HADAMARD'S THEORY OF GEODESICS
ON SURFACES OF NEGATIVE CURVATURE.

Neville C. Hunsaker

A Minor Thesis presented to the Faculty of the
Rice Institute in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

July 1941
Degree conferred, 1943.
I.

Introduction

We will say that a curve C on a surface S is a geodesic if at each point of C the osculating plane of C contains the normal line of S.

This definition admits at once a physical interpretation of a geodesic. Consider a particle moving on a surface under the influence of no forces other than pressure. The resultant acceleration of a particle moving on a surface lies always in the osculating plane, and since in this case the only acceleration is normal to the surface, this normal lies in the osculating plane. We conclude that the particle moves on a geodesic curve.

Before developing the subject further, we need to make more precise the meanings of the words "surface" and "curve".

We will suppose that our surface S is such that any finite portion of it may be divided into a finite number of regions in which the Cartesian coordinates of any point of S may be expressed in terms of functions with several derivatives of two independent parameters, say
\[ \chi = \chi(u, v), \quad y = y(u, v), \quad z = z(u, v) \]
where the Jacobians
\[ J_1 = J\left(\frac{\partial \chi}{\partial u, v}\right), \quad J_2 = J\left(\frac{\partial \chi}{\partial u, v}\right), \quad J_3 = J\left(\frac{\partial y, z}{\partial u, v}\right) \]
not all vanish identically. Moreover the set of singular points, that is to say the points where \( J_1, J_2, \) and \( J_3 \) vanish simultaneously, will be supposed null in the finite part of the plane.

If the parameters \( u \) and \( v \) are restricted to be functions of a third variable through the equations
\[
\begin{align*}
\mu &= \mu(t) \\ v &= v(t)
\end{align*}
\]
a curve is described on the surface. If the element of arc length, or linear element \( ds \) on this curve, is computed by
\[
\sqrt{d\tau^2} = \sqrt{d\xi^2 + d\eta^2 + d\zeta^2}
\]
one obtains
\[
\sqrt{d\tau^2} = \sqrt{E d\mu^2 + 2 F d\mu d\nu + G d\nu^2}
\]
where
\[
\begin{align*}
E &= \chi_u^2 + \gamma_v^2 + \zeta_{\mu}^2 \\
F &= \chi_u \chi_v + \gamma_u \gamma_v + \zeta_u \zeta_v \\
G &= \gamma^2 + \zeta^2
\end{align*}
\]
The use of Lagrange's identity yields
\[
J_1^2 + J_2^2 + J_3^2 = E G - F^2
\]
whence part of our hypothesis regarding the surface is that
\[
E G - F^2 \neq 0
\]

We will suppose further that the Gaussian curvature of the surface is negative except for at most a finite number of points where the curvature may be zero. The Gaussian curvature \( K \) at a point of a surface \( S \) is the product of the principal normal curvatures of \( S \) at the point, that is to
say the product of the maximum and minimum values of the normal curvatures at the point.

Perhaps the two best known surfaces of negative curvature are the hyperbolic paraboloid (a simply connected surface) and the hyperboloid of one sheet (order of connectivity, 2).

From the foregoing definition we may deduce a property of our surface, namely, a surface of negative curvature necessarily extends to infinity. Indeed suppose the equation of the surface is $Z = f(x, y)$. If for a certain $(x, y)$ $Z$ has a maximum value, the principal normal curvatures at this point cannot be oppositely directed, and hence the curvature is not negative. And if the curvature is zero at an isolated point this point cannot be a point of maximum.

By use of the definition of a geodesic at the beginning of this paper, one may obtain the differential equation of geodesics on a surface, namely: (1)

\[
2(EG-F^2)(\mu'' - \nu''\mu') = (E F_v + F F_{\mu} - 2E F_{\mu})(\mu')^3 + (3F E_v + G E_{\mu} - 2 F F_{\mu} - 2 E G_{\mu})(\mu'')^2(\nu')
- (3 F G_{\mu} + E G_v - 2 F F_v - 2 G E_v)\mu'(\nu'')^2 - (G G_{\mu} + F G_v - 2 G F_{\nu})(\nu')^3
\]

where the differentiation is with respect to any parameter along the curve.

From the fundamental existence theorem of differential equations it follows that one may prescribe arbitrarily the

value of the solution of the above equation at any point and also the value of the first derivative at that point. Thus we have the valuable conclusion: **Through every point on a surface there passes a unique geodesic with a given direction.**

It is interesting to note that if for curves \( u = u(t), \ v = v(t) \) on a surface one considers the calculus of variations problem of minimizing the integral (2)

\[
J = \int_{t_0}^{t_1} \sqrt{E \mu^2 + 2 F \mu^2 \nu' + G \nu'^2} \, dt
\]

that the equation of the extremals turns out to be precisely the differential equation given above. One may conclude that for two points sufficiently close together the shortest distance on the surface between the two points is on an arc of a geodesic curve.

