FINSLER MANIFOLDS

By

THOMAS EDISON BONNER

Bachelor of Science
Lincoln University
Jefferson City, Missouri
1943

Master of Arts
University of Illinois
Urbana, Illinois
1949

Master of Science
University of Illinois
Urbana, Illinois
1953

Submitted to
the faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the
requirements for the degree
of
DOCTOR OF EDUCATION
July, 1967

FINSLER MANIFOLDS

Thesis Approved:

N B Deal
Thesis Adviser

Vernon Troxel

Milton E Berg

D D Durham
Dean of the Graduate College

PREFACE

The aim of this dissertation is to present an expository account of the concept of the Finsler manifold.

As the idea of the n-dimensional differentiable manifold and the idea of the Finsler metric can be traced to an 1854 lecture by Riemann, the first chapter is given to Riemann's introduction of these concepts presented in modern terminology.

The second chapter is a dicussion of the concept of distance as developed by Paul Finsler in 1918. Like the first chapter, this discussion is concerned primarily with local geometry.

In the third chapter the theory of chapter II is restricted to two-dimensional manifolds, and a study of such geometric ideas as geodesics, orthogonality and area is developed, bringing the student to the threshold of the modern approach, begun by Busemann.

The fourth chapter is concerned with aspects of global geometry and culminates in a brief discussion concerning fibre bundles.

While Finsler geometry does have practical significance in modern physics, this aspect is not covered here.

On the part of the undergraduate mathematics major, for whom this exposition is written, basic notions of advanced calculus, linear algebra and point set topology are presumed.

Two remarks should be made concerning notations: First, the use of superscripts with the same meaning ordinarily attached to subscripts should be distinguishable from the use of superscripts as exponents by the context; second, any term in which repeated indices appear is to indicate a summation from 1 to n on the index or indices rereated. For example, $F_{x^i}X^i$ means $F_{x^1}X^1 + F_{x^2}X^2 + \dots + F_{x^n}X^n$, i.e., $\frac{\partial F}{\partial x^1} X^1 + \dots + \frac{\partial F}{\partial x^n} X^n$.

I wish to acknowledge my indebtedness to Professor R. B. Deal for his guidance in the preparation of this thesis and for the inspiration that led me to the study of Finsler geometry; I wish to acknowledge equal indebtedness to Professors Vernon Troxel and Milton Berg for their guidance in the formal presentation of the subject matter; and I wish to express my gratitude to Miss Mary Winstead, Dr. Robert Poe and Miss Vicki Bruyr for their very significant contributions.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

In a very real sense the idea of the Finsler manifold had its inception in the famous 1854 dissertation of Bernhard Riemann entitled "Concerning the Hypotheses Which Lie at the Base of Geometry." In this paper Riemann laid the foundation for the study of the n-dimensional differentiable manifold and for the study of a generalized metric space which included Euclidean spaces as special cases and which became known one half century later as Finsler spaces.

The geometry introduced in Riemann's paper may be characterized by three properties: it was n-dimensional; it was infinitesimal; and it was Non-Euclidean. Although all three properties had already been introduced by other famous mathematicians, Riemann in his 1854 dissertation applied all at the same time.

The study of a space having more dimensions than three had been introduced by Arthur Cayley in 1843 and by Herman Grassman in 1844.

The study of infinitesimal geometry had been well developed by scholars beginning with Christian Huygens in 1673 and continuing up to the geometers of Riemann's own time. The names which can hardly be omitted in this development would include Gaspard Monge, Leonhard Euler and Carl Friedrich Gauss.

Finally, the study of Non-Euclidean geometry had already brought forth the work of Johann Bolyai, Nilolai Lobachevski, and Gauss. But their work in this area was characterized by the full use of the synthetic method of elementary geometry.

What Riemann achieved in his 1854 dissertation was a further development of all three ideas in one unified theory. He began by postulating the n-dimensionality of space, he applied the method of investigating the behaviour of the infinitesimal parts of the space in order to gain knowledge of the whole, and with this he created the basis for the concept now known as the differentiable manifold.

Riemann's Manifold

The first part of Riemann's three-part dissertation was devoted to the concept of the n-dimensional manifold.

> As is well known geometry assumes the concept of space as well as the basic concepts for constructions in space as something given in advance. Only nominal definitions arise from these concepts while the real essential difinitions appear in the form of axioms. The relationship between these axioms consequently remains hidden; we do not see whether or in what way their relationship is necessary or possible.[1]

For example, by assuming that space is three-dimensional, infinite, unbounded, flat (in the sense that Euclid's Fifth Postulate is everywhere valid) and containing struc-

[1] Georg Bernhard Riemann, Gesammelte Mathematische Werke (New York, Dover, 1953), p. 272.

tures which are assumed to be infinite such as straight lines as well as structures such as spheres, the geometer may be restricting himself too much. Are all these assumptions so related that all are necessary? Or is it possible that space could be three-dimensional and unbounded without being infinite and flat? And why is it that the measure relation, the distance, between two points on a subset of space - on a sphere in E^3, for example - is more complicated than the measure relation between points in the space itself?

Riemann felt that answers to such questions could not be found under the existing axioms simply because the axioms themselves contained some assumptions on the basis of self-evidency rather than logical necessity. To remedy the situation he suggested that space be considered as a set of points having one fundamental property - namely, that each point be determined by the results of n independent measurements.

This set of points may be a discrete set or a continous set. If it is discrete, then the measure relations possible on it would in some way be connected to the cardinality of the set. If it is continuous, it must be considered amorphous, and the measure relations possible must be reached by a study of the subsets which comprise the space.

The concern of this dissertation is with the latter case. Therefore, a first goal will be to establish machinery. The path to be taken toward this end will be that of

Riemann translated into the terminology developed by post-Riemannian achievements in topology.

The first restrictions to be put on the space are restrictions which make possible the rigorous application of the theory of topology. Toward this end, consider a set M of points such that each point of M is determined by n independent measurements and also such that M itself is a separable Hausdorff topological space. (A space M is Hausdorff if for any two points of M there exist disjoint open sets each containing exactly one of the points; M is separable if it contains a countable basis for its topology.)

Consider now the space R^n with its standard product topology. All the properties known about this space are available to the mathematician. What is needed is a method of relating the topological space M to the topological space R^n so that properties of M can be deduced through a study of R^n. The most apparent way of setting up a relation between these two topological spaces is to require that for each point of M there is an open neighborhood of that point that is homeomorphic to some open subset in R^n. A separable Hausdorff topological space which satisfies this requirement is called a topological manifold.

The restriction to the class of topological manifolds is justified by the facts that this class is very vast and that the tools for working with this class are already available. But even with the limitation to the class of topological manifolds, the task of introducing concepts of differentiability is not an easy one. It is accomplished by relating differentiability on a manifold to differentiability in R^n, the space of real n-tuples with all its known properties.

Consider, first, two open neighborhoods U_1 and U_2 of a point P of M. Each of these is mapped homeomorphically into R^n. Let ϕ_1 and ϕ_2 be the corresponding homeomorphisms.

Then ϕ_1 maps $U_1 \cap U_2$ into some open set in R^n and ϕ_2 maps $U_1 \cap U_2$ into some open set in R^n. I.e., $\phi_2(U_1 \cap U_2) \subset R^n$.

$$\phi_2^{-1} : \phi_2(U_1 \cap U_2) \rightarrow U_1 \cap U_2 \subset M$$

$$\phi_1 \circ \phi_2^{-1} : \phi_2(U_1 \cap U_2) \rightarrow \phi_1(U_1 \cap U_2) \subset R^n$$

In brief, $\phi_1 \circ \phi_2^{-1}$ is a mapping of a set in R^n into a set in R^n. For such mappings the notion of differentiability is well defined in courses of advanced calculus. Consequently some machinery for differentiability should result if one considers only those functions ϕ_α and ϕ_β such that $\phi_\alpha \circ \phi_\beta^{-1}$ and $\phi_\beta \circ \phi_\alpha^{-1}$ are differentiable functions of n real variables in the sense of advanced calculus.

A differentiable manifold is defined, therefore, to be a topological manifold whose local homeomorphisms $\phi_\alpha : U_\alpha \rightarrow R^n$, where $\{U_\alpha \mid \alpha \in \Lambda\} = M$, satisfy the condition that for every $\alpha, \beta \in \Lambda$ the mapping $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha (U_\alpha \cap U_\beta) \rightarrow \phi_\beta (U_\alpha \cap U_\beta)$ is such that all of its first partial derivatives exist and are continuous. If the mapping is such that it has continuous partial derivatives of all orders less than or equal to r, the manifold is called a differentiable manifold of class C^r.

It is to be expected that this new structure, this differentiable manifold of class c^r, will be amenable to a concept of differentiation. I.e., from the differentiability of the function $\phi_\alpha \circ O_\beta^{-1}$ whose domain is in R^n one hopes to define differentiability for some function f whose domain is in M itself. To do this, a concept of local coordinates is introduced.

Let P be a point in M with an open neighborhood $U_\alpha(P)$. The homeomorphism ϕ_α maps $U_\alpha(P)$ into some open set V_α in R^n. Hence ϕ_α maps the point P into an n-tuple $(x^1, \ldots x^n)$ in R^n. In other words, the n coordinates of $\phi_\alpha(P)$ are the n real variables $x^1, \ldots x^n$. They are functions of $\phi_\alpha(P)$ and are called the local coordinates of P.

The open sets U_α are called coordinate neighborhoods.

Let $U_\beta(P)$ be another open neighborhood of P with corresponding homeomorphism ϕ_β. Then the n coordinates of $\phi_\beta(P)$ are n real variables $y^1, \ldots y^n$. Since $P \in (U_\alpha \cap U_\beta)$, and since the mapping $\phi_\beta \circ \phi_\alpha^{-1}$ is a homeomorphism of $\phi_\alpha(U_\alpha \cap U_\beta)$ onto $\phi_\beta(U_\alpha \cap U_\beta)$, these two sets of local coordinates of P are related. (See figure 1.) These two different expressions $(x^1, \ldots, x^n)$ and $(y^1, \ldots y^n)$ for the point P are continuously related.

This means that the homeomorphism is given by the n equations expressing the coordinate functions y^i in terms of the n-tuple $(x^1, \ldots x^n)$. Briefly $y^i = y^i(x^1, \ldots x^n)$.

These functions defining the map $\phi_\beta \circ \phi_\alpha^{-1}$ are differentiable functions of the x^i since the mapping $\phi_\beta \circ \phi_\alpha^{-1}$

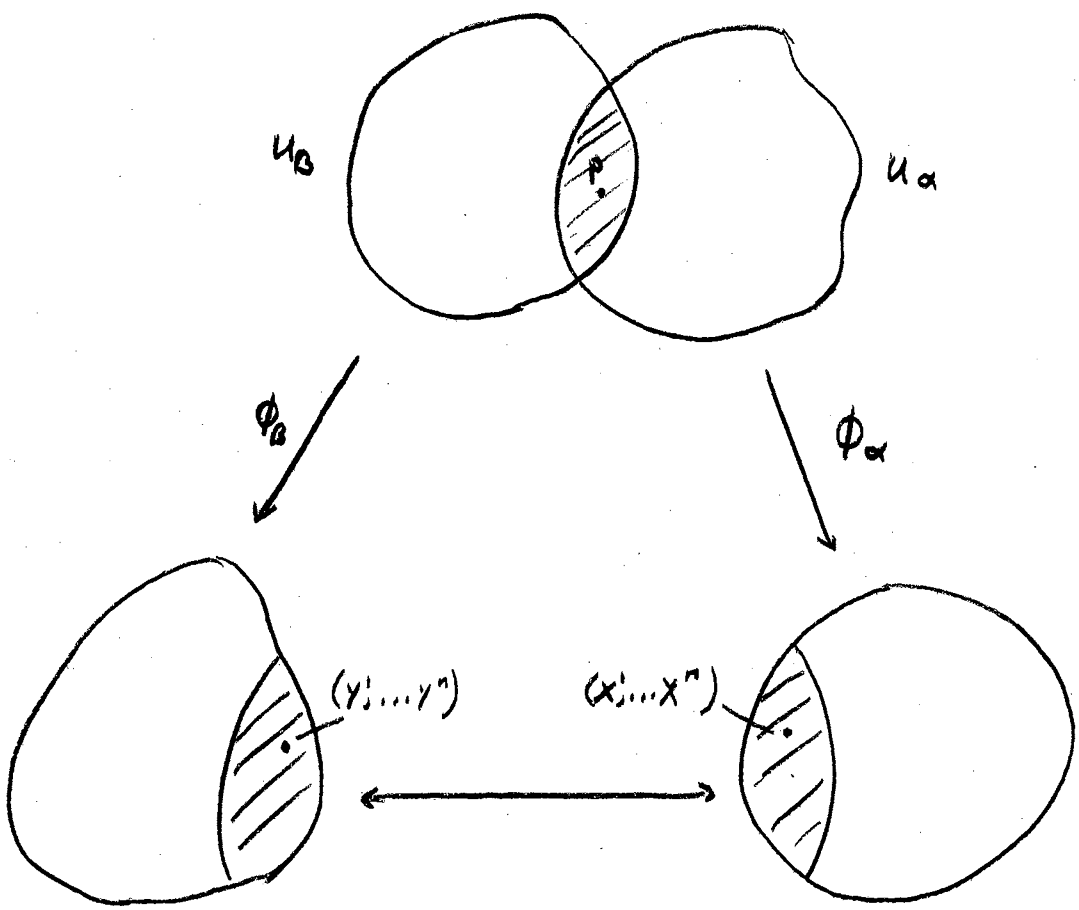
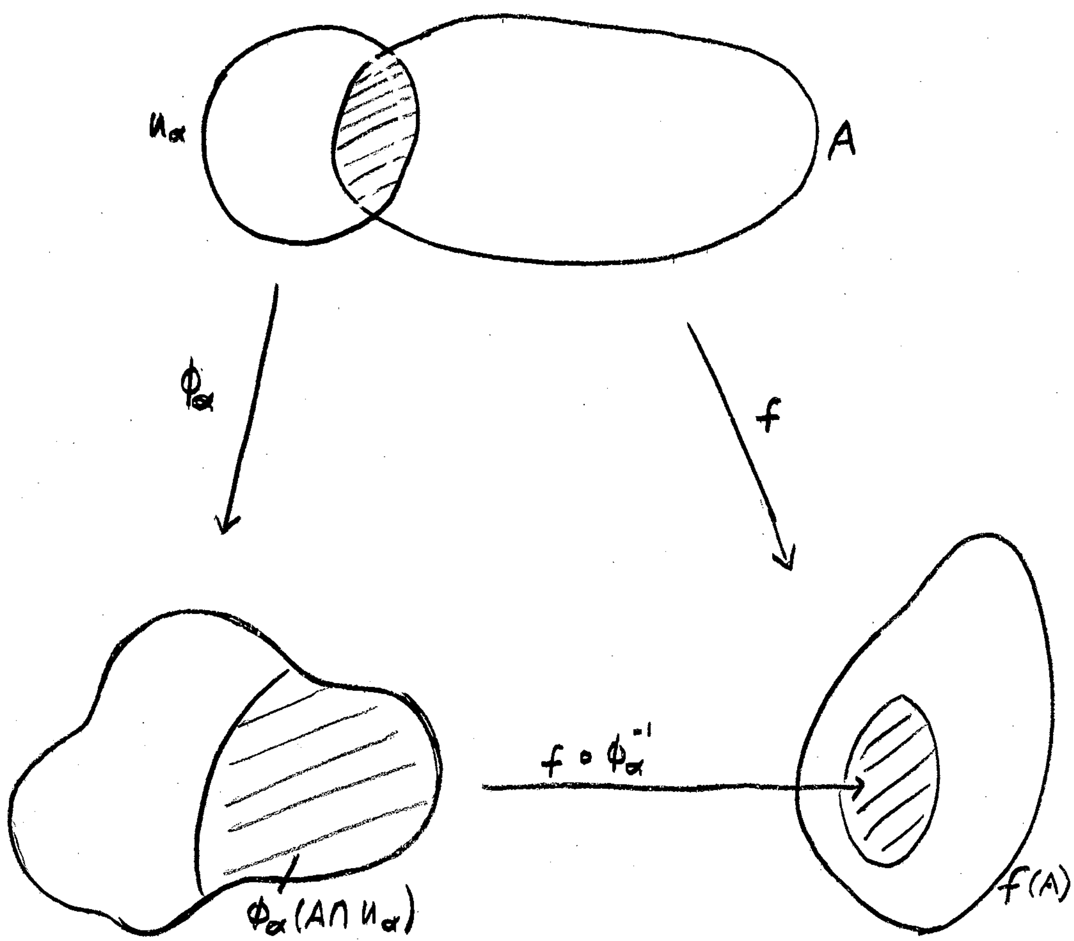


Figure 1



u_α
A
ϕ_α
f
$f \circ \phi_\alpha^{-1}$
$\phi_\alpha(A \cap u_\alpha)$
f(A)

Figure 2

is by postulate differentiable.

