SOME REMARKS CONCERNING CARATHÉODORY'S METHOD OF "EQUIVALENT" INTEGRALS IN THE CALCULUS OF VARIATIONS

BY

H. RUND

(Communicated by Prof. J. F. KOKSMA at the meeting of December 20, 1958)

§ 1. *Introduction*

We shall be concerned with parametric problems in the calculus of variations as exemplified by an integral of the type

$$I = \int F(x^i, \dot{x}^i)\, dt, \quad \left(\dot{x}^i = \frac{dx^i}{dt};\ i = 1, 2, \ldots, n\right), \tag{1.1}$$

where it is assumed that the integrand $F(x^i, \dot{x}^i)$ is of class C^2 with respect to its $2n$ arguments and positively homogeneous of the first degree in the variables $\dot{x}^i$. If the function $F(x^i, \dot{x}^i)$ does not only assume positive values, it is often of utmost importance to know whether or not it is possible to replace the integrand in (1.1) by a new function $F^*(x^i, \dot{x}^i)$, which is always positive [1]) and such that the extremals of the new problem are identical with those of (1.1). For instance, if we put [2])

$$F^*(x^i, \dot{x}^i) = F(x^i, \dot{x}^i) + \frac{\partial S}{\partial x^i}\dot{x}^i, \tag{1.2}$$

where $S = S(x^i)$ is some function depending on the x^i only, the extremals corresponding to the integrand $F^*(x^i, \dot{x}^i)$ coincide with those of (1.1), since the addition of an exact differential to the integrand obviously cannot affect any extremals. Problems in the calculus of variations whose integrals are related according to (1.2) are called *equivalent*.

However, it may not always be possible to find a suitable function $S(x^i)$ which ensures that $F^*(x^i, \dot{x}^i)$ be positive. This is indicated by the following result, due to CARATHÉODORY [3]).

Theorem: If the integrand $F(x^i, \dot{x}^i)$ is defined for all line-elements

[1]) We remark that in the theory of Finsler spaces, based on integrals of the type (1.1), it is invariably assumed that the integrand is positive. It will be evident that this is indeed a very strong assumption, which indicates the need for a more general geometrical theory.

[2]) The summation convention (according to which repeated indices imply summation over the range 1 to n) will be used throughout.

When x^i or $\dot{x}^i$ appear as suffixes, they denote partial differentiation with respect to the variable shown.

[3]) See [1], § 1 for the case $n = 2$; for the general case [2], p. 243.

$(x^i, \dot{x}^i)$ of a region of the space R_n of the x^i, and if at some point $x^i_{(0)}$ of R_n there exists a strong positive regular line-element $(x^i_{(0)}, \dot{x}^i_{(0)})$, it is possible to construct equivalent integrands which are positive in a neighbourhood of the point $x^i_{(0)}$.

The essential condition, then, appears to be the existence of some strong positive regular line-element $(x^i_{(0)}, \dot{x}^i_{(0)})$, i.e. that at $x^i_{(0)}$ the Weierstrass excess function be positive for all values $\dot{x}^i \neq k\dot{x}^i_{(0)}$, $(k \geqslant 0)$:

$$(1.3) \qquad E(x^j_{(0)}, \dot{x}^j_{(0)}, \dot{x}^j) \equiv F(x^j_{(0)}, \dot{x}^j) - F_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)})\,\dot{x}^i > 0.$$

In § 2 we shall show that this condition imposes a very severe restriction on the types of integrands which may be replaced by positive equivalent integrands. In fact, suppose that $F(x^i, \dot{x}^i)$ satisfies the symmetry condition

$$(1.4) \qquad F(x^i, -\dot{x}^i) = F(x^i, \dot{x}^i),$$

such integrands being of particular significance in the calculus of variations since (1.4) ensures that the value of the integral (1.1) is independent of the direction of integration. It will be seen that for integrands of this type the process exemplified by (1.2) is superfluous, while for a large class of integrands condition (1.3) cannot be fulfilled, this restriction being effective in view of a converse to Carathéodory's theorem.

