle defect*** is given by

$$\delta(\triangle ABC) = 180 - |\angle CAB| - |\angle ABC| - |\angle ACB|.$$

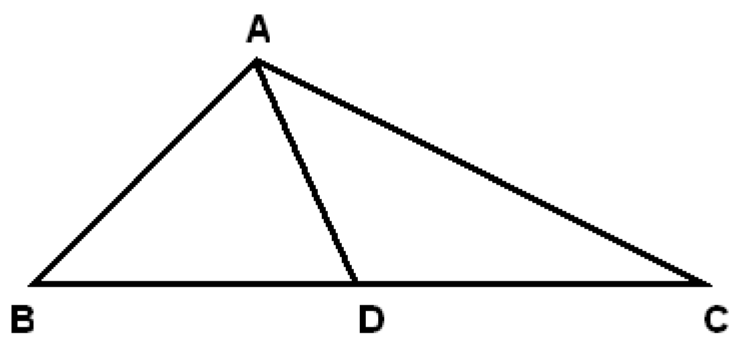
By Theorem 2, ***in a hyperbolic plane the angle defect of*** $\triangle$**ABC *is a positive real number which is always strictly between* 0 *and* 180.**

The Hyperbolic Angle – Angle – Angle Congruence Theorem

We have already seen that in spherical geometry there is a complementary notion of ***angle excess***, and the area of a spherical triangle is proportional to its angle excess. There is a similar phenomenon in hyperbolic geometry: ***For any geometrically reasonable theory of area in hyperbolic geometry, the angle of a triangle is proportional to its angular defect.*** This is worked out completely in the book by Moïse. However, for our purposes we only need the following property which suggests that the angle defect behaves like an area function.

<u>**Proposition 4.** (Additivity property of angle defects)</u> *Suppose that we are given* $\triangle$**ABC** *and that* **D** *is a point on* **(BC)**. *Then we have*

$$\delta(\triangle ABC) \;=\; \delta(\triangle ABD) \;+\; \delta(\triangle ADC).$$



Proof. If we add the defects of the triangles we obtain the following equation:

$$\delta(\triangle ABD) \;+\; \delta(\triangle ADC) \;=\; 180^\circ - |\angle DAB| - |\angle ABD| - |\angle ADB| + 180^\circ - |\angle CAD| - |\angle ADC| - |\angle ACD|$$

By the Supplement Postulate for angle measure we know that

$$|\angle ADB| \;+\; |\angle ADC| \;=\; 180^\circ$$

by the Additivity Postulate we know that

$$|\angle BAC| \;=\; |\angle BAD| \;+\; |\angle DAC|$$

and by the hypotheses we also know that $\angle ABD = \angle ABC$ and $\angle ACD = \angle ACB$. If we substitute all these into the right hand side of the equation for the defect sum $\delta(\triangle ABD) + \delta(\triangle ADC)$, we see that this right hand side reduces to

$$180^\circ - |\angle CAB| - |\angle ABC| - |\angle ACB|$$

which is the angle defect for $\triangle$**ABC.**■

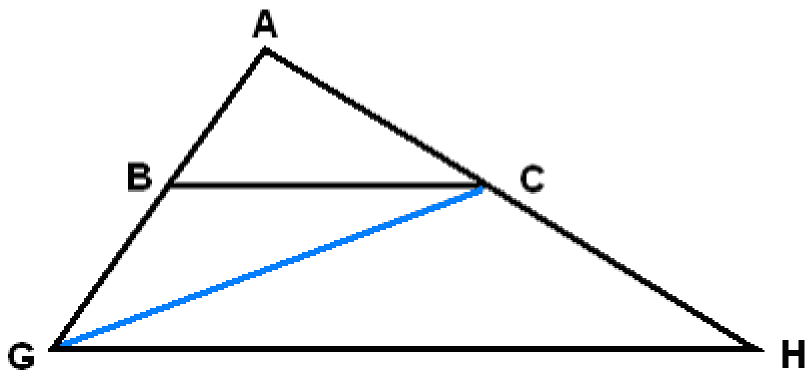
The next result yields a striking conclusion in hyperbolic geometry, which shows that the latter ***does not have a similarity theory comparable to that of Euclidean geometry.***

Theorem 5. (Hyperbolic $\mathbf{AAA}$ or *Angle – Angle – Angle Congruence Theorem*) *Suppose we have ordered triples* **(A, B, C)** *and* **(D, E, F)** *of noncollinear points such that the triangles* $\triangle$**ABC** *and* $\triangle$**DEF** *satisfy* $|\angle\mathbf{CAB}| = |\angle\mathbf{FDE}|$, $|\angle\mathbf{ABC}| = |\angle\mathbf{DEF}|$, *and* $|\angle\mathbf{ACB}| = |\angle\mathbf{DFE}|$. *Then we have* $\triangle\mathbf{ABC} \cong \triangle\mathbf{DEF}$.

Proof. If at least one of the statements $d(\mathbf{B},\mathbf{C}) = d(\mathbf{E},\mathbf{F})$, $d(\mathbf{A},\mathbf{B}) = d(\mathbf{D},\mathbf{E})$, or $d(\mathbf{A},\mathbf{C}) = d(\mathbf{D},\mathbf{F})$ is true, then by $\mathbf{ASA}$ we have $\triangle\mathbf{ABC} \cong \triangle\mathbf{DEF}$. Therefore it is only necessary to consider possible situations in which all three of these statements are false. This means that in each expression, one term is less than the other. There are eight possibilities for the directions of the inequalities, and these are summarized in the table below.

CASE	$d(\mathbf{A},\mathbf{B})$?? $d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C})$?? $d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C})$?? $d(\mathbf{B}',\mathbf{C}')$
000	$d(\mathbf{A},\mathbf{B}) < d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) < d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) < d(\mathbf{B}',\mathbf{C}')$
001	$d(\mathbf{A},\mathbf{B}) < d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) < d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) > d(\mathbf{B}',\mathbf{C}')$
010	$d(\mathbf{A},\mathbf{B}) < d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) > d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) < d(\mathbf{B}',\mathbf{C}')$
011	$d(\mathbf{A},\mathbf{B}) < d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) > d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) > d(\mathbf{B}',\mathbf{C}')$
100	$d(\mathbf{A},\mathbf{B}) > d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) < d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) < d(\mathbf{B}',\mathbf{C}')$
101	$d(\mathbf{A},\mathbf{B}) > d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) < d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) > d(\mathbf{B}',\mathbf{C}')$
110	$d(\mathbf{A},\mathbf{B}) > d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) > d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) < d(\mathbf{B}',\mathbf{C}')$
111	$d(\mathbf{A},\mathbf{B}) > d(\mathbf{A}',\mathbf{B}')$	$d(\mathbf{A},\mathbf{C}) > d(\mathbf{A}',\mathbf{C}')$	$d(\mathbf{B},\mathbf{C}) > d(\mathbf{B}',\mathbf{C}')$

Reversing the roles of the two triangles if necessary, we may assume that at least two of the sides of $\triangle$**ABC** are shorter than the corresponding sides of $\triangle$**DEF**. Also, if we consistently reorder $\{\mathbf{A},\mathbf{B},\mathbf{C}\}$ and $\{\mathbf{D},\mathbf{E},\mathbf{F}\}$ in a suitable manner, then we may also arrange things so that $d(\mathbf{A},\mathbf{B}) < d(\mathbf{D},\mathbf{E})$ and $d(\mathbf{A},\mathbf{C}) < d(\mathbf{D},\mathbf{F})$. Therefore, if we take points **G** and **H** on the respective open rays **(BA** and **(BC** such that $d(\mathbf{A},\mathbf{G}) = d(\mathbf{D},\mathbf{E})$ and $d(\mathbf{A},\mathbf{H}) = d(\mathbf{D},\mathbf{E})$, then by $\mathbf{SAS}$ we have $\triangle\mathbf{AGH} \cong \triangle\mathbf{DEF}$.