A result similar to that above is the following: **About any point \( P \) there exists a region in which any point other than \( P \) can be joined to \( P \) by a geodesic arc lying entirely in the region (3).**

I will conclude the introduction to this paper by stating other facts well known in the theory of differential geometry which shall be useful in the subsequent developments.

A useful set of coordinates on the surface \( S \) arises by letting \( v \) represent the distance measured from any fixed

---

point on a geodesic \( L \) to a normal geodesic, and letting \( u \) represent the distance along this normal geodesic. In this system of coordinates the linear element becomes (4)
\[
d_s^2 = d\bar{u}^2 + C^2 d\bar{v}^2
\]
where \( C^2 = G \). Moreover the Gauss formula for curvature yields (5)
\[
\frac{\partial^2 C}{\partial u^2} = -\frac{C}{R R'}
\]
where \( R \) and \( R' \) are the radii of principal normal curvatures. For surfaces of negative curvature this last equation permits us to make the valuable conclusion that in this coordinate system \( C \) is an increasing function of \( u \).

Another theorem due to Gauss which we shall find to be useful is the following: Two geodesics on a surface of negative curvature cannot meet in two points and inclose a simply connected area (6).

Finally we state: If two points on a surface are such that only one geodesic passes through them, the segment of the geodesic measures the shortest distance on the surface between the two points (7).

---

(6) Eisenhart, Differential Geometry, p. 211.
(7) " " " p. 212.
II.

Further Considerations Regarding the Surface S.

We have observed that our surface S cannot be contained in any finite portion of space. We shall now make further hypotheses regarding the nature of S. Let $S'$ be a finite part of the surface limited by closed curves $C_1, \ldots, C_n$. We will suppose that, for a proper choice of these curves, one can by continuous variation of the curves extend them all to infinity without in so doing altering the topological properties of the surface. Thus we may regard the surface as being composed of $n$ independent infinite parts which we will call infinite nappes, and which we may regard as having been generated by the movement of the curves $C_1, \ldots, C_n$ in such a way that each point of each nappe is swept over once and only once.

We will distinguish between two sorts of infinite nappes. If the curve which generates the nappe increases in length indefinitely as it varies we will call the nappe a funnel. If on the other hand for an infinite sequence of positions of the generating curve more and more distant, the perimeter of the curve is bounded, we will call the infinite nappe a chimney.

One can form surfaces of negative curvature with any num-
ber of infinite nappes. One manner of doing this to yield a surface with two chimneys is as follows. Take two fixed points \( P \) and \( P' \) in the \( x, y \) plane. Let \( S \) and \( S' \) measure the distance from any point \( P \) to these points respectively. Then consider the surface formed by ascribing as the \( z \) coordinate of a point corresponding to a given \( (x, y) \) the value given by

\[
z = k \log \frac{S}{S'}
\]

This surface has two vertical chimneys and an infinite horizontal nappe. A sketch of the surface appears in figure 1.

An immediate generalization of this surface with as many chimneys as desired is seen by considering

\[
z = k \log \frac{S_1 S_2 \ldots S_n}{S'_1 S'_2 \ldots S'_n}
\]

To obtain a surface with as many funnels as desired one need only increase the radius of each horizontal section of the
chimneys by a quantity proportional to $Z$.

Let us reconsider the surface corresponding to the equation

$$Z = h \log \frac{S_i}{S'_i}$$

Think of an infinite mirror without thickness in the position $Z = G > 0$ cutting this surface. Consider the part of the given surface below the mirror together with its reflection. This is a surface with one hole and four infinite nappes. An extension of these schemes can result in a surface with any number of holes or infinite nappes.

We now turn to a formulation of some principles for studying curves traced on such surfaces as those whose existence we have just shown.

We will say that two closed contours drawn on a surface are of the same **species** if either can be made to coincide with the other by a continuous deformation. We will say that a certain contour is **simple** if it is of the same species as a bound of the surface (for instance the generating curve of a funnel).

If a certain contour can be made to conform with a second traced in a certain sense $n$ times, we will call the second a multiple of the first. Two contours which are composed of different multiples of the same species of contour are not to
be regarded as essentially distinct. Thus a closed contour which winds several times about a hyperboloid of one sheet is not to be regarded as essentially distinct from the simple contour generating one of the funnels. But if the order of connectivity of the surface exceeds two, the contours distinct from one another are infinite in number. The following sketch implies the truth of this fact, in that \( C_1 \) and \( C_2 \) are distinct contours.

However, there are only a finite number of species of contours of finite length. For in the first place our hypothesis regarding the surface \( S \) assures us that there are only a finite number of simple contours. In the second place if the length of any contour is finite it cannot make an infinite number of loops each deformable into a simple contour.