In § 3 we shall indicate that the application of (1.2) is equivalent to a projective transformation applied to the extremals; this in turn gives rise to certain implications – albeit of a negative nature – with regard to the inverse problem of the calculus of variations.

§ 2. *Applicability of the process* (1.2)

We shall now prove the following

Theorem: Let $F(x^i, \dot{x}^i)$ be a function of class C^2, defined for all line-elements $(x^i, \dot{x}^i)$ of a region G of R_n. If $F(x^i, \dot{x}^i)$ is positively homogeneous of the first degree in the $\dot{x}^i$, while satisfying condition (1.4), and if there exists a line-element $(x^i_{(0)}, \dot{x}^i_{(0)})$ of G for which condition (1.3) is satisfied, then $F(x^i, \dot{x}^i)$ is automatically positive in a neighbourhood of $x^i_{(0)}$.

Proof: If we put $\dot{x}^i = -\dot{x}^i_{(0)}$ in condition (1.3), we immediately deduce by means of (1.4) and Euler's theorem on homogeneous functions that

$$(2.1) \qquad F(x^i_{(0)}, \dot{x}^i_{(0)}) > 0.$$

Let us write (1.3) in the form:

$$(2.2) \qquad [F(x^j_{(0)}, \dot{x}^j) - F(x^j_{(0)}, \dot{x}^j_{(0)})] + [F(x^j_{(0)}, \dot{x}^j_{(0)}) - F_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)})\,\dot{x}^i] > 0,$$

for all $\dot{x}^i \neq k\dot{x}^i_{(0)}$ $(k \geqslant 0)$.

Suppose now that $F(x^i_{(0)}, \dot{x}^i)$ is not positive for all line-elements $(x^i_{(0)}, \dot{x}^i)$.

Then there exists at least one direction ξ^i ($\xi^i \neq 0$) at $x^i_{(0)}$ for which $F(x^i_{(0)}, \xi^i) = 0$, i.e.

$$F(x^i_{(0)}, \lambda\xi^i) = 0 \tag{2.3}$$

for all values of λ, which may be positive or negative according to (1.4). In the tangent space $T_n(x^i_{(0)})$ of R_n at $x^i_{(0)}$ we may define a hyperplane Π by means of the equation

$$F(x^j_{(0)}, \dot{x}^j_{(0)}) - F_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)})\,\dot{x}^i = 0. \tag{2.4}$$

We now have to consider two possibilities:

Case (a): The vector ξ^i (produced if necessary) issuing from the origin of $T_n(x^i_{(0)})$ intersects Π, i.e. there exists a number μ such that $\dot{x}^i = \mu\xi^i$ satisfies (2.4). If these values are substituted in condition (2.2) while (2.3) is taken into account we obtain

$$-F(x^j_{(0)}, \dot{x}^j_{(0)}) > 0,$$

which contradicts (2.1).

Case (b): The vector ξ^i (produced) does not intersect Π. In this case ξ^i is parallel to Π, or transversal to $\dot{x}^i_{(0)}$:

$$F_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)})\,\xi^i = 0.$$

Putting $\dot{x}^i = \xi^i$ in (1.3), we obtain

$$F(x^j_{(0)}, \xi^i) > 0,$$

which contradicts (2.3).

Thus in both cases we arrive at a contradiction, which shows that under the conditions of the theorem $F(x^i_{(0)}, \dot{x}^i)$ cannot change sign, so that our assertion follows directly from (2.1) and the continuity of $F(x^i, \dot{x}^i)$.

An immediate consequence of the above analysis is the fact that a function $F(x^i, \dot{x}^i)$ which satisfies condition (1.4) and which assumes positive as well as non-positive values at $x^i_{(0)}$ cannot possess a strong positive regular line-element at $x^i_{(0)}$. Also, a glance at (1.3) shows that the same conclusion applies to all functions for which the relations

$$F(x^i, -\dot{x}^i) = -F(x^i, \dot{x}^i) \tag{2.5}$$

are satisfied identically, in particular, therefore to all rational integrands $F(x^i, \dot{x}^i)$, these being homogeneous of the first degree in $\dot{x}^i$.