Now one can define differentiable functions on a differentiable manifold M of class c^r. Let A be any open set of M and let f be a real valued function on A. Let P $\in A \cap U_\alpha$ where $\{U_\alpha\}$ are the coordinate neighborhoods covering M.

$$\phi_\alpha (A \cap U_\alpha) \subset R^n$$

$f \circ \phi_\alpha^{-1}$ is defined on $\phi_\alpha (A \cap U_\alpha)$ and has its range in R. If $f \circ \phi_\alpha^{-1}$ is differentiable on the open set $\phi_\alpha(A \cap U_\alpha)$ for every α, then f is called differentiable on the open set A in M. If $f \circ \phi_\alpha^{-1}$ is of class c^r on $\phi_\alpha(A \cap U_\alpha)$, then f is said to be of class c^r on A. (See figure 2). Again it is important to stress that f is defined on the space M but $f \circ \phi_2{}^{-1}$ is defined on a set in R^n the differentiability of this is studied in advanced calculus.

In this definition of the differentiability of a real-valued function f defined on an open set A of M it is to be noted that the differentiability of f at P does not depend on any particular U_α containing P. However, the open set A of M on which f is differentiable does depend on f. This means that the set of all functions which are differentiable on some open set in M do not all have the same domain in M. More will be said about this set of differentiable functions in a later chapter.

An Example

A familiar example of a separable Hausdorff space is the Euclidean plane itself; the topology of this plane induces onto any circle defined in the plane a topology such that the circle itself is a separable Hausdorff space. Let M be such a circle, let P be any arbitrary fixed point on M, and let Q be the point diametrically opposed to P.

If the radius of the circle is denoted by the constant K and if the radian measure of an arc beginning at P is denoted by θ, all points P_θ of M can be determined by the values of θ. Let $U_1 = M \setminus P$, the set of all points of M except P. Let $U_2 = M \setminus Q$. Clearly, $U_1 \cup U_2 = M$.

The mapping ϕ_1 of the point P_θ onto the real number $\frac{\theta - \pi}{\pi}$ where $0 < \theta < 2\pi$ takes all of U_1 onto the real open interval (-1,1). This mapping is a homeomorphism since it is one to one, open, continuous and onto. Similarly, the mapping ϕ_2 of each point P_θ onto the real number $\frac{\theta}{\pi}$ when $-\pi < \theta < \pi$ is a homemorphism of U_2 onto the open interval (-1, 1).

ϕ_1^{-1} maps (-1,1) onto U_1, and consequently it maps the open set $(-1,0) \cup (0,1)$ onto $U_1 \setminus Q$. ϕ_2 then maps $U_1 \setminus Q$ (which is identically equal to $U_2 \setminus P$) onto $(-1,0) \cup (0,1)$. In other words, $\phi_2 \circ \phi_1^{-1}$ is a homeomorphism of the real set $(-1,0) \cup (0,1)$ onto itself.

This example selected for its simplicity indicates the idea that $\phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha (U_\alpha \cap U_\beta) \to \phi_\beta (U_\alpha \cap U_\beta)$ is a homeomorphism. In this case, $U_\alpha \cap U_\beta$ is the set $M \setminus (P \cup Q)$, all the points of M except P and Q. $\phi_\alpha (U_\alpha \cap U_\beta)$ is $(-1,0) \cup (0,1)$; $\phi_\beta (U_\alpha \cap U_\beta)$ is $(-1,0) \cup (0,1)$; and $\phi_\beta \circ \phi_\alpha^{-1}$ is $\phi_2 \circ \phi_1^{-1}$ which maps $(-1,0) \cup (0,1)$ homeomorphically onto itself.

Now the set $(-1,0) \cup (0,1)$ is an open set on the Euclidean line, which can be designated as the x — axis. In terms of coordinates on this axis $(-1,0)$ is $-1 < x < 0$ and $(0,1)$ is $0 < x < 1$.

Let $\phi_1(P_\theta) = x \in (-1,0)$

Then $\frac{\theta - \pi}{\pi} = x$ and $0 < \theta < \pi$

Hence $\theta - \pi = \pi x$, and $\theta = \pi x + \pi = \phi_1^{-1}(x)$

And $\phi_2 \circ \phi_1^{-1}(x) = \frac{1}{\pi}\theta = \frac{1}{\pi}(\pi x + \pi) = x + 1$

Let $\phi_1(P_\theta) = x \in (0,1)$

Then $\frac{\theta - \pi}{\pi} = x$ and $\pi < \theta < 2\pi$

Again $\theta - \pi = \pi x$ and $\theta = \pi x + \pi$

But ϕ_2 is defined on $-\pi < \theta < \pi$, and P_θ corresponding to $\theta = \pi x + \pi$ is the same P_θ which corresponds to $\theta = \pi x + \pi - 2\pi = \pi x - \pi$.

Therefore, $\theta = \pi x - \pi = \phi_1^{-1}(x)$

$$\phi_2 \circ \phi_1^{-1}(x) = \frac{\theta}{\pi} = \frac{\pi x - \pi}{\pi}$$

$$= x-1$$

Conclusion:
$$\phi_2 \circ \phi_1^{-1}(x) = \begin{cases} x+1 & \text{for } -1 < x < 0 \\ x-1 & \text{for } 0 < x < 1 \end{cases}$$

The point P_θ was mapped by the function ϕ_1 onto the real number x. This real number x was mapped by the function $\phi_2 \circ \phi_1^{-1}$ onto another real number - either x+1 or x-1. This mapping is differentiable. Consequently the topological manifold M satifies the requirement necessary to be a differentiable manifold.

Let f be a real-valued function defined on an open set G of M. Let P_θ be any point of G. Then $P_\theta \in G \cap U_1$ or $P_\theta \in G \cap U_2$. Without loss of generality, assume the latter. Then since ϕ_2^{-1} maps some open set of (-1,0) $\cup$ (0.1) onto the open set $G \cap U_2$, and since f then maps this set $G \cap U_2$ onto some open set in the real number line R, the composite function $f \circ \phi_2^{-1}$ is a mapping of an open set of real numbers into the set of real numbers. If this composite is differentiable, then f itself is differentiable. The function f which maps each P_θ for $0 < \theta < \pi$ onto the real number $2\theta^2$, for example, is simply the differentiable mapping of the real interval (0,1) onto the real interval $(0, 2\pi^2)$ ordinarily expressed as $y = 2\pi^2 x^2$.

An important feature which should be noted in the consideration of the above example is that the radius k of the given circle played no part. In fact, the same discussion holds without change if M is an ellipse of any eccentricity.

Another important idea to bear in mind is closely connected with the last remark and is the major concern of this dissertation. It is the fact that nothing has been mentioned which would in any way enable one to find the distance between two given points of M. Perhaps all that could be said, and this because of the familiar simple space chosed as M, is that such distances exist and are bounded.

Riemann's Line Element

The second part of Riemann's dissertation was entitled "Possible Measure Relations on an n-dimensional Manifold Under The Hypothesis that Lines Possess Length Independent of Their Positions." The significance of the given hypothesis becomes obvious when one recalls from lessons on affine geometry that there exist spaces in which a selected unit of length along one line does not determine the unit of length along other lines. In such spaces there is no way to compare lengths of segments of lines unless the lines have the same or the exact opposite directions.

Riemann was concerned with the problem of putting a metric geometry on a manifold in such a manner that the length of any line segment could be compared to that of any other.

In the previous section of this chapter the aim was to put a differentiable structure onto a topological manifold. But in all that section no mention was made of a notion of distance between points on the manifold M. One has topological transformations from M to R^n, and there is a standard way for defining a metric on R^n, but this in no way implies a corresponding metric on M itself.

Riemann's method for introducing the metric property to his manifold was to define distance for points which are close enough together. He considered a fixed point P on the manifold with local coordinates $x^1, x^2, \ldots x^n$. If Q is a nearby point, then the coordinates of Q can be taken to be $x^1 + dx^1$, $x^2 + dx^2$, ... $x^n + dx^n$ where the dx^i are infinitisimal changes in the respective coordinates. If now dS is going to be the distance from P to Q what properties should dS have? Obviously dS should be a function of P and Q, but it should have other properties as well.

The dS is called the element of length or the line element in the sense that the length of a curve is to be defined by means of the definition of the dS.

It should be noted that, unlike Euclid, Riemann did not assume that the properties of distance between any two points would be the same as that between two points close together. Thus, in Riemann's approach, if he had wanted to postulate the Pythagorean Theorem, he would have postulated it only for restricted neighborhoods. He did not want to postulate that a line is infinite in length or that it is not. He could not make such postulates because of the very nature of his infinitesimal geometry. What then should be the properties of length in the small?

Following are the properties which Riemann suggested as requirements for length.

1) The length of a line element should be positive; it should be independent of the sense traversed.

2) The length of a line element should be independent of its position. This means that the length should be invariant under a change of coordinate systems.

3) The length of a line element should remain unchanged if all points undergo an infinitely small change of

position: and when all the quantities dx^i increase in the same proportion, the length of the line element itself changes in this same proportion. This amounts to the requirement that the length should not depend on the parametrization.

If hypothesis (1) holds, then clearly the length dS must be positive even when the quantities dx^i are all changed in sign, but by (3) dS must be multiplied by a constant k whenever each dx^i is multiplied by the constant $k > 0$. Hence the line element dS must be homogeneous of the first degree in the dx^i, but the former condition together with this requirement implies that dS cannot be linear in the dx^i. To illustrate the restrictions thus imposed on the nature of the function dS, consider the following situations when $n = 2$, P is (x,y) and Q is $(x + dx, y + dy)$.

If dS were defined to be $dx + dy$, then the distance between (x,y) and $(x - dx, x - dy)$ would be $-dx -dy$, a negative quantity if $dx + dy$ is positive. Thus dS cannot be given by the expression $dx + dy$.

If dS were defined to be $dx + x$, then 2 dS must be 2 $dx + x$ by the third condition. But then, $dS = dx + x/2$ contradicting the fact that $dS = dx + x$. Thus, no expression of the type $dx + x$ can be used to define dS.

Briefly, if one accepts all the Riemann conditions for the linear element, many forms involving the variables x^i and dx^i must immediately be regarded as impossible since the permissible forms must be homogeneous of the first degree in dx^i.

Consider the form $dS^2 = (g_{ij}\ dx^i\ dx^j)$ where the g_{ij}'s may be functions of x^i but not of dx^i. Since dS^2 is homogeneous of the second degree in dx^i, dS is homogeneous of the first degree. Similarly, $dS^4 = (g_{ij}\ (dx^i)^2\ (dx^j)^2)$ is homogeneous of the fourth degree. While all such forms are permissible, those such as $dS^3 = (g_{ij}\ (dx^i)^2\ dx^j)$ are not permissible because a change of sign of all dx^i implies a change of sign of dS.

Riemann called $dS = \left(g_{ij}\ dx^i\ dx^j\right)^{1/2}$ the simplest general case. He explored it in great depth, and it is the form usually associated with him. However, the other possibilities did not escape his attention.

> ....dS is possibly the square root of a positive homogeneous function of the second degree in the quantities dx^i whose coefficients are continuous functions of the x^i. For the case when the position of a point is expressed in rectangular coordinates,
>
> $$dS = \left[\sum_{i=1}^{n} (dx^i)^2\right]^{1/2}$$
>
> This case is therefore a special case of the simplest general case. For the next simplest case ... dS can be expressed as the fourth root of a biquadratic differential form. The investigation of these more general types would not require any essentially different principles ... but it would contribute comparatively little new to the theory of space. [2]

With this passage Riemann discovered what are now called Finsler spaces and, underestimating their importance, left their development and study to later mathematicians.

---

[2]Ibid. p. 278

A Note on Riemann's Differential Form

Consider on the Euclidean plane a rectangular coordinate system. The variables x and y are called cartesian coordinates of the plane referred to a selected pair of rectangular axes. Clearly any portion of the plane is covered by two families of straight lines - those parallel to the x-axis and those parallel to the y-axis. Any point of the plane lies on exactly one straight line of each family.

Now remove the restriction that the coordinate system be rectangular. Then any two families of curves which cover the plane in such manner that any point on the plane lies on exactly one curve of each family form a new coordinate system. Denote the curves of one family as U^1 curves and those of the other as U^2 curves. The variables U^1 and U^2 are called curvilinear coordinates of the plane. An obvious example of such a coordinate system is the polar-coordinate system.

Consider two points on this plane given in the rectangular system by $P(x,y)$ and $Q(x + dx, y + dy)$ and in the curvilinear system by $P(U^1, U^2)$ and $Q(U^1 + dU^1, U^2 + dU^2)$.

Let the distance in the first system be given by $dS^2 = (dx)^2 + (dy)^2$. The aim is to express this function as one of U^1, U^2, dU^1 and dU^2 under the assumption that the variables x and y are functions of U^1 and U^2. I.e., $x = f^1(U^1, U^2)$ and $y = f^2(U^1, U^2)$ and these are independent. In this case the Jacobian of the transformation from one coordinate system to the other is non-zero, and the inverse transformation exists: $U^1 = g^1\,(x^1, x^2)$ and $U^2 = g^2\,(x^1, x^2)$.

Therefore,

$$dx = \frac{\partial f^1}{\partial U^1}\, dU^1 + \frac{\partial f^1}{\partial U^2}\, dU^2$$

and

$$dy = \frac{\partial f^2}{\partial U^1}\, dU^1 + \frac{\partial f^2}{\partial U^2}\, dU^2$$

$$dS^2 = \left(\frac{\partial f^1}{\partial U^1}\, dU^1 + \frac{\partial f^1}{\partial U^2}\, dU^2\right)^2 + \left(\frac{\partial f^2}{\partial U^1}\, dU^1 + \frac{\partial f^2}{\partial U^2}\, dU^2\right)^2$$

$$= \left[\left(\frac{\partial f^1}{\partial U^1}\right)^2 + \left(\frac{\partial f^2}{\partial U^1}\right)^2\right](dU^1)^2 + 2\left[\frac{\partial f^1}{\partial U^1}\,\frac{\partial f^1}{\partial U^2} + \frac{\partial f^2}{\partial U^1}\,\frac{\partial f^2}{\partial U^2}\right] dU^1\, dU^2 + \left[\left(\frac{\partial f^1}{\partial U^2}\right)^2 + \left(\frac{\partial f^2}{\partial U^2}\right)^2\right](dU^2)^2$$

While the element of length in cartesian coordinates had the seemingly simple form $dS^2 = (dx)^2 + (dy)^2$, it assumes this above complicated form in the curvilinear coordinate system. The cartesian form represented a Euclidean line-element; the latter form is Riemannian. It can be written as $dS^2 = g_{ij}\, dU^i\, dU^j$ where $i = 1, 2$; $j = 1, 2$ and

$$g_{ij} = \frac{\partial f^1}{\partial U^i}\,\frac{\partial f^1}{\partial U^j} + \frac{\partial f^2}{\partial U^i}\,\frac{\partial f^2}{\partial U^j}$$

This transfer from Euclidean geometry to Riemannian by way of generalizing from cartesian coordinates in the Euclidean plane to curvilinear coordinates is a fundamental relation between the two geometries.

