AN ABSTRACT OF THE THESIS OF

Joseph Michael Kincaid for the Master of Arts Degree

in Mathematics presented on December 21, 1985

Title: CONSTRUCTIBLE NUMBERS IN DOUBLE ELLIPTIC GEOMETRY

Abstract approved: Marion P. Emerson

Constructible numbers serve to connect the areas of algebra and geometry. In the Euclidean plane, the straight edge and compass can be used to create line segments with the length of any finite combination of arithmetic operations or square roots. This set of numbers has the algebraic properties of an Archimedean ordered field.

Double elliptic geometry differs from Euclidean geometry in that it requires that all lines meet and are hence finite in length. Constructions in this geometry are isomorphic to constructions on a sphere.

The constructions presented in Euclid's Elements can be redone on the sphere to demonstrate the differences between the two geometries. When this is done, the line segments that are constructed are not the same lengths as the line segments in Euclidean geometry. Instead, this set of segments has properties based upon trigonometric identities, specifically ones involving the cosine function.

Though the numbers are different, theorems concerning the numbers resemble theorems concerning the numbers in Euclidean geometry. This resemblance between the theorems can be generalized to other geometries as well.

CONSTRUCTIBLE NUMBERS IN DOUBLE ELLIPTIC GEOMETRY

A thesis

submitted to

The Division of Mathematical and Physical Sciences

Emporia State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Joseph Michael Kincaid

December, 1985

Marion P. Emerson

Approved for the Major Department

Harold E. Durst

Approved for the Graduate Council

DEDICATION

This thesis is dedicated to the memory of the one man who made this and many, many other projects possible. My father, I thank you, for teaching me how to learn.

TABLE OF CONTENTS

Chapter I

Introduction

The idea of constructible numbers is an important link between two apparently different areas of mathematics. In geometry, Euclidean constructions play an important role in the history and development of Euclidean geometry. In algebra, multi-quadratic extension fields over the rational numbers play an important role in proofs of several theorems. Together, these two concepts lead to constructible numbers and using these numbers some very interesting results can be proved.

This paper will explore the concept of constructible numbers in double elliptic geometry - a non-Euclidean geometry. The approach used will consist of first defining constructions in this geometry, then examining some specific constructions more closely to see how they work, and finally demonstrating that a specific set is closed under the operations of construction in this geometry. The paper will discuss applications and properties of this set of numbers in comparison to constructible numbers in Euclidean geometry.

A few theorems of absolute, synthetic, or metric geometry are referred to without proof or explanation. For some theorems, however, the proof or perhaps the concept involved is important in the development of non-Euclidean constructions. In these cases a brief explanation is given.

Notation in this paper is fairly standard notation for the topics involved. In a few cases, the notation is explained as it is initially used, but generally, it is assumed to be known. In any event, the context will give an explanation if the text does not.

The ideas presented in this paper are motivated by a number of works on constructible numbers in Euclidean geometry, but no work has yet been found on non-Euclidean constructions.

THE EUCLIDEAN PLANE

The Euclidean plane is the logical result of five axioms. Euclid put forth these axioms in the first book of the *Elements* as follows:

1. To draw a straight line from any point to any point.
2. To produce a finite straight line continuously in a straight line.
3. To describe a circle with any center and distance.
4. That all right angles are equal to one another.
5. That, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which are the angles less than two right angles.

Euclid called these postulates. In more modern English, the first four read as follows:

1. Through any two points, there exists one and only one straight line.
2. Any line segment may be extended indefinitely in either direction.

3. Given a point and a line segment, there exists a circle with the point as center and the length of the segment as radius.

4. Any right angle is congruent to any other right angle.

Although it can be shown that everything after the word "meet" in the fifth axiom can be proven from the preceding assertions and hence could be left off, there really aren't other ways to phrase it and retain the intended meaning. There are, though, other axioms that are logically equivalent to the fifth axiom. One such axiom, known as Playfair's Axiom reads as follows:

Through a point not on a given line there passes not more than one parallel to the line.

Because this axiom is easier to read and understand, it often replaces Euclid's fifth axiom in modern-day textbooks. Because it is equivalent, though, nothing is lost logically by replacing Euclid's fifth axiom with this one.

<u>EUCLIDEAN</u> <u>CONSTRUCTIONS</u>

The first three axioms of Euclid are essentially existence axioms describing points and lines in the plane that do exist. Specifically, they present a method whereby, given the existence of a minimum of points in the plane, an infinite number of other points and an infinite number of lines may be shown to exist. The process of creating these other points and lines is known as Euclidean constructions.

In the original spirit of the words, Euclidean constructions are defined to be those constructions that are possible using only an unmarked straight edge and a collapsible compass. These tools are "ideal" tools which "draw" straight lines and circles as exact sets of points. This ideal sense of the tools is a necessary product of the axiomatic nature of the material of the plane.

The straight edge is used to draw the line through any pair of points (Axiom I). It can also be used to extend a line indefinitely in either direction (Axiom II), but it cannot be used to measure the length of any line segment, or even to tell whether two segments are congruent. There is no metric given in Euclid's postulates. In the Euclidean plane, since lines are infinite in extent, the "ideal" straight edge should also be considered infinite in extent.

The compass is traditionally used to draw the circle through a given point with center at another given point (Axiom III). Because the compass is collapsible, it may not be directly used to construct the circle through a third point with the same radius. However, in the Euclidean plane, the combination of straight edge and compass does allow for a series of constructions whereby a segment is copied onto a ray. The construction of the circle follows directly from the construction of this segment.

If the real number system and a coordinate system for the Euclidean plane are given, then the straight edge and compass can be used to perform algebra. That is, given a line segment of length 1 and two segments of length a and b, the straight edge and compass can be used to construct segments of lengths:

$$a + b \qquad ab \qquad 1/a \qquad b/a \qquad a^{1/2}$$

The first of these is trivial and the middle three are important to this paper mostly in that they involve similar triangles. The final one, however, is just tricky enough that it merits a brief explanation.

In order to extract a square root of a given segment PQ, one must first construct R such that P-Q-R (read "Q is between P and R") and QR = 1. This segment is then bisected, and a semicircle drawn with PR as diameter. The perpendicular segment to PR at Q will intersect the circle at S. The segment QS has a length equal to the square root of PQ. This fact is easily demonstrated using similar triangles PQS and SQR.

Being able to construct segments with each of these lengths is equivalent to saying that given any two numbers there exists a third number of each of the above types. Since the operations of addition and multiplication for line segments are defined and discussed in terms of the corresponding operations for real numbers, each of the properties in the real number system is inherited by this system. By virtue of the above constructions, any of the rational numbers can be constructed by a finite number of constructions using the straight edge and compass. Further, since the square root operation is allowed, any number that can be written as the result of a finite number of additions, multiplications, and square root extractions can be constructed.

The rational numbers, Q, form an algebraic structure known as a field. Generally speaking, fields have all the properties normally associated with numbers including such ones as commutative property of addition, associative property of multiplication, and the closure property for both operations. When the square root operation is added to the operations, the resulting set of numbers form what is known as a multi-quadratic extension field, F, over the rational numbers, Q.

Because this field is closed under the operations associated with Euclidean constructions, any segment that can be constructed using a straight edge and compass has a length in this field. This closure property is used primarily to demonstrate that certain constructions are not possible using straight edge and compass. A popular example is the question of the existence of a general technique for trisecting an angle. It can be shown that $\cos 20^{\circ}$ is not in the field F. Hence, there are no constructible lines that form an angle of 20°. Since a 60° angle is easily formed in an equilateral triangle, and a general trisection technique would be able to trisect this 60° angle giving a 20° angle, the fact that there is no 20° angle leads directly to the fact that there is no general trisection technique.

It merits pointing out that the facts mentioned thus far are dependent upon the properties of the Euclidean plane. The constructions are performed in the plane and the numbers in the field F are results of these constructions. This implies that the proof that there is no general method of trisecting an angle is dependent upon the characteristics of the Euclidean plane. It follows as a natural question to ask what happens when the plane is not Euclidean and that is the topic of this paper.

Chapter II

Definitions

DOUBLE ELLIPTIC GEOMETRY

Double elliptic geometry differs from Euclidean geometry at the fifth postulate. The point is most easily stated in terms equivalent to Playfair's Axiom:

Every line through a point not on a given line intersects the given line.

The results of this change in the axioms are dramatic. In fact, the new geometry is at first inconsistent. To repair this, straight lines, circles, and even points shall be re-examined carefully in light of this new axiom. The axioms shall then be modified to accept this new axiom and allow for a consistent geometry.

First, however, a model is necessary to adequately discuss a geometry so different from the usual intuitive conceptualization of the real world. The sphere, seen as a two-dimensional surface in a three-dimensional Euclidean space, is conceptually identical to double elliptic geometry, and hence provides the necessary model. Since every property on the sphere has a corresponding property in double elliptic geometry, and vice versa, it is sufficient to examine spherical geometry and then translate the results to double elliptic geometry. This method of attack has the chief advantage of providing a means for discussion of intuitively difficult concepts.

On the sphere, the shortest distance between points is the arc of a great circle. Hence, by extending this arc, a "line" in spherical geometry is a great circle of the sphere. Although this fits some of the intuitive notions of a line, if also poses an immediate problem. Every great circle intersects every other great circle on the sphere, as desired, but it does so twice: once at each end of a diameter of the sphere. Thus, in order to retain the familiar properties of lines and points, some compromises must be made. One option is to allow lines to intersect in two points and accept the fact that through two antipodal points (the two opposite points where two lines intersect) there exist infinitely many lines. The other option is to call the two points where the lines intersect the same point, in effect, redefining the word "point". The latter method results by definition in every pair of lines meeting in precisely one "point", produces what is called single elliptic geometry and is modeled by the geometry on a modified hemisphere. It is the former method, however, that is actually spherical geometry, and since it is easier to work in, it is the geometry this paper will be confined to.

It should also be noted that the lines in spherical geometry are not infinite in extent. Instead each line is bounded in space, hence finite in extent. Although, this is different from the lines in Euclidean geometry, it is not an axiom of Euclid's that is contradicted, but one of the underlying assumptions. The only requirement in the axioms is that it be possible to extend each line segment continuously in each direction. This is certainly possible on a sphere using the concept of a great circle as a line. This difference will not pose a problem with the logical system, but will cause a problem when the theorems of Euclidean geometry are examined in the double elliptic space.