By hypothesis and construction we know that the angular defects of these triangles satisfy $\delta(\triangle\mathbf{AGH}) = \delta(\triangle\mathbf{DEF}) = \delta(\triangle\mathbf{ABC})$. We shall now derive a contradiction using the additivity property of angle defects obtained previously. This yields the following equations:

$$\delta(\triangle\mathbf{AGH}) \;=\; \delta(\triangle\mathbf{AGC}) \;+\; \delta(\triangle\mathbf{GCH})$$

$$\delta(\triangle\mathbf{AGC}) \;=\; \delta(\triangle\mathbf{ABC}) \;+\; \delta(\triangle\mathbf{BGC})$$

If we combine these with previous observations and the positivity of the angle defect we obtain

$$\delta(\triangle\mathbf{ABC}) \;<\; \delta(\triangle\mathbf{ABC}) + \delta(\triangle\mathbf{BGC}) + \delta(\triangle\mathbf{GCH}) \;=\; \delta(\triangle\mathbf{AGH}) \;=\; \delta(\triangle\mathbf{DEF})$$

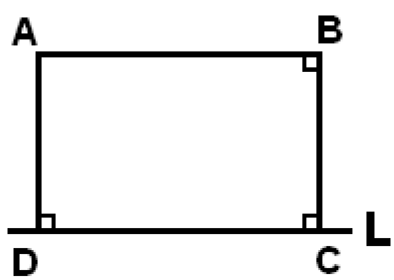
which contradicts the equation $\delta(\triangle\mathbf{DEF}) = \delta(\triangle\mathbf{ABC})$. The source of this contradiction is our assumption that the corresponding sides of the two triangles do not have equal lengths, and therefore this assumption must be false. As noted at the start of the proof, this implies $\triangle\mathbf{ABC} \cong \triangle\mathbf{DEF}$.■

One immediate consequence of Theorem **6** is that in hyperbolic geometry, ***two triangles cannot be similar in the usual sense unless they are congruent.*** In particular, this means that we cannot magnify or shrink a figure in hyperbolic geometry without distortions. This is disappointing in many respects, but if we remember that angle defects are supposed to behave like area functions then this is not surprising; we expect that two similar but noncongruent figures will have different areas, and in hyperbolic (just as in spherical!) geometry this simply cannot happen.

The Strong Hyperbolic Parallelism Property

The negation of Playfair's Postulate is that there is ***some line*** and ***some external point*** for which ***parallels are not unique.*** It is natural to ask if there are neutral geometries in which unique parallels exist for ***<u>some but not all</u>*** pairs **(L, A)** where **L** is a line and **A** is an external point. The next result implies that no such neutral geometries exist.

<u>Theorem 7.</u> *Suppose we have a neutral plane $\mathscr{P}$ such that for some line* **L** *and some external point* **A** *there is a unique parallel to* **L** *through* **A**. *Then there is a rectangle in $\mathscr{P}$.*



<u>Proof.</u> Let **D** be the foot of the perpendicular from **A** to **L**, and let **C** be a second point on **L.** Let **M** be the line in the plane of **L** and **A** such that **M** is perpendicular to **L** at **C.** Then **AD** and **M** are lines perpendicular to **L** and meet the latter at different points, so that **AD** and **M** are parallel. Next let **B** be the foot of the perpendicular to **M** from the external point **A.** The lines **AB** and **L** are distinct since **A** does not lie on **L,** and since they are both perpendicular to **M** it follows that **AB** and **L** are also parallel. Since we have **AB** || **CD** and **AD** || **BC,** it follows that **A, B, C, D** are the vertices of a convex quadrilateral.

If **N** is the perpendicular to **AD** through the point **A** in the plane of **L** and **A,** then we know that **N** is also parallel to **L.** Therefore the uniqueness of parallels to **L** through **A** implies that N must be equal to the line **AB** constructed in the previous paragraph; thus we know that **AB** is perpendicular to **AD,** and therefore it follows that the convex quadrilateral ☐**ABCD** is a rectangle.■

<u>Corollary 8.</u> *If $\mathcal{P}$ is a neutral plane such that for some line* **L** *and some external point* **A** *there is a unique parallel to* **L** *through* **A***, then Playfair's Postulate is true in* $\mathcal{P}$*.*

<u>Proof.</u> By the theorem and the results of the previous section, we know that the angle sum for every triangle in $\mathcal{P}$ is equal to $180°$. On the other hand, if Playfair's Postulate does not hold in $\mathcal{P}$, then by Theorem 2 we know that the angle sum for every triangle is less than $180°$. Therefore Playfair's Postulate must hold in $\mathcal{P}$; in other words, for every line **M** and external point **B** there is a unique parallel to **M** through **B.**■

Asymptotic parallels

We have already noted that Playfair's Postulate is equivalent to the following statement:

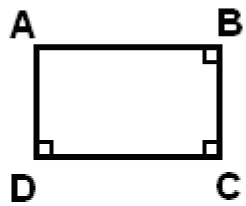
<u>EQUIDISTANCE OF PARALLELS.</u> *Let* **L** *and* **M** *be parallel lines in a neutral plane* $\mathcal{P}$, *let* **X** *be a point on one of the lines, and let* **Y(X)** *be the foot of the perpendicular from* **X** *to the other line. Then the distance* **η(X)** *from* **X** *to* **Y(X)** *is the same for all choices of* **X.**

It is natural to ask what can be said about the distance function **η(X)** if **L** and **M** are parallel lines in a hyperbolic plane $\mathcal{P}$. Thus far all of our ***<u>explicit</u>*** examples of parallel lines in hyperbolic planes have been pairs for which there is a common perpendicular (although we have not necessarily proven this in all cases). Our next result describes the behavior of **η(X)** for such pairs of parallel lines.

<u>Theorem 9.</u> *Let* **L** *and* **M** *be parallel lines in a hyperbolic plane* $\mathcal{P}$, *and suppose that* **L** *and* **M** *have a common perpendicular. Then* **L** *and* **M** *have a unique perpendicular, and if* **C** *and* **B** *are points of* **L** *and* **M** *such that* **BC** *is perpendicular to both lines, then the minimum value of* **η** *is realized at* **C** *and* **B.**

In other words, the distance between two such lines behaves somewhat like the distance between two skew lines in Euclidean 3 – space.

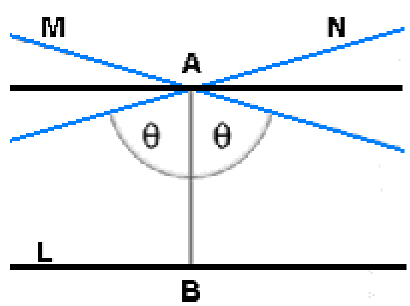
<u>Proof.</u> Let **A** be a second point of **M,** and let **D** be the foot of the perpendicular from **A** to **L;** then **D** and **C** are distinct (otherwise **A, B, C** are collinear, so that **M** = **BC,** which is impossible since **M** is parallel to **L** and **C** lies on **L**), and **A, B, C, D** form the vertices of a Lambert quadrilateral with perpendicular sides at the vertices **B, C** and **D.** By the results and exercises on neutral geometry from the previous section, we have d(**B, C**) ≤ d(**A, C**).



We claim that in fact $d(\mathbf{B}, \mathbf{C}) < d(\mathbf{A}, \mathbf{D})$. If equality held, then by SAS we would have $\triangle\mathbf{ADC} \cong \triangle\mathbf{BCD}$. This in turn would imply $d(\mathbf{A}, \mathbf{C}) = d(\mathbf{B}, \mathbf{D})$, which would further imply $\triangle\mathbf{DAB} \cong \triangle\mathbf{CBA}$ by SSS, so that $|\angle\mathbf{DAB}| = |\angle\mathbf{CBA}| = 90°$. Thus the Lambert quadrilateral is a rectangle, and since rectangles do not exist in a hyperbolic plane we have a contradiction. Therefore we must have strict inequality as claimed, and accordingly the shortest distance between the two lines is the distance between **B** and **C** on the common perpendicular.■

Not every pair of parallel lines in a hyperbolic plane has a common perpendicular. The other parallel line pairs form an important class of ***asymptotic parallels*** for which the function η does not reach a minimum value but can be made less than an arbitrarily small positive real number (hence the lines approach each other asymptotically much as the hyperbola $y = 1/x$ asymptotically approaches the x – axis defined by $y = 0$). To describe such lines, suppose that $(\mathbf{L}, \mathbf{A})$ is a pair consisting of a line **L** in a hyperbolic plane $\mathscr{P}$ and a point **A** which is in $\mathscr{P}$ but not on **L**, and let **B** be the foot of the perpendicular to from **A** to **L.** We then have the following result, which is obtainable by combining several separate theorems in Sections $24.1 - 24.4$ of Moïse:

<u>**Theorem 10.**</u> *In a hyperbolic plane $\mathscr{P}$, let* **L** *be a line, let* **A** *be a point not on* **L**, *and let* **B** *be the foot of the perpendicular from* **A** to **L.** Let Ψ *be the set of all points* **X** *in* $\mathscr{P}$ *such that* **XA** *is parallel to* **L** (*hence* **X** *cannot lie on* **AB**). *Then the set of all angle measures* $|\angle\mathbf{XAB}|$, *taken over all* **X** *in* Ψ, *assumes a* ***minimum positive value*** $\Pi(\mathbf{A}, \mathbf{B})$ *which is always* ***strictly less than*** $90°$.■



(Source: http://en.wikipedia.org/wiki/Hyperbolic_geometry)

In the drawing, the line **M** is given by **AC**, where $|\angle\mathbf{CAB}| = \Pi(\mathbf{A}, \mathbf{B})$. It follows that **M** is parallel to **L**, and the angle θ between **AB** and **M** (measured counterclockwise from **AB**) is as small as possible (*i.e.*, if the angle is smaller, then the line will meet **L**). Such a line in hyperbolic geometry is called a ***critically parallel*** (or ***asymptotically parallel***, or ***hyperparallel***) ***line***; in some books or papers such lines are simply called [hyperbolic] parallel lines. Similarly, the line **N** that forms the same angle θ between **AB** and itself but clockwise from **AB** will also be hyperparallel, but there can be no others. All other lines through **A** parallel to **L** form angles greater than θ with **AB**, and these are called ***ultraparallel*** (or ***disjointly parallel***) lines; this turns out to be the same as the class of line pairs which have common perpendiculars. Since there are an infinite number of possible angles between θ and 90 degrees, and each value will determine two lines through **A** that are ultraparallel to **L**, it follows that ***we have an infinite number of ultraparallel lines to* L *passing through* A.**

__Notation.__ The number Π(**A, B**) is called the (Lobachevsky) ***critical angle*** or ***angle of parallelism*** for **L** and **A**, and it plays a fundamentally important role in hyperbolic geometry. As suggested above, a great deal of information about this number is contained in Moïse; for example, the value only depends upon $d = d$(**A, B**), and the ***Bolyai – Lobachevsky Formula*** states that

$$\Pi(x) = 2\tan^{-1}(e^{-x})$$

where $x = d/k$ for some positive "curvature constant" we shall call k. The need to include the curvature constant k reflects the fact that similar triangles are always congruent in hyperbolic geometry, and in the 1824 letter from Gauss to Taurinus there are some comments about this constant:

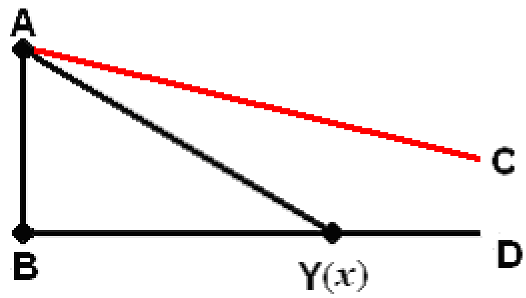
> I can solve every problem in it [non – Euclidean geometry] with the exception of the determination of a constant, which cannot be designated ***a priori.*** The greater one takes this constant, the nearer one comes to Euclidean geometry, and when it is chosen infinitely large the two coincide. ... If it [non – Euclidean geometry] were true, there must exist in space a linear magnitude, determined for itself (but unknown to us). ... If this non – Euclidean geometry were true, and it were possible to compare that constant with such magnitudes as we encounter in our measurements on the earth and in the heavens, it could then be determined ***a posteriori.*** Consequently, in jest I have sometimes expressed the wish that the Euclidean geometry were not true, since then we would have ***a priori*** an absolute standard of measure.

__Further results on parallel lines in hyperbolic geometry.__ One can now state the following more complete description of parallelism in hyperbolic geometry:

__Theorem 11.__ *Let $\mathcal{P}$ be a hyperbolic plane. Given a line* **L** *in $\mathcal{P}$, and a point* **A** *not on* **L**, *there are exactly two lines through* **A** *which are critically parallel to* **L** *and infinitely many lines through* **A** *that are ultraparallel to* **L.**■

The previously stated asymptotic property of critical parallel lines is given by the next result:

__Theorem 12.__ *Suppose we are given points* **L, A, B, X** *as above in $\mathcal{P}$ such that* $|\angle$**CAB**$|$ = Π(**A, B**). *Let* **D** *be a point of* **L** *on the same side of* **AB** *as* **C.** *Given a positive real number x, let* **Y**(x) *be the unique point on* (**AD** *which satisfies* d(**A, Y**(x)) $= x$, *and let* $\sigma(x)$ *be the distance from* **Y**(x) *to the foot of the perpendicular to* **AC.** *Then the function* $\sigma(x)$ *is strictly decreasing and the limit of* $\sigma(x)$ *as* x *approaches* $+\infty$ *is equal to* 0.■



In the preceding result the function $\sigma(x)$ is defined for nonnegative values of x, and one can extend the definition of the function to negative values of x by first taking to be the unique point $\mathbf{Y}(x)$ on the ***opposite ray*** **[BD**$^{\mathbf{OP}}$ such that $d(\mathbf{A}, \mathbf{Y}(x)) = x$, and then setting $\sigma(x)$ equal to the distance from $\mathbf{Y}(x)$ to the foot of the perpendicular to **AC.** We then have the following result:

<u>Complement 13.</u> *In the setting above, the function $\sigma(x)$ is strictly decreasing for all real values of x, and the limit of $\sigma(x)$ as x approaches $-\infty$ is equal to $+\infty$.*■

The preceding results imply that the graph of the function $\sigma(x)$ resembles the graph of the familiar function $f(x) = e^{-x}$.

To complete the discussion, we shall state the corresponding result for ultraparallel lines.

<u>Proposition 14.</u> *Suppose we are given disjoint lines* **AC** *and* **BD** *in the neutral plane in $\mathcal{P}$ such that* **AB** *is perpendicular to both* **AC** *and* **BD**, *and assume that* **C** *and* **D** *lie on the same side of* **AB.** *Let $\mathbf{Y}(x)$ and $\sigma(x)$ be defined as in the preceding results. Then the function is an even function with $\sigma(x) = \sigma(-x)$, there is a minimum value at $x = 0$, the function is strictly increasing for positive values of x and strictly decreasing for negative values, and the limit of $\sigma(x)$ as x approaches $+\infty$ or $-\infty$ is equal to $+\infty$.*■

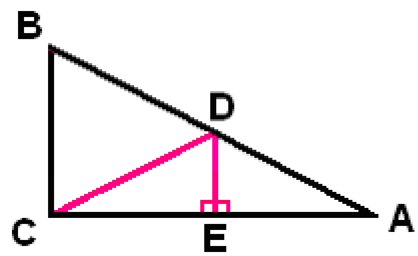
In these cases the graph of the function $\sigma(x)$ resembles the graphs of the familiar functions $f(x) = x^2$ and (more accurately) $f(x) = ½(e^x + e^{-x})$.

Proofs of the statements about critical parallels and ultraparallels can be found in many books covering hyperbolic geometry. References for this and other material will be given in the next section.

Appendix – Solved exercises in neutral and hyperbolic geometry

Here are some further examples of problems similar to the exercises for Unit **V** with complete solutions.

