

ON GEOMETRIES IN WHICH CIRCLES ARE THE SHORTEST LINES*

BY

CARL EBEN STROMQUIST

HAMEL[†] has considered those geometries in which the straight line is the shortest distance between two points. The axioms at the basis of these geometries are, except for some minor changes, the well known axioms of HILBERT, and are such that by means of them a one to one correspondence can be established between the points in the plane and pairs of numbers (x, y) . "Length" of any rectifiable curve $y = y(x)$ between two given points (x_1, y_1) , (x_2, y_2) is defined[‡] as an integral,

$$\int_{x_1}^{x_2} g(x, y, p) dx, \quad p = dy/dx.$$

The nature of the geometry will then depend on the particular form of the integral chosen.

HAMEL then studies the geometries for which these axioms are satisfied and for which the integral shall have such a form that its value taken from x_1 to x_2 along a straight line shall be less than that along any other rectifiable curve joining the two points (x_1, y_1) and (x_2, y_2) .

HAMEL's methods apply, with but slight modifications, to the problem we wish to consider in this paper; viz., the study of a geometry in the upper half-plane whose elements satisfy the axioms given by HAMEL, but in which the last stated restriction on length shall be replaced by the following:

* Presented to the Society February 28, 1903, under a different title. Received for publication October 28, 1905.

† HAMEL, *Ueber die Geometrien in denen die Geraden die Kürzesten sind*, dissertation, Göttingen, 1901; in revised form, *Mathematische Annalen*, vol. 57 (1903), pp. 231-264.

‡ HAMEL defines the length of a straight line between two points as a function of the coördinates of the points and shows that it can be expressed as an integral taken along the straight line. In his dissertation he then tacitly assumes that length along any curve is expressed by the same integral (p. 13); in the article, in the *Annalen* already referred to, this assumption is definitely stated (p. 241). Another method would be to define length along any curve as the limit of the length of inscribed polygons, as is usual in Euclidean geometry, and then to prove that the limit would be the value of the same definite integral which expressed the length along the straight line segment, taken along the limiting curve. It seems simpler to start with the above definite integral as the definition of length of any curve.

The shortest distance between two points shall be the segment of that circle joining the points whose center lies on the boundary line (x -axis) of the half plane.

This leads to an inverse problem in the calculus of variations. For by this last restriction the shortest lines are given, and the problem is to find the integral

$$\int_{x_1}^{x_2} g(x, y, p) dx$$

which will be minimized by these shortest lines. There will in general be an infinite number of solutions to such a problem and our object is to determine the most general function g satisfying these assumptions.

In addition to the problem stated above we shall consider (in section 2) the geometries in which transversals are perpendicular to their extremals; in particular, HAMEL's geometries and the geometries in which circles are the shortest distances. In section 3 examples of particular cases of the "circle geometry" will be given.

§ 1. Determination of the integrand g .

To determine the definition of length in the geometries defined above we must then find the most general function $g(x, y, p)$ such that the solution of the problem of rendering

$$(1) \quad l = \int_{x_1}^{x_2} g(x, y, p) dx$$

a minimum is given by the function y defined by

$$(2) \quad (x - a)^2 + y^2 = c^2,$$

i. e., the so-called extremals (minimizing curves of our integral) must be the circles with centres on the x -axis, whose differential equation is

$$(3) \quad y'' = \phi(x, y, p) = -\frac{1 + p^2}{y}.$$

From the calculus of variations we know* that the first necessary condition that the curve $y = y(x)$ shall minimize the integral

$$\int_{x_1}^{x_2} g(x, y, p) dx$$

is that $y = y(x)$ should satisfy EULER's equation

$$y'' \frac{\partial^2 g}{\partial p^2} + p \frac{\partial^2 g}{\partial y \partial p} + \frac{\partial^2 g}{\partial x \partial p} - \frac{\partial g}{\partial y} = 0.$$

* BOLZA, *Lectures on the Calculus of Variations*, p. 22.