On the other hand it may be of interest to note what happens on another transformation to a new coordinate system. The dS^2 should remain fixed for fixed nearby points P and Q but how is the form $g_{ij}\, dx^i\, dx^j$ affected?

Repeating the properties of dS^2, it is homogeneous of degree two in the n variables dx^i; it is positive for all real values of dx^i not all zero; it is zero only when all dx^i are zero; it is a real number and is invariant under any change of coordinate systems.

These restrictions on the form $g_{ij}\, dx^i\, dx^j$ imply restrictions on the coefficients g_{ij}.

Let $dS^2 = g_{ij}\, dx^i\, dx^j$ in one coordinate system and $g_{\alpha\beta}\, dx^\alpha dx^\beta$ in another coordinate system. Since dS^2 is invariant under such transformations, $g_{ij}\, dx^i\, dx^j = g_{\alpha\beta}\, dx^\alpha$ dx^β. But

$$dx^\alpha\, dx^\beta = \left(\frac{\partial x^\alpha}{\partial x^i}\, dx^i\right) \left(\frac{\partial x^\beta}{\partial x^j}\, dx^j\right)$$

Therefore,

$$g_{ij}\, dx^i\, dx^j = g_{\alpha\beta}\, \frac{\partial x^\alpha}{\partial x^i}\, \frac{\partial x^\beta}{\partial x^j}\, dx^i \quad dx^j .$$

Accordingly, the coefficients g_{ij} in one coordinate system must become

$$g_{\alpha\beta}\, \frac{\partial x^\alpha}{\partial x^i}\, \frac{\partial x^\beta}{\partial x^j}$$

in the other system. This rule describing how the g_{ij} be-have under coordinate transformations is of basic importance.

Let g denote the determinant of the matrix (g_{ij}). Assume that $g = 0$. Then the system of n linear equations $g_{\alpha\beta}\, dx^{\alpha} = 0$ whose determinant of coefficients is g would have a non-trival solution $(dx^1, dx^2 \ldots dx^n)$. In other words dS^2 would be zero even though some dx^i is not zero. This contradiction to restrictions already placed on dS^2 implies that g cannot be zero.

Finally, consider the case when $g_{ik} \neq g_{ki}$ for some fixed i and k. The sum of the two terms $g_{ik}\, dx^i\, dx^k + g_{ki}\, dx^k\, dx^i$ can be expressed as $(g_{ik} + g_{ki})\, dx^i\, dx^k$ which in turn is equal to $(1/2)(g_{ik} + g_{ki})\, dx^i\, dx^k + (1/2)(g_{ik} + g_{ki})\, dx^k\, dx^i$. In this final form the coefficient of $dx^i\, dx^k$ is equal to that of $dx^k\, dx^i$. For this reason the given quadratic form can always be expressed as a symmetric form.

In summary, Riemann's linear element dS^2 is a homogeneous-of-the-second-degree, positive definite, real valued differential form. This is characterized by the coefficients g_{ij} whose matrix is non-singular, symmetric and positive definite. And if the space being considered is such that a rectangular coordinate system may be used, the dS^2 becomes the ordinary dS^2 of Euclidean geometry.

Examples of common forms studied in Riemannian geometry are:

$$dS^2 = a^2 \left[(du^1)^2 + \sin^2 u^1\, (du^2)^2\right]$$

$$dS^2 = \frac{(du^1)^2 + (du^2)^2}{\left\{1 + a\left[(u^1)^2 + (u^2)^2\right]\right\}}$$

Note the homogeneity and the fact that g_{ij} is a function of the variables du^1 and du^2 only.

CHAPTER II

LOCAL FINSLER GEOMETRY

The aim which Riemann had set for himself was to ascertain the forms for the element dS in a general finite dimensional space in which every curve has a length derived from an infinitesimal line element and independent of its position in space. He reached his famous form $dS = (g_{\alpha\beta} dU^{\alpha} dU^{\beta})^{1/2}$ by postulating certain conditions to be satisfied by length between two neighboring points.

If conditions other than Riemann's are postulated, other equally valid notions of the element of length result. For instance, if one does not accept the restriction that distance between two points be always positive, the resulting geometry turns out to be of consequence today both theoretically and practically.

The observation which Riemann himself made concerning the possibility of using some function other than $dS = (g_{\alpha\beta} dU^{\alpha} dU^{\beta})^{1/2}$ led to a series of efforts along this line by many mathematicians. In 1918 Paul Finsler succeeded in developing a more general function for the line element which led to a geometry including that of Riemannian geometry as a special case. A significant change in the conditions imposed on the element of length was that it may in fact, depends on the way the measurement is taken.[1]

Finsler's F-function

Finsler began his work with the assumptions that the length of a curve in an n-dimensional space should be given by the integral of an essentially arbitrary function of the coordinates used to express the curve and of their first derivatives and that this function should be a real-valued function.

Every curve of the space is to be considered as a set of points with a positive direction attached. Two curves with the same point set but with different senses of direction are to be considered as different curves. In any parametric representation of the curve, the parameter is assumed to be such that the curve is given its positive sense. This means that the direction of the curve is that for which the parameter is increasing. For example, the interval $0 \leq x \leq 1$ is a point set for two different curves - one whose parametric representation is $x = T$, $0 \leq T \leq 1$, and the other whose parametric representation is $x = 1 - T$, $0 \leq T \leq 1$. The direction of the first is from 0 to 1 while that of the second is from 1 to 0.

Let the equations of an arc C in the n-dimensional space be given by $x^i = x^i(U)$ where $i = 1, \ldots n$ and $U_1 \leq U \leq U_2$

---

[1]Paul Finsler, Uber Kurven and Flachen in Allgemeimen Raumen, Birkhauser and Basil, Switzerland, 1951.

The assumption is that arc length S along this curve is given by the value of an integral of the form

$$S = \int_{U_1}^{U} L(U, x, \dot{x})\, dU$$

where x denotes the vector $(x^1, \ldots x^n)$; $\dot{x}$, the vector $(\dot{x}^1, \ldots \dot{x}^n)$; and $\dot{x}^i$, the derivative of x^i with respect to the parameter U.

The first task is to find a necessary and sufficient condition that the value of the integral shall be independent of the parametric representation of the curve along which the integral is taken.

If a new parameter $t = U - c$ is used, the integral becomes

$$\int_{t_1}^{t} L(t,\ x,\ \frac{dx}{dt})\ dt.$$

Differentiate both integrals with respect to t. If the integrals are equal, these results should be equal. For the first,

$$\frac{d\int_{U_1}^{U} L\ dU}{dt} = \frac{d\int_{U_1}^{U} L\ dU}{dU} \quad \frac{dU}{dt} = L(U,x,\dot{x});$$

and for the second the derivative is obvious. Therefore

$$L(U,x,\dot{x}) = L(t,x,\frac{dx}{dt}) = L(t,x,\dot{x}).$$

This implies that for equality the integral function can not contain the parameter explicitly. Thus, the integral must be of the form

$$S = \int_{U_1}^{U} F(x,\dot{x})\ dU.$$

Now consider the new parameter $t = U/k$ where k is a positive constant. If

$$\int_{U_1}^{U} F(x,\dot{x})\, dU = \int_{t_1}^{t} F(x,\frac{dx}{dt})\, dt,$$

the derivatives with respect to t are equal.

$$F(x,\dot{x})\frac{dU}{dt} = F(x,\frac{dx}{dt})$$

$$\frac{dU}{dt} = k, \text{ and } \frac{dx}{dt} = \frac{dx}{d(U/k)} = \frac{dx}{1/k\, dU} = k\,\frac{dx}{dU} = k\dot{x}$$

Therefore $F(x,\dot{x})k = F(x,k\dot{x})$.

The necessary and sufficient condition that the value of the integral be independent of the parameter is that the integral be a function which does not contain the parameter explicitly and which is positively homogeneous in the variable $\dot{x}$.

If arc length is to be always positive for any real nondegenerate arc, then $F(x,\dot{x})$ must be positive.

These two conditions were among the earlier requirements suggested by Riemann. Finsler reached his, however, by methods of the calculus of variation. His third requirement for his F-function is the Legendre condition for problems in parametric form. This condition is equivalent to the assertion that the quadratic form

$$\frac{\partial^2 F^2(x,\dot{x})}{\partial x^i\, \partial x^j}\, \xi^i\, \xi^j$$

be positive definite for all variables ξ^i.

Using these three conditions Finsler built his generalization of Riemannian geometry. These three axioms together with four consequences needed for this section are listed below for convenience.

Axiom 1. $F(x,\dot{x}) > 0$ if all $\dot{x}$ are non-zero.

Axiom 2. $F(x,k\dot{x}) = kF(x,\dot{x})$ for any number $k > 0$.

Axiom 3. $\frac{\partial^2 F^2(x,\dot{x})}{\partial \dot{x}^i \, \partial \dot{x}^j} \xi^i \xi^j > 0$ unless all ξ^i are zero.

Consequences:

1. $\dot{x}^i \frac{\partial F}{\partial \dot{x}^i}(x,\dot{x}) = F(x,\dot{x})$

2. $\dot{x}^i \frac{\partial^2 F(x,\dot{x})}{\partial \dot{x}^i \, \partial \dot{x}^j} = 0$

3. $F(x,\dot{x} + x') \leq F(x,\dot{x}) + F(x,x')$ where x' denotes the derivative with respect to some parameter other than **u**.

Proofs of these consequences will be given in the next chapter.

The set of 2n variables $(x,\dot{x})$ is called Finsler's line element. This line element indicates a ray beginning at the point x and having direction numbers given by the $\dot{x}$. Moreover, this ray is tangent at x to any curve through x whose parametric equations have derivatives $\dot{x}$.

As examples consider the following distinct plane curves:

$$\begin{cases} x = 2t \\ y = t^2 \end{cases} \qquad \begin{cases} x = 3S - 1 \\ y = S^2 \end{cases} \qquad \begin{cases} x = 2U \\ y = U^2+2U-2 \end{cases}$$

All pass through the point (2,1) when $t = 1$, $S = 1$ and $U = 1$ respectively.

In the first case the tangent line at (2,1) is given by the direction numbers $\dot{x} = 2$, $\dot{y} = 2t$ evaluated at $t = 1$. Thus the direction numbers are $\{2,2\}$.

In the second case the direction numbers of the tangent line are $\{3,2\}$; and in the third case, $\{4,4\}$.

The first and third curves have the same tangent line at (2,1) while the second curve does not.

The Finsler line element for the first curve at the given point is the 4-tuple (2,1, 2,2). They were determined by fixing the point and fixing the curve. In summary, $(x,\dot{x})$ is a function of a given curve and a given point on it

The F-function is assumed to be defined for all line elements in a sufficiently small region, and

$$S = \int_{T_0}^{T} F(x,\dot{x})\, dt$$

is therefore a sort of distance function for the length of curves through x. I.e., if x^i and $x^i + dx^i$ are nearby points, the distance between them can be defined to be $F(x^i, dx^i)$.

The third consequence stated above is very important. It states, in fact, that the Finsler F-function is convex in the $\dot{x}^i$. Recall that a set S of points is called convex if it contains the entire segment of the straight line connecting any two of its points. Let the elements of a convex region S be the n-tuple $(x^1, \ldots x^n)$. A function f on S is said to be a convex function if it is defined everywhere on S and if for all pairs of points $(x_1^1, \ldots x_1^n)$ and $(x_2^1, \ldots x_2^n)$ in S the following inequality holds:

$$f\left[(1-\theta)\; x_1^i + \theta x_2^i\right] \leq (1-\theta) f(x_1^i) + \theta f(x_2^i)$$

where $0 \leq \theta \leq 1$.

It should be noted here that Riemann's linear element satisfies all three postulates of Finsler, and consequently any function dS in Riemannian geometry is a Finsler F-function.

Elements of Minkowskian Geometry

It is known that an n-dimensional vector space is simply an n-dimensional affine space with one point specified as fixed. Any linear transformation of the vector space is a linear transformation of the affine space leaving the specified point fixed. Such affine spaces are called centered affine spaces, and the proper group of transformations for the study of such spaces is the group of all non-singular homogeneous linear transformations. Consequently, in the study of vector algebra, the fundamental tool employed is the transformation given by $y' = AY$ where y' and y are vectors and A is a non-singular n x n matrix. This is precisely the same thing as saying that vectors are invariant under non-singular homogeneous linear transformations.

If a metric is to be put on a vector space, there is no prior stipulation that the metric should be Euclidean; also it may be possible to put Euclidean metrics on all the lines through the origin(where the origin is the zero element of the vector space and also the point specified to be the center of the affine space) without getting a Euclidean metric for the whole space.

In Figure 3 of page 30 two vector spaces of dimension two are indicated. In each case $\overrightarrow{OA}$ and $\overrightarrow{OB}$ are selected to be unit vectors. The Euclidean length of $\overrightarrow{OA}$ need not be equal to that of $\overrightarrow{OB}$.

A more basic idea is that there is no way to compare lengths along OC with lengths along OA until a unit is selected on OC. But ratios of lengths along OC can be compared to ratios of lengths along OA. In Figure 3, if CA is parallel to xy, then the ratio of the length of OC to that of Ox must be equal to the ratio of the length of OA to that of Oy. This is the basic notion on which is built Minkowskian geometry. Ratios of certain lengths - not lengths themselves - are invariant in affine space.

Consider in the Euclidean plane an arbitrary but fixed convex set S which is symmetric with respect to a fixed interior point O. Every ray through O intersects S in exactly one point. Denote the Euclidean distance between any two points x and y on the plane by e(x,y).

Minkowski created a new space out of this by assigning a new metric determined as follows: take the ray from O parallel to xy. Call the intersection of this ray with the

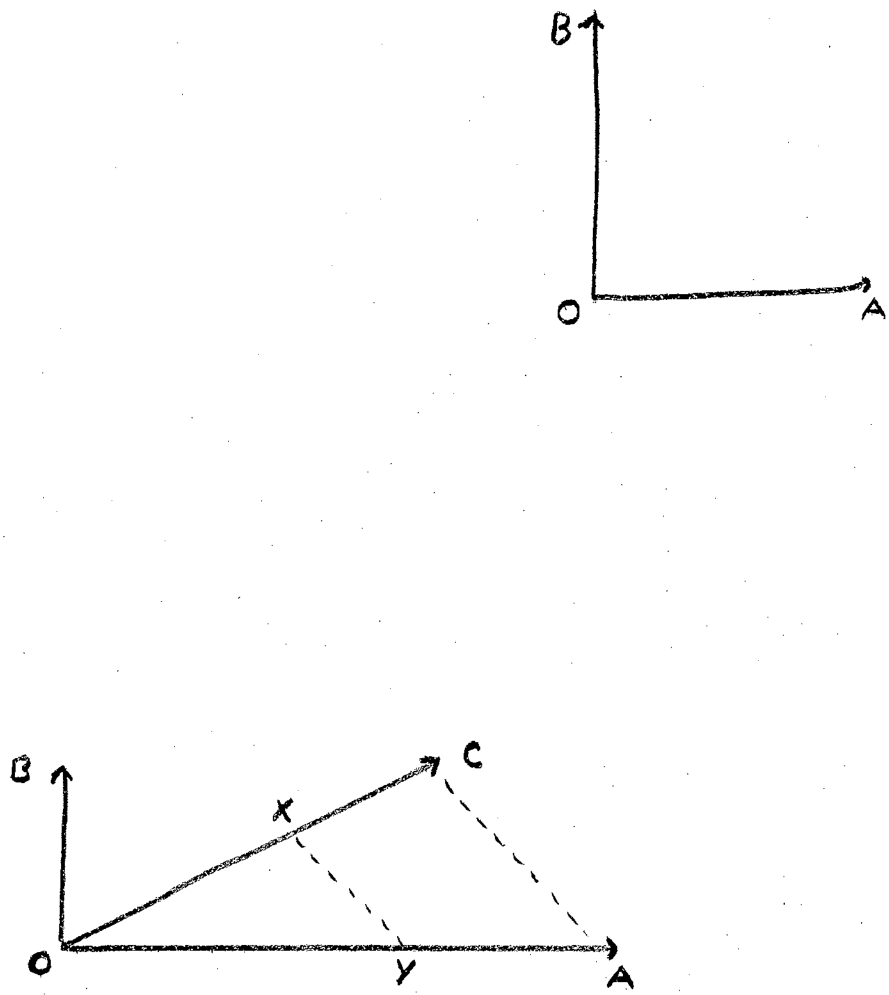


Figure 3

given convex set P. Define the distance from x to y to be the ratio of the Euclidean distance e(x,y) to the Euclidean distance e(O,P). If x = y, define the distance to be zero. This new metric is denoted by m(x,y).

