"Über die Enveloppen der Extremalen eines Feldes in mehrdimensionalen Räume," Bull. Soc. Math. Hellen. **4** (1923), 23-31.

## On the envelopes of the extremals of a field in a multidimensional space

By C. CARATHÉODORY (in Athens)

Translated by D. H. Delphenich

**1.** – If a variational problem that belongs to the function:

(1) 
$$f(x_i, \dot{x}_i, t)$$
  $(i = 1, 2, ..., n)$ 

is given in an (n + 1)-dimensional space whose coordinates might be denoted by  $x_1, ..., x_n, t$  then it is known that not every *n*-parameter family of extremals will define a field, but certain conditions must be fulfilled that have as a consequence the fact that the extremals of the field must intersect a one-parameter family of *n*-dimensional surfaces transversally, as they must from the definition of a field.

From a theorem that includes the generalization of a famous discovery by *Malus* in geometrical optics (1808), and for that reason can also be given the name of that author, in order for that condition to be fulfilled in all of space, it is sufficient that they should be fulfilled on an *n*-dimensional surface in  $R_{n+1}$  that intersects our *n*-parameter family of extremals, but *nowhere contacts it*.

In what follows, we will exhibit the condition that must be fulfilled on an *n*-dimensional surface  $\Phi(x_i, t) = 0$  that contacts an extremal of our family at each of its points in order for those extremals to define a field. The theorem that we have in mind is already known in one case:

Namely, in the theory of surfaces, one shows that a two-parameter family of lines can define the normal congruence to a family of surfaces if and only if it contacts one, and therefore each, of its two focal surfaces along a family of geodetic curves  $(^1)$ .

We will now see that in a completely-analogous way, the extremals of our family will define a field if and only if their envelope on the surface  $\Phi = 0$  represents a field of extremals of the variational problem that arises from our original variational problem by adjoining the auxiliary condition that  $\Phi = 0$ .

<sup>(&</sup>lt;sup>1</sup>) **Darboux**, *Théorie des Surfaces*, § 441 (t. II, pp. 263).

2. – As geometrically trivial as that theorem might seem in the calculus of variations, its proof will still lead to some quite non-obvious calculations when one does not go to work with sufficient care.

We will appeal to an algorithm that I have described in a chapter on the calculus of variations in the soon-to-appear new edition of Riemann-Weber's *Partiellen Differentialgleichungen der Physik* (Braunschweig, Vieweg) that was edited by von Mises, and for that reason, I will treat it only briefly here.

3. – We first introduce canonical variables into our variational problem by the equations:

(2) 
$$y_i = f_{\dot{x}_i}, \qquad H(x_i, y_i, t) = -f + \sum_{i=1}^n \dot{x}_i y_i.$$

An arbitrary family of curves that simply cover a region in our  $R_{n+1}$  can be represented most simply by a system of differential equations:

(3) 
$$\dot{x}_i = \varphi_i(x_j, t)$$
  $(i, j = 1, ..., n)$ .

We now introduce the functions of position:

(4) 
$$(i) = f_{x_i}(x_j, \varphi_j, t), \qquad (t) = -H(x_i, (i), t)$$

$$(i, j = 1, 2, ..., n),$$

which can serve to define our family of curves in their own right, just like the functions  $\varphi_i$ . For that reason, they shall be called the *canonical coordinates* of the family of curves.

One can express the condition for our family of curves to define a field of extremals for our problem by saying that one postulates the existence of a function  $S(x_i, t)$  that satisfies the equations:

(5) 
$$S_{x_i} = (i), \qquad S_t = (t) \qquad (i = 1, 2, ..., n).$$

Namely, it follows from those equations, first of all, that the function S satisfies the Jacobi-Hamilton partial differential equation:

$$S_t + H(x_i, S_{x_i}, t) = 0,$$

and secondly, that the family of surfaces S = const. is intersected transversally by our curves.

If one then introduces the notations:

(6) 
$$[ij] = \frac{\partial(i)}{\partial x_j} - \frac{\partial(j)}{\partial x_i} \quad \text{and} \quad [it] = -[ti] = \frac{\partial(i)}{\partial t} - \frac{\partial(t)}{\partial x_i}$$

then equations (5) will show that *our family of curves defines a field of extremals if and only if the square brackets* [*i j*] and [*i t*] *mean functions that all vanish identically.* 

Upon differentiating and considering the equations above, one will further find the identity:

$$\dot{y}_i + H_{x_i} = \sum_{j=1}^n [i j] \dot{x}_j + [it]$$
,

which is fulfilled by any arbitrary family of curves. It follows from this that the necessary and sufficient condition for our family of curves to consist of nothing but extremals can be expressed by the equations:

(7) 
$$\sum_{j=1}^{n} [i j] \dot{x}_{j} + [i t] = 0 \qquad (i = 1, 2, ..., n).$$

The  $\dot{x}_i$  in it are calculated from equations (3). One can also calculate  $\dot{x}_i$  with the help of the functions (*i*), since it follows with the help of (2), (3), and (4) that:

$$\dot{x}_i = H_{y_i}(x_j, (i), t)$$
.