<u>PROBLEM 1.</u> *Suppose that we are given a right triangle $\triangle$**ABC** in the hyperbolic plane $\mathcal{P}$ with a right angle at* **C**, *and let* **E** *denote the midpoint of* **[AB].** *Prove that the line* **L** *perpendicular to* **AC** *through* **E** *contains a point* **D** *on* **(AB)** *and that $d(\mathbf{B}, \mathbf{D})$ is greater than $d(\mathbf{A}, \mathbf{D}) = d(\mathbf{C}, \mathbf{D})$.*



<u>SOLUTION.</u> First of all, by Pasch's Theorem we know that the perpendicular bisector **L** either contains a point of **[BC]** or of **(AB).** However, since **AC** is perpendicular to both **BC** and **L** we know that the first option cannot happen, and therefore the line **L** must contain some point **D** of **(AB).** By SAS we have $\triangle$**DEA** $\cong$ $\triangle$**DEC,** and therefore it follows that d(**A, D**) = d(**C, D**). Furthermore, we have $|\angle$**DAE**$|$ = $|\angle$**DCE**$|$. By the additivity property for angle measurements, we have

$$|\angle \mathbf{DAE}| + |\angle \mathbf{DCB}| = |\angle \mathbf{DCE}| + |\angle \mathbf{DCB}| = 90^\circ$$

and if we combine this with $\angle$**DAE** = $\angle$**BAC**, $\angle$**CBD** = $\angle$**CBA**, and the hyperbolic angle – sum property

$$|\angle \mathbf{BAC}| + |\angle \mathbf{CBD}| < 90^\circ$$

we see that $|\angle$**DBC**$|$ < $|\angle$**BCD**$|$. Since the larger angle is opposite the longer side, it follows that d(**B, D**) > d(**C, D**) = d(**A, D**).■

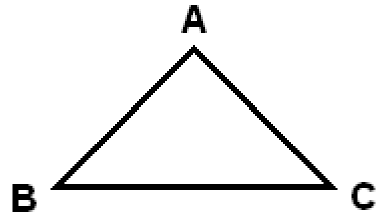
<u>PROBLEM 2.</u> *In the setting of the previous problem, determine whether* $|\angle$**BAC**$|$ *is less than, equal to or greater than* ½ $|\angle$**BDC**$|$.

<u>SOLUTION.</u> We have $|\angle$**ADC**$|$ = $|\angle$**CDE**$|$ + $|\angle$**EDA**$|$ because the midpoint **E** lies in the interior of $\angle$**ADC**, and since $\triangle$**DEA** $\cong$ $\triangle$**DEC** it also follows that $|\angle$**ADC**$|$ = $2\,|\angle$**EDA**$|$. By the supplement property for angle measures we have $|\angle$**BDC**$|$ + $|\angle$**ADC**$|$ = 180°. Therefore we also have ½ $|\angle$**BDC**$|$ + $|\angle$**EDA**$|$ = 90°. On the other hand, the hyperbolic angle – sum property implies that $|\angle$**BAC**$|$ + $|\angle$**EDA**$|$ < 90°. Therefore we have $|\angle$**BAC**$|$ + $|\angle$**EDA**$|$ < ½ $|\angle$**BDC**$|$ + $|\angle$**EDA**$|$, and if we subtract the second term from each side of this inequality we conclude that $|\angle$**EDA**$|$ < ½ $|\angle$**BDC**$|$.■

<u>PROBLEM 3.</u> *Suppose that we are given a right triangle* $\triangle$**ABC** *in the neutral plane* $\mathscr{P}$ *with a right angle at* **C**, *and let* **F** *denote the midpoint of* **[AB].** *Show that if* **F** *is equidistant from the vertices, then* $\mathscr{P}$ *is Euclidean.*

<u>SOLUTION.</u> If **F** is equidistant from the vertices, then **EF** is the perpendicular bisector of **[AC]**, and hence we must have **F** = **D.** However, by the first problem we know **D** is ***<u>not</u>*** equidistant from the vertices if the plane $\mathscr{P}$ is hyperbolic, and therefore $\mathscr{P}$ must be Euclidean.■

PROBLEM 4. *Suppose that we are given an isosceles triangle* $\triangle$**ABC** *in the neutral plane* $\mathcal{P}$ *with* $d(\mathbf{A},\mathbf{B}) = d(\mathbf{A},\mathbf{C})$ *and* $|\angle \mathbf{BAC}| > 60°$. *Prove that* $d(\mathbf{B},\mathbf{C}) > d(\mathbf{A},\mathbf{C}) = d(\mathbf{A},\mathbf{B})$.



Discussion. The drawing depicts an isosceles right triangle. As such, we know that its hypotenuse is longer than either of its legs, and this is in fact true in neutral geometry. The object of the exercise is to prove a more general result which is also true in neutral geometry.

SOLUTION. By the Saccheri – Legendre Theorem we have

$$|\angle \mathbf{BAC}| + |\angle \mathbf{ABC}| + |\angle \mathbf{ACB}| = |\angle \mathbf{BAC}| + 2\,|\angle \mathbf{ABC}| \leq 180°$$

and since $|\angle \mathbf{BAC}| > 60°$ it follows that $2\,|\angle \mathbf{ABC}| < 120°$, so that $|\angle \mathbf{ABC}| < 60°$. Since the larger angle of a triangle is opposite the longer side, we have $d(\mathbf{B},\mathbf{C}) > d(\mathbf{A},\mathbf{B})$, and the final part of the conclusion follows because the right hand side is equal to $d(\mathbf{A},\mathbf{C})$. ■

Note. In Euclidean geometry there is a companion result for isosceles triangles: *If* $|\angle \mathbf{BAC}| < 60°$, *then* $d(\mathbf{B},\mathbf{C}) < d(\mathbf{A},\mathbf{C}) = d(\mathbf{A},\mathbf{B})$. — This is true because the angle – sum property in Euclidean geometry implies that $|\angle \mathbf{ABC}| > 60°$ if $|\angle \mathbf{BAC}| < 60°$. However, the companion result does not hold in hyperbolic geometry. In fact, under these conditions for a fixed value of $|\angle \mathbf{BAC}|$ it is possible to construct triangles in a hyperbolic plane for which $|\angle \mathbf{ABC}| = |\angle \mathbf{ACB}|$ is arbitrarily small.

V.5 : Further topics in hyperbolic geometry

> I have developed this geometry to my own satisfaction so that I can solve every problem that arises in it with the exception of the determination of a certain constant which cannot be determined ***a priori.***
>
> Gauss, previously cited letter to Taurinus, 1824

We have noted that the geometry and trigonometry of a hyperbolic plane were worked out completely in the early 19th century; more precisely, the formulas for rectangular and polar coordinate systems, trigonometry, and measurements and volumes are just as complete as they are for the Euclidean plane even though they are a usually great deal more complicated. A detailed treatment of this material is beyond the scope of this course. However, a great deal of additional information about hyperbolic geometry appears in Greenberg; we shall also list several classic references for this material and one online reference which develops a considerable amount of the subject.

H. S. M. Coxeter, ***Non – Euclidean Geometry*** (6^{th} Ed.), Mathematical Association of America, Washington, DC, 1998. ISBN: 0–883–85522–4.

R. Bonola, ***Non – Euclidean Geometry*** (Transl. by H. S. Carslaw). Dover, New York, 1955. ISBN: 0–486–60027–0.

K. Borsuk and W. Szmielew, ***Foundations of Geometry*** (Rev. English Transl.). North Holland, Amsterdam (NL), 1960.

W. T. Fishback, ***Projective and Euclidean Geometry*** (2^{nd} Ed.). Wiley, New York, 1969. ISBN: 0–471–26053–3.

H. E. Wolfe, ***Introduction to Non-Euclidean Geometry.*** Holt, New York, 1945. ISBN: 0–030–07425–8.

http://www.math.uncc.edu/~droyster/math3181/notes/hyprgeom/hyprgeom.html

In this section we shall concentrate on a few topics that are closely related to previously discussed results in Euclidean geometry or are relevant to the remaining sections of this unit.

Concurrence theorems in hyperbolic geometry

In Section **III.4** of these notes, Euclidean geometry was the setting for the proofs of the four triangle concurrence theorems. Two of these results are also true in hyperbolic geometry. The proof of the incenter theorem carries over with no essential changes, and there is also a centroid theorem, but its proof is entirely different from the Euclidean argument. A sketch of the proof for hyperbolic geometry appears in Exercise **K – 19** on page **355** of Greenberg.

