

AN EXISTENCE THEOREM FOR A PROBLEM OF THE CALCULUS OF VARIATIONS IN SPACE

BY

E. GORDON BILL*

It is a fundamental theorem of the theory of functions, that if any function is continuous over a closed region, it will necessarily assume its maximum and minimum values. Similarly for the calculus of variations it is of prime importance to know when the integral

$$J = \int_{t_0}^{t_1} F(x_1, \dots, x_n, x'_1, \dots, x'_n) dt$$

will have associated with it a definite curve along which it assumes its maximum or minimum value, i. e., when the calculus of variations problem will have a solution. In regard to this question HILBERT † in 1899 stated: "Eine Jede Aufgabe der Variationsrechnung besitzt eine Lösung, sobald hinsichtlich der Natur der gegebenen Grenzbedingungen geeignete Annahmen erfüllt sind und nötigenfalls der Begriff der Lösung eine sinngemässe Erweiterung erfährt."

In 1900 Professor HILBERT sketched a proof of this statement for the case of the plane, which Professor BOLZA ‡ and others have since completed.

It is the object of this paper to demonstrate this *a priori* existence theorem for space, i. e., we shall prove that under certain general conditions every problem of the calculus of variations in space admits a solution. Without loss of generality we may consider only the case of a minimum.

The proof will be accomplished in two parts: first, as a result of the researches of BLISS and MASON in these *T r a n s a c t i o n s*, the theorem will be seen to be true "im Kleinen," i. e., where the region considered is sufficiently small; second, it will be demonstrated "im Grossen" by a method whereby the minimizing curve is obtained by applying the existence theorem "im Kleinen" to only a *finite number* of points, which are defined as limiting points of points lying on a sequence of approximating curves.

We shall assume that all curves considered, i. e., all "allowed" curves are

* Presented to the Society, December 29, 1908.

† *Jahresbericht der deutschen Mathematiker-Vereinigung*, vol. 8 (1899), p. 184.

‡ BOLZA, *Vorlesungen über Variationsrechnung*, pp. 419-436.

rectifiable and except at possibly a finite number of points have continuously turning tangents. The integral J considered is defined in the generalized sense of Weierstrass.

It will be convenient to represent the integral

$$J = \int_{t_1}^{t_2} F(x, y, z, x', y', z') dt,$$

when carried over the curve C from the point P_1 to P_2 by $J(C_{P_1P_2})$. If carried over a curve C_i we shall designate it by $J_i(C_{P_1P_2})$.

The term "field sphere" will be used for a sphere of such radius that there exists a set of extremals through its center simply covering the interior of the sphere with the exception of the center itself. Finally the expression "sphere (P, r) " will represent a sphere of radius r and center P .

§ 1. The existence theorem for a small region.

We shall make the following assumptions:

I. $F(x, y, z, x', y', z')$ is of class C''' for all points (x, y, z) in a region of space R' and for $x'^2 + y'^2 + z'^2 \neq 0$, and satisfies the condition

$$F(x, y, z, \kappa x', \kappa y', \kappa z') = \kappa F(x, y, z, x', y', z') \quad (\kappa > 0),$$

in the same domain.

II. $F_1(x, y, z, \alpha, \beta, \gamma)^* > 0$ for (x, y, z) in a finite, closed region R entirely interior to R' , and for all values of α, β, γ such that

$$\alpha^2 + \beta^2 + \gamma^2 = 1.$$

III. $F_{x'x'}, F_{y'y'}, F_{z'z'} \geq 0$ for values $(x, y, z, \alpha, \beta, \gamma)$ in the same region.

This set of hypotheses is sufficient to prove the existence of a solution of the minimum problem "im Kleinen," i. e.,

THE AUXILIARY THEOREM. *Positive constants d, e can be determined such that there passes from any point P_1 interior to R , to any other point P_2 also interior to R exactly one extremal arc E of class C'' , provided that the distance $P_1 P_2$ is less than d . This extremal is of length less than e , lies entirely within the sphere (P_1, d) , has no multiple point, and gives to J a smaller value than that of any other curve of class D' joining P_1 and P_2 and lying in the sphere (P_1, d) .*

The validity of this theorem follows from the above mentioned work of Bliss and Mason, namely from the following three theorems:

* For the definition and properties of these functions see the paper by BLISS and MASON, *The properties of curves in space which minimize a definite integral*, Transactions of the American Mathematical Society, vol. 9 (1908), p. 440.