We may conclude that no problems in the calculus of variations corresponding to such classes of functions can be replaced by equivalent positive definite problems. This follows directly from *the converse of Carathéodory's theorem:* "If there exists a function $S(x^i)$ such that $F^*(x^i, \dot{x}^i)$ as defined by (1.2) is positive for all line-elements $(x^i, \dot{x}^i)$ of a region G of R_n (provided not all the $\dot{x}^i$ vanish simultaneously), then there exists at least one strong positive line-element of $F(x^i, \dot{x}^i)$ at each point x^i of G".

This assertion is easily established as follows: in accordance with the usual definition of the Weierstrass excess function (left hand side of (1.3)) we define

$$E^*(x^j_{(0)}, \dot{x}^j_{(0)}, \dot{x}^j) = F^*(x^j_{(0)}, \dot{x}^j) - F^*_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)})\,\dot{x}^i. \tag{2.6}$$

Differentiation of (1.2) yields

$$F^*_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)}) = F_{\dot{x}^i}(x^j_{(0)}, \dot{x}^j_{(0)}) + S_{x^i}(x^j_{(0)}).$$

If this result, together with (1.2), is substituted in (2.6), we find

$$E^*(x^j_{(0)}, \dot{x}^j_{(0)}, \dot{x}^j) = E(x^j_{(0)}, \dot{x}^j_{(0)}, \dot{x}^j). \tag{2.7}$$

Hence the excess function is unaffected by the transition (1.2), and it follows that F has a strong positive regular line-element if F^* possesses one. But $F^*(x^i, \dot{x}^i) > 0$ by hypothesis, so that the set $F^*(x^i_{(0)}, \dot{x}^i) \leqslant 1$ is bounded in $T_n(x^i_{(0)})$. It is therefore immediately obvious from a purely geometrical point of view that this set possesses supporting hyperplanes, each of which defines a line-element of the required type [4]. – [Alternatively, one may approach this problem analytically by actually constructing a strong positive regular line-element. Consider extreme values of $E^*(x^i_{(0)}, \dot{x}^i_{(0)}, \dot{x}^i)$ subject to $F^*(x^i_{(0)}, \dot{x}^i) = 1$, $(x^i_{(0)}, \dot{x}^i_{(0)})$ being an arbitrary initial line-element. If, apart from the trivial solution corresponding to $\dot{x}^i = \dot{x}^i_{(0)}$, these extreme values are positive, $(x^i_{(0)}, \dot{x}^i_{(0)})$ is of the required type. Otherwise, let $\dot{x}^i_{(l)}$ be the direction giving the least extreme value, and consider extreme values of $E^*(x^i_{(0)}, \dot{x}^i_{(l)}, \dot{x}^i)$, again subject to $F^*(x^i_{(0)}, \dot{x}^i) = 1$. It is easily verified that these extreme values are positive (apart from that given by $\dot{x}^i = \dot{x}^i_{(l)}$), so that $(x^i_{(0)}, \dot{x}^i_{(l)})$ is a strong positive regular line-element].

Before concluding this section, we should refer to a theorem due to Damköhler and Hopf [5] according to which a function $S(x^i)$, giving rise to an equivalent positive definite problem, may be constructed in a region G, provided there exists a pair of fixed points $x^i_{(1)}$, $x^i_{(2)}$ in G such that the integral (1.1) taken from $x^i_{(1)}$ to $x^i_{(2)}$ along any (piece-wise continuous) path in G possesses a lower bound. Clearly the converse of this statement is also valid: and since we have shown that (1.3) is also a necessary condition, it follows that the condition of Damköhler and Hopf is equivalent to that of Carathéodory's theorem [6].

---

[4]) See [2], p. 244.

[5]) See [3], p. 13.

[6]) It should be pointed out that in [3] it is not assumed that the integrand F is of class C^2; strictly speaking, then, the equivalence of those conditions should be stated subject to the differentiability conditions of Carathéodory's theorem.