Let us turn our attention from closed contours to open curves or lines on \( S \). Having two fixed points \( a \) and \( b \) on the surface \( S \), we will say that two curves joining them are of the same type if one can be shifted into the other by continuous deformation. It is interesting to note that if \( a \) and \( b \) coincide there may be an infinite number of closed curves thru \( a \), all of the same species, but when these are considered as
curves joining a to itself so that a must remain fixed, they are of different types. This fact is illustrated in figure 1 where two curves are shown such that if a were allowed to move they could be deformed one into the other, but with a fixed they are curves of different types.

Analogously to the case of closed contours, for the present case of curves joining two points, we may note that in a multiply connected surface the curves joining two fixed points are of infinitely many types, but the types of curves of finite length are finite in number.

Regarding the joining of a fixed point a and a curve L we say that if b and c are two points on L that curves ab and ac are of the same type if ab can be continuously deformed into ac as b goes to c.

Notice that if the line L is closed there are an infinite number of ways of going from b to c. Thus in a hyperboloid of one sheet all the paths which go from a given point to a closed contour are reducible one to the other.
III.

Distribution of Geodesics on $S$

A first important result which we may now obtain is the following:

Theorem 1. To each type of line joining two points $a$ and $b$ on $S$ there corresponds one and only one geodesic arc joining the points.

Join the points $a$ and $b$ by any curve of a given type. Let $m$ be a point on this curve in a sufficiently small neighborhood of $a$ that it is possible to join $a$ and $m$ by a geodesic arc which is deformable into the arc $am$ of $ab$. As $a$ varies continuously from $a$ to $b$ if there continues to be a geodesic joining $am$ the theorem is true. But suppose this is not possible. Let $B$ be the point whose distance from $a$ measured along $ab$ is the upper limit of the distances from $a$ to $m$ for which there is a geodesic (of the given type) joining $am$. It follows that there is a geodesic $aB$, but for some point $D$, beyond $B$, as close to $B$ as you please, there is no geodesic (of the given type) joining $a$ to $D$.

But let the geodesic $aB$ make an angle $\alpha$ with some line through $a$. If one takes the geodesic through $a$ determined by the point $a$ and the angle $d+\epsilon$, and if one lets $v$ represent the
distance along the geodesic $aB$ starting from $a$, and $u$ the normal distance from any point on $a$ to the new geodesic, then $u$ is a continuous function of $v$ and $\epsilon$, and for $v$ fixed $u$ approaches zero as $\epsilon$ approaches zero. Then for a proper determination of $\epsilon$ there is a geodesic $aD$ which becomes coincident with the geodesic $aB$ as $D$ approaches $B$. Then $B$ is not the upper bound of the distances for which there is a geodesic of a given type joining $am$. From this contradiction we infer the proof of the first part of the theorem.

We can be sure that the geodesic of a certain type joining two points is unique because, first, any two lines of a given type connecting two points inclose a simply connected area and, second, a theorem given in the introduction assures us that on a surface of negative curvature two geodesics cannot meet in two points and inclose a simply connected area. Another conclusion follows without difficulty.

**Theorem 2.** To each type of line joining a point $a$ to geodesic $L$ there corresponds one and only one geodesic normal to $L$.

New considerations will need to be introduced in order to obtain our next important result.

**Theorem 3.** When a point describes a geodesic $L'$, its normal geodesic distance (of any given type) to a geodesic $L$ cannot have a maximum.

For coordinates let $v$ represent the distance from a fixed
point on $L$ to the geodesic normal $u$ drawn from $M'$ on $L'$ to $L$. In this coordinate system the linear element becomes (1)

$$ds^2 = du^2 + C^2 dv^2$$

Suppose that as $M'$ varies on $L'$ the quantity $u$ attains a maximum when $M'$ is at $M_1'$. Take two points $P_1'$ and $P_2'$ on either side of $M_1'$ on $L'$ and such that an arc $P_1'P_2'$ may join these points and have these properties: (1) the arc is deformable into the segment of $L'$ between $P_1'$ and $P_2'$ and (2) the distance, $u$, from $L$ to points on this line is less than the distance from $L$ to the segment $P_1'M_1'P_2'$ on $L'$.

Then compute the length of the arc $P_1'P_2'$:

$$\nu = \int_{P_1'}^{P_2'} \sqrt{du^2 + C^2 dv^2}$$

Since $C$ is an increasing function of $u$, the value of this integral is smaller than the value computed similarly for $s$ along the segment $P_1'M_1'P_2'$ of $L'$. But then $L'$ is not a geodesic.

Corollary I. If, as a point travels a geodesic line $L'$ in a certain sense, its geodesic distance to another line $L$ is ever increasing, it continues to increase indefinitely.