(a)* The two parameter family of extremals

$$x = \varphi (s; x_i, y_i, z_i, \alpha, \beta, \gamma),$$

$$y = \psi (s; x_i, y_i, z_i, \alpha, \beta, \gamma),$$

$$z = \chi (s; x_i, y_i, z_i, \alpha, \beta, \gamma)$$

through a fixed point $P_i (x_i, y_i, z_i)$ of the region R , simply cover the interior of a sphere (P_i, ρ_i) with exception of the point P_i itself.

It is to be noted that the functions φ, ψ, χ and their first and second derivatives with respect to s are of class C' for (x_i, y_i, z_i) in $R, 0 \leq s \leq e$ and $\alpha^2 + \beta^2 + \gamma^2 = 1$; where s is the length of arc measured along an extremal from the fixed point P_i , and α, β, γ are the direction cosines of the tangents to the extremals at P_i .

(b)† The radii ρ_i of these field spheres are not evanescent for any set of points of R .

(c)‡ The existence of a field about a given extremal arc, which itself is one of the field, is, together with the assumptions II and III, a sufficient condition for the existence of a strong minimum.

The truth of the statement that E has no multiple point follows from the fact that if r represents the rectilinear distance from P_1 to a point of E interior to (P_1, d) , the derivative dr/ds is positive for $0 \leq s \leq e, (x_1, y_1, z_1)$ in R , and $\alpha^2 + \beta^2 + \gamma^2 = 1$.

From the three known theorems above, we are assured of the existence of a minimizing arc $E_{P_1 P_2}$ of the integral J , provided that the distance from P_1 to P_2 is sufficiently small.

§ 2. The existence theorem in general:

In addition to the assumptions made on F and allied functions in §1, we make the following assumptions and call the totality I–V, assumptions (A):

IV. $F(x, y, z, \alpha, \beta, \gamma) > 0$ for (x, y, z) in R and $\alpha^2 + \beta^2 + \gamma^2 = 1$.

V. R is "extremal convex."§

By the aid of these hypotheses (I–V) we can demonstrate the existence of a curve which minimizes J , i. e., we shall prove:

THE PRINCIPAL THEOREM: *Any two points A and B of the region R can be joined*

* BILL, *The construction of a space field of extremals*, Bulletin of the American Mathematical Society, vol. 15 (1909), p. 374; and BLISS and MASON, *Fields of extremals in space*, Transactions of the American Mathematical Society, vol. 11 (1910), pp. 328–332.

† BLISS and MASON, loc. cit., vol. 11, p. 332.

‡ BLISS and MASON, loc. cit., vol. 9.

§ BOLZA, loc. cit., p. 276.

by at least one curve E , which is an extremal of class C'' , and for which the integral

$$J = \int_A^B F(x, y, z, x', y', z') dt$$

is less than or equal to the value given to J by any other curve C of class D' lying in R and joining A and B .

As the region R is closed, it follows from assumptions I and IV that for every point (x, y, z) of R and $\alpha^2 + \beta^2 + \gamma^2 = 1$ we have the inequalities

$$(1) \quad F(x, y, z, \alpha, \beta, \gamma) > m > 0,$$

where m is a properly chosen constant.

Hence if C is any curve joining A and B and lying in R , L its length from A to B , and \overline{AB} the rectilinear distance from A to B , it follows that

$$J(C_{AB}) \geq m \cdot L \geq m \cdot \overline{AB} > 0,$$

and hence there exists a lower limit $K > 0$ of the values $J(C_{AB})$ taken over the totality of all allowed curves C joining A and B and lying in R . It follows then from the definition of a lower limit, that from these curves C we can pick out an infinite set of curves

$$\mathfrak{C}: C_1, C_2, C_3, \dots$$

lying in R and such that

$$(2) \quad \lim_{i \rightarrow \infty} J_i(C_{AB}) = K, \quad J_r(C_{AB}) > J_s(C_{AB}) \text{ if } r < s.$$

We shall next define an important set of points $P_i^{(k)}$ ($k = 0, 1, \dots, n$) where $P_i^{(0)} \equiv A$ and $P_i^{(n)} \equiv B$. These are points dividing C_i into n parts such that

$$(3) \quad J_i(C_{P_i^{(k)} P_i^{(k+1)}}) = \frac{1}{n} J_i(C_{AB}).$$

Among these points consider the points $P_i^{(1)}$, i. e., all the *first* points of division of C_i . As this point set is infinite and limited, it has at least one limiting point $P^{(1)}$ in R . We can then pick out from the curves of \mathfrak{C} , the set

$$\mathfrak{C}_1: C_1^{(1)}, C_2^{(1)}, C_3^{(1)}, \dots$$

such that

$$\lim_{i \rightarrow \infty} P_i^{(1)} = P^{(1)},$$

and hence a set whose first points of division have a *single* limiting point $P^{(1)}$.