DARBOUX* has pointed out that, conversely, if EULER's equation has as extremals the solutions of $y'' = \phi(x, y, p)$, the function $g(x, y, p)$ must satisfy the differential equation

$$(4) \quad \phi \frac{\partial^2 g}{\partial p^2} + p \frac{\partial^2 g}{\partial y \partial p} + \frac{\partial^2 g}{\partial x \partial p} - \frac{\partial g}{\partial y} = 0.$$

If we differentiate this equation with respect to p we obtain the following linear differential equation for $M = \partial^2 g / \partial p^2$:

$$(5) \quad \frac{\partial M}{\partial x} + p \frac{\partial M}{\partial y} + \phi \frac{\partial M}{\partial p} + M \frac{\partial \phi}{\partial p} = 0,$$

where ϕ is defined by equation (3).

The most general solution of equation (5) in our case is †

$$(6) \quad M = y^2 \cdot W[y^2(1 + p^2), x + yp],$$

where W is an arbitrary function of the arguments $y^2(1 + p^2)$ and $x + yp$.

It follows that g must have the form

$$(6) \quad g = y^2 \int_c^p \int_c^p W[y^2(1 + p^2), x + yp] dp dp + p \cdot w_1(x, y) + w_2(x, y),$$

where w_1 and w_2 are functions of x and y alone which must be so restricted that g satisfies equation (4).

Since the differential equation from which g is obtained is equation (4) differentiated with respect to p , the result of substituting g in (4) must result at most in a function of x and y alone. If we call $K(x, y)$ the function resulting from substituting

$$g_0 = y^2 \int_c^p \int_c^p W dp dp$$

in (4), then it is readily seen that the result of substituting g in (4) gives

$$K(x, y) + \frac{\partial w_1}{\partial x} - \frac{\partial w_2}{\partial y} = 0.$$

Moreover since the substitution of g_0 in (4) must result in a function free from y' we may evaluate it by giving y' a particular value, say $y' = 0$, from which it follows that $K(x, y)$ must have the form $-y \cdot W(y^2, x)$.

Hence our last equation becomes

$$(7) \quad -y W(y^2, x) + \frac{\partial w_1}{\partial x} - \frac{\partial w_2}{\partial y} = 0.$$

* DARBOUX, *Leçons*, vol. 3, p. 53. In solving EULER's equation for our problem we have followed closely the method given by DARBOUX.

† See, for instance, FRICKE, *Analytisch-Funktionentheoretische Vorlesungen*, p. 467.

If \bar{w}_1 and w_2 are particular values of w_1 and w_2 respectively, which satisfy (7), then the most general values of w_1 and w_2 are given by

$$w_1 = \bar{w}_1 + \frac{\partial u}{\partial y}, \quad w_2 = \bar{w}_2 + \frac{\partial u}{\partial x},$$

where u is an arbitrary function of x and y . We may then choose one of the functions \bar{w}_1 and \bar{w}_2 arbitrarily and the other can be determined by equation (7). Let, then, $\bar{w}_2 = 0$, and we have

$$\bar{w}_1 = y \int_{x_0}^x W(y^2, x) dx.$$

The function g now becomes

$$(8) \quad g = y^2 \int_c^p \int_c^p W[y^2(1+p^2), x+yp] dp dy + yp \int_{x_0}^x W(y^2, x) dx + p \frac{\partial u(x, y)}{\partial y} + \frac{\partial u(x, y)}{\partial x}.$$

This is the most general solution of the problem, where W and u are arbitrary functions of the arguments indicated.

We then know:

1) That the circles (2) are extremals, i. e., they satisfy EULER's equation when g is given by equation (8).

2) That a family of these circles through a point in the upper half plane does not have an envelope in this region, and therefore the condition known in the calculus of variations as JACOBI's condition is satisfied.