The straight line being finite implies that there exists a maximum distance in double elliptic geometry. This maximum distance is the distance between two antipodal points and shall be denoted as 2q. This distance is named 2q in reference to an interesting property of perpendiculars. Since any two lines will meet in the plane, two lines perpendicular to the same line will meet. In fact, all perpendiculars to the same line will meet at the same point. This point is called the pole of the line and the line is called the polar of the point. An example is the north pole and the equator on the globe. Every line of longitude is a great circle, perpendicular to the equator, and passes through the north pole. For completeness' sake, it should be noted that each line has two poles, one on each side.

Straight lines being finite has one more implication that needs to be pointed out. Because there does exist a maximum distance, there will also exist an absolute measure of distance. Hence, the very structure of the space itself forces a metric to be used. This metric will be used implicitly in almost everything that follows. An arbitrary coordinate system with the center of the sphere at the origin is assumed. The actual coordinates themselves will be discussed later as the need arises.

CONSTRUCTIONS IN DOUBLE ELLIPTIC GEOMETRY

In order to perform constructions in double elliptic geometry, traditional physical notions of a straight-edge and compass must be abandoned in favor of more abstract meanings of the terms. A straight-edge in this system shall be defined so as to draw the great circle on the sphere. This definition allows both of the necessary tasks to be performed: The great circle between any two points may be constructed using this straight edge and any arc of a great circle may be extended completely around the sphere. Each of these has a characteristic different from Euclidean geometry that has already been pointed out. First, between two antipodal points there does not exist a unique straight line. Second, a line is no longer infinite in extent, but finite. The great circle still allows the line segment to be extended indefinitely as the second axiom requires.

The great circle on the sphere, in addition to being the shortest distance between two points, is also the intersection of the sphere with a plane through the center of the sphere. This definition of the great circle has an advantage over any other in that the algebra associated with planes is simpler than the algebra associated with circles. Hence, if the intersection of two great circles is desired, the intersection of their respective planes can be found and then the intersection of that line with the sphere will give the desired intersection points.

In order to define a compass on a sphere, the circle it draws needs to be viewed more abstractly. The circle in the plane is the locus of points in the given plane equidistant from a given point. Since the locus of points equidistant from a given point, but not restricted to the plane, is the surface of a sphere, the operation of construction using a compass is actually the operation of intersecting a sphere with the plane. The actual size of this sphere itself is completely arbitrary, bounded only from below by the size of the circle to be constructed.

Moving to spherical geometry, the circle on the original sphere may be viewed as the intersection of the sphere with the surface of another sphere. Thus, the compass of spherical geometry is another sphere that can be used to find intersections with the original sphere. Again, the radius is arbitrary. The work is simplified, though, if it is taken to be the radius of the original sphere.

There is an important difference between the circles of double elliptic geometry and the circles of Euclidean geometry that stems from the fact that lines are now finite in extent. Since the maximum absolute distance, $2q$, is the distance between two antipodal points, there cannot be a radius given greater than this distance. Further, if the given radius is half this distance, then the circle so constructed is actually a straight line.

Another way of simplifying the work involved is to properly choose the coordinate system. If it is given to construct a circle with a certain point and a certain radius, the x-axis should be placed so as to pass through the point given as center of the circle. If this is done, the center of the sphere used for constructing the circle will also be on the x-axis. The equation of the original sphere in three dimensions is:

$$(1) \qquad x^2 + y^2 + z^2 = r^2$$

since its center is at (0, 0, 0) and its radius is r. The equation of the new sphere is:

$$(2) \qquad (x - x')^2 + y^2 + z^2 = r^2$$

since its center is at (x', 0, 0) and its radius is r. Considering these as a system of equations to be solved for the intersection points, it is readily apparent that subtracting (1) from (2) gives:

$$(3) \qquad (x - x')^2 - x^2 = 0$$

and then simplifying:

$$(4) \qquad x = \frac{x'}{2}$$

From this, given r and hence x' from the conditions for the circle, x will have a constant value. In three dimensions, a constant x, with no conditions on y and z, is the equation for a plane perpendicular to the x-axis. Therefore, given that x is a constant, it follows directly that the circle on the sphere will be the intersection of the sphere with the plane $x = x'/2$.

It remains to define x' in terms of the original conditions given for constructing the circle: the center and radius as measured on the sphere. The center is placed at (r, 0, 0), since the x-axis passes through that point. The radius of the circle is still an arbitrary length at this point in the discussion and shall remain arbitrary until later. If the radius of the circle is called p to distinguish it from the radius of the sphere r, then the problem is to express x in terms of p. At first, this question seems rather involved but since p is measured from the x-axis, x is simply the projection of the radius of the sphere onto the x-axis. Algebraically,

(5) $x = r\cos(p/r)$

is the equation for the plane that will intersect the sphere in the desired circle. If the radius of the sphere is taken to be 1, then the equation is simply $x = \cos(p)$. Nothing is lost by making this assumption.

Treating these circles as intersections of planes with spheres is more than just a parlor game. In the Euclidean plane, the formulas for distance and ratios are familiar to the high school geometry student. In any non-Euclidean space, however, the formulas are usually not familiar to even the graduate student in mathematics. Further, since distance is to be measured on a sphere, the planes can be used to find intersections of curves that would otherwise be algebraically tedious.

Since both types of constructible curves, i.e. lines and circles, can be expressed as the intersection of planes with the sphere, the constructions performed in spherical geometry can be viewed in terms of the intersections of various planes with the sphere. For some constructions, this is more work than it is worth, but when the lengths themselves become important, this tool is very handy and can be used often. Further, when the lines make unfamiliar angles and form odd shapes, the visualization of the problem is facilitated greatly by these planes.

Chapter III

The First Few Constructions

The first twenty-eight propositions of the first book of Euclid's Elements are performed without the benefit of the fifth postulate. They do, however, use some underlying assumptions that contradict the double elliptic axioms. Hence, they are not necessarily valid in double elliptic geometry. A number of those propostions continue to be valid under certain conditions, though, and it is those propositions that will provide a starting point for the constructions in double elliptic geometry.

While examining Euclid's constructions to see which ones remain valid in double elliptic geometry, it will be periodically necessary to examine some constructions unique to this new geometry. None of these double elliptic constructions appeared in the Elements because their nature is unique to elliptic geometry.

EQUILATERAL TRIANGLES

The first of Euclid's proposition is:

(I) On a given finite straight line to construct an equilateral triangle.

The demonstration of this given in the Elements consists of first constructing the circles at each endpoint with radius equal to the length of the segment. These two circles will intersect at two points. Choose one and construct the lines through this point and the endpoints of the segment. These two lines together with the original segment form the desired equilateral triangle.

In spherical geometry this construction is not always possible. If this construction is considered in terms of the planes of the previous chapter, it is easiest to discuss. For the sake of simplicity, the x-axis shall be placed through one of the points, say point A. The plane drawing the circle with center at A will then have equation:

(6) $x = \cos(k)$

where k is the distance from A to the other point B.

The equation of the plane creating a circle through A with center at B is not near as simple. In terms of an x'-axis through B, the equation is simply:

(7) $x' = \cos(k)$

as before. However, in order to use these together, they both must be written in terms of one variable, x. The x'-axis makes an angle of k with the x-axis. If the y-axis and z-axis are placed so that the x'-axis is in the xy-plane, then the equation of the x'-axis in terms of x and y would be:

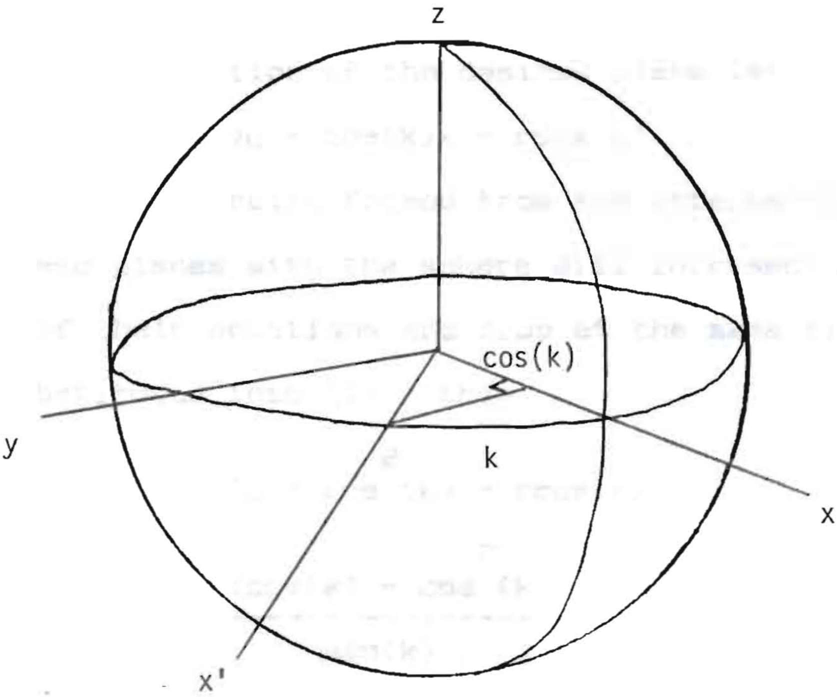
(8) $\cos(k)y - \sin(k)x = 0 \quad ; \; z = 0$

Hence, a plane perpendicular to this axis would have equation:

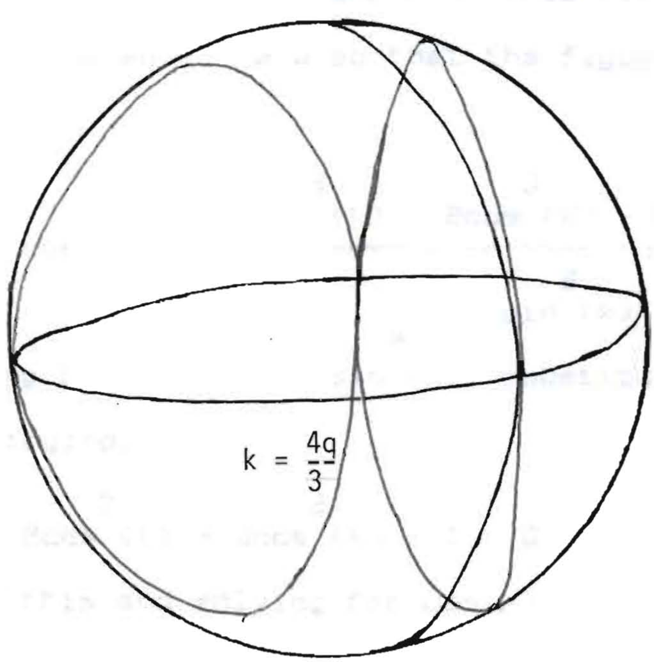
(9) $\sin(k)y + \cos(k)x = D$

where D is a constant telling where the plane lies. In this case, it is desired that the plane pass through (r, 0, 0). Substituting these coordinates for x and y, it is obvious that

(10) $D = r\cos(k)$



(6)



(16)

Hence, the equation of the desired plane is:

(11) $\sin(k)y + \cos(k)x = r\cos(k)$

The two circles formed from the intersection of each of these planes with the sphere will intersect whenever both of their equations are true at the same time. If (6) is substituted into (11), then

(12) $\sin(k)y + \cos^2(k) = r\cos(k)$

and hence,

(13) $$y = \frac{r\cos(k) - \cos^2(k)}{\sin(k)}$$

This equation, together with (6) and that there are no restrictions on z, form the equation for the line in three dimensions where the two planes intersect. In order for the two circles to interect, this line must intersect the sphere. Substituting (6) and (13) into (1) with r taken to be 1, and z taken to be 0 so that the figure lies in the xy-plane:

(14) $$\cos^2(k) + \frac{\cos^4(k) - 2\cos^3(k) + \cos^2(k)}{\sin^2(k)} = 1$$

Multiplying both sides by $\sin^2(k)$, substituing $1 - \cos^2(k)$, and simplifying:

(15) $2\cos^3(k) - 3\cos^2(k) + 1 = 0$

Factoring this and solving for cos(k):

(16) $\cos(k) = -1/2$ or $\cos(k) = 1$

These two solutions represent the locations where the two circles are tangent to each other. The first of these occurs 2/3 of the way from one antipodal point to another. The second occurs when the two points are simultaneous. For all radii given between these two, the circles will intersect on the sphere. For any radii given greater than 4q/3, there will be no intersection points for the circles.

This has immediate important implications. First, in any construction that follows, the first question to be asked will always be: does the construction require that two circles intersect when the distance between their centers may not be less than 4q/3? If it does, then either the construction will not be possible, the construction will be restricted to points closer than this distance, or a new way to perform the construction will need to be found.

Second, since an equilateral triangle with sides equal to 4q/3 will have a perimeter equal to 4q, or twice the maximum distance, the triangle itself is a great circle. Hence, every triangle on a sphere will confine itself to a single hemisphere. Another way to visualize this is by focusing on two points and the great circle through them. Any third point that will be used to form a triangle will have to be in one hemisphere or another. The result obtained above regarding the equilateral triangles then becomes an obvious truism.

<u>ANTIPODAL</u> <u>POINTS</u>

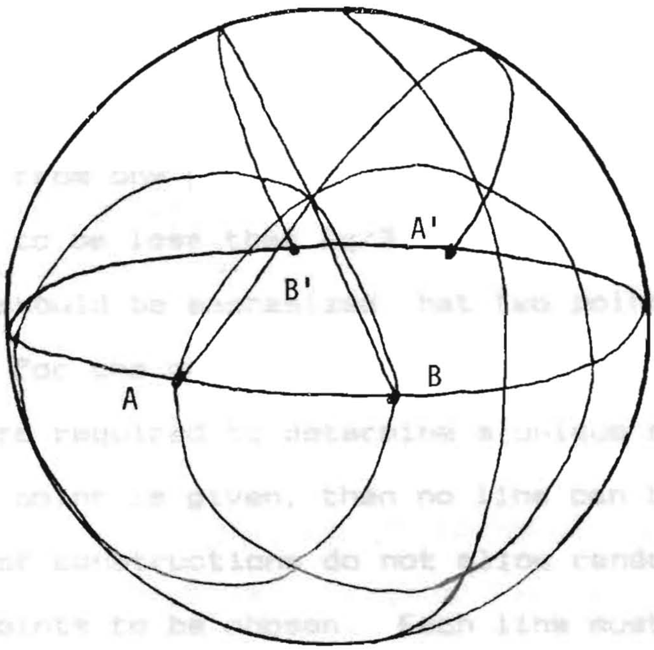
The first construction special to double elliptic geometry will be:

(A) Given two non-antipodal points, to find the points antipodal to each.

The method is simple given the above construction. First, the equilateral triangle mentioned above is to be constructed. Then the three sides of the triangle should be extended completely around the sphere. Since each pair of lines will intersect a second time in the antipodal points, the desired points will appear as intersections of these lines. In accordance with the restrictions on construction (I), it is important that the two given points be no farther apart than 4q/3.

If the two points are farther apart than 4q/3, a different method must be employed. This second method is more general than the previous one, and in fact is applicable in more cases. The antipodal points are found one at a time as follows: The circle with center at either and radius equal to their distance is first drawn. Since this circle is also the circle with center at the antipodal point of the first center, the question of finding the antipodal point becomes one of finding the other center of this circle.

The line segment under consideration is then extended completely around to provide a diameter of the circle. The desired point at the center of the circle is the midpoint of this segment. To find this midpoint, two equilateral triangles are constructed on the segment representing the diameter and their apexes joined. The equilateral



Constructing Antipodal Points

triangles are always possible in this case since if the original points were more than 4q/3 apart, then the distance from one point to the antipodal point of the other is going to be less than 2q/3.

It should be emphasized that two points must be given in order for the construction to be possible. Since two points are required to determine a unique straight line, if only one point is given, then no line can be drawn. The "rules" of constructions do not allow random lines or random points to be chosen. Each line must be determined by two points and each point must be determined by the intersection of two lines or a line and a circle or two circles.

Chapter IV

The First Few Numbers

At this point, constructions have been performed to construct an equilateral triangle and to locate an antipodal point. The letter q has been used freely to represent half of the maximum distance and underlying all of the previous work has been the concept of a metric on a sphere. Before any further work is done, all of this needs to be summarized explicitly in terms of the operations and numbers resulting from this work.

Axiomatically speaking, no constructions can be done until the first few points are given. Hence, it will be assumed that two points on the sphere are located as given points. From these the remainder of the points and lines shall follow.

Because there exists an absolute metric on the system, the location of these points is critical to the set of numbers that follows. For example, if the two given points are antipodal points and hence their distance is the maximum distance, then no number greater than one can be constructed. This is an important distinction from Euclidean geometry where arbitrary points and an arbitrary unit of length would suffice. For the purpose of simplifying the work involved, the given points shall be assumed to be a unit apart and the sphere shall be assumed to have unit radius. Note that from experience in analytical and coordinate geometry, this implies immediately that q is $\pi/2$.

THE INTEGERS ON THE SPHERE

Given the two initial points, x_0 and x_1, the line through them may be constructed completely around. Also, the circle with center at x_1 and radius x_0x_1 may be constructed. By repeatedly constructing a new circle with center at x_n and radius x_nx_{n-1}, a point corresponding to each of the integers may be constructed.

However, this does not imply that the integers themselves are constructible. For a number to be constructible, a segment with that number as length needs to be constructible. Since π is the maximum distance, it is not possible to construct a length for an integer greater than π.

But obviously points corresponding to the integers have been constructed. The question then is what lengths do these integers correspond to? If the point corresponding to 4 is taken as an example, the answer can be made readily apparent.

The point 4 is constructed by marking off a unit distance four times. However, after the distance is marked off three times, the point is now approximately 0.14 units ($\pi - 3$) away from the antipodal point of the starting point. Therefore, when the final unit is marked off, this segment will pass through the antipodal point and make an appearance on the opposite side of the sphere, approximately 0.86 units ($4 - \pi$) away from the antipodal point. Since the distance is to be measured from the starting point of the segment in the shortest manner possible, the distance corresponding to the number 4 will be $2\pi - 4$, or approximately 2.28. Similarly, the number 5 corresponds to $2\pi - 5$, 6 corresponds to $2\pi - 6$, 7 corresponds to $7 - 2\pi$, etc.

This mapping from the positive integers to these points can be written thus:

(17) $f : k \longmapsto |k - 2n\pi|$

where the value of n is chosen such that $f(k) < \pi$ for all k. Multiples of 2π are chosen since all distances are measured from the starting point of 0, or alternatively, $2n\pi$. π itself does not correspond to 0, but instead to the maximum distance on the sphere.

<u>ADDITIVE INVERSES OF THE POSITIVES</u>

In defining the distance corresponding to negative numbers the formal definition of negatives should be kept in mind in order to retain a consistent system. Negative integers are properly considered to be the additive inverses of the positive integers. In this sense, the point that should be called -1 must meet the following criterion: Its distance added to the unit distance must result in a total distance of 0. If the distances are considered to be directed, then the meaning of this is immediately clear. The distance corresponding to -1 will be simply a unit distance in the opposite direction of the original unit distance.

COMPLEMENTS OF THE INTEGERS

There remains a distance, constructible with the procedures we have discussed thus far, that does not fall into either of the above categories of positive integers or negative integers. It has been shown how to find the antipodal points of any two given points. The distance from one point to its antipodal point is, by definition of the metric, M. This is an important length to have in the set of constructible numbers in this space. It does not exist in the Euclidean set.

With M in the set, a whole new set of numbers becomes possible. First the number M is constructed. Then, from that point, a distance corresponding to any of the integers can be constructed. Suppose 1 is so constructed. Then, by this procedure, the length $M - 1$ has been constructed. This length does not correspond to any of the integers since that would imply that $M - 1 = 2nM - k$ (or $k - 2nM$) for some integer k. Simplifying would give $(2n - 1)M = k - 1$ for some integer k implying that M is an integer. This contradiction forces the conclusion that starting with M instead of with 0, an infinite number of lengths may be constructed, each one corresponding to one of the previous lengths $\pm M$.