On the set of curves \mathfrak{C}_1 are the second points of division $P_i^{(2)}$, which have at least one limiting point $P^{(2)}$, and we now discard all the curves of \mathfrak{C}_1 except the set

$$\mathfrak{C}_2: C_1^{(2)}, C_2^{(2)}, C_3^{(2)}, \dots,$$

which have the property that

$$\lim_{i=\infty} P_i^{(2)} = P^{(2)}.$$

At the same time for this set \mathfrak{C}_2 we must also have

$$\lim_{i=\infty} P_i^{(1)} = P^{(1)},$$

as this latter point set $P_i^{(1)}$ is contained in the set originally designated $P_i^{(1)}$ which had but one limiting point $P^{(1)}$.

Thus we pass from \mathfrak{C}_2 to \mathfrak{C}_3 and finally obtain a set of allowed curves joining A and B and lying in R , namely

$$\mathfrak{C}_n : C_1^{(n)}, C_2^{(n)}, C_3^{(n)}, \dots$$

such that

$$(4) \quad \lim_{i=\infty} P_i^{(k)} = P^{(k)} \quad (k = 0, 1, \dots, n),$$

where $P^{(0)} \equiv A$ and $P^{(n)} \equiv B$.

We can now prove the important theorem:

THEOREM I. *The minimum value of the integral J over all allowed curves, which lie in R and join $P^{(k)}$ and $P^{(k+r)}$, is rK/n .*

In the proof of this and the following theorems we shall adopt the following notation:

$$L_i(k, k+r) \equiv \text{length of } C_i \text{ from } P_i^{(k)} \text{ to } P_i^{(k+r)};$$

$$\overline{P_1 P_2} \equiv \text{the rectilinear distance from } P_1 \text{ to } P_2;$$

$$S_{P_1 P_2} \equiv \text{the straight line determined by } P_1 \text{ and } P_2.$$

By the definition of a lower limit, i can be taken so great that simultaneously

$$\overline{P^{(k)} P_i^{(k)}} < \frac{\epsilon}{3M}, \quad \overline{P_i^{(k+r)} P^{(k+r)}} < \frac{\epsilon}{3M},$$

where M is the finite maximum value of F in R , and ϵ is any positive constant. Hence

$$H \equiv J(S_{P^{(k)} P_i^{(k)}}) < M \cdot \overline{P^{(k)} P_i^{(k)}} < \frac{\epsilon}{3}$$

and

$$Q \equiv J(S_{P_i^{(k+r)} P^{(k+r)}}) < M \cdot \overline{P_i^{(k+r)} P^{(k+r)}} < \frac{\epsilon}{3}.$$

Moreover from equations (3) we shall have, when i is chosen sufficiently large,

$$G \equiv J_i(C_{P_i^{(k)} P_i^{(k+r)}}) = \frac{r}{n} \cdot J_i(C_{AB}) < \frac{r}{n} \cdot K + \frac{\epsilon}{3},$$

and hence finally

$$H + G + Q < \frac{r \cdot K}{n} + \epsilon.$$

Let us for the moment designate by $\min. J(P_1 P_2)$, the minimum value of J carried over all allowed curves joining P_1 and P_2 and lying in R . Then the last inequality shows that

$$\min. J(P^{(k)} P^{(k+r)}) \leq \frac{r \cdot K}{n}.$$

But here the inequality sign cannot hold; for assume the contrary, i. e.,

$$a \equiv \min. J(P^{(k)} P^{(k+r)}) < \frac{r \cdot K}{n}.$$

As we know that

$$b \equiv \min. J(AP^{(k)}) \leq \frac{k \cdot K}{n},$$

$$c \equiv \min. J(P^{(k+r)}B) \leq \frac{(n - k - r) \cdot K}{n},$$

it follows from the properties of lower limits that

$$\min. J(AB) = K \leq b + a + c < K,$$

which is absurd, and hence

$$(5) \quad \min. J(P^{(k)} P^{(k+r)}) = \frac{r \cdot K}{n}.$$

Let us now so choose the constants d , e defined in the auxiliary theorem, that if for any two points P_1 , P_2 in R we have $P_1 P_2 < d$, the extremal joining P_1 and P_2 will lie entirely in R , a choice which is possible from assumption V. In order to complete the proof of the principal theorem, we need to be sure that the curves joining $P^{(k)}$, $P^{(k+1)}$ and $P^{(k+2)}$, and made up of the curves $S_{P^{(k)}P_i^{(k)}}$, $C_{P_i^{(k)}P_i^{(k+1)}}$ and $S_{P_i^{(k+1)}P_i^{(k+1)}}$, lie entirely in the field sphere $(P^{(k)}, d)$. Hence we prove