If we now impose on g the condition 3) that $g_{pp}(x, y, p) > 0$ for x and y in the upper half plane and for any finite p , the above three conditions are sufficient in order that the arcs of our circles shall render the integral $\int g dx$ a strong minimum as compared to its value for curves in a restricted neighborhood.*

Furthermore, through any two points of the upper half plane we can pass one and only one extremal† [i. e., arc of the circles (2)] and hence this arc will actually render the integral a strong minimum not only in comparison to curves in a restricted neighborhood but in comparison to any curves in our region

* BOLZA, loc. cit., p. 96, theor. V. Also see BOLZA for the definitions of *strong minimum*, *absolute minimum*, *neighborhood*, etc.

† For two points on the vertical we can consider this vertical as the limiting case of our circles when the center has receded to infinity along the x -axis. In this case p is not finite, but this can be overcome by taking as the definition of length in this case

$$\int_{y_1}^{y_2} g\left(x, y, \frac{1}{q}\right) q dy$$

where $q = dx/dy$. Or by taking the problem in parameter-representation the above difficulty would be avoided.

joining the two points, i. e., the circles (2) render the integral an *absolute* minimum.

In order that condition 3) shall be satisfied we must impose the following condition on the function W :

$W[y^2(1+p^2), x+yp] > 0$ for x and y in the upper half plane and for every finite p .

The function g can be reduced to a form containing only one integration, if we make the following transformation:

Let us define the function U such that

$$W[y^2(1+p^2), x+yp] = \frac{\cos^2 \theta}{y^3} U(y^2 \sec^2 \theta, x + y \tan \theta),$$

where $p = \tan \theta$, $c = \tan \theta_0$. Then by integration by parts g can readily be reduced to the form

$$(9) \quad g = \frac{1}{y \cos \theta} \int_{\theta_0}^{\theta} \sin(\theta - \tau) \cdot U(y^2 \sec^2 \tau, x + y \tan \tau) d\tau \\ + \frac{\tan \theta}{y^2} \int_{x_0}^x U(y^2, x) dx + p \frac{\partial u}{\partial y} + \frac{\partial u}{\partial x}.$$

We have therefore the following result:

Any geometry in the upper half plane in which the element of length is defined by

$$dl = dx \left\{ \frac{1}{y \cos \theta} \int_{\theta_0}^{\theta} \sin(\theta - \tau) \cdot U(y^2 \sec^2 \tau, x + y \tan \tau) d\tau \right. \\ \left. + \frac{\tan \theta}{y^2} \int_{x_0}^x U(y^2, x) dx + p \frac{\partial u}{\partial y} + \frac{\partial u}{\partial x} \right\},$$

where $U > 0$ for all x and y in the upper half plane and for every finite p , has the circles $(x-a)^2 + y^2 = c^2$ as shortest distances.

Any two points* may be joined by one such circle which has a shorter length, according to the definition of length given above, than any other curve joining the two points.

If we impose HAMEL's strong monodrome axiom—that the length of a curve from A to B shall be the same as the length from B to A —we must further restrict g . Following HAMEL† without giving details we obtain the result: In order that the strong monodrome axiom shall be satisfied it is necessary and sufficient that U shall have the period π considered as a function of θ ; the values of $\partial u / \partial x$ and $\partial u / \partial y$ are given by

* See last footnote above.

† HAMEL, loc. cit., p. 18.

$$\frac{\partial u}{\partial x} = \frac{1}{2y} \int_{\theta_0}^{\theta_0+\pi} \sin \tau \cdot U(y^2 \sec^2 \tau, x + y \tan \tau) d\tau,$$

$$\frac{\partial u}{\partial y} = -\frac{1}{2y} \int_{\theta_0}^{\theta_0+\pi} \cos \tau \cdot U(y^2 \sec^2 \tau, x + y \tan \tau) d\tau - \frac{1}{y^2} \int_{x_0}^x U(y^2, x) dx,$$

and are thus determined aside from a constant. The function g then reduces to

$$(9') \quad g = \frac{1}{2y} \int_{\theta-\pi}^{\theta} \sin(\theta - \tau) \cdot U(y^2 \sec^2 \tau, x + y \tan \tau) d\tau.$$

§ 2. *Transversals perpendicular to their extremals.*

Suppose we have a one parameter set of extremals simply covering a portion of the plane. Let p be the function of x and y which defines the direction of an extremal at the point (x, y) . Then a transversal,* $y = Y(x)$, is determined by the equation

$$(10) \quad g(x, y, p) + \left(\frac{dY}{dx} - p \right) g_p(x, y, p) = 0,$$

where dY/dx refers to the transversals.

