CONVEX REGIONS IN THE GEOMETRY OF PATHS

By J. H. C. WHITEHEAD (*Princeton*)

[Received 15 August 1931]

1. Introduction. A classical theorem in differential geometry asserts the existence of a region C_q containing a given point q in a Riemannian space, such that any point in C_q can be joined to q by one and only one geodetic segment which does not leave C_q. A similar theorem holds for the geometry of paths, and is equivalent to the statement that a normal coordinate-system exists having q as origin. There does not, however, seem to be a proof of the theorem that a region exists in which two points are joined by one, and only one, segment of a path which does not leave the region. Such a region will be called *simple*, because not more than one, and *convex*, because at least one path joins any two points. We shall show that any non-singular point in an affine, or projective, space of paths is contained in a simple, convex region which can be made as small as we please.

Instead of the usual 'point-direction' or 'initial conditions' existence theorem for the differential equations to the paths, we use Picard's 'two-point' or 'boundary value' existence theorem. By this means the theorem is proved as a generalization of the remark that the points in a flat affine space, given in cartesian coordinates by

$$V(y^1,\dots,y^n) \leqslant 0,$$

constitute a convex region, if the quadratic form

$$\frac{\partial^2 V}{\partial y^j \partial y^k} dy^j dy^k$$

is positive definite at each point of the hypersurface

$$V(y^1,\dots,y^n) = 0.$$

Unless otherwise stated, an open region will mean an open region in the arithmetic or number space of n dimensions. That is, a set X containing the cell

$$|x^i - x^i_0| < \delta \qquad (i = 1,\dots,n),$$

for some positive δ, where x_0 is any point in X. A *closed region* $\bar{X}$ will mean the closure of an open region, X. The word region used by itself may mean either an open or a closed region. We deal only with real variables and real functions.

2. An existence theorem. There is a theorem due to E. Picard,* which asserts that differential equations of the form

$$\frac{d^2x^i}{ds^2} = f^i\left(s, x, \frac{dx}{ds}\right) \qquad (i = 1,..., n),$$

admit a unique set of solutions

$$\psi^1(x_0, x_1, s_0, s_1, s),..., \psi^n(x_0, x_1, s_0, s_1, s),$$

satisfying the boundary conditions

$$\left.\begin{aligned} \psi^i(x_0, x_1, s_0, s_1, s_0) &= x_0^i, \\ \psi^i(x_0, x_1, s_0, s_1, s_1) &= x_1^i, \end{aligned}\right\},$$

provided $f^i(s, x, \xi)$ satisfy certain continuity conditions, and s_0, s_1, x_0, x_1 are properly chosen.

We shall have to do with differential equations of the form†

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds} = 0, \tag{2.1}$$

where Γ^i_{jk} are functions of $x^1,..., x^n$. We assume Γ to be defined in the region

$$|x^i| < 2, \tag{2.2}$$

to be bounded, continuous, and to satisfy a Lipschitz condition

$$|\Gamma^i_{jk}(x_1) - \Gamma^i_{jk}(x_0)| \leqslant \Delta \sum_l |x_1^l - x_0^l|, \tag{2.3}$$

x_0 and x_1 being any points in (2.2), and Δ some positive constant. Equations of the form (2.1) have the property that if

$$\psi^i(s) \qquad (s_0 \leqslant s \leqslant s_1),$$

are solutions, so are

$$\phi^i(t) = \psi^i\left\{\frac{s_1 - s_0}{t_1 - t_0}(t - t_0) + s_0\right\} \qquad (t_0 \leqslant t \leqslant t_1),$$

where t_0 and t_1 are constants, arbitrary except that $t_0 \neq t_1$. If λ^i is the maximum of $|\psi'^i(s)|$ as s varies from s_0 to s_1, it follows that

$$\mu^i = \frac{s_1 - s_0}{t_1 - t_0}\lambda^i$$

is the maximum of $|\phi'^i(t)|$ as t varies from t_0 to t_1. This means that there is a certain homogeneity relation between the upper bound

* *Traité d'analyse*, 2nd edition, Paris, 1908, vol. iii, pp. 90–6.