We may now make a distinction between the closed curves on funnels and on chimneys.

Corollary II. There does not exist a closed geodesic about a chimney.

Suppose there were one, namely \( G \) on a certain chimney, let \( u \) be the distance to any other closed curve \( G' \) "farther out" on the chimney. Recalling again that \( C \) increases as \( u \) increases we note that the length of the closed curve \( G' \) as given by

\[
\int \sqrt{du^2 + C^2 dv^2}
\]

increases indefinitely as \( u \) increases. But then we wouldn't have a chimney.

Chimneys then become less interesting than funnels and so in the developments that follow we shall restrict ourselves to surfaces with funnels. With such a limitation on \( S \) we may obtain the following theorem.

Theorem IV. Corresponding to each species of closed contour on \( S \) there is one and only one geodesic.

Take any point \( a \) on the surface and join \( a \) to itself by a geodesic of a given species. According to a remark on page 9, there are only a finite number of types of curves of finite length for which this can be done. Then there is one of minimum length. This length varies continuously with \( a \),
increasing when a recedes to infinity and hence having at least one minimum value when a is in the finite part of the plane. Then there is at least one closed geodesic of the given species.

We can establish that there is only one by the method used in Corollary II. Thus let u be the normal geodesic distance from a point on any closed curve of a certain species to a closed geodesic of that species. The value of 

$$\int \sqrt{\mu^2 + \nu^2}$$

is least when u = 0 and increases as u increases and hence the circumference of any other closed curve exceeds that of the geodesic, and therefore there is only one geodesic of the given species.

As an illustration we note that the hyperboloid of one sheet has only one closed geodesic, namely the shortest curve around its neck.

Having noted the existence of closed geodesics on the surface S we proceed to make the classification of the geodesics more complete.

Suppose we have a geodesic L and a point a' not on L. Take a geodesic a'm of any type joining a' to a point m of L. If as m moves continuously in a certain sense on L the geo-
desic a'm takes a limiting position L', we say that L' is asymptotic to L.

The existence of asymptotes may be argued from the following point of view. Let m as it moves on L take at a certain instant the position a. As m recedes the angle ma'a is continuously increasing, but according to the Gauss theory of geodesic triangles this angle is always bounded by the exterior angle at a of the geodesic triangle maa'.

Suppose on the other hand you have two geodesics, such that as a point m on one recedes, its distance to the other approaches zero; then either of these geodesics is asymptotic to the other.

Indeed let aa' be a geodesic of any type joining L and L' and as m recedes, let mm' be a line of the same type as aa' and let it measure the distance λ between the two lines which by hypothesis is approaching zero, the normal being at m'.

Now in the above sketch one sees

\[ d\lambda = \sqrt{d\lambda^2} \]

So

\[ \lambda > \left( \text{minimum of} \right) \int d\lambda \]

\[ \lambda > \text{minimum of} [P_2 a' - P_1 P_2] d\lambda \]

\[ \lambda > (m'a' - \lambda) d\lambda \]

Then as \( \lambda \to 0, \ d\to 0 \) and the geodesic a'm has as its limit
We have just established that if a geodesic \( L' \) is asymptotic to \( L \) then \( L \) is asymptotic to \( L' \). It follows that if two geodesics are asymptotic to a third they are asymptotic to each other, but **asymptotes drawn from a point to two geodesics not asymptotic one to the other cannot coincide.**

Let it be noted that we have now established the existence not only of closed geodesics but also of geodesics asymptotic to closed geodesics, for neither in the definition of an asymptote nor the results immediately following did we require that \( L \) be open.

We can establish the existence of another class of geodesics through the following considerations. Take a funnel \( F \) and a closed geodesic \( g \) on it. Join a point \( a \) out beyond \( g \) on the funnel to a point \( 0 \) not on the funnel by a geodesic curve \( L \) of a given type. There is only one type of curve joining \( a \) to \( g \), hence the geodesic distance \( u \) from \( a \) to \( g \) is unique and depends only on the position of \( a \). But then as \( a \) recedes on \( L \), \( u \) increases, and since \( u \) cannot have a maximum the geodesic \( L \) is forced to recede in the funnel to infinity.

We observe also that **if a geodesic goes to infinity there are others arbitrarily close to the first which do also.**

In addition to the geodesics on \( S \) which go to infinity and those which are closed or asymptotic to closed geodesics,
there is still another class, composed of geodesics running indefinitely on a finite part of the surface.

To establish the existence of this latter class, some preliminary results will be useful. Let the geodesic $L$ joining $a$ on $F$ to $0$ not on $F$ cut $g$ in $m$. Suppose that $m$ moves indefinitely in one sense on $g$, and then indefinitely in the other sense. The geodesics $\text{Om}$ occupy successively all positions in an angle $A$ whose sides are the two asymptotes which go from $0$ to the closed geodesic $g$.