Geometrically stated,† two transversals to the same set of extremals are two curves which intercept on the extremals arcs along which the integral under consideration has a constant value.

If for any set of extremals the transversals are always perpendicular to the extremals then equation (10) must be an identity in x, y, p when for dY/dx is substituted $-1/p$.

From this it follows readily that g must have the form ‡

$$(11) \quad g(x, y, p) = \psi(x, y) \sqrt{1 + p^2},$$

where ψ is an arbitrary function of x and y alone.

Consider now any geometry in which the extremals are perpendicular to the transversals. The line element must have the form

$$dl = \psi(x, y) \sqrt{1 + p^2}.$$

But by the theory of surfaces § we can always find a surface whose line element has the form

$$ds = \psi(u, v) \sqrt{du^2 + dv^2}.$$

This surface is represented conformally on the plane by the trivial transformation $u = x, v = y$. Also, as is evident from the theorem of GAUSS that

* BOLZA, loc. cit., p. 106.

† BOLZA, loc. cit., p. 172.

‡ This was first pointed out by E. R. HEDBICK in a course of lectures on the calculus of variations (1901-1902).

§ DARBOUX, *Leçons*, livre VII, chap. IV.

geodesics are perpendicular to the curves of constant geodesic distances, any geometry obtained by a conformal transformation of a surface on the plane is one in which extremals are perpendicular to the transversals. Therefore,

The necessary and sufficient condition that a geometry be such that extremals are perpendicular to their transversals is that the geometry be obtained by a conformal transformation of some surface upon the plane.

Let us now consider the result of making this restriction that extremals shall be perpendicular to their transversals in the geometries in which the circles (2) are the shortest distances.

We must then have, from (11) and (8),

$$W[y^2(1+p^2), x+yp] = \frac{\psi(x, y)}{y^2(1+p^2)^{\frac{1}{2}}}.$$

In order that this functional relation shall hold we must have *

$$\begin{vmatrix} \frac{\partial}{\partial x} \left(\frac{\psi(x, y)}{y^2(1+p^2)^{\frac{1}{2}}} \right) & \frac{\partial}{\partial y} \left(\frac{\psi(x, y)}{y^2(1+p^2)^{\frac{1}{2}}} \right) & \frac{\partial}{\partial p} \left(\frac{\psi(x, y)}{y^2(1+p^2)^{\frac{1}{2}}} \right) \\ 0 & 2y(1+p^2) & 2y^2p \\ 1 & p & y \end{vmatrix} = 0,$$

or

$$y\psi_x + (y\psi_y + \psi)p = 0.$$

Hence

$$y\psi_x = 0 \quad \text{and} \quad y\psi_y + \psi = 0.$$

From the second of these equations we see that ψ must have the form $\psi = c/y$, where c is an arbitrary constant, and therefore from (11),

$$g = \frac{c}{y} \sqrt{1+p^2}.$$

This is the geometry obtained by a conformal representation of the pseudosphere on the plane. In fact, the line element on the pseudosphere † is given by $ds^2 = c^2(du^2 + e^{2u}dv^2)$ (where $-1/c^2$ is the constant curvature). By the conformal transformation $v = x$, $e^{-u} = y$ the line element becomes

$$dl^2 = c^2 \frac{dx^2 + dy^2}{y^2},$$

or the form given above. Therefore,

The most general geometry in which the circles (2) are the shortest distances and extremals are perpendicular to their transversals is obtained by a conformal transformation of the pseudosphere upon the plane.

* PASCAL, *Repertorium der Höheren Mathematik*, vol. 1, p. 52.

† DARBOUX, loc. cit., vol. 3, p. 394.