† This note only applies to the restricted geometry of paths as apart from more general theories in which Γ^i_{jk} depend on the direction dx. See J. Douglas, *Annals of Math.* 29 (1928), 143–68.

which must be put on s_1-s_0, and the upper bound which must be put on $|\psi'^i|$. This relation, and the fact that the functions

$$\Gamma^i_{jk}\xi^j\xi^k \tag{2.4}$$

are defined for all values of ξ, are the special features of the equations (2.1), which enable us to prove our theorem.

We first restrict the variables ξ in (2.4) by the conditions

$$|\xi^i| < \lambda, \tag{2.5}$$

where λ is any positive constant. Let M/n^2 be an upper bound for the functions Γ^i_{jk} as x varies in (2.2). Then

$$|\Gamma^i_{jk}\xi^j\xi^k| < M\lambda^2,$$

for values of x in (2.2) and for ξ in (2.5). Further, let

$$\alpha = n^2\Delta, \tag{2.61}$$

$$\beta = 2M/n, \tag{2.62}$$

where Δ is the constant in (2.3).

Then for x_0 and x_1 in (2.2), and ξ_0 and ξ_1 in (2.5), we have

$$|\Gamma^i_{jk}(x_1)\xi^j_1\xi^k_1-\Gamma^i_{jk}(x_0)\xi^j_0\xi^k_0|$$
$$\leqslant |\{\Gamma^i_{jk}(x_1)-\Gamma^i_{jk}(x_0)\}\xi^j_1\xi^k_1|+|\Gamma^i_{jk}(x_0)\xi^j_1\xi^k_1-\Gamma^i_{jk}(x_0)\xi^j_0\xi^k_0|.$$

By (2.3) and (2.61) applied to the first term on the right-hand side, and by the first mean-value* theorem and (2.62), applied to the second term, we have

$$|\Gamma^i_{jk}(x_1)\xi^j_1\xi^k_1-\Gamma^i_{jk}(x_0)\xi^j_0\xi^k_0| \leqslant \lambda^2\alpha\sum_j |x^j_1-x^j_0|+\lambda\beta\sum |\xi^j_1-\xi^j_0|. \tag{2.7}$$

Now let x_0 be any point in the closed region

$$|x^i| \leqslant 1.$$

By Picard's existence theorem, there exists one, and only one, set of solutions to (2.1)

$$\psi^i(x_0, x_1, s_0, s_1, s), \qquad (s_0 \leqslant s \leqslant s_1), \tag{2.8}$$

such that

$$\left.\begin{aligned}\psi^i(x_0,\ldots, s_0) &= x^i_0\\ \psi^i(x_0,\ldots, s_1) &= x^i_1\end{aligned}\right\}, \tag{2.9}$$

provided

$$\left.\begin{aligned}\frac{M\lambda^2(s_1-s_0)^2}{8}+|x^i_1-x^i_0| &< 1\\ \frac{M\lambda^2(s_1-s_0)}{2}+\frac{|x^i_1-x^i_0|}{(s_1-s_0)} &< \lambda\\ \frac{\theta\lambda^2(s_1-s_0)^2}{8}+\frac{\theta\lambda(s_1-s_0)}{2} &< 1\end{aligned}\right\}, \tag{2.10}$$

* $|\partial\Gamma^i_{jk}\xi^j\xi^k/\partial\xi^l| = 2|\Gamma^i_{jl}\xi^j| < 2M\lambda/n.$

where θ is any number such that

$$n\alpha < \theta, \qquad n\beta < \theta.$$

As s varies from s_0 to s_1, $|\psi^i| < 2$ and $|\psi'^i| < \lambda$.