As one changes the type of geodesic $\text{Om}$ the angle $A$ is changed, but it is important for us to note that two different angles $A$ cannot have a side in common. This follows from the statement underlined on page 17. Two geodesics of different types cannot be asymptotic to each other, and hence cannot each be asymptotic to the same third geodesic.

We may now show that the closed geodesics and their asymptotes cannot constitute the totality of geodesics which remain in the finite part of $S$.

Let $0$ be any point in the finite part of $S$ and let $E$ denote the set of geodesic elements emanating from $0$ and remaining in the finite part of $S$. We note that the set $E$ is not null because it includes at least the geodesics from $0$ that are asymptotic to the different closed geodesics on $S$. In other words, $E$ contains the sides of all the angles $A$ with
vertices at 0, but contains no elements interior to the angles A.

Let us recall that if a geodesic from 0 runs to infinity, there are others arbitrarily close to that one that do also. In other words a geodesic which does not belong to E is not a limit geodesic of the set E. On the other hand any element of E belongs also to E'. (Recall that different angles A cannot have a side in common.) Then set E is a perfect set having the power of the continuum. In the next section of this paper we will show that there is only a denumerable infinitude of closed geodesics on S, thus there are only a denumerable infinitude of asymptotes going to closed geodesics among the elements of E. Then there are geodesics in addition to the closed geodesics and their asymptotes which run on the finite part of S.

In the next section of this paper the theory of this class of geodesics will be developed further.

We will conclude this section with some remarks, in a sense philosophical, regarding the dynamical implications of the preceding observations. If you fix your attention on the geodesic element emanating from 0 which forms the direction of a side of an angle A, we note that as close as you please to this element there are others which, on the one hand, lead to
geodesics running to infinity, on the other hand to geodesics confined to the finite region. More generally it is true that any variation in the initial direction of any geodesic running entirely in the finite plane, no matter how slight, is sufficient to alter the final course of the curve in such a way as to make it become on the one hand a geodesic running to infinity, or, on the other hand, a geodesic asymptotic to a closed geodesic.

Suppose then some observations were made by a physicist, say, on a particle moving without restraint upon a geodesic on S. Since these observations are subject to errors (however small), it would be impossible to predict on the basis of the observations whether the particle would continue to run all over the finite part of the surface, or if it would extend through some funnel to infinity, or if it would settle down ultimately to an asymptotic journey about a closed contour.

Similarly, if a particle were moving on what was thought to be a closed contour it might ultimately turn out to be somewhere remote from this closed contour, having in reality been moving on the geodesic asymptotic to the closed contour.
IV.

The Morse M Surface.

In the preceding section we saw that on the surface $S$ corresponding to the species of curves forming a closed contour about any funnel there existed a unique geodesic of that species. Suppose there are $v$ such funnels. Let $g_1, g_2, \ldots, g_v$ be these closed geodesics. From here on let us understand by the surface $S$ the finite surface remaining after cutting the original surface along all these geodesics. We will confine our attention in this section then to geodesics which if continued indefinitely in either sense lie entirely on a finite surface.

Moreover we shall consider surfaces with at least two funnels, and those surfaces with only two funnels, we shall insist, must have at least one hole. A schematic drawing of the simplest kind of surface we will consider is given in the upper half of figure 2.

Let us select a point $P$ on any one of the boundary geodesics, say $g_v$, and cut $S$ along the curves $h_1, h_2, \ldots, h_{v-1}$ joining $P$ on $g_v$ to each of the other geodesics. This cut surface now has a single boundary. If the surface has $p$ holes, we may then form a simply connected surface $T$ by
taking $2p$ non-intersecting closed curves through $P$

$$c_1, c_2, \ldots, c_2p$$

and cutting the given surface along these curves.

Now suppose we have an infinite number of copies of $T$ lying one above the other over $S$. Take any copy of $T$, call it $M_1$, and join $M_1$ to another copy of $T$ over any $c$ or $h$ cut, say $c_1$, joining the positive shore of $c_1$ in $M_1$ to the negative shore of $c_1$ in the other sheet. Then take another $c$ or $h$ cut, say $h_1$, and join the positive shore of $h_1$ in $M_1$ to the negative shore of $h_1$ in a third sheet. Continue this process joining all the positive shores of $c$ and $h$ cuts in $M_1$ to corresponding negative shores of different sheets of $T$. Let the totality of sheets required for this operation be designated by $M_2$. We continue in the same manner to join to each different $c$ and $h$ cut of $M_2$ a different copy of $T$ and not one already a part of $M_2$. This new set of sheets will be called $M_3$. Continue this process indefinitely. The infinitely many sheeted surface spread over $S$ obtained in this manner will be called $M$.