In the case of HAMEL's geometries * g is of the form

$$g = \int_c^p \int_c^p W(p, y - xp) dp dy + p \frac{\partial u}{\partial y} + \frac{\partial u}{\partial x},$$

and if the extremals are to be perpendicular to their transversals, we find, proceeding as above, that

$$g = c \sqrt{1 + p^2}, \quad \text{or} \quad l = c \int_{x_1}^{x_2} \sqrt{1 + p^2} dx.$$

The case $c = 0$ is excluded by our axioms, and for $c = 1$ the geometry is the ordinary Euclidean geometry; $c \neq 1$ gives simply a magnification of the plane, that is, a Euclidean geometry with a different unit of length. We have therefore the following conclusion:

The Euclidean geometry is the most general geometry possible in which the shortest lines are the straight lines and the transversals are perpendicular to the shortest lines.

§ 3. Examples.

Let us now consider some particular examples of the geometries in which the circles are the shortest distances.

Example 1. Let $U = \text{const.} = 1$. From equation (9') we then obtain

$$(12) \quad l = \int_{x_1}^{x_2} \frac{\sqrt{1 + p^2}}{y} dx,$$

and the transversals $y = Y(x)$ must satisfy

$$(13) \quad 1 + p \frac{dY}{dx} = 0.$$

From the equation of the circles (2) we have $p = (a - x)/y$, and if we seek the transversals corresponding to the extremals through a given point, as $(1, 1)$, we shall readily have

$$p = \frac{x^2 - 2x - y^2 + 2}{2y(1 - x)}.$$

Then equation (13) becomes, after integration,

$$(14) \quad (x - 1)^2 + (y - k)^2 = k^2 - 1,$$

where k is a constant of integration. Hence the transversals are circles with centers, not at $(1, 1)$, but at $(1, k)$.

Since the transversals are circles, in order to construct a transversal of "radius" l_1 for the extremals through $(1, 1)$ we need only determine the values of y above and below $(1, 1)$ on $x = 1$ such that the distance from $(1, 1)$ to

* HAMEL, loc. cit., p. 14.