Let
$$\mu = \lambda(s_1 - s_0). \tag{2.11}$$

Then (2.10) may be written

$$M\mu^2/8 + |x_1^i - x_0^i| < 1, \tag{2.121}$$
$$M\mu^2/2 + |x_1^i - x_0^i| < \mu, \tag{2.122}$$
$$\mu^2/8 + \mu/2 < 1/\theta. \tag{2.123}$$

Let
$$|x_1^i - x_0^i| < a,$$

for some positive a less than unity. Then (2.121) and (2.123) are both satisfied for $\mu = 0$. Therefore they are satisfied for every small μ, and for every x_0, x_1 such that

$$|x_0^i| < 1, \qquad |x_1^i - x_0^i| < a. \tag{2.13}$$

Therefore we can find μ_0 such that (2.121) and (2.123) are satisfied, subject to (2.13), for $\mu \leqslant \mu_0$, and such that

$$M\mu_0^2/2 < \mu_0.$$

Then (2.12) are all satisfied for $\mu \leqslant \mu_0$, and

$$|x_1^i - x_0^i| \leqslant 2\delta',$$

for any positive δ' such that

$$\left.\begin{aligned} 2\delta' &< a \\ 2\delta' &< \mu_0 - M\mu_0^2/2 \end{aligned}\right\}.$$

In (2.5) let $\lambda = \mu_0$. Then there is one, and only one, solution (2.8), where $s_0 = 0$, $s_1 = 1$, and x_0 and x_1 are in the closed region

$$|x^i| \leqslant \delta'. \tag{2.14}$$

A set of n functions $\psi^i(x_0, x_1, s)$ of $2n+1$ variables, x_0, x_1, and s, is thus defined for x_0 and x_1 in (2.14) and for

$$0 \leqslant s \leqslant 1.$$

The functions $\psi^i(x_0, x_1, s)$ are continuous in the $2n+1$ variables.

On the assumption that bounded, continuous derivatives $\partial\Gamma/\partial x$ exist, this is a consequence of a general theorem proved by G. A. Bliss.* According to this theorem the functions $\psi^i(x_0, x_1, s)$ are

* *Trans. American Math. Soc.* 5 (1904), 113–25, in particular pp. 114–16. Bliss was considering the one-dimensional case, and the extension of his result from $n = 1$ to any n involves a modification of the last paragraph on p. 115. He proved that a certain derivative $\dfrac{\partial\phi}{\partial\eta_0'}$ does not vanish. For $n > 1$ it is

differentiable. We shall only need their continuity, and this follows from the Lipschitz condition, without assuming the existence* of $\partial\Gamma/\partial x$.

This section may be summarized as follows. Any integral curve of (2.1) given by

$$x^i = \psi^i(x_0, x_1, s), \qquad (s_0 \leqslant s \leqslant s_1), \tag{2.15}$$

where ψ^i satisfy (2.9), is called a path,† and x_0 and x_1 its end-points. Any parameter referred to which a path satisfies (2.1) is called an affine parameter, and the class of affine parameters consists of those, and only those, related to a given one by linear equations

$$t = \alpha s + \beta \qquad (\alpha \neq 0).$$

On each path we define a function

$$\mu(s) = \lambda(s)(s - s_0) \qquad (s_0 \leqslant s \leqslant s_1), \tag{2.16}$$

where $\lambda(s)$ is the maximum of

$$|\psi'^1(x_0, x_1, \sigma)|, \ldots, |\psi'^n(x_0, x_1, \sigma)|,$$

when σ varies from s_0 to s. The function $\mu(s)$ is continuous and strictly monotonic in s. Though it is not invariant under transformations of coordinates $x \to y$, it is invariant under transformations from one affine parameter to another. The path (2.15) will be described as a μ_0-path if

$$\mu(s) \leqslant \mu_0,$$

as s varies from s_0 to s_1. We have shown that:

One and only one μ_0-path has as its end-points a given pair of points in the closed region (2.14), *and the path varies continuously with the end-points.*

necessary to show that a corresponding Jacobian $\left|\frac{\partial \phi^i}{\partial \eta_0'^j}\right|$ does not vanish. This can be done by using the same argument as in the one-dimensional case to show that $\frac{\partial \phi^i}{\partial \eta_0'^j}\lambda^j \neq 0$, where $\lambda^1, \ldots, \lambda^n$ are given constants not all zero. Otherwise the same arguments apply for $n > 1$ as for $n = 1$.