This mapping from the positive integers to these points can be written thus:

(18) $g : k \mapsto | k - (2n + 1)M|$

where the value of n is chosen such that $g(k) < M$ for all

k. Odd multiples of π are chosen since all distances are measured from the starting point of 0, or alternatively, 2nπ and an initial distance of π is added to each one.

THE INTEGERS MOD PI

The numbers thus far constructed can be put into a correspondence with the set of integers mod π. Recall from number theory that the resultant of any integer mod k is a number greater than or equal to 0 and less than the modulus k. This result is equal to the remainder obtained when the number is divided by the modulus. For example, 10 = 3 mod 7 because 10 = 1*7 + 3. For negative numbers, the definition is the same but emphasis is placed on the range allowed for the resultant. Hence, -10 = 4 mod 7 because -10 = -2*7 + 4.

When the modulus is not an integer, the same definition is still used, but the resultant will not always be an integer. For example, 5 = 1/2 mod (3/2) since 5 = 3*(3/2) + 1/2. Note that the quotient resulting from the division is always an integer as, in this example, it is 3.

In the case of mod π, though, an interesting result arises. In each of the above examples, two rational numbers were equivalent modulo another rational number. When π is used as the modulus, there will be no rational numbers congruent to each other. If there were two rational numbers, say, p and q such that p = q mod π, then p = k*π + q where k is an integer. This would imply that there exist rational solutions to the equation π*x + y = 0.

This is known to be false because π is not an algebraic number. Since no rational numbers are congruent mod π, there exists a unique real number in $[0, \pi)$ corresponding to each rational number.

To define the correspondence between the numbers constructed above and the set of integers mod π, it serves to use the two definitions given above for the mappings f and g. Given any integer k, by definition of modulo π, k mod π is equal to f(k) if there exists an integer n such that $0 \leq k - 2n\pi < \pi$. If there is no such n, then k mod π is equal to g(k). As an example, 4 mod π = 4 - π = g(4), while 3 mod π = 3 + 0 * π = f(3)

Note that for f(k) negative numbers were defined to have the same image as their positive counterparts; the only difference was that the negative numbers were considered to be directed opposite from the positive numbers. This definition arose from the desire for the sum of a number k and the additive inverse of k to be 0. For the integers mod π, a negative number is defined by the modulus operator and is not the number that will add to k to give the point associated with 0. Instead, -k will add to k to give π, which, mod π, is 0. Geometrically then, the concept of additive inverse has been changed to include the antipodal point of the starting point as the additive identity as well.

Since these numbers are formed using addition of segments it is not unreasonable to ask if they form some sort of algebraic structure with the operation of addition. Since each of these numbers is a real number, properties of commutativity and associativity are inherited from the real number system. Further, a single point has length 0, hence the additive identity is an element of the integers mod M. Inverses exist in accordance with the definition of negative numbers given above.

The only remaining requirement is closure. Since $K + nM + L + mM = (K + L) + (n + m)M$ is in the proper form for the set for any two integers K and L. It might seem to be an obvious point. However, suppose K and L are both 2. Then n and m are both 0, and $(K + L) + (n + m)M = 4 > M$ and hence is not in the set. This problem can be remedied easily, though, by redefining the addition operation to be addition modulo M. Then the result of every addition will be less than M and the set will be closed under the operation.

These properties classify this set as an abelian (meaning commutative) group under the operation of addition modulo M. This is an important result of the work thus far. The final set of constructible numbers has this set as a subset. Hence, whatever set it may be, it will have a subset that is an abelian group under addition modulo M.

It will become convenient to develop a notation for a number k mod M. In some textbooks, a line is placed over the number to indicate when it is in a congruence class modulo another number. For the purposes of this paper, it will be simpler to use square brackets around a number for this purpose. For example, 5 mod $\mathcal{M}$ shall be denoted [5]. Whenever any other modulus is used, it will be stated explicitly to avoid confusion. A few equations exist where square brackets are used as grouping symbols, but the context is always explicitly clear.

Also, the notation P will be used to denote the entire set of integers modulo $\mathcal{M}$. This name will allow simplification of some of the ensuing text. Hence, P denotes the integers mod $\mathcal{M}$, and [k] is any element of P.

These numbers discussed thus far have been restricted solely to the line through x_0 and x_1. No numbers have yet been constructed through lines other than this, although since the lines themselves are easily constructible, the points would also be easily constructible. The only constructions used thus far have been the circle drawn with center at x_n and having unit radius and the construction of the antipodal point to x_0.

Chapter V
More Constructions

TRANSLATING A DISTANCE

The second proposition in the Elements is:

(II) To place at a given point (as an extremity) a straight line equal to a given straight line.

The demonstration of this is more complex than the demonstration of the previous constructions. Let A be the given point, and BC the given straight line. First, construct the line AB joining point A to the point B. On this segment, construct an equilateral triangle according to the previous propostion. Call the third vertex of the triangle D. Then construct the circle with center B and radius BC. If the segment DB is extended, it will intersect this circle. This length is equivalent to AB + BC. If the circle is drawn with this radius and center at D, then if DA is extended to meet this circle at, say, E, then the segment AE is the desired segment with endpoint A and length BC.

This construction has several points in it where it might not be possible in double elliptic geometry. First, the radius of the circle with center at D is the sum of two other lengths. If the two lengths AB and BC add together to be more than 2q, there will be no circle with the desired radius. Second, the segment DA must intersect the circle when extended. In this geometry it is possible for a line to lie entirely in the interior of a circle. Hence, in order to take these exceptions into account, this construction needs to be taken apart and examined step by step for either possible flaws or alternate routes.

If the two lengths do add together to more than 2q, this construction can still be performed with a slightly different technique. First, according to (A) find the point A', antipodal to A. Then extend the segment BC completely around and locate B' antipodal to B, using construction (A) again. There are several cases to consider.

1) $AB > q$; $BC < q$	4) $AB > q$; $BC = q$
2) $AB > q$; $BC > q$	5) $AB = q$; $BC = q$
3) $AB < q$; $BC > q$	6) $AB = q$; $BC > q$

In cases 1) and 4) the length of the segment A'B will be less than q. In fact, the two lengths A'B and BC will now add together to less than the maximum distance, and hence, there is no problem performing the construction with A' instead of A. After the point D' with A'D' = BC is found, then the point D antipodal to D' will be the point such that the segment AD has length equal to BC and has A as an extremity.

In cases 3) and 6) the length of segment B'C will be less than q and hence this segment may easily be copied onto the point A'. Note that it is A' and not A. This is necessary for two reasons. First, A is at most q away from B, so it is at least q from B'. Hence, the situation would be similar to cases 1) and 4) which require that the antipodal point A' be used. Second, by constructing the complementary distance on A', the point D' that is found will actually be the point D which is required. Thus it is eventually better to use the antipodal point A' for the intermediate construction.

In case 2), however, the point A' is farther than q away from B', since A is farther away than q from B. In this case, A is closer to B' and hence the distance B'C can be constructed on A. It is desired, though, to have the distance BC. If the segment D'A is extended around to the antipodal point of D, then the desired length will be the complement of D'A or simply DA.

The final case, 5), is the simplest. In this case, both distances are equal to q, so the line through A and B will have the desired distance necessary. Hence, construct AB and the question is answered.

The importance of this construction stems from the fact that by using this construction as a tool, it is possible to translate any length as if the compass being used were not collapsible. This tool shortens many of the constructions in the plane.

<u>BISECTING AN ANGLE</u>

In the <u>Elements</u>, Euclid's proposition 9 describes how to bisect any given angle. To do this, he chooses a point at random on one of the rays of the angle, draws the arc with that as radius, and then constructs an equilateral triangle using the endpoints of that arc as the endpoints of the segment serving as base. Although a point is chosen at random, this point can be made unit distance and then the construction follows straight through without any problems.

In spherical geometry, this construction carries over exactly. Since the unit length is part of the constructible lengths, a segment corresponding to that distance can be marked off on any ray by the procedure for translating a distance described above. Then, the arc can be drawn and the equilateral triangle constructed. This will always be possible since the unit length is less than 4q/3.

BISECTING A SEGMENT

In order to perform the construction to find an antipodal point, it was necessary to bisect the diameter of a circle. This is possible whenever the circles intersect by using equilateral triangles. Euclid places this construction at proposition 10 of the Elements, but used a slightly different method. It turns out that a slight modification of his method is applicable to this geometry and provides a means for bisecting a segment of any length.

Euclid begins by constructing an equilateral triangle on his given segment. He then proceeds to bisect the upper angle of the triangle, creating a ray which will bisect the given segment. This construction does not transfer to this geometry exactly. Instead, those certain segments that cannot be sides of equilateral triangles would not allow for bisection according to this method.

A careful inspection will reveal, though, that in order for the ray to bisect the segment all that is required is for the triangle to be isosceles. This type of triangle is easily constructed on any segment. At each endpoint, the line segment can be extended to create a straight angle. Bisect these angles using the method described above. Since two lines will always meet on a sphere, these two may be extended until they meet. Bisect the angle they form. This ray will bisect the given segment.

Chapter VI

More Numbers

There are now several constructions that can be performed on a sphere. Combining these constructions in a variety of ways produces a variety of segments, each with its own length. Since these lengths were all derived in a common way, they will always have certain common characteristics. These unifying factors will allow the set of constructible numbers to be characterized completely.

<u>POWERS OF 1/2</u>

The construction of bisecting a segment, when performed upon a segment of length 1, produces a segment with a new length, in this case 1/2. Since the bisection may be performed upon any segment of length [k], a new segment with a new length will be found. In general, though, it cannot be said that the new segment will be of length [k/2], but only of length [k]/2.

As an example, a segment of length [4] is actually of length $4 - \pi$. Hence, a segment of length [4]/2 will be of length $2 - \pi/2$. This number is not an element of P or else π would be a rational number.

Instead, the length [k]/2 found by bisecting [k] will be equal to $(k + n\pi)/2$ or $k/2 + \pi(n/2)$. If k and n are even, this is an element of P. If k is even and n is odd, this length is equivalent to $k/2 + \pi(n + 1)/2 - \pi/2$. If k is odd and n is even, the length is equivalent to $(k + 1)/2 - 1/2 + \pi(n/2)$. Finally, if k and n are both odd, then the length is equivalent to $(k + 1)/2 + \pi(n + 1)/2 - 1/2 - \pi/2$.