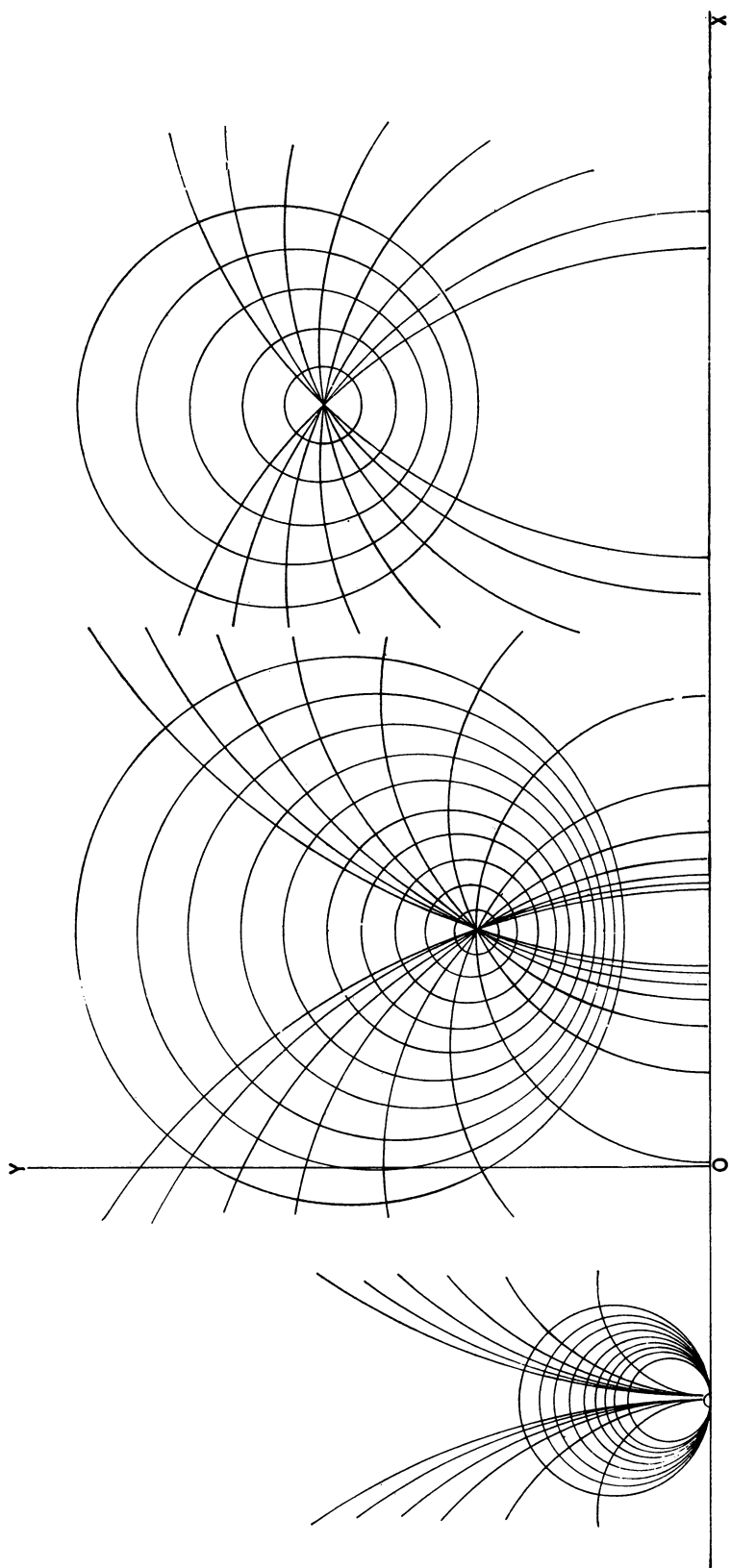


PLATE I.