* Let $\psi^i(x_0, x_1, s)$ be solutions to $\frac{d^2x^i}{ds^2} = f^i(s, x, dx/ds)$, which take on the boundary values x_0 and x_1 for $s = 0$ and $s = b$ respectively. On the assumption that $f^i(s, x, \xi)$ satisfy a Lipschitz condition, the continuity of $\psi^i(x_0, x_1, s)$, in x_0 and x_1, follows by an argument similar to that used by Picard (loc. cit., p. 93) in proving the convergence of approximations to a solution. We do not give this argument because it is quite straightforward, and because the existence of $\partial\Gamma/\partial x$ is necessary to so many theorems in the geometry of paths.

† We describe as a path what would usually be called a segment of a path, including the end-points.

In particular, if $x_0 = x_1$, the only μ_0-path both of whose end-points coincide with x_0 is the 'degenerate' path

$$x^i = x^i_0.$$

Thus a non-degenerate μ_0-path cannot be closed, nor can it have a double point.

3. Simple regions. We shall now show that a positive δ exists, such that not more than one path joins a given pair of points x_0 and x_1 in the region $\bar{X}_\delta$, given by

$$|x^i| \leqslant \delta, \tag{3.1}$$

without leaving $\bar{X}_\delta$. A region having this property will be described as *simple*. Any region, open or closed which is contained in a *simple* region is obviously simple.

Since not more than one μ_0-path joins a given pair of points in (2.14), it will be sufficient to prove the following:

There exists a positive $\delta \leqslant \delta'$, such that any path $\gamma(x_0, x_1)$, joining a given pair of points in $\bar{X}_\delta$ which is not a μ_0-path contains at least one point outside $\bar{X}_\delta$.

Let $$x^i = \psi^i(x_0, x_1, s) \qquad (0 \leqslant s \leqslant 1), \tag{3.2}$$
be any path joining x_0 to x_1, where x_0 and x_1 are in (2.14). From (2.1) we have
$$|\psi''^i(s)| \leqslant M\lambda(s)^2,$$
and therefore*
$$|\psi(x_0, x_1, s) - x_0| \geqslant \mu(s) - \tfrac{1}{2}M\mu(s)^2, \tag{3.3}$$

* Let $f(x)$ be any function defined for $0 \leqslant x \leqslant 1$, whose derivatives $f'(x)$ and $f''(x)$ exist and are continuous in this interval. Let
$$M_1(x) = \max|f'(\xi)|, \qquad M_2(x) = \max|f''(\xi)|,$$
as ξ varies from 0 to x. For a given x between 0 and 1 there is an x_0 such that
$$|f'(x_0)| = M_1(x) \qquad (0 \leqslant x_0 \leqslant x),$$
and by the first mean-value theorem we have
$$f'(x_0) - f'(\xi) \leqslant M_2(x)|x_0 - \xi|,$$
for any ξ between 0 and x. That is to say,
$$f'(\xi) \geqslant f'(x_0) - M_2(x)|x_0 - \xi|,$$
and on integrating both sides from 0 to x and simplifying, we have
$$f(x) - f(0) \geqslant f'(x_0)x - \tfrac{1}{2}M_2(x)x^2.$$
Therefore $$|f(x) - f(0)| \geqslant M_1(x)x - \tfrac{1}{2}M_2(x)x^2,$$
which is the result used in the text. Of course we strengthen the inequality if we replace $M_2(x)$ by a greater function.

where the omission of indices means 'at least one of $|\psi^1 - x_0^1|, \ldots, |\psi^n - x_0^n|$ exceeds the expression on the right-hand side'.