A schematic drawing illustrating the joining of the sheets of the surface $M$ is given in the lower half of figure 2 where the sheets of $M_1$, $M_2$, and part of $M_3$ are shown.

There are several useful inferences we may draw regarding the structure of the surface $M$. (1) A finite number of copies of $M$ constitutes a surface topologically equivalent to the interior and boundary of a simple plane closed curve.
Thus any one of the $\epsilon$ or $h$ curves along which two sheets are joined is a simple curve joining two boundary points of $M$ and divides $M$ into two regions. (3) All the points on a sufficiently restricted neighborhood of any point on $M$ lie in only a finite number of copies of $T$ that are joined together at that point. (There are only a finite number of cuts emanating from $P$.)

As a consequence of this latter observation we obtain immediately the result: A set of points every infinite subset of which has at least one limit point on $M$ cannot have points in more than a finite number of different sheets of $M$.

Indeed suppose the set had points in an infinite number of sheets. Then the subset consisting of one point from each sheet would be an infinite subset and therefore by hypothesis have a limit point on $M$. Then we would have an infinite number of copies of $T$ with points in the neighborhood of a certain point, which is not possible.

A curve on $M$ will be said to correspond to a given curve on $S$ if as a point $P$ traces the curve on $S$ a point on $M$ may move continuously always overhanging $P$ and trace out the curve on $M$.

If a curve on $S$ is closed and deformable into a point, it either crosses none of the $\epsilon$ and $h$ cuts or crosses any
that it does cross an even number of times in opposite senses. Then the corresponding curve on M either is confined to one sheet or crosses particular images of \( \alpha \) and \( h \) curves an even number of times in opposite senses. From the nature of the structure of M we conclude: If a curve on S is both closed and deformable into a point, it corresponds to only closed curves on M, and conversely.

If we take two points on M they overhang a pair of points on S. There is a unique geodesic of a given type joining these points on S. Then we may affirm the following:

**Theorem.** There exists on M one and only one geodesic joining two given points. The existence of one is clear if we consider the image on M of the selected geodesic on S. If it is not unique, take a pair of them on M, constituting then a closed curve. This closed curve on M corresponds on S, as we have just seen, to a curve that is closed and deformable into a point. But this latter implies that there are two geodesics of the same type through a pair of points on S.

In addition we may note:

**Corollary 1.** On M two geodesics can intersect in but one point.

**Corollary 2.** On M a geodesic can have no multiple points.
We shall use the term **linear set of copies of** \( T \) to refer to a finite or infinite set of sheets in \( M \) written as

\[ \ldots T_2, T_1, T_0, T_1, T_2 \ldots \]

in which each symbol represents any particular copy of \( T \), two successive copies being joined along a common \( c \) or \( h \) curve, and any three successive symbols represent distinct copies.

Recalling that any \( c \) or \( h \) curve along which two sheets are joined forms a curve dividing \( M \) into two regions, it is apparent that **no copy of** \( T \) **appears more than once in any linear set**. Moreover, **given any two copies of** \( T \) **of** \( M \), there is just one linear set having one of these for the first, the other for the last member**. One forms such a linear set by successively annexing to the first copy (or any succeeding copy) that copy of \( T \) which adjoins the first copy (or any succeeding copy) along the \( c \) or \( h \) curve that separates the one you have from that part of \( M \) in which the last lies. This procedure can be terminated after a finite number of steps. For join interior points of the first and last copies by a curve \( k \). This curve has at least one point in each copy of the linear set formed according to the above description. Then it constitutes a set of points every infinite subset of which has a limit point on \( M \). We have seen that such a set may appear in only a finite number of sheets.
The uniqueness follows from the fact that in no stage of the formulation of the linear set was there any ambiguity about which was the succeeding copy.

We will next show how to form a set of curves on $M$ which can be put into one to one correspondence with the unending linear sets on $M$.

We will denote by $H$ the set of all the images of any of the $c$, $h$, or $g$ curves of $S$. We shall use the term reduced curve to refer to the continuous curve formed from pieces of the set $H$, and written as

$$\cdots k_{-2}, k_{-1}, k_0, k_1, k_2 \cdots$$

in which each symbol represents any piece of $H$ except an image of $g_v$, and in which no two successive symbols represent the same piece of $H$. A reduced curve without end points is called an unending reduced curve.

In view of the fact that the elements of any reduced curve are taken from the boundary pieces of some copy of $T$ but do not include one piece (the image of $g_v$), and since no two successive elements of a reduced curve consist of the same piece of $H$, it follows that no reduced curve or segment of a reduced curve is closed. Moreover, given any two points on pieces of the set $H$ (neither of which is an interior point to an image of $g_v$), there is one and only one reduced curve joining these points. For beginning with any curve formed
from a continuous succession of pieces of $H$, replace the piece crossing the image of any $g_\nu$ curve in any sheet by the rest of the boundary of that sheet, and then among the curves of this latter class select the one with the fewest pieces of $H$ in it. This curve will, for instance, have no two successive parts formed from the same piece of $H$. The uniqueness is a consequence of the fact that no reduced curve may begin and end in the same point. As a consequence also of this same principle, we may announce that a reduced curve never returns to a copy of $T$ from which it has once departed.