If x is at O and if y is on the convex set, then the ratio of e(x,y) to e(O,P) is clearly unity: any point on the convex set lies at a Minkowski distance of 1 from the origin.

In extending the idea to the affine space of n dimensions, nothing fundamental is changed. The convex set is then a convex hypersurface. The reader should prove that the function M(x,y) as defined here is actually a metric.

It is clear that, on the Minkowskian plane, distance is invariant under translation and that the linear transformation which maps the vector x emanating from O into the vector λx where λ is real also maps every vector y into λy. It is also obvious that if the Euclidean circle is chosen as the convex hypersurface on the Eulidean plane, the resulting Minkowskian geometry coincides with the original Euclidean geometry.

The statement which Riemann made about the use of some metric other than the square root of a quadratic expression involving differentials has application here. Let $(x_1, x_2, \ldots x_n)$ be a point and define its Euclidean distance from the origin to be $\rho(x_i) = \sqrt{\Sigma x_i^2}$.

The locus determined by $\rho(x_i) = 1$ is just the unit circle. However, equally valid meters could be defined by

$\rho(x_i) = \sqrt[4]{\sum x_i^4}$ or by $\rho(x_i) = \sqrt[6]{x_i^6}$ or in fact by $\rho(x_i) = \sqrt[P]{\sum |x_i|^P}$ where $P \geq 1$.

The locus $\sqrt[P]{\sum |x_i|^P} = 1$ is for all $P \geq 1$ a convex set symmetric with respect to the origin. Either of these is satisfactory as a "unit ball" for Minkowskian geometry. Only the first is a quadric, and it alone is the satisfactory unit ball for Euclidean geometry.

The Indicatrix

Let P be a point of M with the n coordinates x^i. If $x^i = x^i(t)$ define a curve C on M through P, the tangent vector to C at P has components $\dot{x}^i = \frac{dx^i}{dt}$. The set of all such vectors tangent to curves through P describe a space denoted by $T_n(P)$ and called the tangent space attached to P.

This space of tangents is a linear space and also an affine space with center corresponding to the values $\dot{x}^i = 0$. P as a point on M has coordinates x^i, but as a point on $T_n(P)$ it is the origin with coordinates $\dot{x}^i = 0$.

The entity which Finsler called his line element has now a new interpretation. The x^i are the n coordinates of an arbitrary point P on M and the $\dot{x}^i$ are the n coordinates of a vector in the tangent plane at P. The 2n coordinates thus determine the points on the various tangent spaces to the manifold. For any fixed tangent space the first n coordinates are fixed while the last n coordinates are variables referring to the various points of the tangent space.

Consider now, at a fixed point $P(x^i)$, the tangent space $T_n(P)$. The equation $F(x^i, \dot{x}^i) = 1$ represents an (n-1) dimensional locus contained in $T_n(P)$. This locus is called the indicatrix.

Now consider the set of all elements of $T_n(P)$ interior to or on the indicatrix. This set of $\dot{x}^i$ satisfying $F(x^i, \dot{x}^i) \leq 1$ (where x^i is fixed) is a convex body with inner points. For let $\dot{x}$ and $\dot{y}$ be two points in $T_n(P)$ satisfying $F(x^i, \dot{x}^i) \leq 1$. Let $\dot{z}^i$ be any point of $T_n(P)$ lying on the join on $\dot{x}^i$ and $\dot{y}^i$. Then $\dot{z}^i$ may be expressed as $\dot{z}^i = (1 - \theta)\dot{x}^i + \theta \dot{y}^i$ where $0 < \theta < 1$.

But the Finsler F-function is convex in the $\dot{x}^i$.

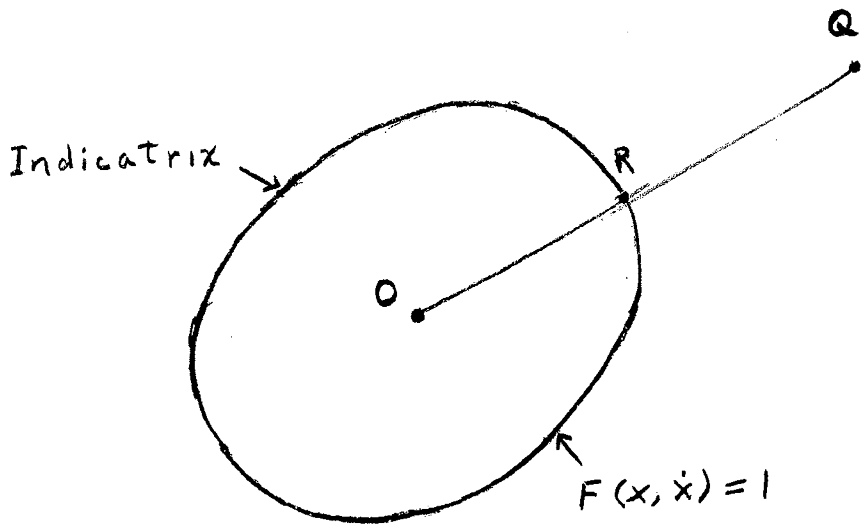
Therefore $F(x^i, \dot{z}^i) \leq F(x^i, (1 - \theta)\, \dot{x}^i) + F(x^i, \theta\, \dot{y}^i)$

$$= (1 - \theta)\, F(x^i, \dot{x}^i) + \theta\, F(x^i, \dot{y}^i).$$

This sum is less than or equal to one since $(1-\theta) < 1$, $\theta < 1$ while $F(x^i, \dot{x}^i) \leq 1$ and $F(x^i, \dot{y}^i) \leq 1$ by virtue of the hypotheses that $\dot{x}^i$ and $\dot{y}^i$ lie inside or on the indicatrix. Thus $F(x^i, \dot{z}^i) \leq 1$ and so z^i is also inside or on the indicatrix. In other words the indicatrix is the boundary of a convex body. It is a closed hypersurface of the space $T_n(P)$.

Let $P(x^i)$ as a point of $T_n(P)$ be denoted by O, the origin of the tangent space. Let Q be any other point $\dot{x}^i$ in $T_n(P)$. Let R be the point of $T_n(P)$ in which OQ (produced if necessary) intersects the indicatrix. Define the length of OQ to be the value $F(x^i, \dot{x}^i)$ which is simply the ratio of OQ to OR. See figure 4 of page 34.

This means that the F-Function assigns a distance



$$\frac{|OQ|}{|OR|} = \lambda \qquad \text{At } Q, \; F(x,\dot{x}) = \lambda$$

Figure 4

from P to be ascertained by comparing the vector OQ with another known vector in the same direction.

The point Q is expressed, of course, in terms of the affine coordinate system of $T_n(P)$ with center at P. OQ is simply a vector and OR is another vector in the same direction. Therefore OQ is some real multiple of OR. This real number is what is meant by the ratio $\frac{OQ}{OR}$. One can evaluate this ratio without knowing the length of OQ beforehand.

In summary, the F-function determines the indicatrix. The tangent space (being an affine space as well as a vector space) is endowed with a means of measuring length from the origin by use of the indicatrix. This in turn will lead to a metric on the tangent space - and thereby to a neighborhood of P in the manifold itself.

Among the postulates for Finsler geometry the one concerning quadratic forms was that $\frac{\partial F^2(x,\dot{x})}{\partial x^i \ \partial x^j} \xi^i \xi^j$ is positive definite.

In Riemann geometry the quadratic form associated with g_{ij} was positive definite.

To derive a corresponding such entity for the Finsler space, define a new set of quantities g_{ij} by $g_{ij}(x,\dot{x}) = \frac{1}{2}$ $\frac{\partial^2 F^2(x,x)}{\partial x^i \ \partial x^j}$.

Using the theorem of the first section of this chapter it can be shown that

$$\frac{1}{2} \ \frac{\partial^2 F^2(x,\dot{x})}{\partial x^i \ \partial x^j} \dot{x}^i \dot{x}^j = F^2(x,\dot{x})$$

Thus,
$$g_{ij}(x,\dot{x})\ \dot{x}^i\ \dot{x}^j = F^2\ (x,\dot{x})$$

Since $F(x,\dot{x})$ gives the length of the vector $\dot{x}$ emanating from x, all lengths on $T_n(P)$ can be stated in terms of the g_{ij}.

Note that the $g_{ij}(x,\dot{x})$ are symmetric in the indices and that they are homogeneous functions of degree 0 in $\dot{x}^i$ (since $g_{ij}(x,\dot{x})\ \dot{x}^i\ \dot{x}^j$ is $F^2\ (x,\dot{x})$ which is known to be homogeneous of the second degree in $\dot{x}^i$).

In terms of these g_{ij}, distance between two neighboring points on the manifold M is given by $dS^2 = g_{ij}(x,\dot{x})\ dx^i\ dx^j$.

The value of this invariant depends not only on the point x at which it is evaluated but also on a vector $\dot{x}$ at that point. This is an essential difference between Finsler and Riemann metrics. However, since the g_{ij} are homogeneous of degree 0 in the components of this vector $\dot{x}$, the distance does not depend on the magnitude of $\dot{x}$.

A Riemann space is clearly a special case of Finsler; it can be thought of as a Finsler space in which the function g_{ij} is independent of direction.

The indicatrix is the (n-1) dimensional locus $g_{ij}(x,\dot{x})\ \dot{x}^i\ \dot{x}^j = 1$ where on any tangent space the g_{ij} assume fixed values for the given coordinate system of that space. In the Riemannian case this locus is a quadric hyperspace - actually an (n-1) dimensional ellipsoid; in the general case this locus need not be a quadric. In fact it may not be symmetric about the origin since the condition $F(x,dx) =$

F(x,-dx) is not one of Finsler's postulates. In any case, however, the indicatrix is convex.

As a example, consider the $F = \frac{\dot{x}^2 + \dot{y}^2}{\dot{y}}$ $(\dot{y} > 0)$. F is positive if x and y are not both zero; F is positively homogeneous of degree 1 in $\dot{x}^i$ since $\frac{(k\dot{x})^2 + (k\dot{y})}{(k\dot{y})} = \frac{K(\dot{x}^2 + \dot{y}^2)}{k\dot{y}}$ = kF; and $F^2_{\dot{x}^i\dot{x}^j}\xi^i\xi^j$ is positive definite because the following determinant and its principal minor are both positive for x , y not both zero.

$$\begin{vmatrix} \frac{2(3\dot{x}^2+\dot{y}^2)}{\dot{y}^2} & \frac{-4\dot{x}^3}{\dot{y}^3} \\ \frac{-4\dot{x}^3}{\dot{y}^3} & \frac{(\dot{y}^4+3\dot{x}^4)}{\dot{y}^4} \end{vmatrix} = \begin{vmatrix} \tfrac{1}{2}F^2{}_{\dot{x}\dot{x}} & \tfrac{1}{2}F^2{}_{\dot{x}\dot{y}} \\ \tfrac{1}{2}F^2{}_{\dot{y}\dot{x}} & \tfrac{1}{2}F^2{}_{\dot{y}\dot{y}} \end{vmatrix} = \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix} > 0$$

This means that the function F satisfies all three necessary conditions of a Finsler function. Note that

$$dS = \sqrt{\frac{2(3\dot{x}^2 + \dot{y}^2)}{\dot{y}^2}(dx)^2 - \frac{8\dot{x}^3}{\dot{y}^3}(dx)(dy) + \frac{(\dot{y}^4 + 3\dot{x}^4)}{\dot{y}^4}(dy)^2}.$$

The coefficients g_{ij} are functions of $\dot{x}$ and $\dot{y}$. But this expression may be written as

$$dS = \sqrt{\left[6\left(\frac{\dot{x}}{\dot{y}}\right)^2 + 2\right](dx)^2 - 8\left(\frac{\dot{x}}{\dot{y}}\right)^3 dx\,dy + \left[3\left(\frac{\dot{x}}{\dot{y}}\right)^4 + 1\right](dy)^2}$$

This shows that the g_{ij} are functions of the ratios $\frac{\dot{x}}{\dot{y}}$ (of the ratios $\frac{dx}{dy}$) which are dependent only on the point (x,y) and the direction at that point. In other words, the functions g_{ij} in this example are functions of point and direction only.

CHAPTER III

TWO-DIMENSIONAL GEOMETRY

Perhaps a study of Finsler geometry for the case when $n = 2$ may serve to illustrate some of the ideas presented in the first chapters of this dissertation. Although some of the concepts valid for $n = 2$ do not generalize for higher dimensions, the machinery for handling these concepts is largely simple enough to be grasped in this introductory stage to Finsler geometry.

In this chapter, then, consider the 2-dimensional Finsler manifold as the universe itself - not as a submanifold of a three-space or of any other space of higher dimension. What can be said of the geometric ideas of length, angle, area, etc., - ordinary concepts in plane Euclidean geometry -when considered in terms of Finsler geometry? Because certain tools will be needed for the computations involved, this chapter begins with a derivation of useful facts that are valid for any natural number n.