Now $\mu(0) = 0$ and $t - \frac{1}{2}Mt^2$ increases steadily as t increases from 0 to $1/M$. Let r be any number less than μ_0 and less than $1/M$. If (3.2) is not a μ_0-path,

$$r < \mu(s) < 1/M$$

for some value of s between 0 and 1. For this value of s we have, from (3.3),

$$|\psi(x_0, x_1, s) - x_0| > 2\delta \tag{3.4}$$

for any δ such that $\quad 0 < 2\delta \leqslant r - \frac{1}{2}Mr^2.$

It follows that any path which is not a μ_0-path has at least one point outside the closed region $\bar{X}_\delta$, given by

$$|x^i| \leqslant \delta. \tag{3.7}$$

As explained above, if we take $\delta \leqslant \delta'$ the region $\bar{X}_\delta$ is simple.

4. Convex regions. A region X open or closed, will be called *convex* if any two points in X are joined by at least one path which does not leave X. We may express this by saying that any two points in X are *visible** from each other in X.

Let X be any open region and $\bar{X}$ its closure. Then a path in $\bar{X}$ will be described as 'in X' if it is contained in X, with the possible exception of either end-point, or both. Two points in $\bar{X}$ will be described as visible from each other in X if they are joined by a path in X, and $\bar{X}$ will be described as *completely convex* if any two points in $\bar{X}$ are visible from each other in X.

Let $$V(x) = 0$$
be the equation of a closed hypersurface V which lies entirely in the closed region $\bar{X}_\delta$, given by (3.7). Let V have the properties:

1. The closed region, $\bar{C}$, given by

$$V(x) \leqslant 0$$

is connected.†

* A term suggested by K. Menger, *Math. Annalen*, 100 (1928), 81. We may think of X as filled with substance which conducts light along paths, all the space except X being opaque.

† That is to say, any two points in $\bar{C}$ can be joined by a continuous curve which does not leave $\bar{C}$.

2. The quadratic form

$$V_{,jk}\,dx^j dx^k = \left(\frac{\partial^2 V}{\partial x^j \partial x^k} - \frac{\partial V}{\partial x^i}\,\Gamma^i_{jk}\right) dx^j dx^k$$

is positive definite at each point x on* V.

Since the closed region $\bar{C}$, consisting of points in and on V, is contained in the simple region $\bar{X}_\delta$, it is simple, and we shall show that it is completely convex.

Let x_0 be any point on V, and

$$x^i = x^i(s) \tag{4.1}$$

a path which touches V at x_0, that is to say,

$$V(x_0) = 0, \qquad \left(\frac{\partial V}{\partial x^j}\right)_{x_0} \xi^j_0 = 0,$$

where $x(s_0) = x_0$, $\xi_0 = \left(\frac{dx}{ds}\right)_0$. Since $x^i(s)$ satisfy (2.1), it follows that

$$V\{x(s+\Delta s)\} = V_{,jk}\,\xi^j_0 \xi^k_0\,\Delta s^2 + \ldots, \tag{4.2}$$

and therefore $V\{x(s_0+\Delta s)\}$ is positive for small values of Δs. That is to say, all points on a tangent path to V, which are near the point of contact, lie outside V.

Let a and b be any two points in $\bar{C}$. Either the point pair (a,b) is visible,† or else the path joining a to b contains at least one point which is outside V. For if a and b are invisible the path ab has at least one inner point x on V, and if ab contains no point outside V it touches V at x. But the possibility of tangency from the inside is excluded by the second condition on V.

It follows that the totality of visible point pairs in V is a closed set in the $2n$-dimensional region

$$|x^i| \leqslant \delta, \qquad |y^i| \leqslant \delta,$$

* For a sufficiently small positive r such a hypersurface is given parametrically by

$$(a)\ x^i = \bar{x}^i - \tfrac{1}{2}(\Gamma^i_{jk})_0 \bar{x}^j \bar{x}^k,$$
$$(b)\ \sum \bar{x}^i \bar{x}^i - r^2 = 0.$$

For the equations (a) define a transformation to a coordinate-system $\bar{x}$ in which the components, $\bar{\Gamma}^i_{jk}$, of the affine connexion vanish at the origin, and in which the equation to V is (b). In the coordinates $\bar{x}$ we have

$$\bar{V}_{,jk}\,d\bar{x}^j d\bar{x}^k = (\delta_{jk} - \sum_i \bar{x}^i \bar{\Gamma}^i_{jk})\,d\bar{x}^j\,d\bar{x}^k,$$

and for all sufficiently small values of r this quadratic form is positive definite at points on the hypersurface V.