Theorem. There is a one to one correspondence between the set of all unending reduced curves and the set of all unending linear sets of $M$, in which each reduced curve corresponds to that linear set in which it is contained.

Indeed an unending linear set contains one and only one unending reduced curve. A reduced curve is formed by taking a continuous succession of boundary pieces of successive copies of $T$, none of which include images of $g_\nu$ or consist of the same boundary piece traversed twice in succession. Moreover, this reduced curve is unique because there is only one path around the boundary of any sheet that does not take you across an image of $g_\nu$. Hence, if there were two reduced curves they would have points in common, but we have just seen that there
is only one reduced curve joining two points of pieces of $H$.

The proof that conversely an unending reduced curve is contained in one and only one linear set, which is an unending linear set, follows the same general plan as a proof to be given later and is omitted here.

We shall now make some developments preliminary to showing that there is a one to one correspondence between the set of all geodesics lying on $M$ and the set of all unending linear sets.

Let $H'$ denote the set of all geodesics on $M$, each of which joins end points of a piece of the set $H$.

We show first that two segments of the set $H'$ arising from different pieces of the set $H$ can have no points in common other than one end point.

Case I. If the two pieces of $H$ have a point in common the same is true of the corresponding pieces of $H'$, but then they share no other point in common, for we have seen that on $M$ two geodesics may intersect in only one point.

Case II. Let the two pieces of $H$ have no points in common, but let one of the geodesics constitute a boundary of $M$. The two geodesics could not be tangent to each other for at any point a geodesic is uniquely determined by a direction.
By hypothesis they don't meet at an end point of either segment and hence if they meet at all one would have to pass off the surface at the point.

Case III. PQ and P'Q' being two pieces of H, let P≠P', Q≠Q' and let neither of the corresponding geodesic segments h, h' form a piece of the geodesic boundary of M.

Take a simply connected region R of M so large that no points of PQ and P'Q' nor of h and h' are boundary points of R unless they are boundary points of M. P, Q, P', Q' lie on the boundary of R in the circular order named. Now h has only its end points on the boundary of R and hence divides R into two regions. If h' crossed R its end points P' and Q' would lie in opposite parts of these two regions.

![Diagram]

and the circular order of points P, Q, P', Q' does not hold.

Recall that the original surface S was cut by certain c, h, and g curves. Each of these is of a certain type or species and hence there corresponds to each a geodesic deformable into the given cut. But the images of these geodesics on M are precisely the curves H' that we have been discussing, for if they were not, there would be distinct geodesics meeting in two points on M. Then it follows from the result of the last paragraph that on S the geodesics corresponding to H' have no points in common other than end
points. If the original surface is cut along these geodesics instead of the arbitrary c and h curves, we will again have a simply connected surface U. And as before, we may build over it the infinitely-leaved surface M, and have a one to one correspondence between the set of unending reduced curves of the set \( H' \) and the set of all unending linear sets of copies of U.

Let \( G \) be an unending geodesic lying entirely on M. We may show that the geodesic \( G \) is contained in one and only one linear set, which set is an unending linear set.

In the first place we remarked that the geodesic \( G \) cannot be infinite in any one copy of U. For a single copy consists of a closed set of points, and the length of the geodesic segment joining any two of these points is a unique continuous function of the points. Then if \( G \) is divided into two parts at any point each of these parts have points in an infinite number of copies of U. Moreover, since two geodesics on M can intersect in but one point, \( G \) can meet any geodesic segment joining different copies of U only once and so it never reenters a sheet from which it has once departed.

Let \( \ldots P_2, P_1, P_0, P_1, P_2 \ldots \) be a set of points on \( G \) in order of their subscripts, and such that any point on \( G \) lies between \( P_n \) and \( P_n \) for \( n \) large enough. Let \( G_n \) be a segment of \( G \) joining \( P_n \) and \( P_n \), and let \( L_n \) be the linear set containing \( P_n \) and \( P_n \) and composed of the fewest possible
copies of $U$.

$G_n$ lies entirely on $L_n$. Since the end points of $G_n$ are on $L_n$ if $G_n$ left $L_n$ it would have to return. But any boundary over which $G_n$ might leave $L_n$ divides $M$ into two parts, and to get back into the set $L_n$, $G_n$ would have to recross this boundary. But $G$ never reenters a sheet from which it has once departed.