Derivation of Auxiliary Facts

<u>Lemma</u> 1. $F_{\dot{x}^i}(x,\dot{x})\dot{x}^i = F(x,\dot{x})$

$$F(x^1,x^2, \ldots x^n, k\dot{x}^1, \ldots k\dot{x}^n)=KF(x^1,\ldots,x^n,\dot{x}^1,\ldots\dot{x}^n)$$

Differentiate with respect to k.

$$\dot{x}^1 F_{K\dot{x}^1} + \dot{x}^2 F_{K\dot{x}^2} + \ldots + \dot{x}^n F_{k\dot{x}^n} = F(x,\dot{x})$$

Set k = 1

$$\dot{x}^1 F_{\dot{x}^1} + \ldots + \dot{x}^n F_{\dot{x}^n} = F(x^1, \ldots x^n, \dot{x}^1, \ldots \dot{x}^n)$$

i.e., $F_{\dot{x}^i}(x,\dot{x})\, x^i = F(x,\dot{x})$

<u>Lemma</u> 2. $F_{\dot{x}^i \dot{x}^j}(x,\dot{x}) X^i = 0$

$F_{\dot{x}^i}\, \dot{x}^i = F$ by property 1 above

Differentiate with respect to $\dot{x}^j$

$$F_{\dot{x}^i} \frac{d\dot{x}^i}{d\dot{x}^j} + \dot{x}^i F_{\dot{x}^i \dot{x}^j} = F_{\dot{x}^j}$$

$\dfrac{d\dot{x}^i}{d\dot{x}^j} = \delta^i_j = 1$ if $x^i = x^j$ and 0 otherwise

Therefore $F_{\dot{x}^i} \dfrac{d\dot{x}^i}{d\dot{x}^j} = F_{\dot{x}^i} = F_{\dot{x}^j}$

$$F_{\dot{x}^i} + \dot{x}^i F_{\dot{x}^i \dot{x}^j} = F_{\dot{x}^i}$$

$$\dot{x}^i F_{\dot{x}^i \dot{x}^j} = 0$$

i.e., $F_{\dot{x}^i \dot{x}^j}(x,\dot{x})\dot{x}^i = o$

<u>Lemma</u> 3. $g_{ij}(x,\dot{x})\, \dot{x}^i \dot{x}^j = F^2(x,\dot{x})$

Find the second partial derivatives of $F^2(x,\dot{x})$ with respect to x^i.

$$F^2_{x^i} = 2\, F\, F_{\dot{x}^i}$$

$$F^2_{\dot{x}^i\dot{x}^j} = 2\,F\,F_{\dot{x}^i\dot{x}^j} + 2\,F_{\dot{x}^j}\,F_{\dot{x}^i}$$

$$\tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}\,\dot{x}^i = F\,F_{\dot{x}^i\dot{x}^j}\,\dot{x}^i + F_{\dot{x}^j}\,F_{\dot{x}^i}\,\dot{x}^i$$

$$\tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}\,\dot{x}^i = F\cdot(0) + F_{\dot{x}^j}\,F = F_{\dot{x}^j}\,F$$

$$\tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}\,\dot{x}^i\dot{x}^j = F_{\dot{x}^i}\,\dot{x}^j\,F = F^2$$

But by definition $g_{ij}(x,\dot{x}) = \tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}(x,\dot{x})$

Therefore $g_{ij}(x,\dot{x})\,\dot{x}^i\dot{x}^j = F^2(x,\dot{x})$.

Lemma 3b. $F_{\dot{x}^i}\,F = \tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}\,\dot{x}^j$

$$F^2_{\dot{x}^i\,\dot{x}^j}(x,\dot{x}) = \frac{\partial}{\partial\dot{x}^j}\left[F^2_{\dot{x}^i}(x,\dot{x})\right]$$

$$= 2\,\frac{\partial}{\partial\dot{x}^j}\,F_{\dot{x}^i}\,(x,\dot{x}) = 2\,F_{\dot{x}^i\,\dot{x}^j}\,(x,\dot{x})$$

$$\frac{\partial}{\partial\dot{x}^j}\,F^2_{\dot{x}^i}\,(x,\dot{x}) = \frac{\partial}{\partial\dot{x}^j}\left[\frac{\partial}{\partial\dot{x}^i}\,F(x,\dot{x})\,F(x,\dot{x})\right] =$$

$$\frac{\partial}{\partial\dot{x}^j}\left[F(x,\dot{x})\,F_{\dot{x}^i}(x,\dot{x}) + F_{\dot{x}^i}\,(x,\dot{x})\,F\,(x,\dot{x})\right] =$$

$$2\left[F(x,\dot{x})\,F_{\dot{x}^i\dot{x}^j}\,(x,\dot{x}) + F_{\dot{x}^i}(x,\dot{x})\,F_{\dot{x}^i}\,(x,\dot{x})\right]$$

$$\tfrac{1}{2}\,F^2_{\dot{x}^i\dot{x}^j}\,(x,\dot{x}) = F(x,\dot{x})\,F_{\dot{x}^i\dot{x}^j}(x,\dot{x}) + F_{\dot{x}^i}(x,\dot{x})\,F_{\dot{x}^j}\,(x,\dot{x})$$

$$\tfrac{1}{2}\,F^2_{\dot{x}^i\,\dot{x}^j}\,(x,\dot{x})\,\dot{x}^j = 0 + F_{\dot{x}^i}(x,\dot{x})\,F\,(x,\dot{x})$$

by lemmas 2 and 1.

$$\tfrac{1}{2} F^2_{\dot{x}^i\dot{x}^j}(x,\dot{x})\,\dot{x}^j = F_{\dot{x}^i}(x,\dot{x})\,F(x,\dot{x})$$

Lemma 4. $F_{\dot{x}^i\dot{x}^j} V^iV^j = 0$ if $V^i = K\dot{x}^i$.

If $V^i = K\dot{x}^i$, $F_{\dot{x}^i\dot{x}^j} V^iV^j = F_{\dot{x}^i\dot{x}^j} K\dot{x}^i V^j =$

$$KV^jF_{\dot{x}^i\dot{x}^j}\dot{x}^i .$$

The last factor is zero by property 2 above.

Therefore, if $V^i = K\dot{x}^i$, then $F_{\dot{x}^i\dot{x}^j} V^iV^j = 0$.

Lemma 5. $F_{\dot{x}^i\dot{x}^j} V^iV^j > 0$ if $V^i \neq K\dot{x}^i$

$F^2_{\dot{x}^i\dot{x}^j}(x,\dot{x})\,(y^i + KW^i)\,(y^j + KW^j) > 0$ if $y^i + KW^i$ are not all zero. This is by the positive-definite postulate. The equation $F^2_{\dot{x}^i\dot{x}^j} y^i y^j +$

$$2 F^2_{\dot{x}^i\dot{x}^j} KW^iy^j + K^2W^iW^j F^2_{\dot{x}^i\dot{x}^j} = 0 \text{ is}$$

quadratic in K and has no real roots if $y^i + KW^i \neq 0$. Its discriminate is negative.

$$4(F^2_{\dot{x}^i\dot{x}^j} W^iy^j)^2 < 4\,(F^2_{\dot{x}^i\dot{x}^j}W^iW^j)\,(F^2_{\dot{x}^K\dot{x}^i}y^Ky^i)$$

$\tfrac{1}{2} F^2_{\dot{x}^i\dot{x}^j} = F_{\dot{x}^i}F_{\dot{x}^j} + F\,F_{\dot{x}^i\dot{x}^j}$ from third step of proof 3 above.

$\tfrac{1}{2} F^2_{\dot{x}^i\dot{x}^j}\,\dot{x}^i\dot{x}^j = F^2$ from sixth step of proof 3 above.

$F_{\dot{x}^i} = \left(\frac{1}{2} F^2_{\dot{x}^i \dot{x}^j} \dot{x}^j\right) \frac{1}{F}$ from fifth step of proof 3 above

$$F_{\dot{x}^i} F_{\dot{x}^k} = \frac{\frac{1}{4} F^2_{\dot{x}^i \dot{x}^j} \dot{x}^J \quad F^2_{\dot{x}^K \dot{x}^L} \dot{x}^L}{F^2}$$

$$\text{Thus, } (\tfrac{1}{2})F^2_{\dot{x}^i \dot{x}^K} = \frac{\frac{1}{4} F^2_{\dot{x}^i \dot{x}^j} \dot{x}^J F^2_{\dot{x}^K \dot{x}^L} \dot{x}^L}{F^2} + F\, F_{\dot{x}^i \dot{x}^K}$$

$$F_{\dot{x}^i \dot{x}^K} = \left[(\tfrac{1}{2})F^2_{\dot{x}^i \dot{x}^K} - \frac{\frac{1}{4} F^2_{\dot{x}^i \dot{x}^j} \dot{x}^J F^2_{\dot{x}^K \dot{x}^L} \dot{x}^L}{F^2}\right](1/F)$$

$$= (\tfrac{1}{2F})\left[F^2_{\dot{x}^i \dot{x}^j} - (\tfrac{1}{2F^2})(F^2_{\dot{x}^i \dot{x}^K} \dot{x}^K)^2\right]$$

$$F_{\dot{x}^i \dot{x}^k} V^i V^k = \frac{1}{2F}\left[F_{\dot{x}^i \dot{x}^k} V^i V^k - \frac{1}{2F^2} \quad (F^2_{\dot{x}^i \dot{x}^k} \dot{X}^k \; V^i)^2\right]$$

$(F^2_{\dot{x}^i \dot{x}^j} \; \dot{x}^i V^j)^2 < (F^2_{\dot{x}^i \dot{x}^j} \; \dot{x}^i \dot{x}^j) \; (F^2_{\dot{x}^K \dot{x}^L} \; V^K V^L)$ if

$V^i + kx^i$ are not all zero.

$$(F^2_{\dot{x}^i \dot{x}^j} \; \dot{x}^i V^j)^2 < (2F^2) \; (F^2_{\dot{x}^K \dot{x}^L} \; V^K V^L)$$

$$(\tfrac{1}{2F^2}) \; (F^2_{\dot{x}^i \dot{x}^j} \; \dot{x}^i V^j)^2 < F^2_{\dot{x}^K \dot{x}^L} \; V^K V^L$$

Therefore $(\frac{1}{2F})\left[F^2_{\dot{x}^K \dot{x}^L} \; V^K V^L - \frac{1}{2F^2}(F^2_{\dot{x}^i \dot{x}^j} \; \dot{x}^i V^j)^2\right] > 0$

I.e., $F_{\dot{x}^i \dot{x}^K} V^i V^K > 0$ if $V^I \neq K\dot{x}^i$

<u>Lemma</u> 6. $\frac{\partial g_{ij}}{\partial \dot{x}^K} \dot{x}^K = 0$

$g_{ij}(x,\dot{x}) = \frac{1}{2}F^2_{\dot{x}^i \dot{x}^j} (x,\dot{x})$ by definition

$$\frac{\partial (g_{ij})}{\partial \dot{x}^K} = \frac{1}{2} \frac{\partial^3 F^2}{\partial \dot{x}^i \partial \dot{x}^j \partial \dot{x}^K}$$

$$\frac{\partial (g_{jK})}{\partial \dot{x}^i} = \frac{1}{2} \frac{\partial^3 F^2}{\partial \dot{x}^j \partial \dot{x}^K \partial \dot{x}^i} = \frac{1}{2} \frac{\partial^3 F^2}{\partial \dot{x}^i \partial \dot{x}^j \partial \dot{x}^K} =$$

$$\frac{\partial (g_{ij})}{\partial \dot{x}^K}$$

Similarly $\frac{\partial (g_{Ki})}{\partial \dot{x}^j} = \frac{\partial (g_{ij})}{\partial \dot{x}^K} = \frac{\partial (g_{jK})}{\partial \dot{x}^i}$

$F(x,\dot{x})$ is positively homogeneous of degree 1 in $\dot{x}^i$.

So F^2 is positively homogeneous of degree 2

$F^2_{\dot{x}^i}$ is positively homogeneous of degree 1

$F^2_{\dot{x}^i \dot{x}^j}$ is positively homogeneous of degree 0

Thus, g_{ij} is positively homogeneous of degree 0 in $\dot{x}^i$.

And $\frac{\partial g_{ij}}{\partial \dot{x}^k}$ is positively homogeneous of degree -1.

Therefore as in Lemma 2, $\frac{\partial g_{ij}}{\partial \dot{x}^K} \dot{x}^K = 0$

Differential Equations of Geodesics

The geodesics of a Finsler space are defined to be the extremal curves satisfying $d\int_{T_0}^{T_1} g_{ij}(x,\dot{x})\ \dot{x}^i\dot{x}^j\ dt = 0$.

The curves satisfying this condition are those that are solutions of the differential equations $\frac{d^2 x^i}{dS^2} + \Gamma^i_{jK}(x,\dot{x})$

$\frac{dx^j}{dS}\ \frac{dx^K}{dS} = 0$ where $\Gamma^i_{jK} = \frac{g^{ir}}{2}\left(\frac{\partial g_{ij}}{\partial x^K} - \frac{\partial g_{jK}}{\partial x^r} + \frac{\partial g_{rK}}{\partial x^j}\right)$.

To prove this assertion let $x^i = x^i(t)$ be a curve of class C^2 defined on $T_0 \leq T \leq T_1$ that is an extremum for the integral $\int_{T_0}^{T_1} F(x,\dot{x})\, dT$. Then, in comparison with all nearby curves $x^i(t) + \epsilon Y^i(t)$ for which $Y^i(T_0) = Y^i(T_1) = 0$. The integral $\int_{T_0}^{T_1} F\ dT$ must be a minimum when $\epsilon = 0$. The derivative of this integral with respect to ϵ must therefore be zero at $\epsilon = 0$. Here the functions $Y^i(T)$ are arbitrary class C^∞ functions.

$$\frac{d}{d\epsilon} \int_{T_0}^{T_1} F(x+\epsilon y,\ \dot{x}+\epsilon \dot{y})\ dT \Big|_{\epsilon=0} = 0$$

Since F is uniformly continous, being of class C^2 on the closed interval $[T_0, T_1]$, $\int_{T_0}^{T_1} \frac{d}{d} F(x+\epsilon y, \dot{x}+\epsilon\dot{y})\Big|_{\epsilon=0} dt=0$

$$\text{I.e., } \int_{T_0}^{T_1}\left(\frac{\partial F}{\partial x^i} Y^i + \frac{\partial F}{\partial \dot{x}^i}\dot{Y}^i\right) dt = 0$$

$$\text{But } \int_{T_0}^{T_1}\left(\frac{\partial F}{\partial x^i} Y^i + \frac{\partial F}{\partial \dot{x}^i}\dot{Y}^i\right) dt = \int_{T_0}^{T_1} \frac{\partial F}{\partial x^i} Y^i\, dt + \int_{T_0}^{T_1} \frac{\partial F}{\partial \dot{x}^i}\dot{Y}^i\ dt$$

Let $u = \frac{\partial F}{\partial \dot{x}^i}$ and $dv = \dot{y}^i \; dt$

Then $du = \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right) dt$ and $v = \int \dot{y}^i \, dt = \int \frac{dy^i}{dt} \, dt = y^i$

Thus, $\int_{T_0}^{T_1} \frac{\partial F}{\partial \dot{x}^i} \dot{y}^i \; dt = \frac{\partial F}{\partial \dot{x}^i} y^i \Big]_{T_0}^{T^1} -$

$$\int_{T_0}^{T_1} y^i \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right) dT$$

Since $y^i(T_1) = y^i(T_0) = 0$, $\int_{T_0}^{T_1} \frac{\partial F}{\partial \dot{x}^i} \dot{y}^i dT = -$

$$\int_{T_0}^{T_1} y^i \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right) dt$$

Therefore $\int_{T_0}^{T_1} \left(\frac{\partial F}{\partial x^i} y^i + \frac{\partial F}{\partial \dot{x}^i} \dot{y}^i\right) dt =$

$$\int_{T_0}^{T_1} \left[\frac{\partial F}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right)\right] y^i \; dt = 0$$

The Fundamental Lemma for the Calculus of Variations is that if R is a region in an n-dimensional space and if f is a continuous function in R with the property that $\int_R f \cdot g = 0$ for every function g of class C in R and vanishing on the boundary of R, then f vanishes identically in R.

$f = \frac{\partial F}{\partial x^i} - \frac{d}{dt} \frac{\partial F}{\partial \dot{x}^i}$ and $g = y^i$ satisfy the hypotheses of the lemma for each value of i in the set 1,2, ... n .