The preceding sentence suggests that the circumcenter theorem and orthocenter theorem do not extend to hyperbolic geometry; in fact, for each statement one can construct triangles in the hyperbolic plane for which the given lines are not concurrent. However, even though the circumcenter theorem fails to hold in hyperbolic geometry, one does have the following result in that setting:

__**Theorem 0.**__ *Suppose we are given* $\triangle$**ABC** *in a hyperbolic plane* **P**, *and let* **L, M** *and* **N** *be the perpendicular bisectors of the sides. Then exactly one of the following is true*:

1. The lines **L, M** and **N** are concurrent (as is always the case in Euclidean geometry).
2. The lines **L, M** and **N** have a common perpendicular.
3. The lines **L, M** and **N** are triply asymptotic; in other words, the lines are pairwise disjoint, but there are ruler functions $\mathbf{f} : \mathbf{L} \to \mathbb{R}$, $\mathbf{g} : \mathbf{M} \to \mathbb{R}$ and $\mathbf{h} : \mathbf{N} \to \mathbb{R}$ for these lines such that

$$\lim_{t \to \infty} d(\mathbf{f}^{-1}(t), \mathbf{g}^{-1}(t)) \quad = \quad \lim_{t \to \infty} d(\mathbf{g}^{-1}(t), \mathbf{h}^{-1}(t)) \quad = \quad 0.$$

Observe that the Triangle Inequality the limit conditions also imply that

$$\lim_{t \to \infty} d(\mathbf{f}^{-1}(t), \mathbf{h}^{-1}(t)) \quad = \quad 0.$$

<u>The Euler line and hyperbolic geometry.</u> In contrast to Euclidean geometry, even if a hyperbolic triangle has both a circumcenter and orthocenter, it does not follow that these points and the centroid are collinear. An example (using the Poincaré disk model for the hyperbolic plane) is described in the following article:

http://josm.geneseo.edu/1-1/00-03.pdf

Isometries in neutral geometry

In Section **II.4** we considered rigid motions of a Euclidean plane. We shall be interested in the corresponding class of mappings on an arbitrary neutral plane:

<u>Definition.</u> Let $\mathscr{P}$ be a neutral plane. A function (or mapping) **T** from $\mathscr{P}$ to itself is said to be an ***isometry*** if it is a $1-1$ onto map such that the following hold:

1. If **x** and **y** are points of $\mathscr{P}$, then **T** preserves the distance between them; in other words, we have $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{T}(\mathbf{x}), \mathbf{T}(\mathbf{y}))$.
2. The function **T** sends collinear points to collinear points and noncollinear points to noncollinear points.
3. If **x, y, z** are noncollinear points of $\mathscr{P}$, then **T** preserves the measurement of the angle they form; in other words, we have $|\angle \mathbf{xyz}| = |\angle \mathbf{T(x)T(y)T(z)}|$.

The statement and proof of Proposition **II.4.3** for Euclidean isometries generalizes directly to isometries of an arbitrary neutral plane.

<u>Proposition 1.</u> *The identity map is an isometry from the neutral plane $\mathscr{P}$ to itself. If* **T** *is an isometry from $\mathscr{P}$ to itself, then so is its inverse* $\mathbf{T}^{-1}$. *Finally, if* **T** *and* **U** *are isometries from $\mathscr{P}$ to itself then so is their composite* $\mathbf{T} \circ \mathbf{U}$.■

In Section **II.4** we proved the following result for the standard Euclidean plane $\mathbf{R}^2$; this result has an important generalization to arbitrary neutral planes.

<u>Theorem 2.</u> <u>(Compare Theorem **II.4.10**)</u> *Let $\mathscr{P}$ be a neutral plane, let* **a, b,** *and* **c** *be noncollinear points in* $\mathbf{R}^2$, *and let* **x, y, z** *be another triple of noncollinear points such that* $|\angle \mathbf{xyz}| = |\angle \mathbf{abc}|$. *Then there is a unique isometry* **T** *of $\mathscr{P}$ such that* $\mathbf{T(b)} = \mathbf{y}$, *the map* **T** *sends the segment* **[ba]** *into the ray* **[yx**, *and the map* **T** *sends the segment* **[bc]** *into the ray* **[yz**.■

A fairly detailed discussion of hyperbolic isometries is given on pages 159 – 170 of Ryan, but for the time being the key point is that all ***neutral planes support extensive families of isometries and are highly symmetric objects.***

Uniqueness theorems for neutral geometries

We have abstractly defined a hyperbolic plane to be a system satisfying certain axioms. However, in mathematical writings (***e.g.***, Ryan), one often sees references to **<u>THE</u>** hyperbolic plane as if there is only one of them, just as we talk about **<u>THE</u>** real number system or **THE** Euclidean plane. In all cases, the reason for this is that all such systems are characterized uniquely up to suitable notions of ***mathematical equivalence***.

Formally, this may be stated as follows:

Theorem 3. (Essential uniqueness of hyperbolic planes) *Suppose that* $(\mathcal{P}, \mathcal{L}, d, \mu)$ *and* $(\mathcal{P}', \mathcal{L}', d', \mu')$ *are hyperbolic planes. Then there is a* $1-1$ *correspondence* **T** *from* $\mathcal{P}$ *to* $\mathcal{P}'$ *with the following properties*:

1. *If* **x** *and* **y** *are arbitrary distinct points of* $\mathcal{P}$, *then there is a positive constant* k *such that* **T** *multiplies the distance between them by* k; *in other words, we have* $k\,d(\mathbf{x}, \mathbf{y}) = d'(\mathbf{T}(\mathbf{x}), \mathbf{T}(\mathbf{y}))$.
2. *The function* **T** *sends collinear points (with respect to* $\mathcal{L}$*) to collinear points (with respect to* $\mathcal{L}'$*) and noncollinear points (with respect to* $\mathcal{L}$*) to noncollinear points (with respect to* $\mathcal{L}'$*)*.
3. *If* **x, y, z** *are noncollinear points of* $\mathcal{P}$, *then* **T** *preserves the measurement of the angle they form; in other words, we have* $\mu\angle\mathbf{xyz} = \mu'\angle\mathbf{T(x)T(y)T(z)}$.

The important point about the $1-1$ correspondence **T** is its compatibility with the data for the two hyperbolic planes. Using this mapping as a "codebook," it is possible to translate every true statement about one of the systems into a true statement about the other, and likewise it every false statement about one system translates into a statement which is also false for the other system.

In principle, the result for hyperbolic geometry was known to mathematicians such as Taurinus, Gauss, J. Bolyai and Lobachevsky, and it reflects their (essentially) complete description of the measurement formulas for non – Euclidean geometry and its associated trigonometry. Proofs of the uniqueness theorem are discussed further in Chapter 10 of Greenberg and Chapter VI (particularly Sections 30 and 31) of the previously cited book by Borsuk and Szmielew.■

If we define a neutral plane to be ***Euclidean*** if Playfair's Postulate is true, then there is a corresponding but slightly stronger uniqueness theorem for Euclidean planes:

Theorem 4. (Essential uniqueness of Euclidean planes) *Suppose that* $(\mathcal{P}, \mathcal{L}, d, \mu)$ *and* $(\mathcal{P}', \mathcal{L}', d', \mu')$ *are Euclidean planes. Then there is a* $1-1$ *correspondence* **T** *from* $\mathcal{P}$ *to* $\mathcal{P}'$ *with the following properties*:

1. *If* **x** *and* **y** *are arbitrary distinct points of* $\mathcal{P}$, *then* **T** *preserves the distance between them*; *in other words, we have* $d(\mathbf{x}, \mathbf{y}) = d'(\mathbf{T}(\mathbf{x}), \mathbf{T}(\mathbf{y}))$.
2. *The function* **T** *sends collinear points (with respect to* $\mathcal{L}$*) to collinear points (with respect to* $\mathcal{L}'$*) and noncollinear points (with respect to* $\mathcal{L}$*) to noncollinear points (with respect to* $\mathcal{L}'$*)*.
3. *If* **x, y, z** *are noncollinear points of* $\mathcal{P}$, *then* **T** *preserves the measurement of the angle they form; in other words, we have* $\mu\angle\mathbf{xyz} = \mu'\angle\mathbf{T(x)T(y)T(z)}$.

Observe that a constant factor k does ***not*** appear in the statement of this result. One way of explaining this difference is that there are similarity transformations in Euclidean geometry with arbitrary positive rations of similitude but in hyperbolic geometry every similarity transformation is an isometry (this reflects the conclusion of the **AAA** Triangle Congruence Theorem; namely, similar triangles are automatically congruent).