If $r > n$ the linear set $L_r$ contains $L_n$. For $L_r$ contains $G_n$ and the latter runs through each copy of $L_n$ or else $L_n$ doesn't have the fewest possible sheets.

Let $L$ denote the linear set composed of all $L_n$ for all integers $n$. Since for $n$ sufficiently large any point of $G$ was contained in some $G_n$, hence in some $L_n$, $L$ contains $G$. But we have seen that if $G$ is divided into two parts each part must have points in an infinite number of copies of $U$. Hence $L$ is an unending linear set.

Finally, $L$ is unique. For any other linear set would have to contain each $L_n$, and hence every copy of $L$. But there is only one linear set joining any two copies of $U$.

The converse of the fact just proved is also true, from whence we infer at once:

**Theorem.** There is a one to one correspondence between the set of all geodesics lying wholly on $M$ and the set of all unending linear sets, in which each geodesic corresponds to that linear set in which it is contained.

And as a consequence of the correspondence between
linear sets and reduced curves previously established we have also:

Theorem. There is a one to one correspondence between the set of all geodesics being wholly on M, and the set of all reduced curves of the set $H'$ in which each geodesic corresponds to that reduced curve that is contained in the same linear set.

We have previously seen that there existed closed geodesics on S. We may now give a new brief proof of that fact.

We will say that two curves on M are congruent if they are images of the same curve on S.

Theorem. If a reduced curve on M consists of successive mutually congruent portions, the geodesic that lies in the same linear set as the given reduced curve consists also of successive mutually congruent portions.

Let a portion of the given reduced curve be carried into a succeeding congruent portion by a transformation T. The geodesic lying in the same linear set is carried into itself by T, or otherwise the one to one correspondence between reduced curves and geodesics breaks down. The successive images of any geodesic segment between two points, one of which is carried into the other by T, which arise through repetitions of T and its inverse yields mutually congruent portions of the given geodesic.
For a new representation of geodesics on $M$ it will be useful to consider what we will call normal segments, namely any one of the geodesic segments

$$c_1, c_2 \ldots c_{\nu}; g_1, g_2 \ldots g_{\nu-1}.$$  
(Note that this is a subset of the set $H'$ in which the geodesic boundary pieces $h_1, \ldots, h_{\nu-1}$ are omitted.)

If $C_m$ represents any sensed normal segment the unending sequence of segments

$$\ldots c_2, c_1, c_0, c_1, c_2 \ldots$$

will be called a normal set $C$ if no two successive symbols represent the same normal segment taken in opposite senses.

It is evident that if we are given any sensed unending reduced curve, its normal segments constitute a normal set, and conversely, given a normal set there can be formed from it only one sensed reduced curve. Due to this one to one correspondence it is natural to say that the normal set $C$ may represent the unending sensed reduced curve and also the sensed geodesic that passes through the same linear set of copies of $U$.

Recall that if a geodesic on $M$ consists of repetitions of a closed geodesic, the unending reduced curve passing through the same linear set consists of repetitions of geodesic segments $H'$. In this case the normal set $C$ representing
the closed geodesic and the unending reduced curve is such that

\[ C_m = C_{m+p} \]

for some positive integer \( p \). The normal set is then said to be \textit{periodic} of period \( p \).

If \( q \) is the smallest period of a periodic normal set \( C \), any \( q \) successive normal segments of the given set \( C \) will be called a \textit{generating set of} \( C \). If \( B \) is a generating set of \( C \), the set \( C \) may be written simply

\[ \ldots BBBBBBB \ldots \]

and any other generating set yielding the same set \( C \) will consist merely of a circular permutation of the sensed normal segments forming \( B \).

Due to the fact that to a given normal set there corresponds just one sensed reduced curve and also therefore one sensed geodesic, and since if the given normal set is periodic the geodesic determined by it is closed (and conversely) the number of distinct closed geodesics on \( S \) depends on the number of periodic normal sets that may be formed.

Then let us examine how an arbitrary generating set of some periodic normal set may be formed. If we take the \( 2p+q-1 \) normal segments and form from them any finite ordered set of sensed segments in which no two successive members in the complete circular order are the same segment taken in
opposite senses and such that the selected ordered set is not composed of successive repetitions of a like set with fewer segments, we clearly do have a generating set of a periodic normal segment. The number of different periodic sets is then seen to be the number of the generating sets that we have just described which are not obtainable one from the other by a circular permutation of their elements. The fact that this number is denumerable permits us to see the proof of the assertion made on page 19, namely, that there is a denumerable infinity of distinct closed geodesics on the surface $S$. 
BIBLIOGRAPHY

J. Hadamard

Les surfaces à courbures opposées et leur lignes géodesiques. *Journal de Mathématique*, 1898.

M. Morse


M. Morse


G. D. Birkhoff


G. Darboux