Therefore $\frac{\partial F}{\partial x^i} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i} = 0$. These are called the Euler-Lagrange equations. They are the differential e - quations of the extremal curves for the integral

$$\int_{T_0}^{T_1} F(x,\dot{x})\, dt.$$

To show that they take the form $\frac{d^2x^i}{dj^2} + \Gamma^i_{jK}(x,\dot{x})$

$$\frac{d\dot{x}^j}{dS}\frac{d\dot{x}^k}{dS} = 0,$$

$$F = (g_{ij}\,\dot{x}^i\dot{x}^j)^{1/2}$$

$$\frac{\partial F}{\partial x^i} = \left[\tfrac{1}{2}(g_{ij}\dot{x}^i\dot{x}^j)^{-1/2}\right]\frac{d}{dx^i}(g_{ij}\dot{x}^i\dot{x}^j) = \frac{\frac{\partial g_{Kj}}{\partial X^i}\dot{x}^K\dot{x}^j}{2(g_{ij}\,\dot{x}^i\dot{x}^j)^{\frac{1}{2}}}$$

$$= \frac{1}{2F}\frac{\partial g_{Kj}}{\partial x^i}\dot{x}^K\dot{x}^j$$

$$\frac{\partial F}{\partial \dot{x}^i} = \frac{g_{ij}\dot{x}^j + \dot{x}^i\frac{\partial}{\partial x^i}(g_{ij}\dot{x}^j)}{2(g_{ij}\,x^ix^j)^{\frac{1}{2}}} = \frac{g_{ij}\dot{x}^j + \dot{x}^ig_j + 0}{2F}$$

$$= \frac{g_{ij}\dot{x}^j}{F}$$

Choose arc length as parameter t. The integrand $F(x,\dot{x})$ is then equal to 1 and $\frac{\partial F}{\partial \dot{x}^i} = g_{ij}\dot{x}^j$.

$$\frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i} = \frac{d}{dt}(g_{ij}\dot{x}^j) = g_{ij}\frac{d\dot{x}^j}{dt} + \dot{x}^j\frac{d(g_{ij})}{dt} = g_{ij}\ddot{x}^j$$

$$+ \frac{dg_{ij}}{dt}\dot{x}^j$$

But $\frac{dg_{ij}}{dt} = \frac{\partial g_{ij}}{\partial \dot{x}^K}\frac{\partial \dot{x}^K}{\partial t} = \frac{\partial g_{ij}}{\partial \dot{x}^K}\dot{x}^K$

$$\frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i} = g_{ij}\ddot{x}^j + \frac{\partial g_{ij}}{\partial x^K}\dot{x}^K\dot{x}^j$$

The Euler-Lagrange equations $\frac{\partial F}{\partial x^i} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i} = 0$ become

$$\frac{1}{2}\frac{\partial g_{Ki}}{\partial x^i}\dot{x}^K\dot{x}^J - g_{ij}\ddot{x}^j - \frac{\partial g_{ij}}{\partial x^K}\dot{x}^K\dot{x}^j = 0$$

$$g_{ij}\ddot{x}^j + \left(\frac{\partial g_{ij}}{\partial x^K} - \frac{1}{2}\frac{\partial g_{Kj}}{\partial x^i}\right)\dot{x}^K\dot{x}^j = 0$$

$$g_{ij}\ddot{x}^j + \frac{1}{2}\left(2\frac{\partial g_{ij}}{\partial x^K} - \frac{\partial g_{Kj}}{\partial x^i}\right)\dot{x}^K\dot{x}^j = 0$$

$$g_{ij}\ddot{x}^j + \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^i}\right)\dot{x}^k\dot{x}^j = 0$$

Define g^{Li} by $g^{Li}g_{ij} = \delta^L_j$

Then $g^{Li}(g_{ij}\ddot{x}^j) + \frac{1}{2}g^{Li}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{kj}}{\partial i}\right)\dot{x}^k\dot{x}^j = 0$

$$\ddot{x}^j + \frac{g^{Li}}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{kj}}{\partial x^i}\right)\dot{x}^k\dot{x}^j = 0$$

$$\frac{d^2x^j}{ds^2} + \Gamma^i_{jk}(x,\dot{x})\frac{dx^k}{ds}\frac{dx^j}{ds} = 0 \text{ where } \Gamma^i_{jk} = \frac{g^{Li}}{2}\left(\frac{\partial g_{ij}}{\partial x^k} + \frac{\partial g_{ik}}{\partial x^j} - \frac{\partial g_{kj}}{\partial x^i}\right)$$

The geodesics in the Finsler space are the curves that satisfy these n differential equations. In the case of interest in this chapter, n = 2, and so there are only two differential equations since each of the indices involved take on the two values 1 and 2.

The Unit Vector

Let the fundamental function for a two-dimensional Finsler manifold be given by $F(x,y,\dot{x},\dot{y})$. Then the arc length S on any curve of the manifold is defined by $dS = F(x,y,\dot{x},\dot{y})dt$ where $\dot{x} = \frac{dx}{dt}$ and $\dot{y} = \frac{dy}{dt}$.

Denoting $\frac{dx}{dS}$ by x' and $\frac{dy}{dS}$ by y' and using the fact that $\frac{dx}{dS} = \frac{dx}{dt} \div \frac{dS}{dt}$, it is clear that $x' = \frac{\dot{x}}{F(x,y,\dot{x},\dot{y})}$ and $y' = \frac{\dot{y}}{F(x,y,\dot{x},\dot{y})}$.

Recall that the length of a vector $(\dot{x},\dot{y})$ in the tangent space at $P(x,y)$ is given by the value of $F(x,y,\dot{x},\dot{y})$. Here the $\dot{x},\dot{y}$ are affine coordinates of the point in $T_n(P)$ denoting a vector whose origin is at the fixed point (x,y). In other words the vector $(\dot{x},\dot{y})$ in $T_2(P)$ having the same direction as the ray determined by $(x,y,\dot{x},\dot{y})$ has as its length $F(x,y,\dot{x},\dot{y})$.

Therefore the vector (x',y') has as its length $k(x,y,x',y') = F(x,y,\frac{\dot{x}}{F}\,\frac{\dot{y}}{F}) = \frac{1}{F}(x,y,\dot{x},\dot{y}) = 1$. This unit vector with components $\frac{\dot{x}}{F}$ and $\frac{\dot{y}}{F}$ will be called the unit vector in the direction of $(x,y,\dot{x},\dot{y})$ or the unit vector of the line element $(x,y,\dot{x},\dot{y})$. In particular, the unit vector of the line element (x,y,x',y') is the vector $(\frac{x'}{F},\frac{y'}{F})$ to be denoted by (ℓ^1, ℓ^2).

In the functions $g_{ij}(x,\dot{x})$ defined in Chapter II the indices take on the values 1 and 2 for the case presently under consideration and the notation $(x^1,x^2,\dot{x}^1,\dot{x}^2)$ has now been replaced by $(x,y,\dot{x},\dot{y})$.

For each point $(\dot{x},\dot{y})$ in $T_2(P)$ one can associate a pair of real numbers (u,v) in various ways, one of which is by the transformation $u = g_{11}\dot{x} + g_{12}\dot{y}$, $v = g_{21}\dot{x} + g_{22}\dot{y}$. In the original notation this association would be expressed by $u_i = g_{ik}(x,\dot{x})\dot{x}^k$. Given a fixed point $(\dot{x},\dot{y})$ in $T_2(P)$, the functions g_{ik} are completely determined and consequently the pair (u,v) is a pair of real numbers. The number pair determined in this manner for the unit vector (ℓ^1, ℓ^2) is denoted by (ℓ_1,ℓ_2). Thus, $\ell_1 = g_{11}\frac{x'}{F} + g_{12}\frac{y'}{F}$ and $\ell_2 = g_{21}\frac{x'}{F} + g_{22}\frac{y'}{F}$ which could be indicated, of course, by

$\ell_i = g_{ik}\,\ell^k$ $(i = 1,2;\ k = 1,2)$. This leads to the relationship $\ell_i = g_{ik}\frac{x'^k}{F} = \frac{g_{ik}x'^k}{F} = \frac{\partial F}{\partial x'^i}$ as shown by lemma 4 of section 1.

The vector (ℓ^1,ℓ^2) is a unit vector in the sense that $\ell_1\ell^1 + \ell_2\ell^2 = 1$.

For $\ell_1\ell^1 + \ell_2\,\ell^2 = \left(g_{11}\frac{x'}{F} + g_{12}\frac{y'}{F}\right)\frac{x'}{F} +$

$$\left(g_{21}\frac{x'}{F} + g_{22}\frac{y'}{F}\right)\frac{y'}{F} = \frac{g_{11}x'^2 + 2g_{12}x'y' + g_{22}y'^2}{F^2}$$

$$= \frac{g_{ik}x'^i x'^k}{F^2} = 1 \qquad (i = 1,2;\ k = 1,2)$$

It was for this reason that the pair of real numbers chosen to correspond to a given vector (x,y) in $T_2(P)$ was chosen in the manner prescribed. In the case of a Euclidean 2-dimensional vector space $\ell_i = \ell^i$, and in order to be a unit vector $(\ell^1)^2 + (\ell^2)^2$ must equal 1. But in the case of Euclidean plane $g_{11} = g_{22} = 1$. $g_{12} = g_{21} = 0$ and $F = \sqrt{x'^1 + y'^2}$. Thus, the vector $(\frac{x'}{F}, \frac{y'}{F})$ is simply the vector $\left(x'/\sqrt{x'^2 + y'^2},\ y'/\sqrt{x'^2 + y'^2}\right)$ which is clearly a unit vector.

In general vector spaces not necessarily Euclidean the set of pairs of real numbers (x_1, y_1) associated with the vectors (x', y') plays an important role.

The Function F_1

Lemma 2 of section 1 takes the following form in the case of n =2: $\dot{x}F_{\dot{x}\dot{x}} + \dot{y}F_{\dot{x}\dot{y}} = 0$, $\dot{x}F_{\dot{y}\dot{x}} + \dot{y}F_{\dot{y}\dot{y}} = 0$. Thus

$$\frac{\dot{x}F_{\dot{x}\dot{x}}}{\dot{y}^2} + \frac{F_{\dot{x}\dot{y}}}{\dot{y}} = 0 \text{ and } \frac{F_{\dot{y}\dot{x}}}{\dot{y}} + \frac{F_{\dot{y}\dot{y}}}{\dot{x}} = 0 \text{ for}$$

$$\dot{x} \neq 0,\ \dot{y} \neq 0.$$

Hence
$$\frac{F_{\dot{x}\dot{x}}}{\dot{y}^2} = \frac{-F_{\dot{x}\dot{y}}}{\dot{x}\dot{y}} = \frac{F_{\dot{y}\dot{y}}}{\dot{x}^2}$$

Let F_1 be the common value of these three ratios.

Recall now that in the deduction of the differential equations for geodesics, use was made of the fundamental lemma for the calculus of variations. In this lemma the continuous function f was taken to be $\frac{\partial F}{\partial x^i} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i}$ and the conclusion was that $\frac{\partial F}{\partial x^i} - \frac{d}{dt}\frac{\partial F}{\partial \dot{x}^i} = 0$. F is $F(x,y,\dot{x},\dot{y}) = F\left(x(t),y(t),\dot{x}(t),\dot{y}(t)\right)$. Therefore

$$F_x = \frac{d}{dt}(F_{\dot{x}}) = \frac{d}{dt}\left(F_{\dot{x}}(x(t),y(t),\dot{x}(t),\dot{y}(t))\right)$$

$$= F_{\dot{x}x}\frac{dx}{dt} + F_{\dot{x}y}\frac{dy}{dt} + F_{\dot{x}\dot{x}}\frac{d\dot{x}}{dt} + F_{\dot{x}\dot{y}}\frac{d\dot{y}}{dt}$$

$$= F_{\dot{x}x}\dot{x} + F_{\dot{x}y}\dot{y} + F_{\dot{x}\dot{x}}\ddot{x} + F_{\dot{x}\dot{y}}\ddot{y} .$$

Similarly $F_y = F_{\dot{y}x}\dot{x} + F_{\dot{y}y}\dot{y} + F_{\dot{y}\dot{x}}\ddot{x} + F_{\dot{y}\dot{y}}\ddot{y}$.

Since $F(x,y,\dot{x},\dot{y})$ is homogeneous of degree one in $\dot{x},\dot{y}$, F_x is also homogeneous of degree one in $\dot{x},\dot{y}$. Thus, by Euler's Theorem on homogeneous functions $\dot{x}F_{x\dot{x}} + \dot{y}F_{x\dot{y}} = F_x$. $F_{x\dot{x}} = F_{\dot{x}x}$ by the continuity assumptions. Making these substitutions in the equation $F_x = F_{\dot{x}x}\dot{x} + F_{\dot{x}y}\dot{y} + F_{\dot{x}\dot{x}}\ddot{x} + F_{\dot{x}\dot{y}}\ddot{y}$, the result is $\dot{y}F_{x\dot{y}} = F_{\dot{x}y}\dot{y} + F_{\dot{x}\dot{x}}\ddot{x} + F_{\dot{x}\dot{y}}\ddot{y}$ or

$$(F_{\dot{x}y} -- F_{x\dot{y}})\dot{y} + F_{\dot{x}\dot{x}}\ddot{x} + F_{\dot{x}\dot{y}}\ddot{y} = o.$$

By the definition of F_1, $F_{\dot{x}\dot{x}} = F_1\dot{y}^2$ and $F_{\dot{x}\dot{y}} = -\dot{x}\dot{y}F_1$.

Hence $(F_{\dot{x}y} - F_{x\dot{y}})\dot{y} + F_1\dot{y}^2\ddot{x} - \dot{x}\dot{y}F_1\ddot{y} = 0$

I.e., $(F_{\dot{x}y} - F_{x\dot{y}}) + F_1(\ddot{x}\dot{y} - \dot{x}\ddot{y}) = 0$ if $\dot{y} \neq 0$.

In the calculus of variations this equality is called the Weierstrass symmetric form of Euler's equation for the parametric problem. It was used by Berwald, Moor and others to aid in the study of certain invariants of a two-dimensional Finsler space. Unfortunately, their study is beyond the scope of this paper. Suffice it to note at this stage one elementary, but important relationship:

$$g_{ik} = \tfrac{1}{2} \frac{\partial^2 F^2}{\partial \dot{x}^i \, \partial \dot{x}^k} \text{ where now } i = 1,2;\ k = 1,2;\ x^1 = x;$$

$$x^2 = y.$$

$$g_{11} = \tfrac{1}{2} \frac{\partial^2 F^2}{\partial \dot{x}\, \partial \dot{x}} = \tfrac{1}{2} \frac{\partial}{\partial \dot{x}} (F^2_{\dot{x}}) = \tfrac{1}{2} \frac{\partial}{\partial \dot{x}} (2FF_{\dot{x}})$$

$$= FF_{\dot{x}\dot{x}} + F_{\dot{x}}F_{\dot{x}} = FF_1\dot{y}^2 + (F_{\dot{x}})^2$$

$$g_{12} = \tfrac{1}{2} \frac{\partial^2 F^2}{\partial \dot{x} \partial \dot{y}} = FF_{\dot{x}\dot{y}} + F_{\dot{y}}F_{\dot{x}} = -FF_1\dot{x}\dot{y} + F_{\dot{x}}F_{\dot{y}}$$

$$g_{21} = g_{12}$$

$$g_{22} = \tfrac{1}{2} \frac{\partial^2 F^2}{\partial \dot{y} \partial \dot{y}} = FF_{\dot{y}\dot{y}} + (F_{\dot{y}})^2 = FF_1\dot{x}^2 + (F_{\dot{y}})^2$$

$g_{11}g_{22} - g_{12}g_{22} = g_{11}g_{22} - g_{12}^{\,2}$ = The determinant of (g_{ij})

That is, the determinant of the matrix (g_{ij}) has this value, which in turn is equal to $(FF_1\dot{y}^2 + F^2_{\dot{x}})(FF_1\dot{x}^2 + F_{\dot{y}}^{\,2}) - (-FF_1\dot{x}\dot{y} + F_{\dot{x}}F_{\dot{y}})^2$. On performing the first obvious simplication this reduces to

$$FF_1\dot{y}^2F_{\dot{y}}^{\,2} + F_{\dot{x}}^{\,2}FF_1\dot{x}^2 + 2FF_1\,\dot{x}\dot{y}F_{\dot{x}}F_{\dot{y}} = F_1F(\dot{x}F_{\dot{x}} + \dot{y}F_{\dot{y}})^2 =$$

F_1FF^2 by Euler's Theorem. Consequently, $|g_{ij}| = g_{11}g_{22} - g_{12}^{\,2} = F^3F_1.$

Orthogonality

The unit vector in the direction of the line element $(x, \dot{x})$ at the point (x,y) has been defined to be the vector (ℓ^1, ℓ^2) where $\ell^1 = \frac{x'}{F}$ and $\ell^2 = \frac{y'}{F}$. The aim now is to find a vector that is a unit vector and that may be considered orthogonal to (ℓ^1, ℓ^2). If (h^1, h^2) is to be such a vector, then as in the previous section, $h_1h^1 + h_2h^2$ must equal 1 to satisfy the unity requirement.