4. – Naturally, as long as n > 1, equations (7) can all be fulfilled without all of the square brackets vanishing. The fact that not every family of extremals defines a field that was mentioned in the introduction will follows from that.

Now, in order to prove Malus's theorem, we assume that equations (7) all exist and then remark that from the definition (6) of our square brackets, the identity:

(8) 
$$[k t]_t + [t j]_k + [j k]_t = 0$$

will likewise exist. Now, it follows from (7) that:

$$[k t] = \sum_{i} [i k] \dot{x}_{i},$$
$$[t j] = \sum_{i} [j i] \dot{x}_{i},$$

and one will then have:

$$[k t]_t + [t j]_k = \sum_{i=1}^n \left( [j i]_k + [ik]_j \right) \dot{x}_i + [i k] \frac{\partial \dot{x}_i}{\partial x_j} + [j i] \frac{\partial \dot{x}_i}{\partial x_k}$$

$$=\sum_{i=1}^{n} [jk]_{i} \dot{x}_{i} + [ik] \frac{\partial \dot{x}_{i}}{\partial x_{j}} + [ji] \frac{\partial \dot{x}_{i}}{\partial x_{k}}.$$

When we now, on the one hand, employ equation (8), but on the other hand remark that the derivatives of [j k] along an extremal of our family can be written:

$$\frac{d[jk]}{dt} = \sum_{i=1}^{n} [jk]_{i} \dot{x}_{i} + [jk]_{t}$$

then we will finally obtain:

(9) 
$$\frac{d[jk]}{dt} = \sum_{i} [ki] \frac{\partial \dot{x}_{i}}{\partial x_{j}} + [ij] \frac{\partial \dot{x}_{i}}{\partial x_{k}}$$

The n(n-1)/2 square brackets [j k] then satisfy a system of linear homogeneous first-order differential equations along each extremal and must vanish along the entire extremal when they are all zero at one point of that extremal.

If all [j k] vanish on a surface  $\Phi = 0$  that intersects our family of extremals then they must be zero in the entire region of space that is covered by our family of curves, and due to (7), all [j t] will vanish. We then have the theorem:

In order for a family of extremals that intersect a surface  $\Phi(x_1, ..., x_n, t) = 0$  to define a field, it is necessary and sufficient that the n(n-1)/2 square brackets [i j] must vanish on that surface.

5. – The condition that the family of extremals in question should intersect the surface  $\Phi = 0$  is necessary: If that surface were a focal surface of the family of curves then, in general, the extremals would lie on one side of  $\Phi = 0$ , and one could not define all of the [*i j*] on  $\Phi = 0$  at all.

We would now like to reduce the case in which  $\Phi = 0$  is a focal surface of our family of extremals to the previous case in which the family of extremals intersects the surface  $\Phi$  by a trick.

In so doing, we would like to assume that the focal surface  $\Phi = 0$  coincides with the plane  $x_n = 0$ . That will imply no loss of generality since we can always transform any arbitrary focal surface into the plane  $x_n = 0$  by introducing curvilinear coordinates.

From now on, the indices *i*, *j*, etc., shall run from only 1 to (n - 1), and in place of the function (*i*), we would then like to write:

(10) 
$$f(x_1, x_n, t; \dot{x}_1, ..., \dot{x}_n)$$
.

**6.** – Now let an *n*-parameter family of curves be given that contacts the plane  $x_n = 0$ . If one denotes the coordinates of a point of contact with a curve of the family by  $x_i^0$ ,  $t^0$  then one can represent the curves of our family by the following equations:

(11) 
$$t = t^0 + \lambda,$$

(12) 
$$x_i = x_i^0 + \lambda \alpha_i(x_j^0, t^0, \lambda) ,$$

(13) 
$$x_n = \lambda^p \beta(x_j^0, t^0) (1 + \lambda \gamma(x_j^0, t^0, \lambda))$$

$$p \ge 2$$
,  $i, j = 1, 2, ..., (n-1)$ ,  $\beta(x_i^0, t^0) \ne 0$ .

The  $\lambda$  in that means the curve parameter, and we have assumed that the contact has order *p*.