In each case listed above, the length is equivalent to an element of P less either 1/2, $\pi/2$, or both. Since 1/2 is constructible by bisecting 1, and $\pi/2$ is constructible by bisecting π, the entire class of numbers formed by bisecting a segment is equivalent to the class of numbers formed by the union of the P with 1/2 and $\pi/2$ and then extending it so as to be closed under the operation of addition.

After a segment is bisected, it can be bisected yet another time. In fact, not only may any finite number of bisections be performed on a segment, but any combination of bisections and additions may be performed. In other words, if this set is called P', then

$$P' = \{[k] + \sum_{i=1}^{K} n_i (1/2)^i + \pi \sum_{i=1}^{L} m_i (1/2)^i \mid [k] \in P\}$$

where K, L, all n, and all m are finite, and all the elements of the set are understood to be modulo π in all cases. If all n and all m are 0 then the number is just [k]. Hence, P is a subset of P'.

P' is closed under the operations of addition and bisection. Further, since P' is a subset of the real numbers, it inherits sufficient properties to form an abelian group under the operation of addition. Since bisection is an unary operation, P' cannot form a group or any similar structure under bisection.

LAWS OF SINES AND COSINES

In any triangle in spherical geometry, there are formulas relating the lengths of sides and the measures of the angles. The first one, known as the spherical Law of Sines is:

$$(19) \qquad \frac{\sin(a)}{\sin(A)} = \frac{\sin(b)}{\sin(B)} = \frac{\sin(c)}{\sin(C)}$$

where a, b, and c are the lengths of the sides and A, B, and C are the measures of the opposite angles. The second is the spherical Law of Cosines:

$$(20) \qquad \cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A)$$

When one of the angles involved is a right angle, the Law of Cosines simplifies considerably. Since the $\cos(\pi/2)$ is 0, the second term on the right side of (20) is 0. Hence, the formula becomes:

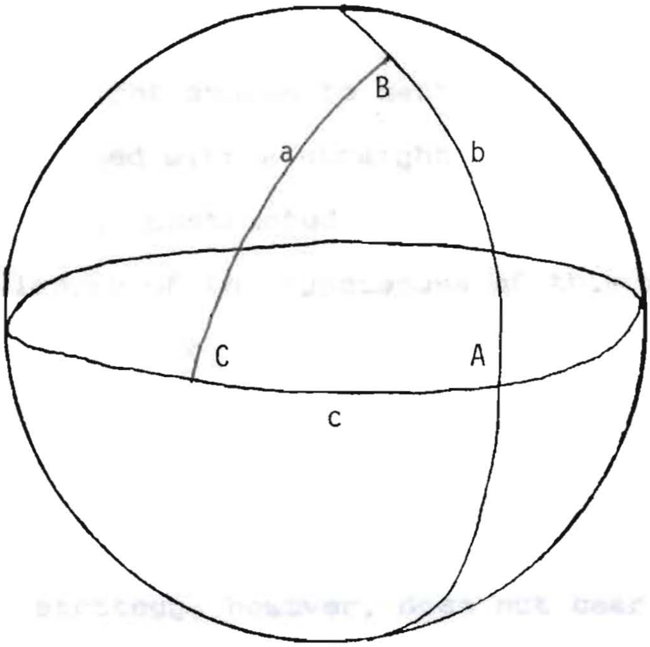
$$(21) \qquad \cos(a) = \cos(b)\cos(c)$$

where a is the length of the side opposite the right angle.

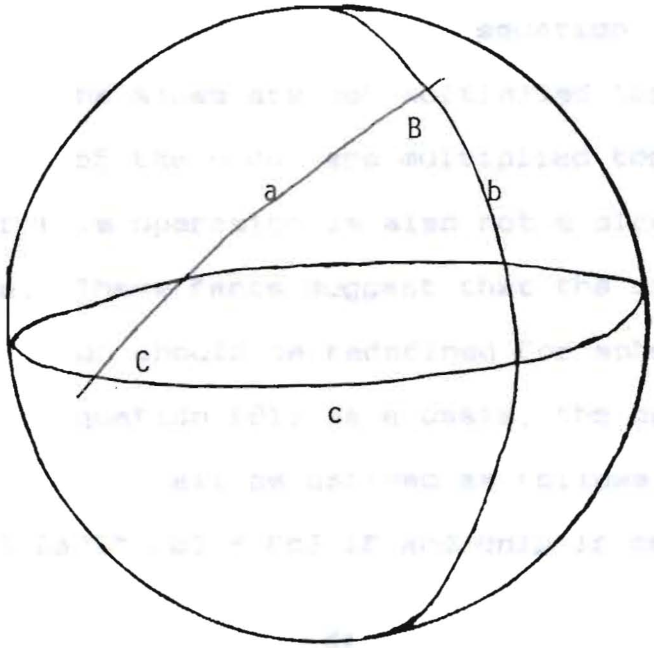
The usefulness of these formulas stems from the connection they give between the lengths of the sides of the triangle. Whenever a triangle is constructed, at least two of the sides and one angle, or two of the angles and one side, will be known. These trigonometric formulas will then provide the lengths of the remaining sides.

COSINE MULTIPLICATION

Since right angles are easily constructed using the method for bisecting a segment, right triangles are among



(19)



(21)

the simplest triangles to construct. Given two segments of any length constructible thus far, the segments can be adjoined at right angles to each other, the remaining endpoints joined with a straight line and a right triangle will be thusly constructed.

The length of the hypotenuse of this triangle will be given in accordance with (21). By taking the Arccosine of both sides of (21), the equation can be solved for a. That is,

(22) $a = \operatorname{Arccos}[\cos(b)\cos(c)]$

This strategy, however, does not bear much fruit. The inverse trigonometric functions are esoteric creatures. Identities simplifying them are few and sparsely applicable. Hence, in order to reach any conclusions about the length of the hypotenuse, it will prove best to use (21) instead of (22).

The most obvious fact about equation (21) is that the lengths of the sides are not multiplied together. Instead, the cosines of the sides are multiplied together. The result of this operation is also not a side, but a cosine of a side. These facts suggest that the operation of multiplication should be redefined for spherical geometry.

Using equation (21) as a basis, the operation of multiplication shall be defined as follows:

(23) $[a] * [b] = [c]$ if and only if $\cos(a)\cos(b) = \cos(c)$

Because the operation is defined in terms of multiplication of real numbers, the properties of this operation are similar to the properties of the original operation of multiplication. First, the operation is associative. Given a, b, and c as lengths on the sphere, a true equation involving their cosines can be written in the real number system with the usual operation of multiplication. That is:

(24) cos(a)[cos(b)cos(c)] = [cos(a)cos(b)]cos(c)

Which leads directly to:

(26) [a] * ([b] * [c]) = ([a] * [b]) * [c]

By a similar argument the operation is commutative.

In order to demonstrate that an identity exists, it needs to be shown that there exists [e] such that for any number [a]:

(27) [a] * [e] = [a]

In the real number system, the multiplicative identity is 1. Hence, if it is desired that cos(a)cos(e) = cos(a), cos(e) will have to be 1. Since the Arccosine of 1 is 0, 0 would be the first logical choice for the multiplicative identity of this new operation. Testing this hypothesis, the expected result occurs. That is, [a] * [0] = [a] is equivalent to cos(a) * 1 = cos(a) which is a true statement for any real number a.

Once the identity is established, the question can be presented of whether or not the set contains inverses for the operation. By definition, an inverse of a number [a] is another number [b] such that [a] * [b] is equal to the identity, in this case [0].

Geometrically, the question is whether a right triangle can be constructed with one side [a] and a hypotenuse of [0]. In this light, it is immediately obvious that no triangle can be constructed with hypotenuse [0]. However, since π is also equivalent to 0 in this system, it remains to be shown that no hypotenuse would ever have a length of π. This, too, is true since any triangle in spherical geometry with a side of length π has, as two of its points, two antipodal points which implies that all three points are collinear. Since collinearity is not allowed of a triangle's vertices, this contradicts the definition of a triangle.

Algebraically, the result is perhaps more easily stated. It is desired for each number [a] to find a number [b] such that cos(a)cos(b) = 1 or cos(b) = 1/cos(a). Written this way, it is obvious that since cos(a) is ≤ 1 for all $a \neq 2k\pi$, $1/\cos(a) \geq 1$. This implies that $\cos(b) \geq 1$, which would be contrary to the range of the cosine function.

COSINE DIVISION

Although inverses do not exist in general, there does exist a division operation of a sort. If [a] < [b] < π − [a], then a right triangle can be constructed with a side of length [a] and a hypotenuse of length [b]. The third side can then be found as:

(28) c = Arccos[cos(b)/cos(a)]

A special case of this occurs in the equilateral triangle used to bisect a segment. Here, the hypotenuse has length

1 and the side has length 1/2.

In a manner similar to the redefinition of multiplication then, division shall also need to be redefined. Using (28) as a model:

(29) [a]\[b] = [c] if and only if cos(c) = cos(a)/cos(b)

where the backwards slash is used to differentiate this division from normal division. Since the cosine of c is defined to be within [-1, 1], [a] and [b] will need to be within the desired range in order for the operation to be defined.

Properties of this division are similar to the properties for real number division. First, whenever this operation is defined, it is an inverse operation for the multiplication operation defined above. Second, it is neither commutative nor associative as would be expected from a division operation.

More importantly, though, this division allows for cancellation in an equation using cosine multiplication. If both sides of an equation have a common factor, then this factor can be cancelled out as in ordinary multiplication and division. This property can be proven using equations for real numbers in a manner similar to the proof that cosine multiplication is associative.

Chapter VII
The Set T

The set of constructible numbers in double elliptic geometry shall be denoted by T. It has been shown that P and P' are subsets of T. It has also been shown that there exist other numbers not in P or P' that are in T. These numbers are found by using the operations of cosine multiplication and division defined in Chapter VI.

It can be proven that the only numbers in T are the ones described in the above chapters. That is, T is the set of all the integers modulo π extended so as to be closed under any finite combination of the operations of addition, bisection, and cosine multiplication and division.