The notion of orthogonality from Euclidean 2-space becomes satisfied in general vector spaces by the requirement that $\ell^1 h_1 + \ell^2 h_2 = 0$.

Let $h_1 = -\sqrt{F^3F_1}\ \ell^2$, $h_2 = \sqrt{F^3F_1}\ \ell^1$

$$h^1 = \frac{-1}{\sqrt{F^3F_1}}\ \ell_2, \quad h^2 = \frac{1}{\sqrt{F^3F_1}}\ \ell_1$$

$h_1h^1 + h_2h^2 = \ell^2\ell_2 + \ell^1\ell_1 = 1$ since (ℓ^1, ℓ^2) is a unit vector.

$$\ell^1 h_1 + \ell^2 h_2 = -\sqrt{F^3F_1}\ \ell^1\ell^2 + \sqrt{F^3F_1}\ \ell^1\ell^2 = 0.$$

Thus, the vector (h^1, h^2) as defined here may be taken to be a unit vector orthogonal to the previously defined vector (ℓ^1, ℓ^2).

Entities such as $\ell^i h_i$ are called invariants. Among the many invariants for Finsler two-dimensional geometry the most important happen to be two scalars for which more background is needed. Their importance makes it essential, however, that they be mentioned here. These are the scalar or curvature k and the principal scalar I.

The first of these two is in a certain way a measure of the "flatness" of the space. It is so defined that if $k = 0$ in an affinely connected two-dimensional space. Then the space is simply the Minkowskian plane.[4]

The other is defined by $I = \frac{1}{2} A_{ijk}h^i h^j h^k$ where $A_{ijk} = \frac{1}{4} F \frac{\partial^3 F^2}{\partial \dot{x}^i \, \partial \dot{x}^j \, \partial \dot{x}^k}$. It has been shown by Moor[5] that a Finsler 2-dimensional space is simply a Riemannian 2-space if and only if $I = 0$.

The reason for interjecting the concepts of these two scalar invariants here is to stress the fact that two-dimensional Finsler geometry (like two-dimensional Riemannian geometry) is not necessarily the geometry of a plane. By means of the afore-mentioned invariants, however, one can ascertain when such is the case. Of course, in any case, the study of the two-dimensional surface is done mainly by way of the two-dimensional tangent space, which is a two-dimensional plane.

Having established the concept of orthogonality of the unit vector (h^1, h^2) with the specific unit vector (λ^1, λ^2) in the direction of the line element $(x, y, \dot{x}, \dot{y})$, mathemati-

---

[4] L. Berwald, "Finsler and Cartan Geometries, III," Ann. Math. (2), 42, pp. 84-112, 1941.

[5] Arthur Moor, "Generalization de Scalaire de Courbure," Canad. J. Math. 2, pp. 307-313, 1950.

cians interested in Finsler geometry next proceeded to an extension of this notion of orthogonality to the case where the given vector is not necessarily in the direction of the line element.

Consider the indicatrix $F(x,\dot{x}) = 1$ which in the two-dimensional case is the curve $F(x,y,\dot{x},\dot{y}) = 1$ on $T_2(p)$. Let $(\dot{x}_0,\dot{y}_0)$ be any arbitrary fixed point on this indicatrix. The equation of the tangent line to the indicatrix at the point $(\dot{x}_0, \dot{y}_0)$ is $F_{\dot{x}}(\dot{x} - \dot{x}_0) + F_{\dot{y}}(\dot{y} - \dot{y}_0) = 0$ where $(\dot{x},\dot{y})$ is a variable point on $T_2(p)$ and where $F_{\dot{x}}, F_{\dot{y}}$ are evaluated at $(\dot{x}_0,\dot{y}_0)$. In the general case this would be expressed as

$F_{\dot{x}^i}(\dot{x}^i - \dot{x}_0^i) = 0$

$$\text{Thus,} \quad F_{\dot{x}^i}(x,\dot{x}_0)\,(\dot{x}^i - \dot{x}_0^i) = 0$$

$$F_{\dot{x}^i}(x,\dot{x}_0)\dot{x}^i - F_{\dot{x}^i}(x,\dot{x}_0)\dot{x}_0^i = 0$$

$$F(x,\dot{x}_0) - F_{\dot{x}^i}(x,\dot{x}_0)\dot{x}_0^i = 0 \text{ by Lemma 1}$$

$$F(x,\dot{x}_0)F(x,\dot{x}_0) = F(x,\dot{x}_0)F_{\dot{x}^i}(x,\dot{x}_0)\dot{x}_0^i$$

$$F(x,\dot{x}_0)F(x,\dot{x}_0) = \tfrac{1}{2} F^2_{\dot{x}^i\dot{x}^j}(x,\dot{x}_0)\dot{x}^j\dot{x}_0^i \text{ by Lemma 3b}$$

$$1 = \tfrac{1}{2} F^2_{\dot{x}^i\dot{x}^j}(x,\dot{x}_0)\dot{x}^j\dot{x}_0^i \text{ since } F(x,\dot{x}_0) = 1$$

Hence, the equation of the line tangent to the indicatrix at the point (x_0,y_0) is $g_{ij}(x,\dot{x}_0)\dot{x}_0^i\dot{x}^j = 1$.

Now consider the locus $F(x,\dot{x}) = |\rho|$ in T_2p where ρ is

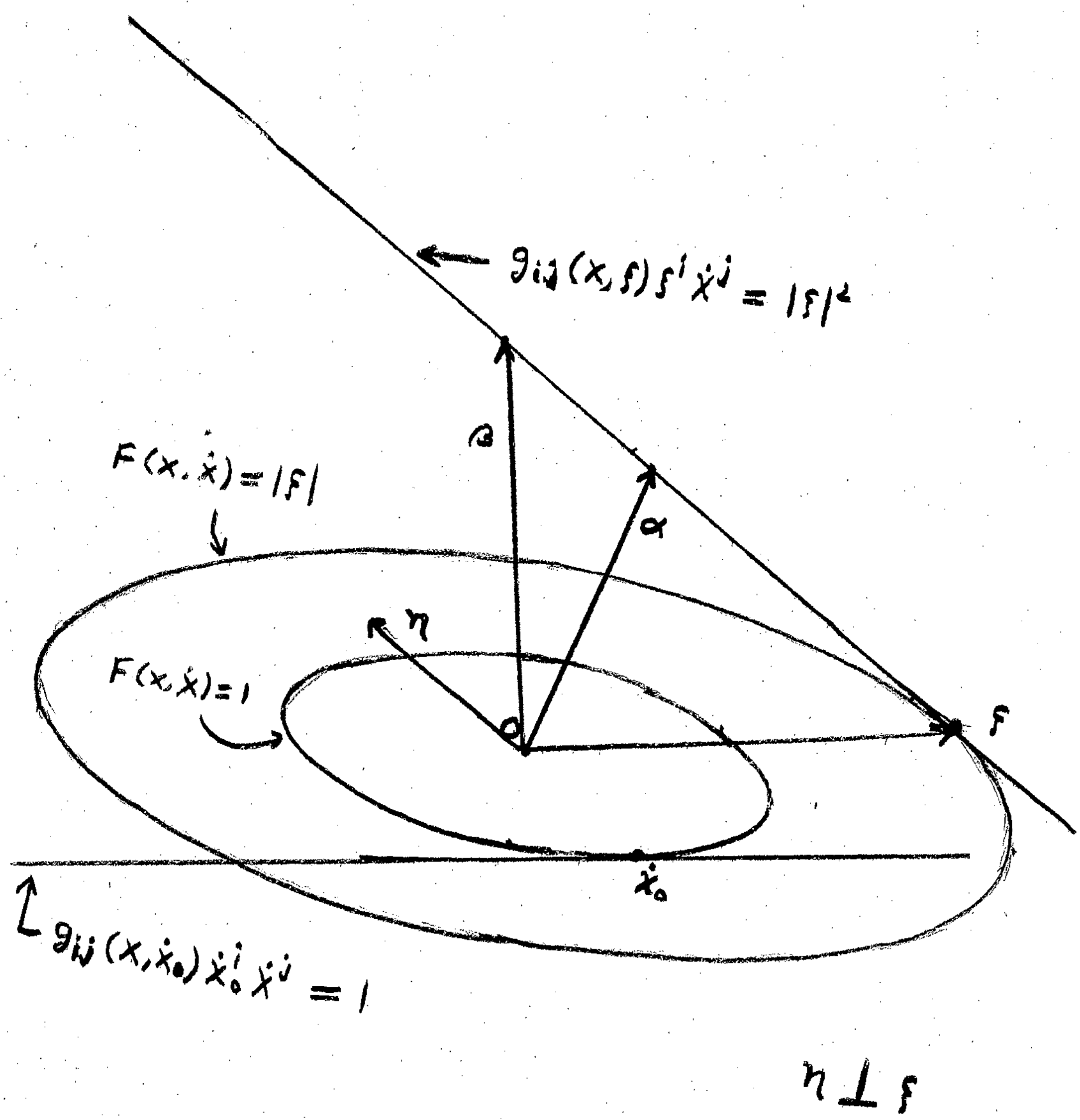

$g_{ij}(x,\xi)\xi^i \dot{x}^j = |\xi|^2$
β
α
$F(x,\dot{x}) = |\xi|$
η
$F(x,\dot{x}) = 1$
o
ξ
$\dot{x}_0$
$g_{ij}(x,\dot{x}_0)\dot{x}_0^i \dot{x}^j = 1$
$\eta \perp \xi$

Figure 5

an arbitrary vector of the space. The equation of this locus can be written as $F^2(x,\dot{x}) = |\xi|^2$ which is the same as $g_{ij}(x,\dot{x})\dot{x}^i\dot{x}^j = |\xi|^2$. Since ξ is a point on this locus, the equation of the line tangent to this locus at ξ is derived in precisely the same manner as for the tangent line to the indicatrix at x_0. Consequently, the equation of the line through ξ tangent to the curve $F(x,\dot{x}) = |\xi|$ at $\dot{x} = \xi$ is $g_{ij}(x,\xi)\xi^i\dot{x}^j = |\xi|^2$. See Figure 5.

Just as the indicatrix plays the part of the unit circle, this new locus $F(x,\dot{x}) = |\xi|$ plays the part of a circle of radius $|\xi|$. And just as in E_2 where any line parallel to a line tangent to a circle at a point p is considered orthogonal to the vector $\overrightarrow{op}$ (if O is the center), so too in $T_2(p)$ any line parallel to the line T tangent to $F(x,\dot{x}) = |\xi|$ at the point ξ is said to be normal with respect to the vector ξ. This is the definition of orthogonality.

Let η be any vector normal with respect to ξ. Then, by definition, η must lie on a line parallel to the line T given by the equation $g_{ij}(x,\xi)\xi^i\dot{x}^j = |\xi|^2$. This implies that $\eta = \beta - \alpha$ where β and α are vectors lying on T. But this means that $g_{ij}(x,\xi)\xi^i\eta^j = g_{ij}(x,\xi)\xi^i(\beta^j - \alpha^j)$

$= g_{ij}(x,\xi)\xi^i\beta^j - g_{ij}(x,\xi)\xi^i\alpha^j = |\xi|^2 - |\xi|^2 = 0.$

I.e., if η is normal with respect to ξ, then $g_{ij}(x,\xi)\xi^i\eta^j = 0$.

It is clear that this equation may be satisfied by vectors ξ and η even when the corresponding equation $g_{ij}(x,\eta)\eta^i\xi^j = 0$ does not hold. For obviously, g_{ij}

(x, ξ) may not equal $g_{ij}(x, \eta)$. In the Riemannian case $g_{ij}(x, \xi)$ is of necessity equal to $g_{ij}(x, \eta)$ because g_{ij} is independent of the direction $\dot{x}$.

Thus, a distinguishing feature of normality in Finsler geometry is that it is not symmetric. If the vector η is normal with respect to ξ, ξ may not be normal to η.

It should not be necessary to point out the fact that the orthogonality discussed at the beginning of the section was symmetric. At that stage the concern was with a unit vector in the direction of the line element. Thus, the line element was taken as fixed; $g_{ij}(x,\dot{x})\dot{x}^i\dot{x}^j = |\xi|^2$ where $(x,\dot{x})$ is fixed is the equation of ellipse; the geometry associated with such convex sets is Riemannian.

Area

Until recently, area on a Finsler manifold was investigated by making full use of the notion suggested in the preceeding paragraph It was defined with respect to an arbitrarily fixed line element.

Thus, the area of a parallelogram formed by the vectors (u,v) and $(\bar{u},\bar{v})$ was simply defined to be $\sqrt{g_{11}\, g_{22} - (g_{12})^2}\,(u\bar{v} - v\bar{u})$ where g_{ij} is taken with respect to the fixed line element $(x,y,\dot{x},\dot{y})$.

For the area of a general region of the surface. It was defined as

$$A = \iint_R \sqrt{g_{11}\, g_{22} - (g_{12})^2}\; dx\, dy$$

By previous computations, this may be written as

$$A = \iint \sqrt{F^3 F_1} \quad dx\,dy$$

In brief, area was defined for Finsler spaces in the way it is defined for Riemannian spaces. This is made possible only because of the selection of a particular line element.

This means that the area is defined with respect to this element and that a different choice for line-element results in a different amount of area.

In the 1950's, Busemann suggested a way to define intrinsic area independent of the choice of line-element. However, an understanding of the Busemann approach involves an understanding of measure - theory.

His conclusions are;

The study of Minkowskian geometry ought to be the first and main step to get to Finsler area----just as area is the first defined in Euclidean space and then extended to Riemannian.

Measure in Minkowski spaces is just as uniquely determined as in Euclidean spaces.

One can not reach a satisfactory notion of area for Finsler spaces by extending the methods of Riemannian geometry.

These three conclusions form the bases for a completely new look at Finsler geometry.

CHAPTER IV

ASPECTS OF GLOBAL GEOMETRY

At every point P of a differentiable manifold M there exists a tangent space. If a Finsler metric is defined on each tangent space, the geometry of the manifold in the neighborhood of the point P under consideration can be investigated. However, it must be stressed that the manifold is not a Finsler manifold unless the metric so assigned on each tangent space varies differentiably as x varies over M.

To investigate this aspect of the problem one is led to the study called the theory of connections. While such a study is not proper here (since it involves mathematics above the level of this work), it may be profitable to discuss the basis on which the study may be made.

Inner - product

It is an important fact that the set of all tangent vectors to a manifold at a point P actually comprises a set that can be given the structure of a vector space. By assigning a metric to this vector space one can approximate distances on the manifold in the neighborhood of P.

The assignment of a metric onto the tangent plane is, therefore, the problem of putting a metric on a vector space. For special vector spaces, such as the type of concern in this work, there is much help available from the norm-inner-product-quadratic-form relationships.