The proof of the Euclidean uniqueness theorem reflects the standard method for introducing Cartesian coordinates into Euclidean geometry, and in principle the details are worked out in Chapter **17** of the previously cited book by Moïse (some material in Section **26.3** is also relevant).■

Euclidean approximations to hyperbolic geometry

For small enough regions on a plane, ordinary experience and the explicit formulas of spherical trigonometry show that Euclidean plane geometry is a very accurate approximation to spherical geometry. The situation for hyperbolic geometry is entirely similar; if we restrict attention to sufficiently small regions, the formulas of hyperbolic trigonometry and geometry show that Euclidean geometry is an extremely accurate approximation and that the degree of accuracy increases as the size of the region becomes smaller. For example, since the angle defect of a hyperbolic triangle determines its area, it follows that the angle sum of a triangle is very close to **180** degrees for all triangles in a very small region of the hyperbolic plane. In both spherical and hyperbolic geometry, as the diameter of a region approaches zero, the formulas of spherical and hyperbolic geometry in the region converge to the standard formulas of Euclidean geometry.

Higher dimensions

Thus far in this unit we have restricted the discussion to Euclidean and hyperbolic planes. In analogy with Euclidean geometry, there is a corresponding theory of ***neutral* 3 – *spaces,*** which satisfy all the axioms of Unit **II** for Euclidean **3 –** space except (possibly) Playfair's Postulate, and ***hyperbolic* 3 – *spaces,*** which are neutral **3 –** spaces that satisfy the negation of Playfair's Postulate. There is some discussion of **3 –** dimensional hyperbolic geometry in the references cited at the beginning of this section. As in the Euclidean case, it is also possible to discuss a theory of hyperbolic n **–** spaces for all $n \geq 4$.

V.6 : Subsequent developments

In Section **2** we indicated how advances in mathematics during the 17th and 18th centuries provided an important background for the work which led to the emergence of non – Euclidean geometry. Mathematical knowledge increased at an even faster pace during the 19th century; one superficial way of seeing this is to compare the amount of space devoted to that period in Kline's ***Mathematical Thought from Ancient to Modern Times*** to the amount of space devoted to the entire period before **1800.** In every area of the subject there were dramatic new discoveries, major breakthroughs in understanding, and substantially greater insight into logical justifications for the many advances of the previous three centuries. Within geometry, there were several major developments in addition to the emergence of non – Euclidean geometry. These include the systematic approach to curves and surfaces by the techniques of differential geometry, the establishment of projective geometry as a major branch of the subject, the explicit study of geometry in dimensions greater than **3**, and the use of algebraic techniques to analyze geometrical constructions by unmarked straightedge and compass which led to proofs that the three outstanding problems from classical Greek geometry (see the discussion near the end of Section **III.6**) could not be solved by means of straightedge and compass alone. We shall see that these other developments in geometry also turned out to have significant implications for non – Euclidean geometry.

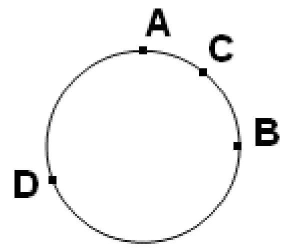
Elliptic geometry

We have noted that the angle sum of a triangle in Euclidean geometry is always equal to **180** degrees, while the angle sum of a triangle in hyperbolic geometry is always less than **180** degrees. On the other hand, we know that the angle sum of a triangle in spherical geometry is always strictly greater than **180** degrees, and thus it is natural to ask if there is a unified setting which includes both neutral geometry and spherical geometry. The crucial steps to constructing such a framework were due to G. F. B. Riemann (1826 – 1866), and his viewpoint led to far – reaching changes in the mathematical, physical and philosophical answers to the question, "What is geometry**?**"

In the next few paragraphs we shall only discuss the aspects of Riemann's ideas that relate directly to unifying spherical and neutral geometry. One key point was his questioning the standard model of a line in which one can find pairs of points whose distance from each other is arbitrarily large, and it is summarized in the following quotation from his writings:

> We must distinguish between ***unboundedness*** and ***infinite extent*** **...** The unboundedness of space possesses **...** a greater empirical certainty than any external experience. But its infinite extent by no means follows from this.

With respect to our setting for neutral geometry, this means that the standard Ruler Postulate should be replaced by a Circular Ruler Postulate which states that every line is in **1 – 1** distance – preserving correspondence with a standard circle of some fixed positive radius (***i.e.***, the radius a is the same for every line)**;** in analogy with our earlier discussion of hyperbolic geometry, the **square** a^2 of this radius can be viewed as a "curvature constant." If we adopt such a Circular Ruler Postulate, then we must also modify the entire discussion of order and separation. One easy way to see this is to observe that there is no reasonable notion of betweenness for three points on a circle. However, there is a decent substitute, for if we are given four points **A**, **B**, **C**, **D** on a circle then there is an obvious concept of ***separation*** for these points. Specifically, we can say that **A *and* B *separate* C *and* D** if each of the two arcs determined by **A** and **B** contains exactly one of the points **C** and **D.**



In the drawing above, the points **A** and **B** separate **C** and **D.**

We shall not attempt to make this precise here, but the previously cited books by Fishback and Coxeter contain further information. Here are two additional references which discuss further topics in elliptic geometry:

H.S.M. Coxeter, ***The Real Projective Plane*** (3rd Ed.), Springer – Verlag, New York, 1992. ISBN: 0–387–97890–9.

http://eom.springer.de/R/r081890.htm

There is still one fundamental issue that requires attention. In neutral geometry there is a unique line containing two points, but the analogous statement in spherical geometry is not necessarily valid because there are infinitely many great circles joining a pair of antipodal points. An idea due to F. Klein provides the usual way of avoiding this problem: Instead of considering the geometry of the sphere, one considers a ***reduced geometry*** whose points are ***antipodal pairs of points on the sphere.*** Such a construction is not really artificial, for *the points and lines of this reduced geometry are identical to the points and lines of the real projective plane* discussed in Unit **IV.**

Klein's motivation for the name ***hyperbolic geometry*** suggests the name ***elliptic geometry*** for the system that one obtains from spherical geometry by identifying pairs of antipodal points as above; sometimes elliptic or spherical geometry is called ***Riemann*** or ***Riemannian geometry***, but in mathematics and physics these terms normally refer to far more general constructions and thus *almost any other terms are preferable*. There is a corresponding name of ***parabolic geometry*** for Euclidean geometry, but this name has never been popular with mathematicians and is rarely used in modern mathematical writings.

Higher dimensions. There are n – dimensional analogs of both spherical and elliptic geometry for every $n \geq 3$.

Redefinition and expansion of geometry

Riemann's unified approach to spherical and neutral geometry was originally presented as illustrations of a far more general approach of geometry. The emergence of non – Euclidean geometry had suggested to Gauss and others that there was more than one "logically permissible" way of looking a space, depending upon which geometric properties one was willing to accept or do without. In addition to hyperbolic geometry, the rapid development of the differential geometry of surfaces at the time was an independent motivation for this viewpoint. Riemann's viewpoint, which also had a place for such systems, abandoned the idea that geometry involved absolute statements about space itself, replacing this with a premise that geometry involves the study of theories of space. In Riemann's approach, one has infinitely many possible theoretical options for describing space.

Even if Euclidean, hyperbolic and elliptic geometry represent only three of many possible theories of space, it is still clear that they represent three especially good theories. Therefore one of Riemann's central aims was to give a criterion for distinguishing these three from the unending list of possibilities. Within his framework, the three classical geometries are characterized by two special properties:

1. The existence of many different geometrical figures isometric to a given one.
2. A real number which is describable as a ***curvature constant.***

For Euclidean geometry the curvature constant is zero, while for hyperbolic geometry it is negative and for elliptic geometry it is positive; in the last two cases, the exact value depends upon the unit of linear measurement one adopts. In elliptic geometry, the square root of the curvature constant is the reciprocal of the radius for the corresponding sphere; the negativity of the curvature constant for hyperbolic geometry is related to Lambert's view of the latter in terms of "a sphere of imaginary radius" (***i.e.,*** the square of the radius is ***negative***).