§ 3. *The transition* (1.2) *as a projective transformation*

We now impose an additional condition, namely that

$$\det \mid F^2_{\dot{x}^i \dot{x}^j}(x^k, \dot{x}^k) \mid > 0.$$

It is then possible to write the differential equations of the extremals in the form [7])

$$\frac{d^2x^i}{ds^2} + 2G^i\left(x^k, \frac{dx^k}{ds}\right) = 0, \tag{3.1}$$

where the special parameter s is defined by

$$ds = F(x^i, dx^i), \tag{3.2}$$

and the functions G^i are given by

$$F^2_{\dot{x}^h \dot{x}^i} G^i = \tfrac{1}{2}(F^2_{\dot{x}^h x^k} \dot{x}^k - F^2_{x^h}). \tag{3.3}$$

We remark that the G^i are positively homogeneous of degree 2 in the $\dot{x}^i$.

Let us now subject (3.1) to a parameter transformation: $\sigma = \sigma(s)$, σ being arbitrary for the present except for the condition $(d\sigma/ds) \neq 0$. The differential equations (3.1) now assume the form

$$\frac{d^2x^i}{d\sigma^2} + 2G^i\left(x^k, \frac{dx^k}{d\sigma}\right) + \frac{dx^i}{d\sigma}\left(\frac{d^2\sigma}{ds^2}\right)\left(\frac{d\sigma}{ds}\right)^{-2} = 0, \tag{3.4}$$

or, alternatively,

$$\frac{dx^j}{d\sigma}\left[\frac{d^2x^i}{d\sigma^2} + 2G^i\left(x^k, \frac{dx^k}{d\sigma}\right)\right] = \frac{dx^i}{d\sigma}\left[\frac{d^2x^j}{d\sigma^2} + 2G^j\left(x^k, \frac{dx^k}{d\sigma}\right)\right]. \tag{3.5}$$

Clearly, in this form the differential equations of the extremals are invariant under parameter transformations.

Alternatively, equations (3.5) remain unchanged if we replace the functions G^i by new functions G^{*i}, the latter being defined by

$$G^{*i}\left(x^k, \frac{dx^k}{d\sigma}\right) = G^i\left(x^k, \frac{dx^k}{d\sigma}\right) - P\left(x^k, \frac{dx^k}{d\sigma}\right)\frac{dx^i}{d\sigma}, \tag{3.6}$$

where $P(x^k, dx^k)$ is an arbitrary scalar function positively homogeneous of the first degree in dx^k. In fact, (3.6) represents the most general modification of the functions G^i which leaves the extremals (3.5) unchanged. The transition (3.6) is called a *projective change.*

We shall now show that *any transition of the type* (1.2) *may be interpreted as being merely a projective change.* This follows from the following considerations:

Corresponding to (3.2) we define a new parameter σ by means of (1.2) by putting

$$d\sigma = F^*(x^j, dx^j) = F(x^j, dx^j) + S_{x^i}(x^j)dx^i. \tag{3.7}$$

---

[7]) See [4], p. 17.

In view of (3.2) we then have

$$\frac{d\sigma}{ds} = 1 + S_{x^i}\frac{dx^i}{ds} = 1 + \frac{dS}{ds}; \tag{3.8}$$

and

$$\frac{d^2\sigma}{ds^2} = \frac{d^2S}{ds^2}. \tag{3.9}$$

If we now replace s by σ in the equations of the extremals (3.1), we find, according to (3.4) in terms of (3.8) and (3.9) that these equations assume the form

$$\frac{d^2x^i}{d\sigma^2} + 2G^i\left(x^j, \frac{dx^j}{d\sigma}\right) + \frac{dx^i}{d\sigma}\frac{d^2S}{ds^2}\left[1 + \frac{dS}{ds}\right]^{-2} = 0. \tag{3.10}$$