In order to prove that the set T is closed under double elliptic constructions, there are four intermediary lemmas that are necessary:

Lemma 1 : The line through two distinct points of T is in T
Lemma 2 : The intersection of two lines of T is in T
Lemma 3 : The intersection of two circles of T is in T
Lemma 4 : The intersection of a line and a circle is in T

In order to prove that the set T is closed under double elliptic constructions, the general equation of a line would need to be developed. Then, the general equation of a circle would need to be developed. Then the final steps of the proof would be to prove each of the above lemmas using the general equations for a line and a circle and the algebraic properties of the set T.

In order to prove the above lemmas, it is necessary to coordinatize the system. This can be done by placing the center of the sphere at the origin of a three dimensional space and placing a metric on the space by defining the radius of the sphere to be 1 as was done previously. An origin, O, shall be defined on the sphere, and from here two coordinates shall be associated with every point on the sphere. Given a point P on the sphere, the first coordinate, p, shall be the distance from P to O. The second coordinate, θ shall be the angle that the line PO makes with a fixed line, to be known as the θ-axis.

In an attempt to simplify much of the ensuing equations, the notation p' shall be used to mean $\pi/2 - p$ for any length p.

EQUATIONS OF LINES

For any line on the sphere, it is desired to show that the equation of the line takes on a certain form in this coordinate system. Choosing any arbitrary line, l, the line will make an angle A with the 'equator', i.e. the polar of the origin. Choosing an arbitrary point x on l, the coordinates of the point will be (p1, θ1).

If a line, say m, is constructed through x and O, it will make a right angle with the equator, since all lines through a point make a right angle with the polar of the point. The line l, this line m, and the equator, say e, will form a right triangle on the sphere. The length of the segment on m is p1'. The length of the segment on e, is a certain amount to be called b. Since for any arbitrary point on the line, one characteristic that will remain constant is the angle the line makes with the equator, an identity relating this to the two legs of the right triangle should be found. This identity is:

(30) tan(A) = tan(pl')/sin(b)

Then, since cot(k) = tan(k'), this becomes:

(31) tan(A) = cot(pl)/sin(b)

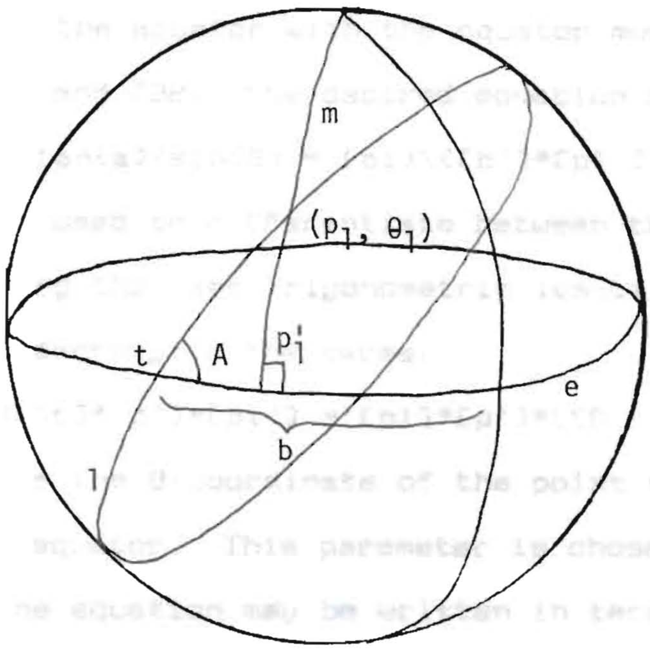
Since it is desired to write this equation entirely in terms of constructible lengths, it remains necessary to show that the length b is constructible. By definition, b is the length from the point where l and e intersect to the point where m and e intersect. If it is assumed that each of these lines is constructible, then geometrically each of the points would be constructible as well. Therefore, the length b is constructible.

One more change to (31) is desired, though. This change uses the identity sin(θ) = cos(θ'). Using this and the definition of tangent to make a substitution in (31):

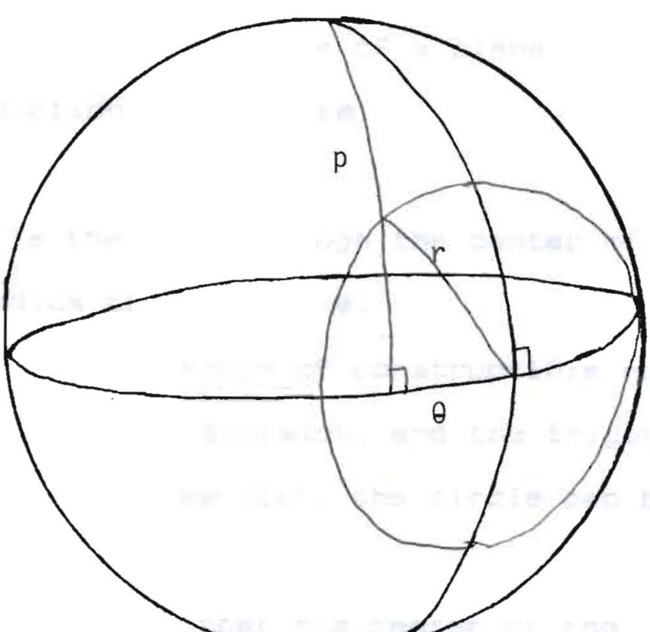
(32) tan(A) = cos(pl)/[cos(b')cos(pl')]

This change is an improvement since the operations are defined in terms of cosines alone.

From here, the equation of the line can be developed. Equation (32) relates the coordinates of a point on a line to the characteristic that is always true for every point on the line: tan(A) is a constant. For any point Q with coordinates (p, θ) on the line, the tangent of the angle formed by the line through Q and the point t where the line



(30)



(36)

intercepts the equator with the equator must be a constant. Using (30) and (32), the desired equation becomes:

(33) tan(a)/sin(B) = [p1]\([b']*[p1'])

where B is used to differentiate between the b used in (32). Using the same trigonometric identities used above and then rearranging the terms:

(34) [p]*[b']*[p1'] = [p1]*[p']*[(θ - θt)']

where θt is the θ-coordinate of the point t where the line meets the equator. This parameter is chosen in place of B so that the equation may be written in terms of p and θ.

EQUATIONS OF CIRCLES

Similar to the development of the equation of a line on the sphere, it is desired to show that for any circle on the sphere, the equation of the circle takes on a specific form. It was shown in Chapter II that a circle on the sphere is the intersection of a plane with the sphere, and has an equation of the form:

(35) x' = cos(p)

where x' is the axis through the center of the circle and p is the radius of the circle.

With the new tools of constructible numbers, cosine multiplication and division, and the trigonometric identities listed earlier, the circle can be put in a new perspective.

First, assume that the center of the circle is at the point (π/2, 0). With this center, the coordinates of an arbitrary point can be shown to form a right triangle with the radius to that point as the hypotenuse. If the coordinates of the arbitrary point on the circle are (p, θ) and the circle has radius r, then the right triangle will satisfy:

(36) $\cos(r) = \cos(\theta)\cos(p')$

When the center is not on the θ-axis, the result easily generalizes by a transformation so that the equation becomes:

(37) $\cos(r) = \cos(\theta - \theta 0)\cos(p')$

where θ0 is the θ-coordinate of the center. If this equation is put in terms of the set T, the result is:

(38) $[r] = [\theta - \theta 0]*[p']$

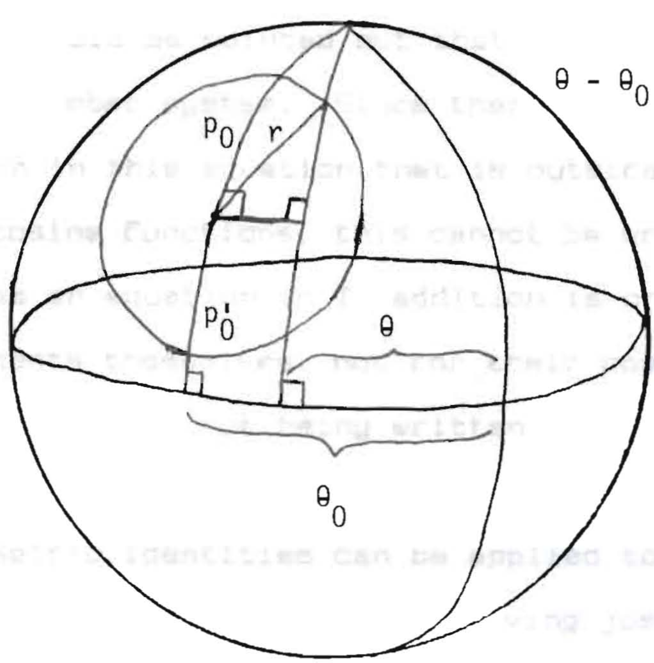
Generalizing to an arbitrary p-coordinate for the center is not as straightforward. There still exists a right triangle with the radius as hypotenuse, but in this case, the leg in the p direction will have length p - p0, where p0 is the p-coordinate of the center of the circle.

The leg in the θ-direction can be found best by the law of cosines. Here the leg is opposite an angle of measure θ - θ0, and the other two sides each have measure p0. Hence, the length of the leg can be found using the equation:

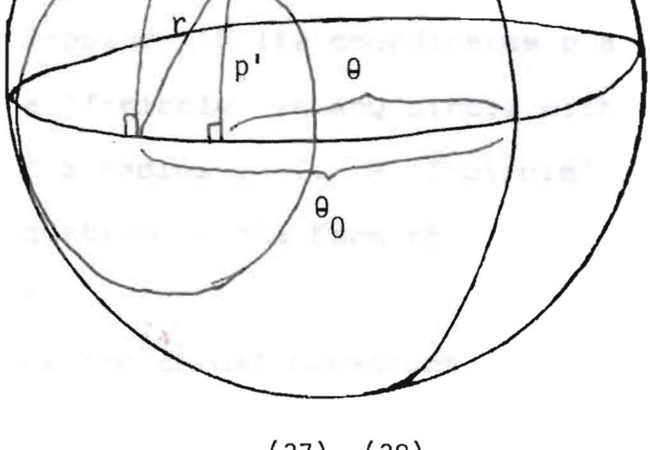
(39) $\cos(\theta k) = \cos^2(p0) + \sin^2(p0)\cos(\theta - \theta 0)$

Using this to write the equation for the hypotenuse of the triangle, the result is:

(40) $\cos(r) = \cos(p - p0)(\cos^2(p0) + \sin^2(p0)\cos(\theta - \theta 0))$



(40)



(37), (38)

It should be pointed out that this is an equation in the real number system. Since there is an addition operation in this equation that is outside of the arguments of the cosine functions, this cannot be written as it stands as an equation in T; addition is only defined for the segments themselves, not for their cosines.