Practicality and convenience. We have already noted that Euclidean geometry is a good approximation to either spherical or hyperbolic geometry if one restricts attention to a very small region. Since the formulas of Euclidean geometry are much simpler than those of the other geometries, for practical purposes it is generally more convenient to work inside Euclidean geometry unless the region under consideration is fairly large. The relative convenience of Euclidean geometry provides one answer to the issue raised in Poincaré's statement which we quoted at the beginning of this unit.

Further advances. Riemann's work opened the door to many new directions in geometrical research. For our purposes, it will suffice to say that the work led to more refined characterizations of the classical non – Euclidean geometries, particularly in the work of H. von Helmholtz (1821 – 1894) and S. Lie (1842 – 1899).

Consistency models in mathematics

Although Gauss, J. Bolyai, Lobachevsky and others ***concluded*** that there was no way to prove Euclid's Fifth Postulate from the other assumptions, they did not actually ***prove*** this fact. Their results gave a virtually complete and apparently logically consistent description of hyperbolic geometry, but something more was needed to ***eliminate***, or at least isolate, ***all doubts*** that ***someone might still succeed in finding a logical contradiction in the system.***

Mathematical statements that something cannot be found are frequently misunderstood, so we shall explain what is needed to show that a mathematical system is at least relatively free from logical contradictions. The discussion must begin on a somewhat negative note: Fundamental results of K. Gödel (1906 – 1978) imply that we can never be absolutely sure that any finite set of axioms for ordinary arithmetic (say, over the nonnegative integers) is totally free from logical contradictions. One far – reaching consequence is that ***there is also no way of showing that any infinite mathematical system is absolutely logically consistent.*** The best we can expect is to show that such a system will be ***relatively logically consistent***; in other words, if there is a logical contradiction in the system, then one can trace it back to a logical contradiction in our standard axioms for the nonnegative integers. The following quotation due to André Weil (pronounced "VAY," 1906 – 1998) gives a whimsical reaction which reflects current mathematical thought:

> God exists since mathematics is consistent, and the Devil exists since we cannot prove it.

The standard way to prove relative logical consistency is to construct a ***model*** for the axioms. Such models are to be constructed using data based upon the standard number systems of mathematics (the nonnegative integers, the integers, the rational numbers or the real numbers); the mathematical descriptions of these number systems show that all of them pass the relative consistency test described in the previous paragraph. If we can construct such a model, then one has the following ***proof for relative logical consistency***: Suppose that there is a logical contradiction in the underlying axiomatic system. Using the model, one can translate every statement about the model for the system into a statement about the mathematical number systems mentioned above, and thus the logical contradiction in the axiomatic system then yields a contradiction about these number systems. In other words, if there is a contradiction in the axioms, then there must also be a contradiction in the standard description of the standard number systems in mathematics.

If we consider the synthetic axioms for a Euclidean plane $(\mathcal{P}, \mathcal{L}, d, \mu)$, the standard model is given by the so – called ***Cartesian coordinate plane***, in which the set $\mathcal{P}$ of points equal to $\mathbf{R}^2$, the family $\mathcal{L}$ of lines is the usual of family subsets defined by nontrivial linear equations in x and y, the distance d between two points is given by the usual Pythagorean formula, and the cosine of μ is given by the standard formula involving inner products. In order to prove this is a model for the axioms, it is necessary to ***prove explicitly*** that ***all the axioms of Unit II are true for the given definitions of points, lines, distance and angle measure.*** Some steps in this process are fairly simple to complete, but others are long, difficult, and not particularly enlightening. It is frequently convenient to split the proof into two parts.

1. Replacement of the axiom system with an equivalent "reduced" one that requires fewer assumptions. (This can be long and difficult.)
2. Verification of the axioms in the "reduced" system.

We shall describe one relatively quick way of carrying out these steps for the Cartesian coordinate model of the synthetic axioms for Euclidean geometry. One particularly concise set of axioms for a Euclidean plane, consisting of only four statements, is given in the following classic paper:

> G. D. Birkhoff, *A set of postulates for plane geometry (based on scale and protractors)*, ***Annals of Mathematics*** (2) **33** (1932), pp. 329 – 345.

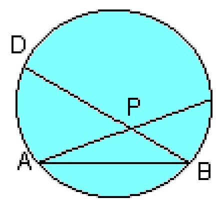
A verification of Birkhoff's postulates for the Cartesian coordinate model is given explicitly in the following online document:

http://www.math.uiuc.edu/~gfrancis/M302/handouts/postulates.pdf

Alternate approaches to verifying the axioms for Euclidean geometry in the Cartesian model appear various sections of the previously cited book by Moïse.

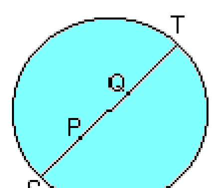
In view of the preceding discussion, it will follow that the Fifth Postulate is not provable from the other axioms if we can construct a model of a neutral plane $(\mathcal{P}, \mathcal{L}, d, \mu)$ which does not satisfy Playfair's Postulate. The first such model was constructed by E. Beltrami (1835 – 1900) in his paper, *Saggio di interpretazoine della geometria non euclidea,* which appeared in 1868. This model is frequently called the ***Beltrami – Klein model***, and it has a natural interpretation in terms of projective geometry, which was covered in Unit **IV** (however, we shall not use anything from that unit here). Much of the discussion below is adapted from the following online site:

http://www.cut-the-knot.org/triangle/pythpar/Model.shtml



The Beltrami – Klein model takes the interior of a circle as the set of ***points*** for a plane; recall that this region does not include points on the circle itself. The ***lines*** are given by open chords connecting points on the circle, with the endpoints excluded. It is not difficult to check that this system satisfies the basic incidence axioms from Section **II.1,** and the drawing above suggests an argument to show that Playfair's Postulate does not hold for points and lines in the Beltrami – Klein model. Specifically, in this picture the Beltrami – Klein lines (= open chords) **AC** and **BD** pass through point **P** and neither meets the open chord **AB.**

Defining the distance and angle measurement for the Beltrami – Klein model is considerably more difficult; we shall only define the distance function and note that one can reconstruct the angle measurement function using the results in the previously cited book by Forder. Since hyperbolic geometry is unbounded, in order to realize it in a bounded region of $\mathbf{R}^2$, it is necessary to define distance so that the distance from one point to another goes to infinity if one is fixed and the other approaches the boundary circle.



Given two points **P** and **Q** in the open disk, suppose that the Euclidean line joining them meets the circle at points **S** and **T.** Then the Beltrami – Klein distance between **P** and **Q** is defined by the following strange looking formula (which is related to the cross ratio that was defined and studied in Unit **IV**):

$$d_{\mathrm{BK}}(\mathbf{P}, \mathbf{Q}) \;=\; \left| \log_e \left(\left(|\mathbf{Q}-\mathbf{S}|\cdot|\mathbf{P}-\mathbf{T}| \right) / \left(|\mathbf{P}-\mathbf{S}|\cdot|\mathbf{Q}-\mathbf{T}| \right) \right) \right|$$

Here $|\mathbf{X}-\mathbf{Y}|$ denotes the ***Euclidean*** distance between **X** and **Y.** It is a routine exercise to check that if **Q** moves away from a fixed **P** staying on the same line, the distance between the two points grows without bound. This curious property of the model sounds somewhat like a line from Shakespeare's play, **Hamlet** (Act **II**, Scene **2**, line **234**):

> I could be bounded in a nutshell, and count myself king of infinite space.

Higher dimensions. There are analogs of the Beltrami – Klein model for hyperbolic n – space in every dimension $n \geq 3$.

Beltrami's model finally gave a definitive answer to questions about the role of Euclid's Fifth Postulate, showing that ***it is impossible to prove it or an equivalent statement from the other usual sorts of axioms.*** In many respects, this outcome is extremely ironic. Many of the early efforts to prove the Fifth Postulate were motivated by a belief that its inclusion was a logical shortcoming of the ***Elements***. For example, the title to Saccheri's work on the subject began with the words which translate to ***Euclid vindicated***, and the following quotation from a letter to J. Bolyai from his father Farkas (Wolfgang) Bolyai (1775 – 1856) expresses a similar view:

> I [also] thought … I was ready to … remove the flaw from geometry and return it purified to mankind.