But since the extremals of F and F^* coincide, these equations also represent the extremals of F^*: hence, in virtue of the significance of the special parameter σ defined by (3.7), equations (3.10) are equivalent to

$$\frac{d^2x^i}{d\sigma^2} + 2G^{*i}\left(x^j, \frac{dx^j}{d\sigma}\right) = 0, \tag{3.11}$$

where the G^{*i} result from F^* according to the analogue of equation (3.3). From (3.10) and (3.11) we then deduce that G^i and G^{*i} are related by a projective change such as (3.6), the function $P(x^k, dx^k/ds)$ being given by

$$2\,P\left(x^k, \frac{dx^k}{d\sigma}\right)\frac{d\sigma}{ds} = 2\,P\left(x^k, \frac{dx^k}{ds}\right) = -\frac{d^2S}{ds^2}\left[1 + \frac{dS}{ds}\right]^{-1}. \tag{3.12}$$

Conversely, suppose that we are given a congruence of extremals which cover a region G of R_n simply, the parameter s being measured along each extremal of the congruence from some initially fixed hypersurface Σ of R_n. Let $P(x^k, dx^k/ds)$ be a given function homogeneous of the first degree in the dx^k/ds, giving rise to a projective change of the type (3.6). *We assert that such a projective change always determines suitable functions* $S(x^i)$, such that a transition (1.2) with a corresponding parameter σ can be found, which ensures that after the projective change the congruence of extremals satisfies the differential equations (3.11).

In fact, if $Q(x^i)$ is any point of G, a unique extremal Γ of the congruence passing through Q is determined. Let Γ intersect Σ in the point $\bar{Q}(\bar{x}^i)$. Let us write

$$\lambda(P) \equiv \lambda(x^i) = -2\int_{\Gamma\bar{Q}}^{Q} P\left(x^k, \frac{dx^k}{ds}\right) ds. \tag{3.13}$$

Integration of (3.12) then gives

$$\frac{dS(x^i)}{ds} = C_1\, e^{\lambda(x^i)} - 1, \tag{3.14}$$

where C_1 is an arbitrary constant, and further

$$S(x^i) = \int_{\Gamma\bar{Q}}^{Q} (C_1\, e^{\lambda(x^i)} - 1)\, ds + C_2. \tag{3.15}$$

This is the required function, for if we substitute (3.14) in (1.2) a function $F^*(x^k, dx^k)$ is obtained, the corresponding parameter σ satisfying the differential equation

$$\frac{d^2\sigma}{ds^2} = -2P\left(x^k, \frac{dx^k}{ds}\right)\frac{d\sigma}{ds} = -2P\left(x^k, \frac{dx^k}{d\sigma}\right)\left(\frac{d\sigma}{ds}\right)^2 \tag{3.16}$$

in view of (3.13) and (3.14). The equation (3.11) then follows directly from (3.16), (3.4) and (3.6).

We remark that in this construction it is necessary that we should confine our attention to a congruence of extremals, for the function (3.15) will, in general, depend on the choice of this congruence.

In conclusion we note the following consequence of our analysis. Suppose that we are given a set of differential equations of the type (3.1), defined by given functions $G^i(x^k, dx^k/ds)$, positively homogeneous of the second degree in dx^k/ds. In general such a set need not represent the extremals of some problem in the calculus of variations. From the above results it follows that a mere projective change cannot affect this aspect of the system (3.1). More precisely: *A given set of equations* (3.1) *which initially does not represent extremals cannot be transformed into a set of differential equations of extremals by means of a projective change.*

This result is of some significance with regard to the geometry of paths, for in the literature on this subject occasional attempts have been made to change paths into geodesics by means of projective transformations.

Department of Mathematics,
University of Natal, South Africa

LITERATURE

1. Carathéodory, C., Über die starken Maxima und Minima bei einfachen Integralen. Math. Ann. **62**, 449–503 (1906).
2. ———, Variationsrechnung und partielle Differentialgleichungen erster Ordnung. Leipzig und Berlin (1935).
3. Damköhler, W. und E. Hopf, Über einige Eigenschaften von Kurvenintegralen und über die Äquivalenz von indefiniten mit definiten Variationsproblemen. Math. Ann. **120**, 12–20 (1947).
4. Cartan, E., Les espaces de Finsler. Actualités **79**, Paris (1934).