If V is an n-dimensional vector space over R, an inner product on V is a mapping of V x V into R having the properties that it is symmetric, positive-definite and bilinear. Let $\langle\ ,\ \rangle$ denote an inner product, let x,y and z be vectors in V, and k_1 and k_2 be elements of R. Then,

$$\langle x,x\rangle = 0 \text{ if } x = 0 \text{ and } \langle x,x\rangle > 0 \text{ if } x \neq 0$$

$$\langle k_1 x + k_2 y, z\rangle = k_1 \langle x,z\rangle + k_2 \langle y,z\rangle$$

$$\langle x, k_1 y + k_2 z\rangle = k_1 \langle x,y\rangle + k_2 \langle x,z\rangle$$

$$\langle x,y\rangle = \langle y,x\rangle$$

The quadratic form determined by the inner product is the function which assigns to each vector α the scalar $\langle\alpha,\alpha\rangle$, and this scalar is called $\|\alpha\|^2$, the square of the norm.

If G is any n x n matrix over R which satisfies $G = G^T$ and $X^T G X > 0$ if $X \neq 0$, then G is the matrix of some inner product on V. This inner product is defined by $\langle\alpha,\beta\rangle = Y^T G X$ where X and Y are coordinate matrices of vectors α and β with respect to some basis.

Therefore, given a symmetric, positive definite, nonsingular quadratic form, an inner product can always be determined. In the case of Riemann metrics this form involved the coefficients $g_{ij}(x)$ while in Finsler metrics it involved $g_{ij}(x,x)$. In either case at each point P of the manifold there is a tangent space on which is defined a particular inner product.

Restrictions on the Manifold

While at each point P of a differentiable manifold there exists a tangent space on which an inner product can be defined, the inner product on T_pM may not be the inner product on other tangent spaces. And since the inner product on $T_p(M)$ depends on P in that values of $F(x,\dot{x})$ are found by means of evaluations at P, there is nothing in what has been said to permit the comparison of such quantities evaluated at P with similar quantities evaluated at another point Q.

To apply properties of R^n in each local coordinate system on M has been one step; to piece these systems together in a meaningful way is another. This is a basic problem in a discussion of the manifold as a whole. If the manifold is to be of such type that distance between points is to exist , certain additional properties need be postulated. No piecing together of coordinate systems can cover the whole manifold if as a topological space the manifold is not connected originally. Similarly it is difficult to speak of distance between points if the manifold is such that the shortest path connecting two points does not exist. And again, what meaning could be attached to distance on a manifold from one point to another point when both points coincide but lie on different "sides".

In order to eliminate such problems, in the remainder of this chapter, it will be assumed that the manifolds under discussion are connected, orientable, and locally compact.

A set M is connected if and only if there exists no continuous mapping $f : M \rightarrow R$ such that f(M) consists of exactly two points.

M is locally compact if and only if, given any point P in M and any neighborhood U of P, there is a neighborhood V of P contained in U whose closure $\bar{V}$ is compact, i.e., every open covering of $\bar{V}$ contains a finite subcovering of $\bar{V}$.

M is orientable if there exists a covering U_α with mappings $\phi_\alpha : U_\alpha \rightarrow R$ such that all the differentiable homemorphisms $\phi_\beta \circ \phi_\alpha^{-1}$ (as defined in Chapter I) have positive jacobians.

Fibre Bundles

While the details necessary to prove statements made in the sequel must be left for advanced study, the plausibility of the statements from an intuitive point of view may be sufficient to present an idea of the nature of fiber bundles.

Consider an ordinary sphere S in Euclidean space E^3. This sphere is of course, a 2-dimensional differentiable manifold. Now consider the set T of all the non-zero vectors tangent to the sphere. How many measurements are necessary to distinguish any one element? Two must locate on the sphere the origin of a particular vector; second, one must ascertain its length. In brief, four measurements are essential to fix a given vector. For this reason the complete set T of non-zero vectors tangent to the sphere is itself a 4-dimensional manifold.

Let U be any open neighborhood of a point P on S. Since S is a differentiable manifold, there is a homeomorphism ϕ such that ϕ^{-1} maps an open set V of E^2 onto U. Let $\tilde{V}$ be the open set in E^4 which consists of all points (P_1, P_2, P_3, P_4) for which (P_1, P_2) is in V. Now define a new function θ which maps V into T in the following manner: $\theta(P_1, P_2, P_3, P_4) = P_3 \vec{x}_1 + P_4 \vec{x}_2$ where $\vec{x}_1$ and $\vec{x}_2$ are the unit tangent vectors to the coordinate curves of U at $\phi^{-1}(P_1,P_2)$. In this way a one-to-one correspondence is set up between all 4-tuples beginning with the ordered pair P_1,P_2 and all tangent vectors emanating from points in U.

Thus, associated with the original differentiable manifold S there can be constructed another topological manifold T. In the construction a very natural relation arises between certain sets of elements of T and points of S. The set of those elements of T which emanate from the same origin in S is one-to-one correspondence with the set of points S, since there is one and only one tangent plane at each point P. of S.

Another fact about the set of elements of T which have the same origin in S is that it is homeomorphic to any other set of elements of T with common origin in S. This is a restatement of the assertion that all the tangent planes to S are homeomorphic. Let F_0 denote any one fixed tangent plane. Let h_p be any homeomorphism mapping the tangent plane T_p onto F_0. Let k_p be any homeomorphism mapping F_0 onto T_p. Then $k_p \circ h_p$ is a transformation of F_0 onto itself, it is an automorphism. It could be a simple rotation of F_0 or, it could be a centered affine transformation of F_0. At any rate it is intuitively plausible that there is a group of automorphisms of F_0 which contains all such composite transformations as $k_x \circ h_x$ where h_x maps T_x onto F_0 homeomorphically and k_x maps F_0 onto T_x homeomorphically.

In this example three topological spaces are involved simultaneously: The sphere S, the 4-dimensional manifold T and the special tangent space F_0. There is also a relation between S and T given by the correspondence of all elements of T to their origins which are elements of S. And then there is a group structure given on F_0 and on each T_p.

The sphere S is called the base space; the manifold T is called the tangent bundle; the set F_0, which is homeomorphic to every T_p is called the fibre; the affine or rotation group of homeomorphisms of F_0 onto itself is called a transformation group in the fibre; every non-zero tangent vector to the sphere is a point in the tangent bundle; and the mapping of any such tangent vector into its origin is called a projection .

M is covered by a family $\{U_\alpha\}$ of neighborhoods. If

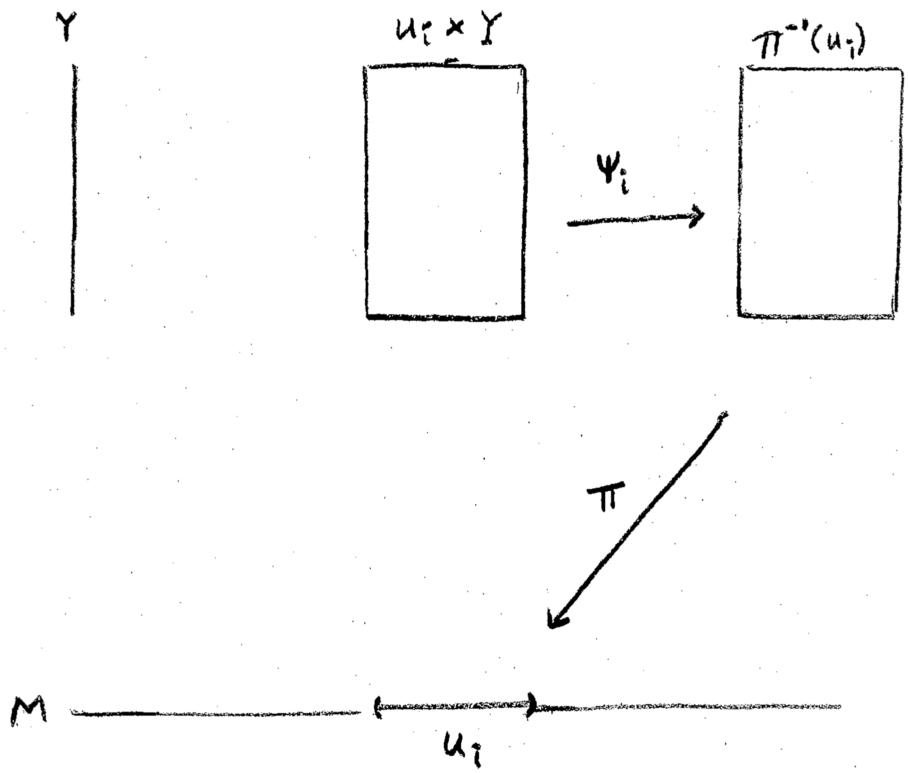


Figure 6

is one such neighborhood and $p \in U$, then $\pi^{-1}(p)$ must be T_p itself and should also be topologically the same as $p \times F_0$. The set $\pi^{-1}(U)$ should be topologically the same as the set formed by $U \times F_0$. And finally, any homeomorphism of $p \times F_0$ into itself is simply a homeomorphism of F_0 into itself.

Among these properties pertinent for a rigorous definition of a fibre bundle the topological equivalence of $\pi^{-1}(U)$ and $U \times F_0$ is most fundamental. It is the key to the essence of fibre bundles. Consider a curve on the sphere S. This curve can be fixed by knowing all the points x of S which lie on it. But for each point x the curve picks out a specific tangent vector y in $T_x(S)$. The set of all such pairs (x,y) should be a graph of the curve on the space of all tangent planes to the sphere. This idea leads to the idea of trying to create a sort of cartesian product of the two topological spaces g and T. If this could be done, then graphs of any function of points on S to vectors in T could be considered in much the same way as graphs of functions in the cartesian product. The fibre bundle is thus a generalization of the cartesian product - it may not be a cartesian product of the base space with the fibre, but locally it is precisely this.

At any rate if one uses the machinery of the fibre bundle, one may be able to ascertain whether a function from S into T is a differentiable function of x. Moreover, the same type of structure created by tangent bundles may hold for other types of bundles.

The definition of fibre bundle which follows is that given by S. S. Chern.[4]

A fibre bundle F is a topological space having the properties that:

1) There exists a continuous mapping π of F onto another topological space M.

2) There exists a family of neighborhoods which cover M and if U is a neighborhood of the family, the inverse image $\pi^{-1}(U)$ is a topological product. I.e., there exists a homeomorphism ψ_U depending on U such that $\psi_U(\pi^{-1}(U)) = U \times F_0$ and $\psi_U(\pi^{-1}(p)) = p \times F_0$ for every $p \in U$. F_0 is a definite topological space which is homeomorphic to $\pi^{-1}(p)$.

3) If U and V are two such neighborhoods of M and if $p \in U \cap V$, then the mapping $\psi_V(\psi_U^{-1}(p \times F_0)$ is a homeomorphism of F_0 into itself. This homeomorphism belongs to a group in F_0 given in advance.

Using the theory of fibre bundles, Auslander[5] was able to prove that a non-Riemannian Finsler metric could be defined on any differentiable manifold.

---

[4]S. S. Chern, "Some New Viewpoints in Differential Geometry in Large", Bull. Amer. Math. Soc., 52, 1946, pp. 7-8.

[5]Louis Auslander, "On Curvature in Finsler Geometry", Trans. Amer. Math. Soc., 79, 1955, pp. 378-381.

CHAPTER V

SUMMARY AND CONCLUSION

In the preparation of this brief introduction to the theory of Finsler manifolds three aims have been attempted. The first was to raise the question in the mind of the undergraduate mathematics major of the possibility of geometric beings not embedded in a preconceived Euclidean space. This idea is expressed today in the concept of the differentiable manifold.

The second aim was to raise the issue of generalized metric spaces. After so many years of early training in the use of the ordinary Euclidean metric, the fact that this metric is defined rather than natural is sometimes forgotten. The development of the basic notions regarding the Finsler metric was used in the dissertation to strengthen the awakening that metric geometry need not be Riemannian.

The third aim, to introduce the definition of the fibre bundle at the undergraduate level, was decided upon in the belief that the reason behind this complicated machinery can be appreciated even before the machinery itself is understood.

These concepts are not elementary, and there has been no attempt here to make them appear so.

Recall the basic ideas.

In the first place, a differential manifold is a topological space that is locally homeomorphic to a Euclidean n-dimensional sphere and that is endowed with a structure which enables concepts of differentiability to be defined by way of mappings in Euclidean space.

When the tangent space at each point of the manifold is given a metric, this local metric may be one which depends on direction. The Finsler metric is such. It can be defined on the tangent space whenever there is a function F(x,x) satisfying certain conditions. This metric associates an inner product to the tangent space.

For the study of global metric geometry and in order for the metric defined on the tangent space to be of use in this study it is necessary that the inner product vary differentiably over the manifold. The problem is therefore to find means to relate functions defined on one tangent space to similar functions on other tangent spaces. One way of doing this is through the creation of fibre bundles.

All of these ideas are of current value in present day geometry. That these topics can be taught at the undergraduate level in some honest form is an assumption upon which the significance of this dissertation rests. It is hoped that this introduction will lead some college teacher or his student by way of the accompanying bibliography to a study of the Finsler manifold.

BIBLIOGRAPHY

Auslander, Louis. "On Curvature in Finsler Geometry" Trans. Amer. Math. Soc., Vol. 79. 1955, pp. 378-388.

Auslander, Louis and Mackenzie, Robert. Introduction to Differentiable Manifolds, McGraw-Hill, New York, 1963.

Berwald, L. "Two Dimensional Finsler Spaces with Rectilinear Extremals", Annals of Math., Vol. 42, 1941 pp. 84-112.

Bliss, Gilbert. Lectures on the Calculus of Variations, University of Chicago Press, 1946

Busemann, Herbert. "The Foundations of Minkowskian Geometry". Com. Math. Helv., Vol. 24. 1950, pp.156-186

Busemann, Herbert. "The Geometry of Finsler Spaces", Bull. Amer. Math. Soc., Vol. 56. 1950, pp. 5-15

Chern, Shiing-Shen. "Some New Viewpoints in Differential Geometry in the Large". Bull. Amer. Math. Soc., Vol. 52, 1946. pp. 1-30

Finsler, Paul. Über Kurven and Flachen in Allgemeinen Raumen, Birkhauser-Basil, Switzerland, 1951

Goffman, Casper. Calculus of Several Variables. Harper and Row, New York, 1965

Hicks, Noel. Notes on Differential Geometry. Van Nostrand Mathematical Studies #3, Princeton, 1965

Laugwitz, Detlef. Differential and Riemannian Geometry, Academy Press, New York, 1965

Moor, Arthur. "Generalization de Scalaire de Courbure", Canad. J. Math. 2, 1950, pp. 307-313.

Riemann, Georg Bernhard. Gesammelte Mathematishe Werke, Leipzig, 1892, reproduced by Dover Pub., New York, 1953. pp. 272-287

Rund, Hanno. The Differential Geometry of Finsler Spaces. Springer-verlag, Berlin, 1959

Willmore, T. J. An Introduction to Differential Geometry. Oxford University Press, London, 1959

VITA

Thomas Edison Bonner

Candidate for the Degree of

Doctor of Education

Thesis: FINSLER MANIFOLDS

Major Field: Higher Education

Biographical:

Personal Data: Born in St. Louis, Missouri, September 29, 1921, the son of Gamaliel and Olivia Bonner.

Education: Attended elementary and secondary schools in St. Louis, Missouri; received the Bachelor of Science degree from Lincoln University (Missouri), with a major in Mathematics in June, 1943; received the Master of Arts degree from the University of Illinois, with a major in Education, in August, 1949; received the Master of Science degree from the University of Illinois, with a major in Mathematics, in June, 1953; attended Washington University for advanced work in mathematics during the academic year 1961 - 62; completed requirements for the Doctor of Education degree in July, 1967.

Professional experience: Taught mathematics in the Public School system of St. Louis, 1949 - 1951: was instructor in mathematics at Southern University, Baton Rouge, Louisiana, 1953 - 1955; was assistant professor of mathematics at Kansas State Teacher's College, Emporia, Kansas, 1960 - 1964.