For the sake of simplicity, we assume that the functions *t*,  $\alpha$ ,  $\beta$ ,  $\gamma$  that enter into our equations are all analytic.

If we substitute the values (11) and (12) for *t* and *x<sub>i</sub>*, resp., in the functions  $\beta(x_i, t)$  and consider the fact that  $\beta \neq 0$  then we will get an equation of the form:

(14) 
$$\beta(x_i, t) = \beta(x_i^0, t^0)(1 + \lambda \varepsilon(x_j^0, t^0, \lambda)) .$$

7. – We now introduce a new variable v by the equation:

(15) 
$$x_n = v^p \beta(x_i, t) .$$

With the help of (13), (14), and (15), we will then get:

$$v_p = \lambda^p \frac{(1 + \lambda \gamma(x_i^0, t^0, \lambda))}{(1 + \lambda \varepsilon(x_i^0, t^0, \lambda))} ,$$

and we infer from this upon taking the  $p^{th}$  root that:

(16) 
$$v = \lambda (1 + \lambda \eta (x_j^0, t^0, \lambda)).$$

The system of equations (11), (12), and (16) represents our family of curves in the transformed space of  $x_i$ , t, v. We see that the curves of family *do not contact* the plane v = 0 and are regular at the points where they intersect that plane.

**8.** – We shall denote the transform of the function (10) after the new coordinate *v* is introduced by  $\overline{f}(x_i, v, t, \dot{x}_i, \dot{v})$ . Since it follows from (15) that:

(17) 
$$\begin{cases} \dot{x}_n = p v^{p-1} \beta \dot{v} + \dot{v}^p \beta, \\ \dot{\beta} = \sum_i \beta_{xi} \dot{x}_i + \beta_i, \end{cases}$$

we can write:

(18) 
$$\overline{f} = f(x_i, v^p \beta, t; x_i, (p v^{p-1} \beta \dot{v} + v^p \dot{\beta}))$$

If we let  $\overline{(i)}$ ,  $\overline{(v)}$ , and  $\overline{(t)}$  denote canonical coordinates of our family of curves in the coordinates, as in § 3, and let (*i*), (*n*), (*t*) denote the canonical coordinates of the same family of curves for the original problem then we can infer from (18), while recalling equations (4), that:

(19) 
$$\overline{(i)} = (i) + (n)v^p \beta_{x_i},$$

(20) 
$$\overline{(v)} = (n) p v^{p-1} \beta.$$

Now, in order to calculate  $\overline{(t)}$  in the same way, we remark that when we employ equations (2) and (4), it will follow from  $\overline{f} = f$  that:

$$\overline{(t)} + \sum_{i} \overline{(i)} \dot{x}_{i} + \overline{(v)} \dot{v} = (t) + \sum_{i} (i) \dot{x}_{i} + (n) \dot{x}_{n} .$$

With the help of (17), (19), and (20), we will then find that:

(21) 
$$(t) = (t) + (n)v^{p}\beta_{t}.$$

**9.** – We can now likewise calculate the square brackets that we introduced in § **3** in the new variables. Namely, when we partially differentiate equation (19) with respect to  $x_j$  and consider (15), it will follow that:

(22) 
$$(i)_{j} = (i)_{j} + (i)_{v} v^{p} \beta_{x_{j}} + (n)_{j} v^{p} \beta_{x_{i}} + \cdots$$

We have not written out the terms that are symmetric in i and j in that. If we switch i and j in that last equation and subtract the equation that is obtained in that way from (22) then that will give:

(23) 
$$\overline{[i\,j]} = [i\,j] + v^p \beta_{x_j} + (n)_j v^p \beta_{x_i} + \cdots$$

We obtain the relation:

(23) 
$$[it] = [it] + v^p \beta_{x_i} + ([in]\beta_t + [nt]\beta_x)$$

in a completely-analogous way. In the same way, (19) and (20) [(20) and (21), resp.] imply the equations:

(25) 
$$\overline{[iv]} = p v^{p-1} \beta[in]$$

(26) 
$$\overline{[vt]} = p v^{p-1} \beta[nt] .$$

10. – It follows from equations (23) - (26) then every field of extremals of our original problem will again correspond to a field of extremals of our transformed problem, but it is also immediately clear that families of extremals can likewise be transformed into such things completely independently of whether they do or do not define a field. One can also easily confirm that by calculation when one forms the expression that are analogous to the expressions on the left-hand side of (7). Namely, it will follow from the latter equations when one recalls equations (17) that:

(27) 
$$\sum_{j=1}^{n-1} \overline{[i\,j]} \dot{x}_j + \overline{[i\,v]} \dot{v} + \overline{[i\,t]} = \sum_{i=1}^{n-1} [i\,j] \dot{x}_j + [i\,n] \dot{x}_n + [i\,t] + v^p \beta_{x_i} \left( \sum_{i=1}^{n-1} [n\,j] \dot{x}_j + [n\,t] \right),$$

(28) 
$$\sum_{j=1}^{n-1} \overline{[v\,j]} \dot{x}_j + \overline{[v\,t]} = p \, v^{p-1} \left( \sum_{i=1}^{n-1} [n\,j] \dot{x}_j + [n\,t] \right).$$

**11.** – From the remark at the end of § 7, our functions  $\overline{(i)}$ ,  $\overline{(v)}$ ,  $\overline{(t)}$ , and therefore the functions  $\overline{[i j]}$ ,  $\overline{[i t]}$ ,  $\overline{[i v]}$ , and  $\overline{[v t]}$ , as well, are regular on the plane v = 0, which corresponds to the focal surface  $x_n = 0$  of our family of curves. Equations (25) and (26) show that this does not need to be the case, as opposed to what is true for the functions [i n] and [n t]. However, when we calculate the latter quantities from (25) and (26) and substitute them in (23) and (24) then we will get:

(29) 
$$[ij] = \overline{[ij]} + \frac{v}{p \cdot \beta} (\overline{[iv]} \beta_{x_j} + \overline{[vj]} \beta_{x_j}) ,$$

(30) 
$$[ij] = \overline{[it]} + \frac{v}{p \cdot \beta} (\overline{[iv]} \beta_t + \overline{[vt]} \beta_{x_i}) .$$

We then see that the quantities [i j] and [i t] are regular for not only v = 0, but also the fact that equations  $[i j] = \overline{[i j]}$  and  $[i t] = \overline{[i t]}$  are true on that plane.

12. – Thus, if the family of curves that is represented by equations (11) to (13) defines a field of extremals then the functions [i j], [i t] must also vanish for  $x_n = 0$ .

Conversely, we now assume that our family of curves consists of nothing but extremals and that the equations:

(31) 
$$[ij] = 0$$
,  $[it] = 0$   $(i, j = 1, 2, ..., n-1)$ 

are true for  $x_n = 0$ . It will then follow from equations (29) and (30) that the quantities [i j] and [it] likewise vanish for v = 0. It will further follow from (17) that the left-hand side of that equation is zero and from (16) and (11) that we always have  $\dot{v} = 1$  for v = 0. We then see that for v = 0, we must have [iv] = 0 for those functions. However, from Malus's theorem (§ 4), our family of extremals must then define a field in the transformed problem, and therefore also on the original problem.

We then have the result that:

A family of extremals that possesses the plane  $x_n = 0$  as a focal surface will define a field if and only if equations (31) are all true for  $x_n = 0$ .

13. – We shall now examine the variational problem that arises when we set  $x_n = 0$  in (10) (and naturally also set  $\dot{x}_n = 0$ ), and remark that the canonical coordinates of a line element for the new problem will be the same as the canonical coordinates of the same line element for our original problem when we merely subtract (*n*) from the coordinates. Namely, if we set:

$$\varphi(x_i,t;\dot{x}_i) = f(x_i,0,t;\dot{x}_i,0)$$

then we will not only have that the equation  $\varphi \dot{x}_i = t_{\dot{x}_i}$  is true for every *i*, but also that:

$$\varphi = \sum_{i=1}^{n-1} \varphi_{\dot{x}_i} \dot{x}_i = f - \sum_{i=1}^{n-1} f_{\dot{x}_i} x_i - f_{\dot{x}_n} \dot{x}_n ,$$

since  $\dot{x}_n = 0$  here. It then follows from this that the condition (31) says simply that the envelope of our family of curves (11) – (13) [i.e., the integral curves of the system of differential equations  $\dot{x}_i = \alpha_i (x_j, t, 0)$ ] that is found on the plane  $x_n = 0$  defines a field of extremals of the variational problem that corresponds to the functions  $\varphi(x_i, t, \dot{x}_i)$  on the plane  $x_n = 0$ .

With that, the assertion that was expressed at the end of § 1 is proved.

14. – The theorem that was just proved gives a means for constructing a field of extremals in the space of  $x_1, ..., x_n$ , t that has the given surface  $\Phi = 0$  for its focal surface. One must next

construct an arbitrary field  $\Gamma$  of extremals of the variational problem that arises from the original problem by adjoining the auxiliary condition  $\Phi = 0$  on that surface  $\Phi = 0$  and then consider the extremals in  $R_{n+1}$  that contact the curves of  $\Gamma$ .