THEOREM II. *The number of divisions n can be chosen so large that*

$$L_i(k, k+1) < \frac{d}{2}, \quad L_i(k+1, k+2) < \frac{d}{2}, \quad \text{and} \quad L_i(k, k+2) < d.$$

The value of n which it is sufficient to take is any value satisfying the inequalities

$$(6) \quad n > \frac{2J_1(C_{AB})}{m \cdot d} \geq \frac{2J_i(C_{AB})}{m \cdot d} \quad (i = 1, 2, \dots)$$

on account of the property (1) of the function F . For we have

$$J_i(C_{P_i^{(k)}P_i^{(k+1)}}) \geq m \cdot L_i(k, k+1),$$

$$J_i(C_{P_i^{(k+1)}P_i^{(k+2)}}) \cong m \cdot L_i(k+1, k+2),$$

$$J_i(C_{P_i^{(k)}P_i^{(k+2)}}) \cong m \cdot L_i(k, k+2).$$

Hence by (2) and (6) we have

$$L_i(k, k+1) \cong \frac{J_i(C_{P_i^{(k)}P_i^{(k+1)}})}{m} \cong \frac{J_i(C_{AB})}{m \cdot n} < \frac{d}{2},$$

$$L_i(k+1, k+2) \cong \frac{J_i(C_{P_i^{(k+1)}P_i^{(k+2)}})}{m} \cong \frac{J_i(C_{AB})}{m \cdot n} < \frac{d}{2},$$

$$L_i(k, k+2) \cong \frac{J_i(C_{P_i^{(k)}P_i^{(k+2)}})}{m} \cong \frac{2J_i(C_{AB})}{m \cdot n} < d.$$

Hence by the definition of limiting points we can choose i so large that the above mentioned curves joining the points $P^{(k)}$, $P^{(k+1)}$ and $P^{(k+2)}$ in pairs have lengths less than d and hence lie entirely in the field sphere $(P^{(k)}, d)$. At the same time we can replace in the above inequalities the points $P_i^{(k)}$, $P_i^{(k+1)}$, $P_i^{(k+2)}$ respectively by the points $P^{(k)}$, $P^{(k+1)}$, $P^{(k+2)}$, and can hence apply to these latter points the results of our auxiliary theorem.

This leads us to the fundamental theorem:

THEOREM III: *The integral J carried over the extremal arcs $E_{P^{(k)}P^{(k+1)}}$ and $E_{P^{(k)}P^{(k+2)}}$ has the values K/n and $2K/n$ respectively.*

It will suffice in order to prove the above theorem to consider simply the value

$$J(E_{P^{(k)}P^{(k+1)}}).$$

Let i be so chosen that

$$\overline{P_i^{(k)}P^{(k)}} < \frac{\epsilon}{3M}, \quad \overline{P_i^{(k+1)}P^{(k+1)}} < \frac{\epsilon}{3M}.$$

Then we have the inequality

$$J(S_{P^{(k)}P^{(k)}}) + J_i(C_{P_i^{(k)}P_i^{(k+1)}}) + J(S_{P_i^{(k+1)}P^{(k+1)}}) < \frac{K}{n} + \epsilon,$$

provided that i is taken sufficiently large, which shows that

$$J(\bar{C}_{P^{(k)}P^{(k+1)}}) \cong \frac{K}{n},$$

where \bar{C} denotes the curve

$$S_{P^{(k)}P^{(k)}} + C_{P_i^{(k)}P_i^{(k+1)}} + S_{P_i^{(k+1)}P^{(k+1)}}.$$

But we are sure that the extremal $E_{P^{(k)}P^{(k+1)}}$ will not give to J a value greater than the one given by the curve \bar{C} , which by THEOREM II also lies in the field sphere $(P^{(k)}, d)$, and by THEOREM I we know that $E_{P^{(k)}P^{(k+1)}}$ cannot give to J a value less than K/n .

Hence finally we have shown that

$$J(E_{P^{(k)}P^{(k+1)}}) = \frac{K}{n}, \quad J(E_{P^{(k)}P^{(k+2)}}) = \frac{2K}{n}.$$

We can now complete the proof of the principal theorem by demonstrating

THEOREM IV: *The arcs of extremal $E_{P^{(k)}P^{(k+1)}}$ ($k = 0, 1, \dots, n-1$) form a continuous curve of class C'' joining A and B and over which the integral J assumes its minimum value K .*

It is certain that the point $P^{(k+1)}$ must fall on the arc $E_{P^{(k)}P^{(k+2)}}$ as otherwise we should have two curves $E_{P^{(k)}P^{(k+2)}}$ and $E_{P^{(k)}P^{(k+1)}} + E_{P^{(k+1)}P^{(k+2)}}$, lying in the field sphere $(P^{(k)}, d)$, over which J assumes its minimum value $2K/n$, and this is impossible.