The equation not being written in terms of T does not, however, imply that the points are not constructible. Trigonometric identities can be applied to the above equation to put it in a form involving just multiplication of cosines of constructible lengths. This is not done because it would hinder, rather than help, the ensuing discussions.

T IS CLOSED UNDER DOUBLE ELLIPTIC CONSTRUCTIONS

In proving this assertion, the following definitions shall be necessary: A constructible line shall be defined to be any line whose equation in the form of (34) has only factors in the set T. This line will be called a 'T-line' or 'in T'. Similarly, a constructible point can be defined to be a 'T-point' if its coordinates p and θ are in T. Finally, a 'T-circle' is any circle with a 'T-point' as center and a radius in T. A 'T-circle' has been shown to have an equation in the form of (40) where each factor is again in T.

Before the actual construction lemmas are proven, it will prove beneficial to prove a quick algebraic lemma about this set T.

Lemma O : Given $\emptyset \in T$, the length θ such that $\tan(\theta) = \cos(\emptyset)$ is also an element of T.

First choose any two point θ and $\emptyset$ such that $\tan(\theta) = \cos(\emptyset)$ and both θ and $\emptyset$ are elements of T. $\emptyset = \pi/2$ and $\theta = 0$ will suffice. Starting with the given equation, trigonometric identities can be used to demonstrate that:

$$(41)\quad \frac{\sin^2(\emptyset)}{1 + \cos^2(\emptyset)} = \cos(\theta)$$

Then, since T is closed under cosine multiplication, there exists a and b such that (41) can be written as:

$$(42)\quad \frac{\cos(a)}{1 + \cos(b)} = \cos(\theta)$$

Using trigonometric identities and the closure of T once more, the denominator of this can be changed to 2cos(c) for some $c \in T$. Using this equation to solve for 1/2, the result would be:

$$(43)\quad 1/2 = \cos(\theta)\cos(c)/\cos(a)$$

Proceeding to the actual points $\theta 1$ and $\emptyset 1$ that satisfy the hypothesis of the lemma, a similar derivation to the one above may be performed. If this derivation is halted immediately before (43), the result would be:

$$(44)\quad \cos(a1)/2\cos(c1) = \cos(\theta 1)$$

Then, using (43) to make a substitution for 1/2, the equation would become:

$$(45)\quad \frac{\cos(a1)\cos(\theta)\cos(c)}{\cos(a)\cos(c1)} = \cos(\theta)$$

Thus, $\cos(\theta)$ can be written as a finite combination of cosine multiplication and division involving only elements of T. Therefore, $\theta \in T$.

Lemma 1 : The line through two distinct points of T is in T

Use r and s to denote the two points of T. Since r and s are T-points, their coordinates (pr, θr) and (ps, θs) are in T. By extending the line segment through these two to create the desired line, l, l either is the equator or meets the equator at some point t. If l is the equator the proof is finished. From here, then, assume that it is not.

Using the points r and s, the tangent of the angle A formed by l and the equator can be written as:

(46) $\tan(A) = \cot(pr)/\sin(\theta r - \theta t)$

and

(47) $\tan(A) = \cot(ps)/\sin(\theta s - \theta t)$

where t is the point where the line l meets the equator as in the above derivation.

For any arbitrary point on the line, say x, the equation of the tangent of A will be similar to (46) and (47) except that the coordinates of x will be used in place of the coordinates of r (or s).

Combining the equation for x and (46) so as to eliminate A, and then multiplying to elimate all fractions, the equation of the line could be written as:

$$(48)\quad [px]*[(\theta r - \theta t)']*[pr'] = [pr]*[px']*[(\theta x - \theta t)']$$

This equation is in the form of (34) and therefore represents a T-line.

Lemma 2 : The intersection of two lines of T is in T

Given two arbitrary T-lines, they must intersect since all lines on the sphere intersect. If the equations of the lines are given, then they can be solved as simultaneous equations for the point(s) of intersection.

Assume the equations are as follows:

$$(49)\quad \begin{aligned} &[p]*[b1']*[p1'] = [p1]*[p']*[(\theta - \theta 1)'] \\ &[p]*[b2']*[p2'] = [p2]*[p']*[(\theta - \theta 2)'] \end{aligned}$$

Solving these for p and θ requires a few steps. First, the two equations should be multiplied together so that the left side of the first is taken times the right side of the second and vice versa. Note that this produces a factor of [p] and one of [p'] on each side. Using the cancellation property of cosine division, these factors can be removed. From this point, the equation should be put in terms of real numbers and cosines in order to facilitate the work. Terms of the equation are then rearranged making use of the identity for the sine of a sum until finally the equation reads:

$$(50)\quad \tan(\theta) = \frac{\tan(B)\sin(\theta 2) - \tan(A)\sin(\theta 1)}{\tan(B)\cos(\theta 2) - \tan(A)\cos(\theta 2)}$$

Here, A and B are the angle made by the first and the second line respectively with the equator.

At first, it may seem that this method of approach is a digression from the usual method of reducing these equations. Though this may be true, it is a necessary digression in order to properly isolate θ and p and then demonstrate that they are constructible lengths.

The terms of (50) are each a tangent multiplied by either a sine or a cosine. If complementation of the arguments of the cosines are used, then it can be said that each is a tangent times a sine. Recalling the identity at (30) and the fact that θ1 is a side adjacent to A, it can be shown that there exists a length Ø1 such that tan(A)sin(θ1) = tan(Ø1). This length is the side opposite A in a triangle consisting of the angle A, the side θ1, and a right angle at the other end of side θ1. A similar argument presents the existence of Ø2, and of Y1 and Y2.

Using these (50) becomes:

$$(51)\quad \tan(\theta) = \frac{\tan(\emptyset 2) - \tan(\emptyset 1)}{\tan(Y2) - \tan(Y1)}$$

Writing these in terms of sines and cosines and then simplifying the complex fraction using a trigonometric identity:

$$(52)\quad \tan(\theta) = \frac{\sin(\emptyset 2 + \emptyset 1)\cos(Y2)\cos(Y1)}{\cos(\emptyset 2)\cos(\emptyset 1)\sin(Y2 + Y1)}$$

Since T is closed under cosine multiplication and addition, there exist sufficient elements of T to simplify this equation. Let Y3 be the element of T such that [Y1]*[Y2] = [Y3]. Let Y4 be the element Y2 + Y1 - π/2.

Defining Ø3 and Ø4 similarly, (52) can be written as:

$$(53)\quad \tan(\theta) = \frac{\cos(\text{Ø}4)\cos(Y3)}{\cos(\text{Ø}3)\cos(Y4)} = \frac{\cos(\text{Ø})}{\cos(Y)}$$

for some constructible lengths Ø and Y.

There are two cases to consider here. First, if Ø ≥ Y then cosine division is defined for these two numbers and it can then be said that there exists C such that tan(θ) = cos(C). If θ < Y then cosine division is defined for the reciprocal of the fraction on the right and there exists C in T such that cot(θ) = tan(θ') = cos(C). In either case, θ or θ' will be in T, which immediatiely implies that θ is in T.

To demonstrate that p is also in T, the result that θ is in T can be substitued back into either of the equations at (49). From here, using the closure of T, this can be written as:

(54) [p]*[a] = [b]*[p']

Here as in (53) [a] ≥ [b] or [b] < [a] so that cosine division is defined in at least one direction. Hence, this can be put in the form for Lemma 0, whence p ∈ T.

Lemma 3 : The intersection of two circles of T is in T

Two arbitrarily chosen T-circles will have an equation in accordance with (40) as follows:

$$\cos(r1) = \cos(p - p1)(\cos^2(p1) + \sin^2(p1)\cos(\theta - \theta1))$$

(56)

$$\cos(r2) = \cos(p - p2)(\cos^2(p2) + \sin^2(p2)\cos(\theta - \theta2))$$

Attempting to solve these as systems of simultaneous equations leads almost immediately to impossibly tedious manipulations that don't actually lead anywhere very quickly. Hence, another approach is necessary.

It is at this point that the planes discussed in Chapter III will prove to be the most useful. Recall that in constructing equilateral triangles on the sphere, the points where two circles intersected were found. This process made use of planes in the three dimensional rectilinear space that the sphere is imbedded in. The equations of the planes were written in terms of the distance k between their centers, and due to the nature of the construction involved, each circle passed through the center of the other.

Not all of these factors are present in the current situation, i.e. each circle does not necessarily pass through the center of the other. In fact, they may only intersect in one point. Nevertheless, the equations for the planes can be written in terms of the radii, r1 and r2, of the two circles.

First, a coordinate system needs to be established. In a fashion similar to the derivation in Chapter III, the x-axis shall be chosen so as to pass through the center of the first circle. Then, the y-axis will be chosen so that the center of the second circle is in the xy-plane, the equation of the first plane will be:

(57) $x = \cos(r1)$

exactly as in Chapter III.

Through a procedure similar to the one in Chapter III, the equation for the other circle can be written and then a substitution made, and the expression simplified so as to solve for y. When this is done, the x and y coordinates of the point(s) of intersection are found to be:

$$x = \cos(r1)$$

(58)

$$y = \frac{\cos(r2) - \cos(k)\cos(r1)}{\sin(k)}$$

In order to translate these results back into the sphere, the transformation equations for spherical coordinates are used. Hence, in terms of p and θ coordinates with the z-axis passing through the origin,

$$\cos(r1) = \sin(pz)\cos(\theta z)$$

(59)

$$\frac{\cos(r2) - \cos(k)\cos(r1)}{\sin(k)} = \sin(pz)\sin(\theta z)$$

These equations can then be solved manipulated to give:

$$\tan(\theta z) = \frac{\cos(r2) - \cos(k)\cos(r1)}{\sin(k)\cos(r1)}$$

(60)

$$\sin(pz) = \cos(r1) / \cos(\theta z)$$

where the first equation implies that θz is in T and this, with the second equaion, implies that pz is in T.