† A pair of points in $\bar{C}$ will be described as visible, if one is visible from the other in the open region C, and invisible otherwise.

which is the product of $\bar{X}_\delta$ with itself. For let (a, b) be any point pair on the boundary of the visible point pairs, and let

$$(a_1, b_1), \qquad (a_2, b_2), \ldots, \qquad (a_\alpha, b_\alpha), \ldots$$

be a sequence of visible point pairs converging to (a, b). Each of the paths

$$x^i = \psi^i(a_\alpha, b_\alpha, s)$$

lies in C, by the definition of visibility. On each of the paths a_α, b_α, let the parameter s be chosen so that

$$\left.\begin{aligned} \psi^i(a_\alpha, b_\alpha, 0) &= a^i_\alpha \\ \psi^i(a_\alpha, b_\alpha, 1) &= b^i_\alpha \end{aligned}\right\}.$$

Let the parameter on the path ab be similarly chosen. Then for each value of s between 0 and 1, the sequence of points

$$\psi(a_1, b_1, s), \qquad \psi(a_2, b_2, s), \ldots$$

converges to

$$\psi(a, b, s),$$

as follows from the continuity of $\psi^i(u, v, s)$ in the variables u and v. Therefore no point on the path ab lies outside V, and by the preceding paragraph the point pair (a, b) is visible. Therefore the totality of visible point pairs is closed.

Further, if (a, b) is any visible pair of points, both of which are inside V, we have

$$V\{\psi(a, b, s)\} < 0,$$

for any s between 0 and 1. By the uniform continuity of $\psi^i(a, b, s)$ we have

$$V\{\psi(a+\Delta a, b+\Delta b, s)\} < 0,$$

for all small values of Δa and Δb. Therefore the set of all visible point pairs (a, b), where a and b are both in the open region C, is open. That is to say, the set of all visible point pairs in C is both open and closed relative to the $2n$-dimensional region, $C \times C$, consisting of all point pairs in C.

Since C is connected, it follows that $C \times C$ is connected. Moreover, the degenerate point pair (a, a) is visible, where a is any point in C. Therefore the set of visible point pairs is not empty, and C is convex.* Since the set of visible point pairs in the closed region $\bar{C}$ is closed, it follows that $\bar{C}$ is completely convex.

Let U be any open region in an affine space of paths, and let P be

* If a non-vacuous sub-set of a connected set X is both open and closed, relative to X, it is the set X itself. This follows at once from our definition, and is often taken as the defining property.

a given point in U. A coordinate-system exists in which a cell contained in U is represented by the region (2.2) and P by some point inside V. We have, therefore, the theorem:

If U is any open region in an affine space of paths, and if P is any point in U, there is a simple and completely convex closed region containing P and contained in U.

Another statement of this theorem is: *an affine space of paths has a set of convex regions for a fundamental set of neighbourhoods.**

In particular the theorem applies to the geodesics in a Riemannian space. *A Riemannian space with a positive ds^2 has a set of convex spheres for a fundamental set of neighbourhoods.* For any point may be taken as the origin of normal coordinates in which the locus given by

$$y^i y^i - r^2 = 0,$$

for a small enough r, is a sphere and may be taken as V.

The theorem obviously applies to projective as well as to affine spaces of paths.

* This statement refers to a topological space with an affine connexion defined at each point. For a set of axioms describing such spaces see O. Veblen and J. H. C. Whitehead, *Proc. National Academy of Sciences*, 17 (1931), 551–61, or Chap. VI of a forthcoming Cambridge Tract by the same authors.