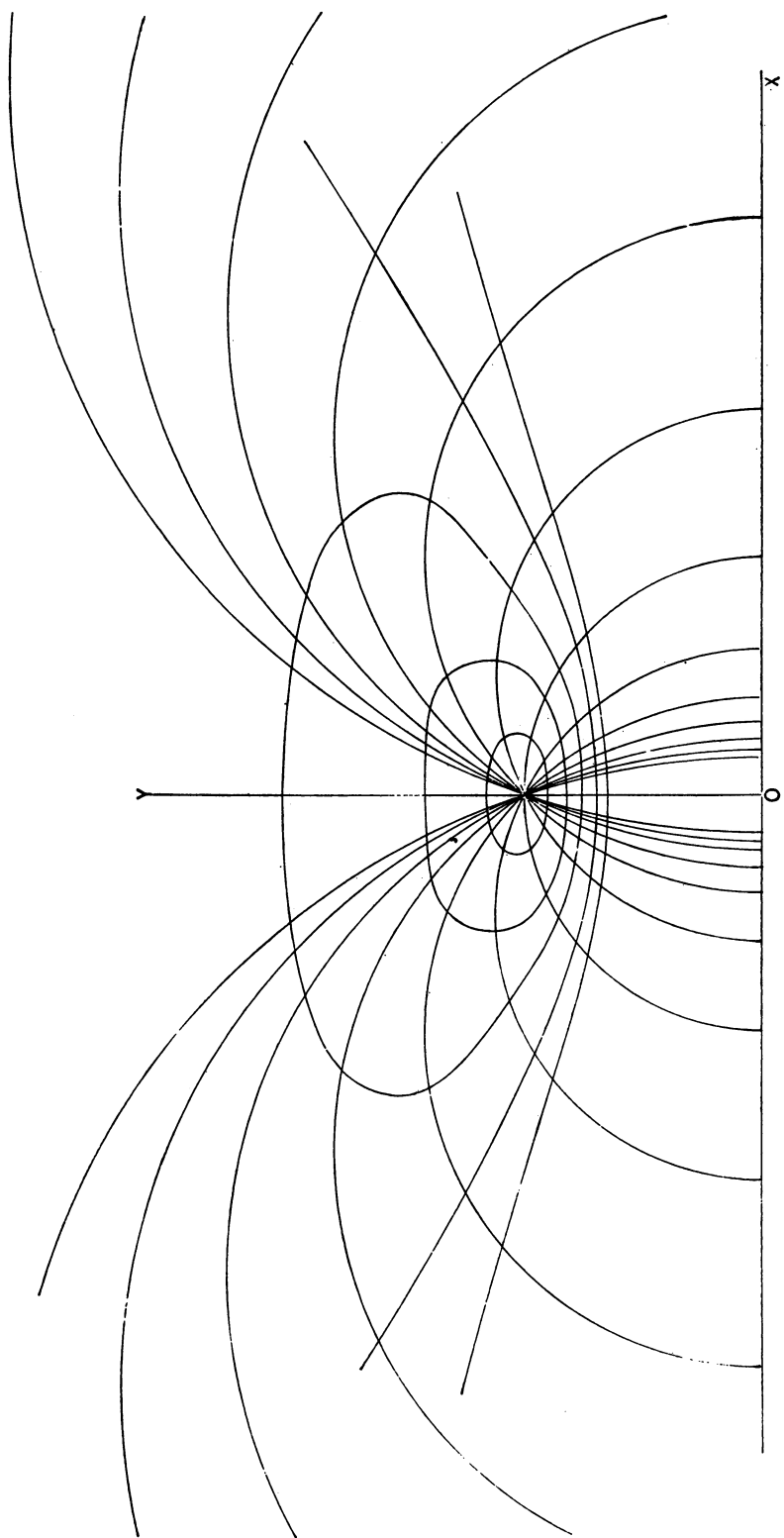


PLATE II.

each of these points is l_1 and then construct the circles (14) through these points and with centers on $x = 1$.

For any line $x = \text{const.}$, equation (12) reduces to

$$l = \log y \Big|_{y_1}^{y_2}$$

which for $y_1 = 1$ and $y_2 = y$ gives

$$(15) \quad l \Big|_1^y = \log_e y.$$

Hence to lay off a given distance l_1 from $(1, 1)$ along $x = 1$ we need only solve for y from $\log_e y = l_1$.

By this method Plate I has been constructed.

Example 2. Let

$$U = \frac{1}{y^2(1+p^2)} = \frac{\cos^2 \theta}{y^2}.$$

In this case we obtain

$$(16) \quad l = \int_{x_1}^{x_2} \frac{1+2p^2}{3y^3(1+p^2)^{\frac{3}{2}}} dx.$$

Substituting the values of y and p from the equation of the circles (2) and integrating we have

$$(17) \quad l = \frac{1}{3c} \left[\frac{x-a}{c^2 - (x-a)^2} \right]_{x_1}^{x_2}.$$

Let us construct the transversals corresponding to the extremals through $(0, 1)$. The differential equation for the transversals in this case is somewhat complicated. Instead of attempting to solve it, and to construct the transversals as in the last example, we proceed as follows. If we wish to lay off a length l_1 along any of the extremals (except $x = 1$) from $(0, 1)$, we let $x_1 = 0$, $x_2 = x$ in (17) and solve for x . This gives

$$(18) \quad x = a - \frac{c^2 - a^2}{2\{3cl_1(c_2 - a_2) - a\}} \pm \sqrt{c^2 + \frac{(c^2 - a^2)^2}{4\{3cl_1(c^2 - a^2) - a\}^2}}.$$

By means of this equation we can compute x for any given circle (2) through $(0, 1)$ and for a given l_1 .

For $x = \text{const.}$, $dx = 0$ and (16) becomes

$$l \Big|_{y_2}^{y_1} = - \frac{1}{3y^2} \Big|_{y_1}^{y_2},$$

or

$$(19) \quad l \Big|_y^\infty = \frac{1}{3y^2} \quad (\text{for } y > 0),$$

by which a given length l_1 can be laid off along $x = 1$.

Plate II has been constructed by use of (18) and (19).

PRINCETON, N. J.,

October, 1905.