The work is not yet finished, though. This pz and θz correspond to an origin through the z-axis (call this point z). It remains to demonstrate that this implies that the coordinates of the point of intersection (call it r) with respect to the original origin, O, are in T.

First, in order to show that the coordinates of the z are in T, the line through the centers of the circles should be drawn. Since the centers have coordinates in T, by the previous lemma, this line is in T. There exists a perpendicular to this line through O. Since this line is in T, the length of this segment to this line is in T. But this segment is the complement of the p-coordinate of z. Hence, this coordinate is in T.

To show that the θ-coordinate of z is in T, consider the triangle formed by z, O, and the point where the θ-axis crosses the polar of z (the line through the centers). Each of the lengths of these sides is in T, being the p-coordinate of z, π/2, and the p-coordinate of the intersection of two lines in T. Further, the angle at O is the θ-coordinate of z. Then, the law of cosines can be used to write an equation for this angle in terms of the three sides, which is sufficient for the angle to be in T.

Once it has been established that z is in T, the proof that the point of intersection of the circles is in T follows by a transformation of coordinates.

Construct the triangle with vertices at the point of intersection, O, and z. Using the law of cosines, the cosine of the p-coordinate of r can be written in terms of other lengths known to be in T. Specifically:

(61) $\cos(p) = \cos(pz)\cos(k) + \sin(pz)\sin(k)\cos(\theta k)$

where pz is given by (60), k is the p-coordinate of z, and

θk is the angle made by the line through z and the point of intersection and the line through z and O.

In a similar fashion, the triangle with vertices at z, the intersection of the θ-axis with the equator, and the intersection of the line through O and r with the equator can be drawn and the law of cosines applied to this triangle to show that the θ-coordinate is in T.

Lemma 4 : The intersection of a line and a circle is in T

Choosing any arbitrary line and circle in T, if they intersect, it is desired to show that the coordinates of the point of intersection are in T. This lemma is actually the most straight forward of all of the lemmas.

The first thing to be done is to reorient the way the problem is posed. Since any line is the set of points equidistant from the pole of the line, any equation of a line can be written in the form of an equation of a circle. Apply lemma 3, and the proof is complete.

These four lemmas together imply that T is closed under double elliptic constructions. Assume that x is any length constructible using a straight edge and compass on the sphere. Then x is the result of a finite number of those operations, and hence is the result of a finite number of the algebraic operations of the set T applied to the two points given to be a unit distance apart. Therefore, x is in T.

CHARACTERISTICS OF T

That the set T forms an algebraic structure is without doubt. There exists elements in the set, an operation on those elements, and axioms for that operation. Just what type of structure it is though, remains to be explained.

It was discussed in Chapter IV and then again in Chapter VI that because T is a subset of the real numbers closed under addition, enough of the properties of the real number system were inherited to allow the set to be an abelian group under addition. Since that time more numbers have been added by way of cosine multiplication and division. At each stage, though, the familiar property of closure under addition was always retained. Therefore, the set remains an abelian group under addition.

The operation of cosine multiplication was presented in Chapter VI as an alternative to normal multiplication. In Chapter VI the properties of this operation were discussed and it was pointed out that the operation was associative, commutative, and had 0 as an identity. The set is also closed under these operations. These properties are not of themselves sufficient to classify the set as a particular structure under multiplication, but when combined with the addition operation some interesting results occur.

First, a close examination of the addition operation and the multiplication operation reveal that both operations have the same identity. That is, not only $0 + x = x$, $\forall\, x \in T$, but also $[0]*[x] = [x]$, $\forall\, x \in T$. This fact prohibits T from being an integral domain or a field of any sort.

Second, the usual distributive law does not hold. That is, $[a]*([b] + [c]) \neq [a]*[b] + [a]*[c]$. There are forms of the distributive law for this system, and, in fact, many ways of distributing elements were used in the above proofs. Each of these laws is based upon a trigonometric identity of some sort. Some of the distributive laws are listed here:

$$[c]*[a + b] = [c]*[a]*[b] - [c]*[a']*[b']$$

$$[c]*[(a + b)'] = [c]*[a']*[b] + [c]*[a]*[b']$$

$$[a]*[c] + [a]*[d] = 2[a]*[(1/2)(c + d)]*[(1/2)(c - d)]$$

$$[a]*[c] - [a]*[d] = 2[a]*[((1/2)(c + d))']*[((1/2)(d - c))']$$

Other trigonometric identities could be chosen and other distributive identities could be derived.

Chapter VIII
Generalizations

<u>I VS E</u>

Where E denotes the set of constructible numbers in the Euclidean plane, the set I can be contrasted with E in hopes of producing some result that will better clarify the differences between the Euclidean plane and the double elliptic plane.

Besides the obvious difference that they are different numbers, other differences present themselves. First, the numbers in the double elliptic plane are bounded on the real number line. In Euclidean geometry, this is not true. This boundedness is due directly to the fact that straight lines are finite in extent in double elliptic geometry.

Second, the two sets are of different order types. The constructible numbers in the Euclidean plane satisfy the Archimedean postulate. This postulate states that for e and M any two positive integers, there exists a positive integer n such that ne > M. In the double elliptic plane, this postulate cannot even be stated properly without first defining the terms more explicitly. Is it required that a positive number of duplications of the segment e be a longer segment than the segment n? Or, instead is the multiplication operation to be the cosine multiplication defined above, in which case positive lengths may have negative cosines and other complications arise? In any event, because we are dealing with a modular system, order is not a necessary attribute of the set.

Third, a close comparison of the equations for lines and circles given above with the equations for lines and circles in the Euclidean plane reveals similar characteristics. An equation of a line in standard form in the plane has certain characteristics that speak to the points the line passes through. In the same way, the equations for the lines given above also speak to the points they pass through. The equation of a circle in the Euclidean plane has the radius and center as parameters. This is also true in the double elliptic plane as shown above.

Finally, the operation of modulo π can be expressed another way. If the length, say p, is to be evaluated mod π, then this can be done by evaluating the Arccos[cos(p)]. This identity points towards more fundamental truths about the number system. The set T has been the set in question from the beginning, but if the operations and properties of T can be expressed in terms of the cosines of the elements of T more easily than in terms of T, then perhaps a new set, cos(T), should be the one to be studied.

APPLICATIONS

At this point in time it is difficult to extrapolate to the applications of this set, T. The applications of the set E are mostly theoretical, i.e. E is used mostly to prove other theorems. Since few questions or theorems regarding double elliptic geometry are present today, there may well be very few applications.

One application of T that could very well prove to have far reaching implications is in the comparison of T with E. It has been known for some time that the number system and the properties associated with it are inherent in the geometric structure of the Euclidean plane. This is a fundamental fact used in the applications of E. However, by demonstrating the existence of a new number system associated with a differently structured plane, perhaps a deeper understanding of the connection between the system and the geometry will result. From this understanding, other tools, similar to E and its relationship to the plane may result.

It should be pointed out that, although the results proven above hold true on the sphere in general, whenever a specific number was mentioned as a numerical example, that example was possible due to the metric placed on the system. The unit of length could have been chosen to be π. This would have changed some of the results and left some others unchanged. Alternatively, the unit of length could have been chosen to be what is now an extremely short length. In this case, if the choice was made small enough, the geometry on the sphere would be a very close approximation to Euclidean geometry implying that the number system would be a close approximation to the rational numbers (if close approximation has any meaning in that sense).

OTHER GEOMETRIES

Almost all of the results in this paper that differed from the results in the Euclidean plane were based upon the identity for a right triangle involving the cosines of the sides. In the Euclidean plane, this identity takes on the form of the Pythagorean theorem, from whence the field E arises.

The natural question regarding hyperbolic geometry then is what sort of number system would arise using these same methods there. The formula in hyperbolic geometry analogous to the one used here is:

(62) $\cosh(c) = \cosh(a)\cosh(b)$

If multiplication is defined in terms of this equation, then a logical hypothesis is that the number system would carry properties in it based upon the hyperbolic cosine function in the same way that this number system carries properties based upon the cosine function. The demonstrations of these properties would follow from the geometric axioms in the same way that the properties in this paper followed from the geometric axioms of the double elliptic plane.

There remains single elliptic geometry to consider. This geometry is different from the double elliptic geometry in that antipodal points are identified so that two lines intersect in one and only point. Straight lines are finite, and many of the properties of this geometry are the same as in double elliptic geometry. The hypothesis for this geometry is that the number system of double elliptic geometry will carry over to single elliptic geometry with few modifications. The reason for this assertion is that the lengths in double elliptic geometry were each taken modulo π implying that the antipodal point of the origin also had coordinate 0. Hence, antipodal points were, in a manner of speaking, identified. There are sure to be some modifications since the maximum distance is $\pi/2$ instead of π, but the algebraic structure will most likely be very similar.

It is known that each of these geometries is a special case of projective geometry without any metric at all. If these concepts are extended to this geometry, perhaps some light will be shed upon the number systems in general.

Bibliography

[1] Adler, Claire Fisher. Modern Geometry. New York, NY: McGraw-Hill Book Co., 1958.

[2] Coolidge, Julian Lowell. The Elements of Non-Euclidean Geometry. London, England: Oxford University Press, 1927.

[3] Euclid. Elements. Trans. by Sir Thomas L. Heath. : Encyclopaedia Britannica, 1952.

[4] Gans, David. An Introduction to Non-Euclidean Geometry. New York, NY: Academic Press, 1973.

[5] Golos, Ellery B. Foundations of Euclidean and Non-Euclidean Geometry. New York, NY: Holt, Rinehart and Winston, Inc., 1968.

[6] Greenberg, Marvin Jay. Euclidean and Non-Euclidean Geometries. :W. H. Freeman and Company,

[7] Hillman, Abraham P. and Alexanderson, Gerald L. A First Undergraduate Course in Abstract Algebra. Belmont, CA: Wadsworth, Inc., 1985.

[8] Moise, Edwin E. Elementary Geometry from an Advanced Standpoint. Reading, MA: Addison-Wesley Publishing Co., 1963.

[9] Spiegel, Murray R. Mathematical Handbook. New York: McGraw-Hill Book Company, 1968

[10] Wolfe, Harold E. Non-Euclidean Geometry. New York, NY: Holt, Rinehart and Winston, 1945.