In fact, the ***real*** vindication of Euclid took place with the construction of Beltrami's example, which showed that something like the Fifth Postulate is logically indispensable for the development of classical Euclidean geometry and indicates a very respectable level of insight on Euclid's part into the logical structure of deductive geometry.

Following the construction of the Beltrami – Klein model, several other models were also described, and some will be described or referred to in the next section.

Foundational and philosophical consequences

In the introduction to this unit, we noted that the emergence of non – Euclidean geometry had a strong impact on the philosophy and foundations of mathematics. In the next few paragraphs we shall describe this impact in greater detail.

Background. At the beginning of the 19^{th} century, Euclidean geometry was viewed as a reliable foundation for mathematics. Its importance for geometry is evident, but it was also important for algebra. The reasons for this are described in the following passage from M. Kline's ***Mathematics and the Physical World*** (Corrected reprint of the 1959 Ed., Dover, New York, 1981. ISBN: 0–486–24104–1):

> As of 1800, mathematics rested upon two foundations: The number system and Euclidean geometry. … Mathematicians would have emphasized the latter because many facts about the number system, and about irrational numbers especially, were not logically established nor clearly understood. Indeed, those properties of the number system that were universally accepted were still proved by resorting to geometric arguments, much as the Greeks had done 2500 [possibly more like 2100] years earlier. Hence, one could say that Euclidean geometry was the most solidly constructed branch of mathematics.

These ideas are explicit in following two quotations from the writings of Isaac Barrow (1630 – 1677):

> Geometry is the basic mathematical science, for it includes arithmetic, and mathematical numbers are simply the signs of geometrical magnitude.

> Geometry is certain [contrary to the infinitesimal calculus] because of the clarity of its concepts, its unambiguous definitions, our intuitive assurance of the universal truth of its common notions, the clear possibility and easy imaginability of its postulates, the small number of its axioms ...

This viewpoint is also implicit in Isaac Newton's monumental work ***Principia***, which uses Euclidean geometry as its logical foundation.

At the beginning of this unit, we included a quotation from Kant reflecting his view of Euclidean geometry as description of ***a priori*** truths, just like the fundamental rules for arithmetic. His viewpoint on such ***a priori*** truths is reflected in the following passage from ***The Story of Philosophy*** (Pocket Books, Simon and Schuster, New York, 1991. ISBN: 0 – 671 – 73916 – 6) by Will Durant (1885 – 1981):

> We may believe that the sun will "rise" in the west tomorrow, or that … fire will not burn stick, but we cannot for the life of us believe that two times two will ever make anything else than four. Such truths are true before experience … they are absolute and necessary; it is inconceivable that they should ever become untrue. … These truths derive their necessary character from the inherent structure of our minds, from the natural and inevitable manner in which our minds must operate.

As suggested by the quotation from Gauss at the beginning of this unit, the discovery of non – Euclidean geometry and the logical independence of the Fifth Postulate provided compelling evidence that the standard axioms for Euclidean geometry are not ***a priori*** truths.

The preceding developments had several implications. One was a need to give a new description of geometry, and this was done along the lines indicated in the following quotation from Kline's ***Mathematics in Western Culture*** (Oxford University Press, New York, 1964. ISBN: 0–195–00714–X):

> A [geometric] mathematical space now takes on the nature of a scientific theory. … The creation of the new geometries … forced recognition of the fact that there could be an "**if**" about mathematical systems. ***If*** the axioms of Euclidean geometry are truths about the physical world then the theorems are. But … we cannot decide on ***a priori*** grounds that the axioms of Euclid, or of any other geometry, are [empirical] truths [about the physical world].

A second implication was the need to replace the role of Euclidean geometry as a foundation for mathematics by something else; actually, the discoveries related to the Fifth Postulate were just one of many factors which forced mathematicians to look more carefully at the foundations of the subject during the 19th century and to find solid logical justifications for the spectacular advances the subject had made during the preceding three centuries. By the end of the 19th century the modern approach to the foundations of mathematics had essentially been outlined with **(1)** the development of set theory, **(2)** the simple axiomatic characterization of the positive integers due to G. Peano (1858 – 1932), and **(3)** the formal construction of the real number system in terms of the rational numbers and characterization of the real numbers due to R. Dedekind (1831 – 1916). Each of these stands as a major achievement for separate reasons. In particular, Peano's axioms effectively answered questions by philosophers such as John Stuart Mill (1806 – 1873) about the ***a priori*** nature of arithmetic, and Dedekind's work finally resolved basic questions about irrational numbers which had been unanswered ever since the Pythagoreans discovered that the square root of **2** is irrational.

> The value of non – Euclidean geometry lies in its ability to liberate us from preconceived ideas in preparation for the time when exploration of physical laws might demand some geometry other than Euclidean.
>
> G. F. B. Riemann
>
> We have Einstein's space, De Sitter's space, expanding universes, contracting universes, vibrating universes, mysterious universes. In fact, the pure mathematician [or the modern theoretical physicist] may create universes just by writing down an equation, and indeed if he is an individualist he can have a universe of his own.
>
> J. J. Thomson (1856 – 1940) [discoverer of the electron]

Although the emergence of non – Euclidean geometry raised immediate questions whether the physical universe satisfies the axioms of geometry, the real impact of these developments on physics did not begin for some time. We have noted that Euclidean geometry provides an excellent approximation to hyperbolic and elliptic geometry in small regions, and until the end of the 19th century experimental observations and classical physics were consistent with the mathematics of Euclidean geometry. However, near the end of the century physicists found that classical physics did not provide adequate explanations for some key experimental observations, and this led physicists to consider new mathematical models which would conform more closely to experimental results. Efforts by H. Lorentz (1853 – 1928) and G. FitzGerald (1851 – 1901) to explain the results of one important experiment led to a generalization of Riemann's geometric structures (***Lorentzian geometry***) that was a precursor to the Theory of Special Relativity introduced by A. Einstein (1879 – 1955) in 1905. Further extensions of Riemann's ideas led to the mathematical theory of space – time that underlies General Relativity Theory. Many other systems that can be called "theories of space" also appear in many contexts of 20th century (and present day) physics.

We shall conclude this section by discussing two points about Einstein's work and its relation to non – Euclidean geometry that are frequently misstated or misunderstood.

<u>Is the geometry of relativity theory a non – Euclidean geometry?</u> The answer to this question depends upon ***how one defines non – Euclidean geometry.*** One basic point in relativity theory is that the presence of mass warps or curves the structure of space – time. In Euclidean geometry there is no curvature whatsoever, and thus it is clear that the geometry of space – time cannot be Euclidean. Furthermore, since the distribution of mass in the universe varies from place (and time) to place (and time), the curvature of space – time is also variable. In the classical non – Euclidean geometries (hyperbolic and elliptic), the curvature is nonzero but the same at all points. This means that ***the geometry of space – time is neither Euclidean, hyperbolic, nor elliptic.*** Therefore the answer to the question at the beginning of this paragraph depends upon whether non – Euclidean means *<u>anything that is not Euclidean</u>* (in which case the answer is **YES**) or means *only the classical examples* of hyperbolic and elliptic geometry (in which case the answer is **NO**).

What was Einstein's role in developing the mathematics of relativistic geometry? The basic mathematical framework for relativistic geometry had been previously created by others, and Einstein's fundamental insight was to see that this framework was uesful for formulating certain fundamental laws of physics. His chief ***mathematical*** contribution was a geometrical formula relating the curvature properties of space – time to the distribution of matter in the universe (the so – called ***Einstein tensor equation***).

A more detailed account of the historical ties between geometry and physics is beyond the scope of these notes, but a fairly readable and detailed account of the history into the early 20th century is contained in the following book.

> C. Lanczos, ***Space through the Ages: The evolutions of geometric ideas from Pythagoras to Hilbert and Einstein,*** Academic Press, New York, 1970. ISBN: 0–124–35850–0.

Finally, here is an article which discusses several other issues related to this section:

> D. W. Henderson and D. Taimina. "How to Use History to Clarify Common Confusions in Geometry," Chapter **6** in ***From Calculus to Computers: Using Recent History in the Teaching of Mathematics***. Mathematical Assoc. of America Notes No. **68** (2005), pp. 57 – 73.