Extending the same arguments to the points $P^{(k+1)}$, $P^{(k+2)}$, $P^{(k+3)}$ and $P^{(k-1)}$, $P^{(k)}$, $P^{(k+1)}$ we see that all the extremal arcs joining these in pairs form a curve E_{AB} , which has a continuously turning tangent and along which J assumes its minimum value K . From the properties of the functions φ , ψ and χ , the curve E_{AB} is of class C'' .

§ 3. Other assumptions sufficing to insure the existence of an absolute minimum.

The importance of the above existence theorem results partly from the fact, that if we know all the curves which make an integral a relative minimum, and at the same time know in advance the *existence* of a curve making the integral an absolute minimum, then the latter curve will always be found among the former. The existence of an absolute minimum has been proved on the basis of the assumptions (A). It will be interesting in this section to see that the proof can be made under the other quite general circumstances described below under (B) and (C).

HYPOTHESES (B):

- I. The same as (A) I.
- II'. $F_1(x, y, z, \alpha, \beta, \gamma) > 0$ for (x, y, z) in R' and $\alpha^2 + \beta^2 + \gamma^2 = 1$.
- III'. $F_{x'x'}, F_{y'y'}, F_{z'z'} \geq 0$ for (x, y, z) in R' and $\alpha^2 + \beta^2 + \gamma^2 = 1$.
- IV'. $F(x, y, z, \alpha, \beta, \gamma) > m$ for (x, y, z) in R' and $\alpha^2 + \beta^2 + \gamma^2 = 1$, where m denotes a positive constant.
- V'. The region R' shall be the whole of space.

HYPOTHESES (C):

- I. The same as (A) I and (B) I.
- II'. The same as (B) II'.
- III'. The same as (B) III'.
- IV''. For any point (x_0, y_0, z_0) a constant $\lambda(x_0, y_0, z_0) > 0$ can be so chosen that if we set

$$\delta \equiv \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2},$$

the function F will satisfy the inequality

$$F(x, y, z, \alpha, \beta, \gamma) > \frac{\lambda}{1 + \delta}$$

for (x, y, z) in R' and $\alpha^2 + \beta^2 + \gamma^2 = 1$.

V' . The same as $(B) V'$.

In order to use the proof of the theorem derived in the preceding sections, for assumptions (B) and (C) , it will evidently suffice to show that there must exist under these hypotheses a finite closed region of space R_1 within which all the curves, joining A and B , that we need to consider must lie and which itself is interior to a finite closed region R .

As we are considering only the case where the extremum of J is a minimum, it will suffice in order to obtain this restricted region to prove the following theorem:

Under hypotheses (B) or (C) , for each $G > 0$ we can choose a finite closed region R_1 of space, so large that if any allowed curve C joining A and B goes outside of R_1 , then $J(C_{AB}) > G$.

Let a and b denote the minimum distances from $A(x_0, y_0, z_0)$ and $B(x_1, y_1, z_1)$ respectively, to the boundary of any finite closed region containing A and B on its interior.

We then choose under (B) the region R_1 as such that

$$a + b > \frac{G}{m}$$

and under (C) so that

$$a + b > e^{\frac{G}{\lambda(x_0, y_0, z_0)}}.$$

Then if L denotes the length of any curve C joining A and B and going outside of R_1 , we have for the cases (B) and (C) , respectively,

$$(B) \quad J(C_{AB}) \cong m \cdot L > m(a + b) > G;$$

$$(C) \quad J(C_{AB}) \cong \int_A^B \frac{\lambda ds}{1 + d} \cong \int_A^B \frac{\lambda ds}{1 + s} \cong \lambda \log(1 + L) > \lambda \log L > G.$$

where in λ and d we suppose the coördinates of A substituted for x_0, y_0, z_0 .

Thus we have shown that for every given minimum problem, there must exist under hypotheses (B) and (C) a finite closed region of space R_1 within which all curves considered must lie. For the region R we choose any finite closed region including R_1 on its interior.

Finally we choose the constants d and e , defined in the auxiliary theorem, for the region R and choose them so small that any extremal of length $\leq e$ through a point of R_1 will be necessarily within R . This insures that our region R is "extremal convex" and the theorems proven for hypotheses (A) now apply.